

NF-E2 AND USF COOPERATE TO REGULATE THE RECRUITMENT AND ACTIVITY OF
RNA POLYMERASE II IN THE BETA-GLOBIN LOCUS

By

ZHUO ZHOU

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2010

© 2010 Zhuo Zhou

To my parents, grandparents, family and friends who always believe in me.

ACKNOWLEDGMENTS

I would first like to thank my mentor, Dr. Jörg Bungert, for the wonderful experience of working in his lab for my Ph.D. training. Dr. Bungert is a role model of a great scientist. His wisdom, dedication and enthusiasm for science always inspire me. I am grateful that he has been supportive and patient with me and my research in the lab. I really appreciate his long time support, encouragement and patience for me. If there are any achievements I have in these years, I owe them all to Dr. Bungert. He made me from a person naïve in science, to a Ph.D. who has the courage and knowledge to face the challenges in the future.

I also thank all of my committee members, Drs. Michael Kilberg, Thomas Yang, Linda Bloom and Kenneth Berns, for their long time support and inspiration. I am really grateful that they would take time from their busy schedule, participate in all of my committee meetings and give me precious suggestions on my projects. Their insightful suggestions and questions have broadened my mind on science. I would also like to thank my previous committee member, Dr. Brian Burke. He helped me to develop great interest in the cell nucleus related researches and provided me precious advice on my research. Dr. Burke moved to Singapore with his family this September.

I must thank the department chair, Dr. James Flanagan and all the administrative and secretarial staff in our department. They work very hard to make us live easy in this department. I really appreciate the advice from Dr. Flanagan on my first Biochemistry Journal Club presentation. I am also very thankful for all the assistance kindly provided to me over the years by Pat Jones, Elise Feagle, Regina Corns, Miriam Williams, Denise Mesa and Mr. Bradley Moore and Terry Rickey.

I also thank all of the past and present members of the Bungert lab. I am grateful that Drs. Karen Vieira, Padraic Levings, Valerie Crusselle-Davis provided me great help when I first

joined the lab. I thank all the other members, Dr. Felicie Andersen, Dr. Boris Thurisch, Shermi Liang, I-Ju Lin, Tihomir Dodev, Joeva Barrow, Michael Rosenberg, Babak Moghimi, Archana Anantharaman and Dorjan Pantic for their friendships and for all of the serious and fun conversations we had over the years. I wish them all the best with their school and research. I would like to thank Shermi Liang for helping with the orders and making sure things we ordered arrived on time. I would especially thank Joeva Barrow for her extreme patience to help me improving my English for the upcoming interviews. I hope all the current lab members keep having fun working in the lab.

I would like to thank my parents, Wenxue Zhou and Haiyan Zhang, for their long time encouragement and support to lead me in every step of my life. I am grateful that even when they found out I chose to go to Peking University instead of the number 1 medical school in China as they had been expected since I was little, they still respected my decision and gave me all their blessing. I would also like to thank my maternal and parental grandparents, Jin Zhang, Dongshi Zhou and Shuxian Zhang. They taught me to be a good person. I also thank all of my family members, who not only encourage me to peruse a career in science, but also help me to take care of my parents so I can feel free to study in a country 8000 miles away from home.

At last but not least, I would like to thank my friends Dr. Nan Su, Dr. Xiaolei Qiu, Dr. Houda Darwiche, Dr. Hong Li, Gang Liu, Tong Lin, Justin Bickford, Dr. Xingguo Li and Dr. Patrick Corsino. They made my life in graduate school much easier by selflessly offering me help at any time. I thank Courtney Bouldin for help me with my first cover letter for a post-doctoral position application. I especially thank Justin Bickford for helping me editing my thesis within such a short time. Also, I would like to thank Stephen Langevin, Quan Yuan and Qiu Lin for being my great friends.

I am grateful that I have met all of these people above. With all the guidance, support and help from them, I get to become me, today.

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS.....	4
LIST OF FIGURES.....	9
ABSTRACT.....	11
CHAPTER	
1 INTRODUCTION.....	13
Hemoglobin and β -globin Related Diseases.....	13
Sickle Cell Anemia and β -Thalassemia.....	13
Treatments for β -Globin Disorders.....	16
The β -Globin Gene Locus.....	20
Overview.....	20
The Locus Control Region (LCR).....	21
The Adult β -Globin Gene Promoter and 3' Enhancer.....	25
Hemoglobin Switching.....	27
Gene Competition.....	28
Autonomous Silencing.....	28
Switching Factors.....	30
The LCR and the β -Globin Gene Promoter Interaction Models.....	33
The “Active Chromatin Hub” Model for LCR Regulating the β -Globin Gene.....	33
Proteins Involved in the Regulation and Expression of Adult β -Globin Gene Expression.....	35
RNA Polymerase II (RNA Pol II).....	35
GATA-1.....	37
NF-E2.....	39
EKLF.....	41
USF.....	42
TFII-I.....	44
Components of Chromatin Remodeling Complexes.....	47
CTCF.....	47
The Role of the LCR in Nuclear Localization.....	48
Summation.....	48
2 MATERIALS AND METHODS.....	52
Cell Culture and Protein Isolation.....	52
<i>In vitro</i> Embryonic Stem Cell Differentiation.....	53
Western Blotting.....	54
Chromatin Immunoprecipitation (ChIP).....	54
Co-immunoprecipitation (CoIP) and GST Pull-Down Assay.....	55
RNA Isolation, Reverse Transcription, and Real-Time PCR.....	56
<i>In vitro</i> Transfer/Dissociation Assay.....	57

3	REGULATION OF β -GLOBIN GENE EXPRESSION BY CIS- AND TRANS-REGULATORS.....	59
	Introduction	59
	Results.....	64
	Discussion.....	70
4	NF-E2 AND USF COOPERATE TO REGULATE RNA POLYMERASE II IN THE β -GLOBIN GENE LOCUS	83
	Introduction	83
	Results.....	88
	Discussion.....	97
5	CONCLUSIONS AND FUTURE DIRECTIONS.....	115
	<i>In vitro</i> ES Cell Erythroid Differentiation.....	115
	Overview	115
	The Antagonistic Role of USF and TFII-I in β -Globin Gene Regulation.....	116
	Recruitment of Histone Modification Enzymes to the β -Globin Locus by TFII-I.....	117
	The Cooperative Role of NF-E2 and USF in β -Globin Gene Regulation.....	118
	<i>In Vitro</i> Transfer/Dissociation Assay	118
	The Different Role of USF1 and USF2 in β -Globin Gene Regulation	119
	NF-E2 p45 Knockout Mice vs. NF-E2 p45 Null Murine Erythroid Cells	120
	The Model of the LCR Regulated β -Globin Gene Expression	122
	LIST OF REFERENCES	124
	BIOGRAPHICAL SKETCH	151

LIST OF FIGURES

<u>Figure</u>	<u>page</u>
1-1 Schematic of the structure of the human and mouse β -globin gene loci and hemoglobin synthesis during development	51
3-1 Sequential activation of globin gene transcription during <i>in vitro</i> erythroid differentiation of murine embryonic stem cells	76
3-2 Interaction of transcription factors and RNA polymerase II with the β -globin locus during <i>in vitro</i> murine embryonic stem (ES) cells erythroid differentiation	77
3-3 RNA polymerase II (RNA Pol II) is not recruited to the β -globin gene locus in MEF and OP9 cells	78
3-4 Interaction of USF and TFII-I with the β -globin gene locus during erythroid differentiation of murine embryonic stem cells	79
3-5 Interaction of Suz12 with TFII-I in K562 cells.....	80
3-6 The interaction of Suz12 with the β -globin gene promoter decreases with increased β -globin gene expression during erythroid differentiation of mouse embryonic stem cells	81
3-7 TFII-I and Suz12 do not interact with the LCR HS3/2 flanking region	82
4-1 DMSO mediated increase in USF, NF-E2 (p45), and β -globin expression in MEL but not CB3 cells	103
4-2 Lack of NF-E2 (p45) reduces the assembly of protein complexes at LCR HS2 and at the adult β maj-globin gene promoter	106
4-3 Efficient recruitment of Pol II and CTD serine 5 and serine 2 phosphorylation at the β -globin gene locus requires NF-E2	107
4-4 Interactions of USF1 and USF2 with NF-E2 (p45) and Pol II during erythroid differentiation of MEL and CB3 cells	109
4-5 USF and NF-E2 regulate the recruitment and dissociation of RNA Pol II to and from immobilized LCR templates.....	111
4-6 Model of NF-E2 and USF mediated assembly and transfer of elongation competent transcription complexes in the β -globin gene locus.....	112
4-7 Interaction of protein complexes with the LCR HS3/2 flanking region in MEL, CB3 and CB3/NF-E2 cells	113

4-8 Interaction of RNA Pol II with the LCR HS3/2 flanking region or the GAPDH 3' untranslated region (UTR) in MEL, CB3 and CB3/NF-E2 cells 114

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

NF-E2 AND USF COOPERATE TO REGULATE THE RECRUITMENT AND ACTIVITY OF
RNA POLYMERASE II IN THE β -GLOBIN LOCUS

By

Zhuo Zhou

May 2010

Chair: Jörg Bungert

Major: Biochemistry and Molecular Biology

The β -globin gene locus has been extensively studied as a model system for understanding tissue- and developmental stage-specific expression and long-range regulation of genes. In humans, it consists of five β -like globin genes that are expressed exclusively in erythroid cells during development that are under the regulation of a far upstream locus control region (LCR). The LCR contains several DNase I hypersensitive (HS) sites, each of which are 200-400bp in length and separated by 2-4 kbp of flanking DNA. The HS sites contain binding sites for a variety of erythroid-specific and ubiquitously expressed transcription factors, which contribute to the recruitment of multiple transcription regulatory proteins including RNA polymerase II (RNA Pol II) and histone modification enzymes to the LCR through protein-DNA and protein-protein interactions.

By using an *in vitro* assay to differentiate mouse embryonic stem (ES) cells into erythroid lineage, we show that RNA Pol II is recruited to the LCR prior to the erythroid differentiation. In addition, we show that the helix-loop-helix proteins USF and TFII-I play antagonistic roles with USF as an activator and TFII-I as a repressor in β -globin gene regulation. I further show that the repressor function of TFII-I is maintained by recruiting polycomb repressive complex 2 (PRC2) to the locus.

Next, by using an *in vitro* transfer/dissociation assay, I found that, while USF is required for RNA Pol II association to the LCR and NF-E2 is essential to transfer RNA Pol II from the LCR to the β -globin gene promoter, the sufficient transfer of RNA Pol II requires a specific USF binding site at the promoter. Together with the findings that NF-E2, USF and RNA Pol II form a complex in differentiated mouse erythroid cells and NF-E2 directly interacts with USF1, these data suggest that USF may stabilize the binding of NF-E2 at the β -globin promoter to continuously bring RNA Pol II from the LCR to the gene promoter.

Interestingly, I also observed that NF-E2 is required for the recruitment of the CTD serine 5-phosphorylated form of RNA Pol II to the β -globin locus to initiate transcription. The fact that the serine 5-phosphorylated RNA Pol II is first recruited to the LCR but not the β -globin gene promoter at early erythroid differentiation suggests that the LCR initiates the recruitment of a transcriptional competitive protein complex prior to the transcription of downstream β -globin gene.

In summary, in this thesis, I present evidence showing that the β -globin gene is highly regulated by the cooperation of erythroid-specific and ubiquitously expressed transcription factors and their interaction with the proximal (β -globin gene promoter) and distal (LCR) DNA regulatory elements.

CHAPTER 1 INTRODUCTION

Hemoglobin and β -globin Related Diseases

Hemoglobin (1) (Hb), the iron-containing metalloprotein presenting in large quantities in erythrocytes, is the protein responsible for transport of oxygen and carbon dioxide throughout the body in mammals and birds. Functional hemoglobin normally exists as a tetramer, containing two α -like globin and two β -like globin chains ($\alpha_2\beta_2$). Each of the globin chains has an embedded ring-shaped heme group containing a central iron atom, which is responsible for reversible oxygen binding.

Hemoglobin synthesis requires the coordinated production of the two types of globin chains (2) (Figure 1-1B). The ξ -globin from the α -globin locus pairs with ϵ -globin from the β -globin locus during the first six weeks of gestation in primitive erythroid cells that originate in the embryonic yolk sac to form the embryonic hemoglobin (Hb ϵ , $\xi_2\epsilon_2$). After the first ten weeks of embryogenesis, the α -globin chains remain unaltered, while the β -globin chains undergo “Hemoglobin Switching” during development. During fetal liver hematopoiesis, α -globin from the α -globin gene locus pairs with one of two γ -chains, G γ and A γ , from the β -globin locus to form the fetal hemoglobin (HbF, $\alpha_2\gamma_2$). Later during adult bone marrow hematopoiesis, α -globin pairs with the adult δ - and β -globin to form the adult hemoglobin (HbA, $\alpha_2\beta_2$). The expression of the α - and β -like globin genes is closely balanced by unknown mechanisms. Balanced globin gene expression is required for normal red blood cell function.

Sickle Cell Anemia and β -Thalassemia

Currently, $\sim 7\%$ of the world’s population are carriers for different inherited disorders of hemoglobin, making them the commonest human monogenic diseases (2). The worldwide birth rate of people with symptomatic globin disorders is no less than 2.4 per 1000 births (3). Because

of the protective effects of the hemoglobin disorders against malaria (4), these β -globin disorders occur mainly in people from tropical or sub-tropical regions where malaria is or was common, such as Africa, Mediterranean countries, and India, etc. (5).

There are two main types of human beta-globin disorders: hemoglobin structural variants, and thalassemias. There is a third type of disorders called hereditary persistence of fetal hemoglobin (HPFH), displaying heterogeneous defects in the normal switch from fetal to adult hemoglobin production, although it has no clinical importance by itself, the co-inheritance of some forms of HPFH can modify the phenotypes associated with the sickle cell disease or thalassemias.

Hemoglobin structural variants are mainly due to single amino acid substitutions. Currently, over 700 structural variants have been identified. Among them, the most well known example is the sickle cell anemia, which is commonly caused by the hemoglobin variant, HbS (6). In this variant, the hydrophilic amino acid glutamic acid at position 6 of the hemoglobin beta-chain is replaced by the hydrophobic amino acid valine. This mutation creates a hydrophobic pocket on the outside of the hemoglobin structure, which attaches to the hydrophobic region of the adjacent hemoglobin beta-chain, thus causing aggregation of hemoglobin into rigid fibers and leading to deformation of red blood cells, which adopt a sickle like configuration when hemoglobin is deoxygenated. The decrease of the elasticity of the sickle shaped red blood cells causes the blockage of the small blood vessels. This blockage causes severe pain and organ damage and eventually causes severe long term symptoms such as stroke or bone necrosis and sometimes sudden death. The hemoglobin HbS also causes erythroid hyperplasia and accelerated red blood cell destruction (7).

Thalassemias are caused by defective and imbalanced globin gene expression. Over 200 mutations in the β -globin gene have been identified in β -thalassemia patients. Except for a few deletions, these patients normally carry point mutations and loss of one or two bases in their β -globin gene, which reduces or abolishes β -globin gene function at transcriptional, translational, or post-translational levels (8). Due to the absence or inadequate β -globin chain production, the excessive unbound insoluble α -globin chains precipitate in mature red blood cells and erythroid precursors which lead to ineffective erythropoiesis and hemolysis (2,5,7,9). The red blood cells of thalassemia patients are pale and in tear drop shape, and they don't have enough functional hemoglobin to deliver oxygen to the body tissues.

The β -thalassemias have diverse clinical phenotypes, which are divided into three categories: the β -thalassemia major, intermedia, and minor. The most severe one is the β -thalassemia major (also known as Cooley's anemia) (10), in which the patients carry homozygous or compound heterozygous β -globin mutations that lead to severe anemia resulting in death in the first year without blood transfusions. Affected individuals show growth retardation and may have enlarged spleens, livers, and hearts, and severe bone deformities (11), due to the expansion of hematopoietic sites (bone marrow and spleen), which attempt to compensate for the innate anemia. The β -thalassemia intermedia is less severe than β -thalassemia major. The syndromes appear early in the childhood or later in life. The patients have mild to moderate anemia, but they can survive without blood transfusions. Afflicted individuals may still have slow growth and bone deformities but can manage a normal life, with the assistance of occasional transfusions. The β -thalassemia minor (also known as β -thalassemia trait) patients only bear mutations in one of the β -globin alleles. These mutations result in mild or no anemia

and reduced blood cell counts. The patients have an increase level in hemoglobin HbA₂ ($\alpha_2\delta_2$), and a corresponding decrease in hemoglobin HbA ($\alpha_2\beta_2$) (8).

Treatments for β -Globin Disorders

Currently, the standard treatments for β -globin disorders are blood transfusions for β -thalassemia and supportive care (mainly medication) for sickle cell anemia patients.

Blood transfusion is life-saving and meant to correct anemia and to suppress massive erythropoiesis (2,7,9). However, it is a life-long treatment, and will cause iron overload in patients, which is lethal if untreated. Ferritin is the key iron storage protein in the body (12,13). When there is excessive iron in the body that exceeds the capacity and detoxification activity of ferritin, “free” iron starts to accumulate in the blood and tissues. The “free” iron can catalyze the formation of hydroxyl radicals (OH^\cdot), which are highly reactive compounds that attack lipids, proteins, and DNA (13,14). Hepatocytes and cardiac cells are damaged by iron overload, which eventually can be deadly (15,16). Thus, chelation therapy with iron-binding chelators is necessary to remove excess iron from the body, in order to prevent death from iron mediated organ defects (17). Currently, the most widely used iron chelator is desferrioxamine (18). It preferentially binds to iron at a one-to-one ratio. Desferrioxamine significantly prolongs the life of blood transfusion-dependent patients, such as those with β -thalassemia major. However, this drug can only be given parentally and it is expensive and requires a sophisticated medical device for administration (19). Therefore, the prevention and treatment of iron overload are the main challenges of blood transfusion therapy.

Allogeneic bone marrow transplantation is the only current curative option of the β -globin disorders, but for most of the patients, it is hard to identify a human leukocyte antigen (HLA)-matched bone marrow donor (20,21).

An alternative treatment to bone marrow transplantation is represented by β -globin gene therapy, which is to transduce autologous hematopoietic stem cells (HSCs) with highly regulated therapeutic β - or γ -globin gene constructs, and then to introduce these HSCs back into the patients by autotransplantation. This treatment not only provides possible cures for patients who lack matched donors, but also avoids the graft-versus-host diseases (22). The ultimate goal of globin gene transfer is to generate tightly regulated, high-level, erythroid-specific and sustained transgene expression. Retroviral vectors were the first viral vectors used to efficiently transfer the human β -globin gene to the mouse HSCs (23,24). Incorporating a minimal 1 kb fragment of the human β -globin LCR including the core elements of HS2, 3 and 4 into a retrovirus based vector expressing the human β -globin gene increased the expression levels in mouse erythroleukemia cells, but the positional variability of expression was a problem in these studies (25). Also, incorporating larger LCR fragments into retroviral vectors is hard to achieve because the instability of retroviral vectors and frequent genomic rearrangements. Lentiviral vectors have been intensively developed to transfer the human β -globin gene to HSCs. The advantages of the lentiviral vectors are their ability to transduce nondividing cells and the relatively stable virus genome (26). The former one is of great importance, because most of the human HSCs are nondividing. In 2000, May et al. (26) first used a lentiviral TNS9 vector to successfully treat β -thalassemia intermedia mice with a therapeutic human β -globin gene. The TNS9 vector contained a modified human β -globin gene, flanked by an extended promoter sequence and β -globin 3' enhancer element, as well as a 3.2 kb LCR fragment spanning HS2, 3 and 4. This TNS9 vector transduction system maintained high-level and erythroid-specific human β -globin gene expression in mice for up to 40 weeks. The safe use of lentiviral vectors relies on developing robust packaging systems, incorporating tissue-restricted promoters and enhancers,

as well as including other regulatory elements such as insulators into the vectors, to abolish the negative effects of random insertions and inappropriate activation of nearby genes but at the same time promote tissue-specific expression of the β -globin gene (27,28). The cure of β -globin disorders in mice models by using lentiviral vectors based globin gene transfer techniques (26,29-31) encourages the start of human clinical trials. A phase I trial has begun for sickle cell anemia and β -thalassemia patients by using a lentiglobin vector (lentiviral vector for globin gene therapy) to transduce the HSCs of the patients and then returning them back to the patients (32). The lentiglobin vector used contains a human β -globin gene with a single amino acid change at the position 87 (β 87). β 87 functions normally and has been shown to have anti-sickling activity (29). There is no data on efficiency yet.

In sickle cell anemia, the vector-coded β -globin chain has to compete against the endogenous mutated β -globin chain for the binding to the α -globin chains. Therefore, it is of great interest to reactivate γ -globin chain, which has a higher affinity for the α -globin chain than the β -globin chain, to treat sickle cell anemia. In fact, the HFPF patients have much less severe syndromes of β -globin gene disorders, presumably due to compensatory high-level fetal γ -globin expression in adults (33).

In the treatment of hemoglobinopathies, there are currently two mechanisms to induce fetal hemoglobin synthesis. The first mechanism is to indirectly activate fetal γ -globin gene expression through manipulation of the kinetics of erythropoiesis by using cell-cycle S-stage-specific cytotoxic drugs (1). The S-stage-specific drugs prevent γ -globin gene silencing by accelerating the rate of erythroid differentiation and maturation. Among all of these drugs, hydroxyurea (34-36) is used clinically to treat sickle cell anemia patients because it is administrated orally, well tolerated, and has low risk of carcinogenicity. Hydroxyurea treatment

reduces the mortality of adults with severe sickle cell anemia and improves the well-being of the patients by boosting a 15-20% increase in fetal hemoglobin production (37). However, the effects of this drug are transient, not all patients respond to this treatment, and there are also significant variations in the levels of fetal hemoglobin synthesis in the patients. In addition, since hydroxyurea is cytotoxic, it may affect the growth and development of fetus and neonates. The effects of hydroxyurea on β -thalassemia patients are currently controversial. The second mechanism of γ -globin gene induction is mediated by drugs that directly reactivate γ -globin gene expression by modulating the DNA demethylation and histone acetylation at regulatory elements of the β -globin gene locus. 5-Azacytidine (38,39) is a potent inducer of fetal γ -globin gene expression firstly. It demethylates DNA of the β -globin gene locus by inhibiting DNA methyltransferase activity, and also induces rapid erythroid regeneration (39). However, due to its toxicity and potential carcinogenicity, it is limited to treat severe homozygous β -thalassemia (40). Butyrate (41), a four carbon fatty acid, is another inducer of fetal hemoglobin production in patients (42). It functions as an inhibitor of histone deacetylase (HDAC) (43) and ultimately weakens the binding of the histone tail regions to linker and nucleosomal DNA, thus allowing transcription factor access to DNA. However, its exact clinical effectiveness still needs to be determined. Currently, the therapies of reactivating fetal hemoglobin in β -globin disorders are still under investigation. There is a great need for alternative strategies for the treatment of hemoglobin disorders.

Presently, the most promising curative method for β -globin disorders is to either express a normal β -globin gene or reactivate γ -globin gene in red blood cells. This can be achieved by transducing the HSCs or embryonic stem (ES) cells of patients with globin gene containing lentiviral vectors. Therefore, understanding the mechanisms of β -globin gene regulation during

development is necessary for designing precisely regulated efficient globin gene transfer systems.

The β -Globin Gene Locus

Overview

The β -globin genes are among the most studied families of tissue-specific and developmentally regulated genes. The 80 kb β -globin locus embedded within one of many olfactory receptor gene arrays on human chromosome 11p15.4 has been extensively studied over the years as a model system for the developmental and tissue specific regulation of gene expression, molecular mechanisms of enhancer-promoter interactions, properties of insulator, and as a target for gene therapy as a cure for various hemoglobinopathies (1).

The human β -globin genes are expressed exclusively in erythroid cells. Its gene cluster consists of five genes (5'- ϵ -G γ -A γ - δ - β -3' globin genes) arranged in the same order in which they are expressed during development (1) (Figure 1-1A). The ϵ -globin gene is expressed in the embryonic yolk sac during the early stages of embryonic development, up to about week 10 of gestation. The two γ -globin genes, G γ and A γ , are then expressed during fetal liver hematopoiesis. Finally, expression of δ - and β -globin genes becomes predominant after birth and into adulthood during bone marrow hematopoiesis. The δ -globin gene expression level is less than 5% of the β -globin gene, likely due to a mutation of its TATA-box.

The mouse β -globin gene cluster, located on chromosome 7, has a very similar structure compared to the human β -globin locus (5'- $\epsilon\gamma$ - β h1- β maj- β min-3' globin genes) (Figure 1-1A). The $\epsilon\gamma$ - and β h1-globin genes are expressed in the embryonic yolk sac, and the two adult genes, β maj and β min, are expressed predominantly after birth in the fetal liver and adult bone marrow. When aligning globin loci from all of the known mammalian globin gene clusters, it seems that

the fetal expression of the two γ -globin genes in human coincides roughly with the duplication of the genes in primate evolution (44).

The Locus Control Region (LCR)

High-level developmental-stage specific expression of the β -like globin genes is regulated by cis-acting DNA elements located both proximal and far away from the genes. The naturally occurring human Hispanic ($\gamma\delta\beta$)-thalassemia revealed that the locus control region (LCR) is an important regulatory element for β -globin gene expression (45). These thalassemia patients lack mutations within the β -globin gene. However, they bare a 35 kb deletion upstream of HS1 known as the Hispanic-deletion, which suggested that this region contained cis-acting regulatory elements required for proper β -globin gene expression.

The LCR is located at about 50kb upstream of the human adult β -globin gene and contains four erythroid specific DNaseI hypersensitive sites (HS1-4) and a further upstream site (HS5) that is ubiquitously and constitutively hypersensitive and may demarcate the functional boundary of the cluster. Each HS site has a core sequence, about 200-400bp in length, and they are separated from each other by 2-4 kbp flanking DNA (44). In the human genome, there also exists one 3'HS1 site downstream of the adult β -globin gene (46). HS sites are often associated with enhancer function. The LCR HS sites consist of binding sites for multiple ubiquitously expressed or erythroid specific transcription factors such as GATA-1 (47,48), NF-E2 (49), USF (50), EKLF (51), Tal-1 (50), Bach1 (52) etc.. The binding motifs shared by all five 5'HS sites include GATA-1 and NF-E2. Also, the LCR HS sites are known to associate with RNA Pol II (47,53), components of chromatin remodeling complexes and chromatin modification enzymes (such as CBP and components of SWI/SNF and NuRD complexes) (54), and components of the basal transcription machinery (such as TFIIB and TBP (55)). These LCR-associated transcription

regulatory proteins appear to function in a synergistic manner to regulate β -globin gene expression (56).

Studies using transgenic mice carrying the human β -globin locus have shown that in the absence of the LCR, human γ - and β - globin expression levels are low relative to those of the endogenous mouse globin genes (57-59). However, when the LCR (containing HS1-HS5) is linked to the human β -globin gene, all resulting transgenic mice lines express the gene at a levels comparable to the endogenous mouse β -globin gene (60). These results demonstrate that the LCR exerts strong domain-opening activity and confers high-level globin gene expression in a position-independent manner. Experiments in which the five β -globin genes or the LCR were inverted in transgenic mice harboring a yeast artificial chromosome (YAC) containing the full β -globin cluster demonstrated that gene order is important for correct developmental expression and that the LCR functions in an orientation dependent manner (61,62). Minimal LCR DNA sequences that confer position-independent expression of a linked β -globin gene in transgenic mice have been determined to contain HS1-HS4 (63). The inversion of the LCR reduces expression of the linked β -globin genes by more than 70%. The effects observed in this study might be due to human HS5's insulator function. Chromatin insulators are elements that can block enhancer action and/or act as barriers to the inappropriate spread of heterochromatin. They demarcate domains in which an enhancer and the gene it regulates are separated from other expression domains or heterochromatin. The inversion of LCR HS sites positions human HS5 between the LCR and the β - globin genes, possibly insulating LCR's regulatory functions on β -globin gene expression. Such insulating activity, characterized by enhancer blocking and heterochromatin suppressing effects, has been clearly demonstrated in the chicken β -globin cluster for the homologous site (cHS4) (64-66). However, in contrast to the chicken globin gene

locus, HS5 does not demarcate the transition from a DNase I sensitive to DNase I insensitive chromatin configuration. HS5 therefore does not function as a boundary element.

The LCR HS sites have specific functions. The HS2 and HS3 are considered to be the most powerful elements of the LCR. Deletion of HS2 in transgenic mice carrying β YAC severely affected human β -like globin gene expression and formation of HS sites at all the developmental stages (67). HS3 functions as an activator for ϵ - and γ -globin expression. Deletion of HS3 from β YAC transgenic mice led to a significant decrease in ϵ -globin expression at the embryonic stage, while the γ -globin failed to express during the definitive erythropoiesis (68). Neither HS2 nor HS3 deletions affected normal globin gene switching during development. The deletion effects of HS1 and HS4 were less severe. The LCR enhancer activity mainly resides within HS2, 3, and 4 (44,46,69). Although HS1 does not contribute to the LCR enhancer activity, deletion of 5'HS1 in β YAC transgenic mice reduced the expression of ϵ - and γ -globin genes at their specific developmental stages (70). Deletion of HS4 in β YAC transgenic mice disrupted β -globin gene expression at all of the erythroid developmental stages (71), and decreased γ - and β -globin gene expression during the definite erythropoiesis in fetal liver and adult blood (72). HS5 functions as an insulator as described above.

Many studies addressed the contribution of individual HS sites to the LCR function. Although conflicting results exist, which are likely due to different experimental systems (57,73), all studies show that individual or combinations of HS fragments are never able to reach the level of activity of the intact LCR. Even though some experiments suggested interactions between specific LCR HS sites and globin genes (74), the current model suggests that the LCR acts as a holocomplex in which all HS sites interact with each other and that the LCR holocomplex contacts only one globin gene at a time during erythroid development. Deletion of

the LCR from the mouse genome by homologous recombination causes a reduction in β -globin expression to 1-4% of that of the wild type, and there is a major effect on RNA Polymerase II (RNA Pol II) serine 5 phosphorylation and transcription elongation at the β -globin gene, even though the LCR deletion only decreases the assembly of transcription preinitiation complex (PIC) at the β -globin gene by twofold (57,75). Deletion of the LCR does not have a major effect on general DNase I sensitivity across the locus or on the histone acetylation pattern at the β -globin genes (75). These studies indicate that the β -globin LCR functions as a strong enhancer and that the domain-opening activity observed in transgenic studies is not critical for establishing accessible chromatin in the endogenous mouse locus. In contrast, deletion of the LCR from a human chromosome situated in a mouse hybrid-MEL cell line abolishes β -globin gene transcription and the promoter is hyperacetylated, consistent with the idea that the LCR recruits histone modifying complexes including histone acetylases, which are required to modify the distant promoter (73). It is not clear what underlies these different results between mice and human but they could indicate that the mouse and human loci are regulated differently.

Although deletion of a single HS by homologous recombination in mice had only mild effects on globin gene expression, deletion of the full LCR caused phenotypes similar to human ($\gamma\delta\beta$)-thalassemias and is lethal in utero (76). Recently, intergenic transcripts were detected throughout the human β -globin gene locus (77). These erythroid-specific transcripts originate within or upstream of the LCR, terminate near β -globin promoters with subsequent initiation just 3' of the genes and transcription appears to proceed in a unidirectional manner towards the genes (78). Intergenic transcripts are found over the LCR at all developmental stages whereas those around the globin genes are generated in a developmental stage-specific manner (79). A good correlation has been observed between the intergenic transcription and the histone acetylation

regions at the β -globin gene locus during development. Although the significance of intergenic transcription is not known yet, it may play a role in keeping the chromatin conformation in an open configuration. It has been shown that when deleting an endogenous intergenic transcript promoter between the γ - and δ -globin genes, both of the chromatin accessibility and the transcription of downstream β -globin gene is affected (79).

The Adult β -Globin Gene Promoter and 3' Enhancer

The promoter of the human β -globin gene is composed of three major sites, a non-canonical TATA-box (CATAAA) located 25-30 bp upstream of the transcription start site (80), a CCAAT-box at around -76, and a CACC-box at approximately -93 (81-84). Mutations within the TATA-like box disrupt the assembly of the basal transcription machinery, especially the association of the TFIID complex (85), whereas mutations within the CACC-box disrupt the association of Sp1/EKLF, an erythroid specific transcription factor required for γ - to β -globin switching (86). There are five transcription factors that have been reported to bind the β -globin CCAAT-box *in vitro*: GATA-1, NF-Y (CP1), C/EBP β , C/EBP γ and C/EBP δ (87-91). GATA-1 is an erythroid specific transcription factor required for erythroid cell maturation (92), NF-Y is a ubiquitous CCAAT-box binding complex (93), and C/EBP β , C/EBP γ and C/EBP δ belong to the CCAAT/enhancer-binding protein family of basic leucine zipper (bZip) factors. It has been shown that in reporter assays, C/EBP δ cooperates with EKLF, which binds to the CACC-box, to active adult β -globin transcription, whereas C/EBP γ inhibits the activator function of EKLF on the β -globin transcription (91).

Another important sequence contributing to high level β -globin gene transcription is a pyrimidine rich initiator (Inr) element, which is located at the transcription start site (94). The Inr *cis*-element is able to bind to *trans*-acting factors to accurately place the transcription start site for the transcription machinery (95,96), and in some TATA-less promoters, the Inr can itself

mediate the initiation of transcription (97,98). Lewis et al. have shown that the Inr element within the human β -globin gene interacts with the TFIID complex *in vitro*, and that mutations in the Inr reduce the transcription efficiency (94,99). Several reports suggest that USF, TFII-I and YY1 associate with the Inr and stimulate transcription (95,100,101). In addition, Inr can be recognized by RNA polymerase II to facilitate active transcription (102). Interestingly, both of USF and TFII-I bind to the human β -globin Inr (103).

In addition to the Inr element, the β -globin gene promoter contains several E-box motifs (CANNTG) downstream of the transcription start site (103). The E-box/Inr and +60 E-box elements are conserved across species. The E-box/Inr motif overlaps with the Inr element, and the +60 E-box motif is located 60 bp downstream of the transcription start site. Mutations within the +60 E-box or the E-box/Inr elements led to similar levels of transcription reduction *in vitro*. USF and TFII-I bind to the E-box/Inr, whereas USF1 and USF2 associate with the +60 E-box (103).

There is a MARE/AP-1 like element located 24 bp downstream of the transcription start site, which interacts with low affinity with transcription factor NF-E2 (104). Although mutations of this putative NF-E2 binding site yielded no effect on the transcription of the β -globin gene *in vitro*, NF-E2 could also be recruited to the β -globin gene promoter by *trans*-acting factors which bind to the adjacent regions through protein-protein interactions. Moreover, the β -globin gene downstream core element (DCE) is localized within +10 to +45. Mutations within the DCE decreased transcription from the β -globin promoter. It has been shown that the DCE is bound by components of the TFIID complex, TAF1/TAF(II)250, and their association is required for the proper function of the DCE (99).

β -globin 3' enhancer is located at 550-800 bp 3' to the polyadenylation site of the β -globin gene (105). Deletion of this fragment in transgenic mice caused a 10-fold reduction in the expression of the β -globin gene (106). The β -globin 3' enhancer contains four regions, each of which binds to the erythroid specific transcription factor GATA-1 and at least one other non-tissue specific factor (107). Donovan-Peluso *et al.* (105) have shown that the β -globin 3' enhancer activity is erythroid-specific but not developmental stage- or gene-specific.

Hemoglobin Switching

The most intriguing and widely studied event occurring in the β -globin gene locus is hemoglobin switching, which means that suppression of the γ -globin gene is associated with a complementary increase in expression of the adult β -globin gene (108). Studying this process is indeed very meaningful for the therapy of sickle cell anemia and β -thalassemia, because it has been shown that the fetal γ -globin can be reactivated in patients with a defective adult β -globin chain and can functionally substitute for the dysfunctional β -globin gene (109,110).

Understanding the mechanism(s) of hemoglobin switching is expected to contribute to the development of new therapies for hemoglobinopathies. In the transgenic mice studies, in the absence of the LCR, γ - and β -globin are still expressed in a developmental stage-specific manner. This suggests that the regulation of hemoglobin switching is largely directed by gene proximal regulatory elements.

Human β -globin gene expression undergoes two switches from the embryonic ϵ -globin gene to the fetal γ -globin gene and from the fetal γ -globin genes to adult δ - and β -globin genes (111). In contrast to humans, the mouse has a single switch from the embryonic $\epsilon\gamma$ - and $\beta h1$ -globin genes to the adult βmaj - and βmin globin genes (112). Human transgenes in mice only undergo one globin switch, with human ϵ - and γ -globin genes expressed along with mouse

embryonic $\epsilon\gamma$ - and β^H1 -globin genes, and human δ - and β -globin genes expressed along with mouse adult β^{maj} - and β^{min} globin genes (113).

There are two main mechanisms that control the hemoglobin switching, gene competition and autonomous gene silencing.

Gene Competition

Transgenic studies in mice provided evidence for a gene competition mechanism during γ - to β -globin gene switching (114,115). When γ - or β -globin genes alone are linked to the LCR, they both failed to display the proper developmental stage-specific expression patterns. β -globin was constantly active during all the stages of development, and surprisingly, in the embryonic stage, the expression level of β -globin is as high as the β -globin gene level in adults. The γ -globin gene was also expressed in the embryonic as well as, albeit at lower levels, in the fetal/adult stages. Interestingly, however, when the γ - and β - globin genes were linked together (with γ - upstream of the β -globin gene) and placed downstream of the LCR, their developmental stage-specific expression was restored with the γ -globin gene expressed strictly during the embryonic stage of erythropoiesis and the β -globin gene expressed only at the adult stage. The above observations can be explained by the proposed model that γ - and β - globin genes compete for interactions with the LCR (115). In the fetal stage of development, the LCR preferably interacts with the fetal γ -globin gene, resulting in silencing of the β -globin gene. On the other hand, in the adult stage of development, the adult β -globin gene favorably interacts with the LCR, which leads to suppression of fetal γ -globin gene expression.

Autonomous Silencing

Automatic silencing means that the canonical gene and adjacent sequences contain all the elements responsible for turning off gene expression during development. This concept was firstly developed based on the following transgenic mice studies. In transgenic mice carrying the

human ϵ -globin gene, the ϵ -globin gene is only abundantly expressed when it is linked to the LCR. However, the expression of ϵ -globin gene is restricted to the embryonic yolk sac (116). Studies have shown that there are sequences within both proximal and distal ϵ -globin gene promoter responsible for the ϵ -globin silencing process in adults (117,118). Indeed, the fact that mutation of GATA-1, YY1 or SP1 sites in the upstream region of the of the ϵ -globin gene promoter abolished ϵ -globin gene silencing implies that the globin gene silencing involved multiple transcription factors which possibly form a silencing complex (119). In addition, the proximal ϵ - and γ -globin promoters contain direct repeat sequences named DR1 box, which are located near the CAAT box. It has been reported that a factor identical to the orphan nuclear receptor COUP-TF binds to these DR elements (120). Also, a high molecular weight complex, named DRED, binds to DR elements of the ϵ -globin promoter (121). Mutations of the DR elements led to ϵ -globin gene expression in adult mice. These observations indicate that DR elements participate in ϵ -globin gene silencing (120,122). Currently, the assumed mechanism of autonomous silencing is that during definite erythropoiesis, the silencing complex formed on the proximal or distal ϵ -globin gene promoter turns off the expression of the ϵ -gene by inhibiting the interaction between the ϵ -globin gene and the LCR.

The mechanisms of controlling γ -globin gene expression seem to involve both gene competition and autonomous silencing. On the one hand, in the β YAC mice studies, when γ -globin gene was placed upstream of the ϵ -globin gene, it was expressed in both of the embryonic and fetal stages, but silenced in the adult stage (123,124). However, when the -378 to -730 region of the γ -globin gene promoter was deleted, it was observed that the γ -globin gene was expressed in the adult stage (125), suggesting that the -378 to -730 region contains sequences responsible for the autonomous silencing of the γ -globin gene. Furthermore, the fact that a mutation in the

GATA site within this -378 to -730 region resulted in hereditary persistence of fetal hemoglobin (HPFH) in human (126) reinforced the above statement. Moreover, findings in transgenic mice carrying galago/human γ -globin gene hybrid promoters suggest that the CACCC box of the γ -globin gene promoter is required for γ -globin gene silencing (127). On the other hand, the persistent γ -globin expression in the adults of the $\delta\beta$ -thalassemia and HPFH patients is best explained by the lack of gene competition due to the deletion of the adult β -globin gene.

Taken together, it is suggested that expression of the adult β -globin gene is mainly regulated by gene competition; the expression of the ϵ -globin gene is regulated by autonomous silencing; and the expression of the γ -globin gene is regulated by gene competition and autonomous silencing.

Switching Factors

Many transcription factors have been discovered to regulate the stage-specific expression of the γ - and β -globin genes, by binding to the promoter elements. One of the most well-studied of these factors is EKLF (erythroid krüppel-like factor) (86), an erythroid specific transcription factor, which behaves as an adult switching factor. It has higher binding affinity to the adult β -CACCC box than the slightly different fetal γ -CACCC box, and preferentially activates the β -globin gene (128,129). EKLF is necessary for β -globin gene expression and is expressed more abundantly in the adult stage of development compared to the embryonic stage (130). Furthermore, β -CACCC box mutations cause β -thalassemia and EKLF knockout mice die at E14.5 from lethal β -thalassemia (131). Interestingly, in EKLF knockout mice, the embryonic hematopoiesis is not severely affected, but the anemia occurs following the embryonic/fetal to adult globin developmental switch.

Two EKLF homologous proteins, FKLF (also known as KLF11) and FKLF-2 (also known as KLF13), have been identified as candidates for fetal switching factors. FKLF mainly functions

actively at the ϵ -globin gene promoter and FKL2 at the γ -globin gene promoter (132,133). FKL and FKL2 activate γ -globin gene expression predominantly by interacting with the γ -CACCC box. Over-expressing FKL in K562 cells increased the expression of the endogenous ϵ - and γ -globin genes, indicating that FKL functions as a transcriptional activator *in vivo*. In luciferase reporter assays, only FKL2 can activate a variety of erythroid specific promoters containing the CACCC-box plus γ -globin gene promoter elements, including a binding site for GATA-1. However, the expression of FKL2 is not erythroid specific. Interestingly, despite the fact that ectopic expression of FKL in the bone marrow of transgenic mice carrying a basally active γ -globin gene cassette is capable of inducing γ -globin gene expression in adult mice (134), FKL knockout mice appear phenotypically normal, with no apparent impact on lifespan, fertility, viability and development. Also, there is no effect on erythropoiesis in the FKL knockout mice (135). After examining FKL^{-/-} mice crossed with mice expressing human γ -globin gene it was found that the FKL^{-/-, γ -globin} mice express the same levels of the γ -globin gene as the FKL^{+/, γ -globin} mice, suggesting that FKL is not required for γ -globin gene expression in mice (135). FKL2 knockout mice have an enlarged thymus and spleen due to the decreased T cell apoptosis, as FKL2 functions as a repressor of the anti-apoptosis factor Bcl-xL. However, there is no defect in hematopoiesis in FKL2 knockout mice (136). The roles of FKL and FKL2 in switching are not clear.

Another factor that may be involved in hemoglobin switching is stage selector protein (SSP), which binds to the stage selector element (SSE) of the γ -globin gene promoter. In transgenic mice studies, mutations in SSE, located at -50 region of the γ -globin gene promoter, results in an increase in β -globin gene expression only in the early stage of fetal development and a prolonged the period it takes to complete the switch, suggesting that SSE contributes to the

competitive ability of the γ -globin gene promoter during early fetal liver hematopoiesis (137). SSP is a heterodimer, consisting of a ubiquitously expressed CAAT binding protein CP2 and an erythroid-specific protein NF-E4 (138). Loss of NF-E4 binding to the SSE of the γ -globin gene promoter resulted in a loss of γ -globin gene transcription relatively to that of the β -globin gene, indicating that NF-E4 preferentially activates γ - over the β -globin gene promoter, which lacks an SSE (137). In addition, overexpression of NF-E4 in cord blood progenitors resulted in upregulation of human γ -globin gene expression with a concomitantly reduction of β -globin gene expression (138). CP2 knockout mice showed no defects in growth, development, behavior and fertility, and displayed normal hematopoiesis, probably due to the compensation of homologous proteins, such as LBP-1a (139). The NF-E4 knockout mice have not been reported yet.

COUP-TFII (140) may play a role in hemoglobin switching. It is a retinoic acid orphan receptor corresponding to the erythroid specific binding activity of NF-E3 (141,142). COUP-TFII binds to a consensus site in the γ -globin gene promoter that includes the γ -globin CAAT element and may function as a repressor of the γ -globin gene expression (120).

The PYR complex is also a putative switching factor (143,144). It is composed of several factors exhibiting SWI/SNF activity to promote chromatin modifications and remodeling (145). PYR is restricted in expression to adult definitive erythroid cells, where it binds to a pyrimidine-rich DNA element located between the γ - and β -globin genes (146). Deletion of the PYR binding site from a human β -globin gene construct resulted in delayed and prolonged human γ - to β -globin switching in transgenic mice. Thus, PYR possibly functions as an adult stage-specific factor to facilitate the switch by repressing γ - and activating β -globin gene expression. The transcription factor Ikaros is the DNA binding component of the PYR complex (147).

The LCR and the β -Globin Gene Promoter Interaction Models

Recent studies by using chromatin immunoprecipitation (ChIP) and relatively new techniques such as capturing chromosome conformation (3C), fluorescence in situ hybridization (FISH) and RNA-TRAP (tagging and recovery of associated proteins) have revealed that LCR HS1-HS4 (148) are in close physical proximity to the actively transcribed β -globin genes in mouse erythroid cells. The 3C data demonstrate that physical proximity between the LCR and the β -globin promoter is established at a time during differentiation of erythroid cells when the gene is actively transcribed (149).

Currently, there are three models for LCR mediated globin gene activation. (1)Tracking Model: activators and chromatin remodeling complexes are recruited to the LCR first, then migrate to the promoter by tracking along the DNA and activate transcription (150); (2)Looping/Linking Model: a distant LCR and promoter contact each other directly or indirectly through protein/DNA interactions. This interaction is independent of intervening DNA between them and can occur in trans (151,152); (3) Facilitated Tracking Model: activators and chromatin remodeling complexes track along the DNA from the LCR to the promoter without losing contact with the enhancer during the whole process, and eventually establish a loop between enhancer and promoter (153). Considering the recent data, the looping and facilitated tracking models are the favorite models for LCR mediated regulation of the β -globin gene (149), while the tracking model may apply for ϵ -globin gene regulation.

The “Active Chromatin Hub” Model for LCR Regulating the β -Globin Gene

The 3C experiments support the current “Chromatin Hub” model for LCR mediated β -globin gene activation. The 3C results demonstrate that in mouse erythroid progenitor cells that do not express the globin genes, the distal HS sites (3'HS1 and 5'HS-60/-62, which are located -60 and -62 kb from $\epsilon\gamma$ -globin gene) and part of the LCR (HS4-6) interact to form a ‘poised’

chromatin hub (148). In differentiated erythroid cells that express the β -globin gene the remaining LCR HS1–HS3 sites as well as the globin gene promoter participate in long range chromosomal interactions and contribute to the formation of the ‘active chromatin hub’ (ACH). The globin genes switch their association with the ACH depending on their development specific transcription activity. The flanking 3’HS1 and 5’HS–60/–62 regions are not required for β -globin gene expression. If the β -globin promoter is deleted together with HS3 (but not HS2) (154), the ACH fails to form at the adult stage. This is accompanied by decreases in both histone acetylation and DNaseI sensitivity at the LCR and gene promoters. CTCF has been shown to be required for the formation of the “poised” chromatin hub in the β -globin gene locus. However, the low level of globin gene expression in erythroid progenitor cells is not affected. The consequence of deleting the 3’HS1 site (a CTCF binding site) on globin locus configuration was analyzed in ES cells that differentiated *in vitro* into erythrocytes. The data demonstrated that even though the pre-existing loop between 3’HS1 and upstream sites (HS-85, HS-60/-62 and HS5) is disrupted, the LCR HS4/5 maintained identical strong interactions with the β maj-globin gene compared to wild type cells (155). Furthermore, deletion of 3’HS1 and did not perturb globin gene expression. These data demonstrate that the “poised” chromatin configuration in the globin gene locus is not required for proper globin gene regulation. Perhaps, this configuration is important for preventing regulatory elements located in the globin gene locus to affect expression of nearby genes.

The erythroid specific transcription factors GATA-1, Erythroid krüppel-like factor (EKLF) and co-activator FOG-1 (friend of GATA) are required for ACH formation and/or stabilization (156,157). EKLF and GATA-1 could stabilize the loop by binding to the LCR HS sites and the globin gene promoters. Interestingly, deletion of the β -globin promoter itself did not disrupt loop

formation, which indicates that other regions within the gene contribute to interactions in the ACH (154). To form the loop, histone modifications might be required ahead of loop formation, because a relatively flexible chromatin fiber may facilitate loop formation. In this respect, EKLF and GATA-1 could be crucial for directing histone modifications to specific regions in the globin gene locus.

As mentioned before, CTCF is required for loop formation in the erythroid progenitor cells, but is not required for LCR-globin gene contacts established in differentiated erythroid cells.

Proteins Involved in the Regulation and Expression of Adult β -Globin Gene Expression

RNA Polymerase II (RNA Pol II)

RNA Pol II is responsible for the pre-mRNA synthesis in eukaryocytes. During transcription, RNA Pol II also scans for the DNA damage, modifies chromatin and serves as a platform for several mRNA processing factors. RNA Pol II has a unique C-terminal domain (CTD) which consists of 25-52 tandem repeats of the consensus sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser ($Y_1S_2P_3T_4S_5P_6S_7$) depending on the organism (158). Factors recruited to the CTD include the Mediator complex, which regulates transcription initiation, histone methyltransferases, mRNA capping enzymes, and polyadenylation factors that modify the 3' end of mRNA (159,160).

The serine residues (ser2, ser5, and ser7) of the CTD are phosphorylated and dephosphorylated during the transcription cycle and the CTD can be hyperphosphorylated during the transition from a promoter-bound complex to an active elongation complex (161). At the beginning of transcription, the CTD unphosphorylated RNA Pol II (RNAP IIA) assembles into a preinitiation complex on the promoter with general transcription factors (GTFs) TFIIA, TFIIB, TFIID, TFIIE, TFIIIF, and TFIIH (162,163). The CTD of the RNA Pol II is phosphorylated

(RNAP IIO) at the time of transcription initiation by the CTD kinase Cdk7/Kin28, a subunit of TFIIF (164). Phosphorylation occurs specifically at ser5, which is important a process called “promoter clearance” and recruits capping enzymes and the histone H3K4 specific methyltransferase Set1, (165,166). The transcription complex is then paused near the transcriptional start site for pre-mRNA capping. Then another CTD kinase, Cdk9, which is part of the elongation factor P-TEFb mediates CTD Ser2 phosphorylation to allow transcription elongation (167). The histone H3K36 specific methyltransferase Set2 preferentially associates with the hyperphosphorylated RNA Pol II, especially with the Ser2 phosphorylated CTD (168-170). At the end of the transcription cycle, a CTD specific phosphatase removes dephosphorylates the CTD (171,172). By converting RNAP IIO to RNAP IIA, it allows RNAP IIA to reenter the transcription cycle.

Previous studies from Leach et al. (53) and Johnson et al. (47) have shown that RNA Pol II is recruited to the LCR and to the β -major promoter in a GATA-1 dependent manner. Erythroid specific transcription factor NF-E2 p45 subunit is required for the loading of RNA Pol II to the murine β maj-globin gene promoter, but not to the LCR (49). It has been proposed that GATA-1 recruits RNA Pol II to the LCR, and NF-E2 assists RNA Pol II relocation to the β -globin gene promoter (47). Sawado et al. (75) also demonstrated that NF-E2 is recruited equally efficiently to the adult β -globin promoters of LCR deleted and wild type (WT) murine alleles. Deleting the LCR reduces preinitiation complex (PIC) assembly twofold at the β -globin gene promoter, but remarkably decreases RNA Pol II C-terminal domain (CTD) Serine 5 phosphorylation and affects transcription elongation. However, the promoter histone acetylation is not affected by the deletion of the LCR. These results suggest that during β -globin gene activation, the LCR-independent chromatin opening, NF-E2 binding and partial PIC assembly at

the β -globin gene promoter take place prior to the active transcription of the β -globin gene that is mediated by LCR-dependent regulation of RNA Pol II elongation. Thus the LCR appears to function downstream of certain activator recruitment and partial PIC assembly at the β -globin gene promoter. However, the fact that inactive transcription complexes are recruited to the promoter could also indicate that the LCR is required for the recruitment of transcription competent transcription complexes to the globin gene locus.

GATA-1

GATA-1, also known as NF-E1 (Nuclear Factor-Erythroid 1), belongs to the GATA family of transcription factors. It binds to the DNA consensus sequence (A/T)GATA(A/G) by two characteristic GATA family specific C₄ (Cys-X₂-Cys-X₁₇-Cys-X₂-Cys) zinc-finger motifs (173-177). GATA-1 is expressed in both primitive and definite erythroid cells (178,179), megakaryocytes (180,181), mast cells(180), eosinophils (182), and Sertoli cells of the testis (176,183). GATA-1 is essential for normal erythropoiesis. GATA-1 deficient mouse embryonic stem (ES) cells are unable to contribute to the mature red blood cells in chimeric mice (131). Further analysis showed that in these chimeric mice, GATA-1 null erythroid cells have an arrest at the proerythroblast stage (184). This arrest has been further confirmed by the *in vitro* differentiation of GATA-1 deficient ES cells (185). The arrested precursor cells die due to apoptosis (186). These findings are consistent with the fact that GATA-1 deficient embryos die between embryonic day 10.5 (E10.5) and E11.5, due to severe anemia (178).

GATA-1 consists of three functional domains: the N-terminal transactivation domain, the N-terminal zinc finger (N-finger), and the C-terminal zinc finger (C-finger) (187). The C-finger is crucial for recognizing the GATA-1 consensus sequence and binding to DNA. The N-finger is responsible for stabilizing the GATA-1/DNA complex and interacting with cofactors. The interaction between GATA-1 and FOG-1 is mediated through the N-finger (188), but when

GATA-1 interacts with EKLF (175,189) and GATA-1 itself (190), the C-finger also participates in complex formation.

GATA-1 is known to interact with a variety of proteins. One of them is the well known GATA-1 cofactor FOG-1 (friend of GATA), which is coexpressed with GATA-1 in erythroid and megakaryocytes (191). FOG-1 contains nine zinc fingers, which interact with the GATA-1 N-finger through its 6th zinc finger (188). FOG-1 is essential for GATA-1 function, since FOG-1 null mice exhibit a phenotype similar to the one found in GATA-1 knockout mice, but it also appears that FOG-1 has GATA-1 independent functions in megakaryopoiesis (192). Although FOG-1 shows no DNA-binding activity, it can either activate or repress GATA-1 mediated transcription based on the context of the promoter (191,193). Recent data suggest that in erythroid cells, FOG-1 is first induced by GATA-1 (194), then it functions together with GATA-1 to regulate the β -globin gene locus. Another protein that interacts with GATA-1 is EKLF. This interaction is dependent on the presence of DNA, such as promoter (189). The CACCC-box motifs that EKLF recognizes are found in close proximity to the GATA motifs in several promoters and enhancers, including the LCR (175). This suggests that EKLF and GATA-1 may have cooperative functions. It has been shown that GATA-1 can associate with itself *in vitro* (195). GATA-1 dimerization may facilitate interactions between enhancers and promoters.

It is reported that GATA-1 interacts with histone acetyltransferases CBP and p300 both *in vitro* and *in vivo* (196,197). Recruiting histone acetyltransferases to the nucleosome can induce an open chromatin configuration, therefore activating transcription. Since CBP and p300 interact with various transcription factors, they can have a “bridge” function to connect components located at the enhancer and promoter elements (198). Moreover, CBP and p300 can acetylate GATA-1 to stimulate its transcriptional activity (197,199). GATA-1 can also be phosphorylated

(200) and sumoylated (201). Phosphorylated GATA-1 is induced in K562 cells leading to an increase in DNA binding (202). The function of sumoylated GATA-1 is unknown.

GATA-1 was first identified to associate with the β -globin 3' enhancer (203). Now it is known that GATA-1 binds to multiple regulatory regions in both α - and β -globin gene loci (187,204). The GATA-1 null G1E erythroid cell line, which is derived from GATA-1 null embryonic stem cells, expresses low levels of adult α - and β -globin, whereas the embryonic ξ - and ϵ -globin are undetectable (205). In G1E cells, the GATA-2 mRNA is expressed at higher levels compared to wild type cells. This suggests that GATA-2 may partially replace the function of GATA-1 in GATA-1 null cells, since GATA binding sites are essential for the expression of β -globin LCR driven transgenes (206).

GATA-1 is also critical for megakaryocytes development (207), G1/S cell cycle progression (208), and reprogramming of hematopoietic precursors (209,210). Cell cycle control is the most important regulation during hematopoietic differentiation, because the precursors need to proliferate to achieve further hematopoietic maturation, but in order to make the terminal differentiation to occur, the cells must exit the cell-cycle (208). GATA-1 target genes also include erythropoietin (Epo) receptor (EpoR) (211), and Epo signaling is known to be important for proliferation, differentiation and survival of erythroid progenitors (212). Furthermore, Bcl-X_L, a gene encoding an anti-apoptotic protein, is regulated by GATA-1 as well (213).

NF-E2

Transcription factor NF-E2 (Nuclear Factor-Erythroid 2) is found almost exclusively in hematopoietic progenitors, and in the erythroid/ megakaryocyte/mast cell trilineage, and is known to regulate globin gene expression by acting through locus control regions (LCRs) upstream of the α - and β -globin gene clusters (214). NF-E2 also associates with the β -globin gene promoter (49). NF-E2 belongs to the basic-leucine zipper (bZip) family and exists as a

heterodimer. It consists of a hematopoietic-specific subunit p45 and a more ubiquitously expressed small Maf protein (p18) (215,216). Both of these subunits belong to the AP-1 superfamily. The p45 subunit is a member of the cap and collar (cnc) family (217). The cnc domain that includes the bZip region was also found in NF-E2 related proteins such as Nrf1, Nrf2, Bach1 and Bach2. But the function of the cnc domain is unknown. The N-terminus of p45 contains the transactivation domain, which is critical for NF-E2's transcription activator function (214). The p18 small Maf protein binds to NF-E2 p45 through its bZip domain (215,216).

The small Maf subunit p18 lacks the transactivation domain. Its dimerization with the p45 subunit to form the NF-E2 heterodimer is required for NF-E2 DNA binding activity. The DNA binding site of NF-E2 contains a core AP-1 motif, (T/C)***GCTGA(G/C)TCA***(T/C) (AP-1 motif is bold and italicized). A mutation at the second position of the NF-E2 site (***G***) was shown to abolish NF-E2 binding, but not AP-1 binding (216).

NF-E2 binding sites exist in β -globin LCR HS2, HS3, and HS4 (214,218). These binding sequences are closely related to Maf-recognition elements (MAREs). Mutations of NF-E2 binding sites in the LCR HS2, 3, and 4 reduces DNase I hypersensitivity (218,219). HS2 harbors enhancer function to regulate high-level and position independent expression of a human β -globin gene in transgenic mice (ref). The tandem MAREs in HS2 of the β -globin LCR is crucial for its enhancer activity of β -globin gene expression, but is not required for position-independent activation (220,221). The presence of NF-E2 binding sites in HS2 alone is insufficient for high level expression of a linked human β -globin gene in transgenic mice and mouse erythroleukemia (MEL) cells (222). Deletion of HS2 from one allele of the β -globin locus in mice has no significant effect on the timing and extent of the expression of all the β -like globin genes on the

mutated chromosome. However, when homozygous, the deletion of HS2 resulted in normal expression of mouse embryonic genes, but the adult β -globin mRNA reduced 30%. (223),(224).

NF-E2 binding sites play a role in chromatin remodeling and transcription activation of globin genes. The N-terminus of NF-E2 p45 contains an activation domain, which can interact with other molecules such as cAMP-response element-binding protein (CREB)-binding protein (CBP) (225) and TAF_{II}130 (a component of the TFIID complex) (226). Some studies showed that homodimers of small Maf proteins repress transcription (217).

Mice deficient for NF-E2 p45 die shortly after birth due to defects in the megakaryocyte lineage (227), but they have no obvious phenotype in erythropoiesis, which suggests that other bZip proteins can substitute for p45's function in the red cell lineage (228). However, erythroleukemia cells lacking functional NF-E2 p45 fail to produce globin genes (229), which show that NF-E2 p45 is important in erythropoiesis. MafK "knock out" mice have no obvious phenotype, which may suggest that small Maf proteins have functional redundancy (230).

EKLF

EKLF (erythroid krüppel-like factor) functions as an adult switching factor by interacting with the CACCC-box located in the β -globin gene promoter (86). EKLF preferentially activates β -globin gene due to a higher affinity for the β -CACCC box compared to the γ -CACCC box located in the promoter regions of the genes (128,129). In the absence of EKLF, DNase I HS sites do not form in the β -globin promoter and in LCR HS2 and HS3. The γ - to β -globin switch is disrupted and characterized by the persistence of high γ -globin levels and severe reduction of β -globin chains (131,231,232). EKLF functions at least in part by recruiting a SWI/SNF remodeling complex to the LCR and to the human β -globin promoter (233).

USF

The evolutionary conserved upstream stimulating factors (USF), USF-1 (43KDa) and USF-2 (44KDa), belong to the basic-Helix-Loop-Helix-Leucine-Zipper (bHLH-LZ) transcription factor family (234). They are highly related at their C-terminal DNA binding domain (70% identity) and can form homo- or heterodimers to interact with high affinity with the E-box regulatory elements (E-box: CANNTG, in most cases, NN are CG or GC), which are abundant in the genome of eukaryotes. The basic region is involved in DNA interaction with the E-box elements, whereas the HLH and LZ domains are mainly involved in dimerization. The integrity of the LZ domain is important for high-affinity and specific DNA binding (235). USF1 and USF2 only have very limited homologous regions in their N-terminus, which includes the transactivation domain, one of them is the highly conserved USF specific region (USR), which locates upstream of the basic region, and has been shown to be essential for transcription activation (236-238), although how it functions is unknown.

The ubiquitously expressed USF transcription factors play important roles in a variety of transcriptional processes. First, USF has been found to stimulate gene transcription by binding to their cognate E-box motifs (239). Secondly, USF1 interacts with the general and cell-specific transcription factors, such as SP1, Pea3 and MTF1 to cooperate in gene regulation (240-242). Thirdly, USF mediates the recruitment of chromatin remodeling enzymes to DNA, such as histone acetyltransferase PCAF, histone 3 lysine 4 (H3K4)-specific methyltransferase SET7/92 (64), and histone 4 arginine 3 (H4R3)-specific methyltransferase PRMT1 (243). Moreover, USF has been found to regulate topoisomerase III (*hTOP3 α*) gene expression (244). Finally, USF1 directly interacts with TATA-containing and TATA-less promoters. Components of the TATA-dependent preinitiation complex, e.g. TFIID (TBP plus TBP associated factors, TAFs), directly

interact with USF1 (245-249). In TATA-less promoters, USF1 has been found to bind to the pyrimidine-rich initiator (Inr) element near the transcription start site (95,97,250). More interestingly, Ferre-D'Amare et al. (239) used hydrodynamic measurements on the USF/DNA complex and showed that USF can exist as a bivalent homotetramer. Sha et al. (235) have shown that the USF homotetramer could potentially associate with two DNA recognition sites to facilitate DNA looping. Therefore it is important to study the gene regulation mechanism of USF in the context of the β -globin gene locus.

Although USF1 and USF2 are expressed ubiquitously in a variety of tissues (251-253), it has been found that the ratio of USF homodimers and heterodimers is cell type specific (253,254). Even for the USF dimers that have similar DNA-binding properties, they may control different target genes, by interacting with different transcription factors through their distinct N-terminal domains (237,255). In addition, USF1 and USF2 knockout mice are viable and exhibit distinct phenotypes, which indicate that they are not completely redundant (256,257). USF1 null mice are viable and fertile, with only slight behavior abnormalities. USF1 deficient mice exhibit an increase in USF2 expression, which may functionally compensate for the loss of USF1. In contrast, USF2 null mice contain reduced levels of USF1 and have an obvious growth defect: they were 20-40% smaller at birth compared to their wild-type or heterozygous littermates and maintained a smaller size throughout postnatal development. However, USF1/USF2 double knockout mice die early during embryogenesis, suggesting that USF is essential for embryonic development (234).

It was recently shown that chicken USF1/USF2 heterodimers bind a divergent E-box element (CACGGG) in the 5' HS4 insulator sequence of the β -globin locus and recruit the methyltransferase SET7/9 and the histone H3 acetyl transferase PCAF (64). It is thought that

recruitment of these proteins establish a barrier function that prevents the spread of heterochromatin into the β -globin locus.

TFII-I

TFII-I was originally discovered as a basal transcription factor that binds and functions through the initiator (Inr) core promoter element *in vitro* (95). Later it was shown that TFII-I functions as a multifunctional transcription factor with the ability not only to stimulate transcription from TATA-less and Inr-containing promoters (258), but also through binding to unrelated upstream elements (E-box) that are usually recognized by helix-loop-helix (HLH) family proteins such as USF. Importantly, TFII-I has been shown to cooperate with USF to interact with both E-box and Inr elements (95). These observations indicate that TFII-I can function as a basal factor and as an activator, thus it may establish communications between the basal transcription machinery assembled at the promoter and activators associated with the upstream regulatory sites (95). TFII-I is phosphorylated at both serine and tyrosine residues. Tyrosine phosphorylation of TFII-I is required for its transcriptional functions (259). TFII-I is regulated by extracellular signals and translocates to the nucleus (260).

TFII-I is cleaved by thrombin at the amino acid 677 and divided into two domains: the 70 KDa N-terminal domain (p70) which carries the DNA binding activity, and the 43 KDa C-terminal separable activation domain (261). The N-terminal p70 mutant functions as a dominant-negative mutant of the wild-type TFII-I. The C-terminal 280 amino acids, when fused to the DNA binding domain of GAL4, are unable to activate transcription from a promoter which contained five GAL4 binding sites upstream of a TATA-box. Therefore, the C-terminus of TFII-I is only partially responsible for the activation function, and it requires certain regions from the N-terminus of TFII-I to achieve appropriate transcription activation.

The multi-functional properties of TFII-I may rely on its unique primary amino acids structure. TFII-I is composed of six direct reiterated I-repeats, R1-R6, with each of them consisting a putative HLH motif (262). The N-terminal DNA binding domain contains R1-R4, whereas the C-terminal transactivation domain contains R5-R6. The basic region which is necessary for DNA binding locates only within R2. Since each of the I-repeats displays a potential HLH motif, which serves as protein interaction surfaces, it is possible that each of them mediates a distinct protein-protein interaction.

Multiple TFII-I isoforms are generated by alternative splicing in human and mice (263-265). Besides the well characterized 957 amino acids form of TFII-I (referred to as Δ), there are three additional isoforms in human and perhaps one more in mice. These isoforms are named α , β , and γ , and they are 20-41 bps longer than the Δ -isoform. All of these isoforms contain the six I-repeats, the basic region, the N-terminal putative LZ motif, and the nuclear localization signal (297-304 amino acids in the Δ -isoform). It has been shown that these isoforms form homo- and heteromeric interactions, which facilitates nuclear localization (263). The expression patterns of these isoforms in a variety of cell types and species indicates that their functions are non-redundant.

TFII-I is activated by various extracellular signals thus linking signal transduction to transcription (260). In the absence of extracellular signals, TFII-I is phosphorylated at both serine/threonine and tyrosine residues. Tyrosine phosphorylation of TFII-I is required for its transcriptional functions (259). TFII-I is involved in stress response (266), B-cell signaling (267), and other diverse processes. In endoplasmic reticulum (ER) stress conditions, induced by depletion of stored endoplasmic calcium, which leads to increased expression of glucose regulated protein (GRP) genes, it is shown that TFII-I increasingly binds to an upstream element

named the endoplasmic reticulum stress response element (ERSRE) at the promoters of GRP genes (266,268). TFII-I also interacts with ATF6 and NF κ B, two signal induced transcription activators (268).

Both USF and TFII-I can either activate or repress gene expression, depending on which factors they associate with. Although ectopically expressed TFII-I only exhibits moderate effects on Inr-dependent transcription, over co-expression of TFII-I and USF1 dramatically enhances the transcription activity of TFII-I at the Inr (262). TFII-I and USF also co-localized at the AdMLP upstream E-box element (95). Also, it has been shown that TFII-I and USF physically interact both on and off the DNA (262). These facts further support the notion that TFII-I and USF function together to facilitate the communication between upstream activators and the basal transcription complex at the core promoter.

In the β -globin gene locus it was known that USF functions at the human β -globin promoter and at the LCR HS2 enhancer element (269). Hardison et al. (50) showed that mutation of a canonical E-box in the LCR HS2 core region diminishes the enhancer activity in transient transfection assays. Both USF and TFII-I can recruit transcription complexes to TATA-less promoters and stabilize the transcription complex at TATA-containing promoters. Leach et al. (53) showed that in human K562 cells, in which the adult β -globin gene is silenced, TFII-I interacts with the Inr; while in MEL cells, which express the β -globin gene, USF1 and USF2 bind to the +60 E-box region of the β -globin promoter. Crusselle-Davis et al. (269) showed that USF1 functions as a β -globin gene activator. A reduction in USF activity reduced β -globin expression and led to a significant decrease in acetylated H3, RNA Pol II, and cofactor recruitment to the LCR and to the adult β -globin gene. On the contrary, TFII-I functions as a β -globin gene repressor possibly by recruiting co-repressor complexes to the β -globin gene.

Components of Chromatin Remodeling Complexes

Multiple components of chromatin remodeling complexes have been found at the β -globin gene promoter and at the LCR. The SWI/SNF complex is recruited to the β -globin promoter by EKLF transcription factor (233). Purification studies of the complexes binding to the pyrimidine rich sequence upstream of the δ -globin gene have revealed an Ikaros-containing SWI/SNF protein complex called PYR complex (145). Further purification of this PYR complex showed that Ikaros is also associated with the NuRD complex and recruits both SWI/SNF (nucleosome remodeling ATPase) and NuRD (nucleosome remodeling deacetylase) complexes to this same sequence (270). In addition, SWI/SNF and MeCP1 (histone deacetylase complex) complexes have been shown to bind LCR HS2 *in vitro* (54). Recent ChIP studies showed that two components of the MeCP1 complex, BRG1 and hnRNPC, associate with both LCR HS2 as well as with the γ - γ -, $A\gamma$ -, and β -globin gene promoters. Another SWI2/SNF2-related DNA dependent ATPase, HLTF, is recruited to the -115 to -140 region of the β -globin promoter (271). Large 2Mda protein complexes that associate with the LCR and with the β -globin promoter have recently been fractionated from K562 nuclei (54,271). It is tempting to speculate that several of these multiple protein complexes pre-exist *in vivo* and that one role of the LCR is to act as an organizer or bridge to assemble giant protein complexes from pre-existing large multiple protein subcomplexes present at the LCR and the globin gene promoters.

CTCF

CTCF is a zinc finger containing DNA-binding protein, which binds to specific insulator elements. The enhancer blocking activities of chicken HS4 insulator depends on CTCF (66). CTCF also binds to a sequence located downstream of the β -globin locus and to both 5' and 3' of mouse and human β -globin loci. These regions vary in insulator strength, and their function *in vivo* is unclear (272). CTCF's binding sites at 5' HS5 and 3' HS1 in the mouse β -globin locus

participates in ACH formation at both early and late developmental stages. However, as mentioned before deletion of these sites do not affect globin gene expression (155).

The Role of the LCR in Nuclear Localization

In the interphase nucleus, the individual chromosomes form discontinuous structures called chromosome territories (CTs). Transcription occurs at the surface of CTs. The space between CTs is called interchromatin compartment (IC). CTs may represent the relatively more condensed domain of chromosomes, hosting inactive or inaccessible genes. Actively transcribed genes are brought to the surface of CTs where they become accessible to protein complexes located in the ICs. Transcription complexes are believed to be located in the ICs where they form transcription factories (273).

The LCR might play a role in the organization of genes within nuclei. FISH analysis of the localization of the endogenous murine β -globin locus from different mature stage fetal liver cells showed that the β -globin LCR is required for efficiently re-locating the β -globin locus away from the nuclear periphery into the nuclear interior (274). It was also shown that the LCR is required to position the β -globin locus close to transcription factories containing Serine 5 hyperphosphorylated RNA Pol II. Interestingly, β -globin transcription initiates at the nuclear periphery but transcription levels increase along with the globin locus moving toward the nuclear interior. These results suggest that the LCR plays an important role in moving the β -globin locus to transcriptionally active RNA Pol II complexes.

Summation

The human and murine β -globin gene loci have been model systems for the study of eukaryotic gene regulation and chromatin structure for a long time. The tissue- and developmental stage specific expression of the five β -like globin genes is regulated by a locus

control region (LCR), which is composed of several DNase I hypersensitive sites in erythroid cells.

Previous studies have shown that the LCR interacts with many proteins involved in transcription, such as RNA Pol II, TBP, NF-E2, USF, TFII-I etc. (49,55,269), and the LCR is required for high-level β -globin gene expression (75). The trans-acting factors also play an important role in β -globin gene regulation. It has been shown that a variety of transcription factors and chromatin modification enzymes can be recruited to the β -globin gene locus directly and indirectly during erythroid cell maturation (145,275-277). Thus our laboratory has proposed the model that the LCR may function as a transcription organizing center, which recruits transcription factor complexes and RNA Pol II, and regulate β -globin gene expression by transferring these activities to the globin gene promoters (55).

It has been shown previously that the erythroid specific transcription factor, NF-E2, is essential for the recruitment of RNA Pol II to the β -globin gene promoter but not the LCR. The absence of NF-E2 leads to a significant decrease or abolishment of β -globin gene expression (49,228). The ubiquitously expressed transcription factor USF has been shown to be partially required for the association of RNA Pol II at both the LCR and the β -globin gene promoter (53). Over-expression of a dominant-negative form of USF leads to a decrease in β -globin gene expression (269). Interestingly, both NF-E2 and USF are bound to the LCR and the β -globin gene promoter (49,269), and they also interact with multiple histone methyltransferases (such as MLL2 (277), G9a (276), PRMT1 (243), etc.) which may help to open the condensed chromatin structure to facilitate β -globin gene transcription.

The goal of this study is to examine the mechanism(s) by which the LCR assembled elongation competent transcription complexes at the globin gene promoters. A particular focus

will be the examination of the relationship between NF-E2 and USF in β -globin locus regulation. Finally, it has been reported that TFII-I functions as an inhibitor of calcium entry into the cells (278), an additional focus of this work is to examine if calcium entry has effects on β -globin gene regulation.

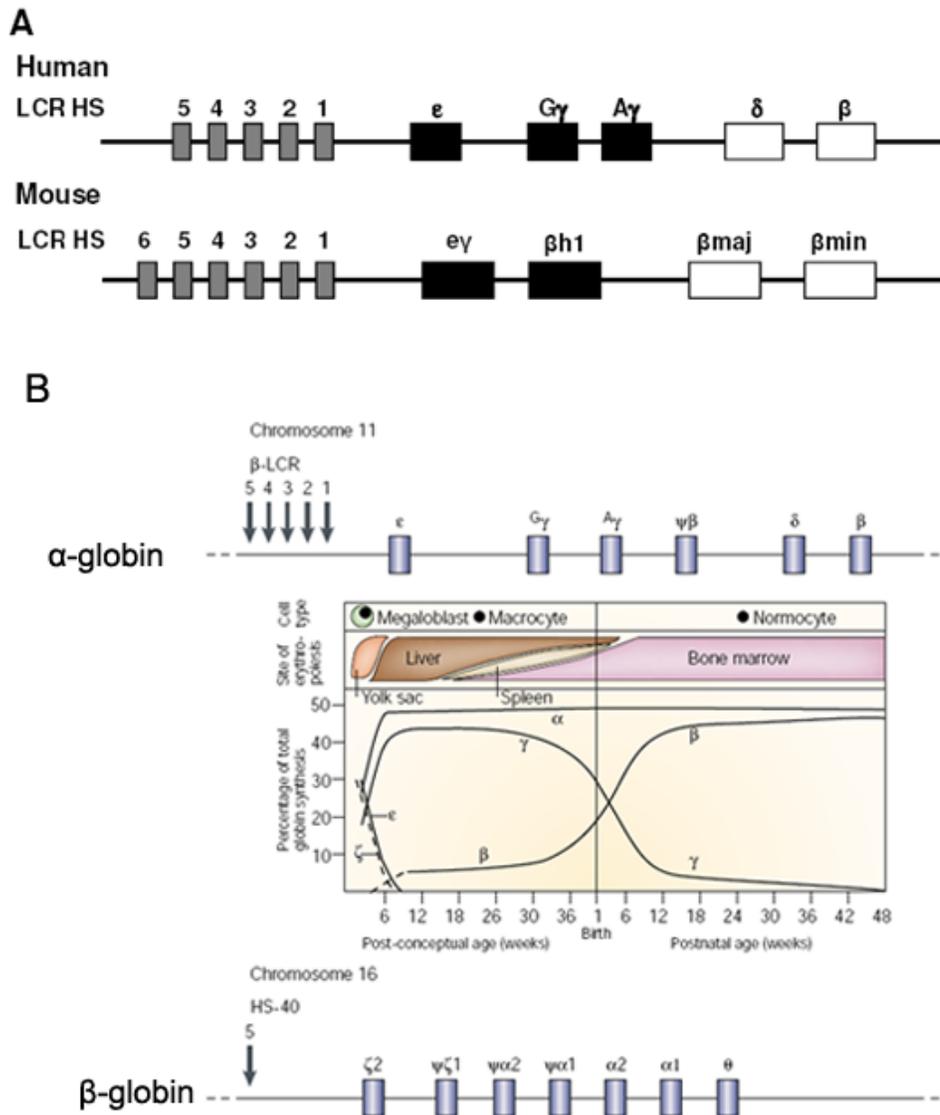


Figure 1-1. Schematic of the structure of the human and mouse β -globin gene loci and hemoglobin synthesis during development. (A) Structure of the human and mouse β -globin gene loci. The β -globin gene loci consist of five (human) or four (mouse) genes that are expressed during development. The murine $\epsilon\gamma$ - and $\beta h1$ -globin genes are expressed in embryonic erythroid cells and the βmin - and βmaj -globin genes are expressed in fetal and adult erythroid cells (152). The human ϵ -globin gene is expressed at the embryonic stage, the $G\gamma$ - and $A\gamma$ -globin genes are expressed at the fetal stage, and the δ - and β -globin genes are expressed at the adult stage. All the genes are regulated by a locus control region (LCR) located far upstream of the genes and composed of several erythroid-specific DNase I hypersensitive (HS) sites. (B) Developmental stage-specific sequential activation of human embryonic, fetal and adult globin genes. ((B) is adapted from Weatherall *et. al.* (5)).

CHAPTER 2 MATERIALS AND METHODS

Cell Culture and Protein Isolation

MEL cells were grown in Dulbecco's Modified Eagle Medium (DMEM) (Cellgro) supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin antibiotic mixture. CB3 and CB3/NF-E2 (stable CB3 cell line expressing ectogenic NF-E2 p45) cells were obtained from Dr. Paul Ney (St. Judes Children Research Hospital) and grown in DMEM (Gibco, cat. # 11885) supplemented with 10% fetal bovine serum, 200mM L-glutamine, and 1% penicillin/streptomycin antibiotic mixture. For DMSO induced cells, 1×10^5 cells/ml MEL or CB3 cells were incubated with 1.5% DMSO for 72 hours. K562 cells were grown in RPMI 1640 medium (Cellgro) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin antibiotic mixture.

Mouse embryonic fibroblasts (MEFs) were grown in DMEM (Gibco) and supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin antibiotic mixture. OP9 cells were grown in OP9 media (α -modified minimum essential media (α -MEM) (Gibco, cat. # 11900-024) supplemented with 20% fetal bovine serum and 2% penicillin/streptomycin antibiotic mixture).

All cells were grown in 5% CO₂ at 37°C.

Nuclear Extracts used in co-immunoprecipitation experiments were prepared as described by Leach et al. (103). Whole Cell Extracts used in the in vitro dissociation assay were prepared as described by Bungert et al. (245). Whole cell extracts used in western blotting experiments were prepared by incubating PBS washed cells with RIPA buffer (50mM Tris-HCl (PH 7.4), 100mM NaCl, 10mM EDTA, 0.25% Na-Deoxycholate, 1% NP-40, 0.1% SDS, and protease inhibitors cocktail, Roché) at 4°C, and spinning slowly on a rotating wheel for 20min. The cell lysate was centrifuged at 14,000rpm for 15min at 4°C and the cell debris was removed.

The cDNA for NF-E2 (murine p45 tethered to human MafG) (103,275) and A-USF (a USF dominant-negative mutant, a gift from Dr. Charles Vinson, NIH) (238) were subcloned into pET-28a(+) and pET-19b vector (Novagen), respectively, and the recombinant His-tagged NF-E2 and A-USF proteins were expressed in *E.coli* and purified by His Band Purification Kit (Novagen) according to the manufacturer's protocol.

All protein concentrations were measured by the Bradford protein assay.

***In vitro* Embryonic Stem Cell Differentiation**

Mouse embryonic stem (ES) cells were *in vitro* differentiated into hematopoietic cells on OP9 stromal cell monolayers as described by Kitajima *et al.* (279). Briefly, ESD3 cells (ATCC, CRL-1934) were maintained on a monolayer of mouse embryonic fibroblasts (MEFs) in ES media (DMEM (high glucose) (Gibco, 12430-054), 15% fetal bovine serum, 0.1mM 2-mercaptoethanol, 1x MEM non-essential amino acids solution (Cellgro, 25-025-CI), 1x EmbryoMax ES cell qualified nucleosides (Chemicon, ES-008-D) and 1% penicillin/streptomycin antibiotic mixture.). Immediately before differentiation assay, ES cells were passaged twice to minimize the presence of MEF cells in the culture and grown on gelatinized flasks in ES media supplemented with 10^3 Units/mL ESGRO (LIF) (Chemicon, ESG1106). Leukemia inhibitory factor (LIF) was added to prevent ES cell differentiation.

At Day 0, $1-2 \times 10^5$ /mL ES cells were seeded onto a confluent monolayer of OP9 cells grown in 6-well plates in OP9 media as described in the "Cell Culture and Protein Isolation" section. From Day 3, erythropoietin (Epo, 1 Unit/mL) was added into the culture medium throughout the rest of the ES differentiation procedure to prevent erythroid progenitors from apoptosis. At Day 5, the differentiated ES cells were treated with trypsin and $1-2 \times 10^5$ /mL cells were seeded onto a new confluent monolayer of OP9 cells grown in 6-well plates. OP9 cells tend to aggregate into cell clumps and precipitate onto the bottom of the cell suspensions. To

remove the OP9 cells from ES/OP9 cell mixture, the trypsinized cells were kept at room temperature for 1-2 min, and only the cells from the supernatant were collected for further seeding. The same procedure was performed for Day 8 cells. $1-2 \times 10^5$ /mL differentiated ES cells were trypsinized and reseeded onto a new confluent monolayer of OP9 cells grown in 6-well plate for further differentiation until Day 12. During the entire ES cell differentiation process, the medium was changed every two days.

At Day 0, 5, 8 and 12, an aliquot of cells were collected for RNA extraction and ChIP analysis. The RT-PCR primers of murine β -actin, Rex-1, $\epsilon\gamma$ - and β maj-globin genes were described previously (280).

Western Blotting

Western blotting experiments were performed as described by Leach et al. (103). A total of 60 μ g of whole cell extract, unless otherwise noted, was loaded onto 4-20% Ready gel (Bio-Rad). The proteins were visualized by ECL Plus chemiluminescence (Amersham Pharmacia Biotech). The following antibodies were used: GAPDH (FL-335; sc-25778), NF-E2 (C-19; sc-291), NF-E2 p18 (MafK) (C-16; sc-477), USF1 (C-20; sc-229), USF2 (C-20; sc-862), TFIIB (C-18; sc-225), CBP (A-22; sc-369), p300 (N-15; sc-584), goat anti-mouse IgG-horseradish peroxidase (sc-2005) (all purchased from Santa Cruz Biotechnology), goat anti-rabbit IgG-horseradish peroxidase (Kirkegaard & Perry Laboratories) and anti-rabbit IgG Trueblot (eBioscience). These antibodies were used for all the western blotting and for the coimmunoprecipitation experiments. The concentration of the antibodies used followed the manufacturers' guidelines.

Chromatin Immunoprecipitation (ChIP)

The ChIP assay was performed essentially as described by Leach et al. with minor modifications (103). After preclearing the diluted cell lysate with Protein A-Sepharose beads (GE Healthcare), 2 μ g of the appropriate antibody was used for each ChIP sample. For ChIP

assays using mouse IgM antibodies, Dynabeads® Rat anti-Mouse IgM (110.39D, a gift from Scott R. Jamison, Invitrogen) was used instead of Protein A-Sepharose beads.

The antibodies used were the same as those described for the western blotting experiments with the additional use of the following antibodies: normal mouse IgM (Santa Cruz), rabbit IgG (Bethyl Laboratories), RNA polymerase II, clone CTD4H8 05-623 (Upstate); RNA polymerase II 8WG16 monoclonal antibody, RNA polymerase II H14 (Ser5P-Pol II) monoclonal antibody (Covance), and RNA polymerase II CTD repeat YSPTSPS (phospho S2) antibody (Abcam, ab5095). Samples were analyzed by quantitative real Time PCR using MyiQ (Bio-Rad). The murine β maj promoter and HS2 primers for ChIP samples used were described by Crusselle-Davis et al. (269). GAPDH and the HS3/2 flanking region were analyzed using primers described by Levings et al. (280). The ChIP primers for murine GAPDH 3' UTR are: US, 5'-GGCTACAGCAACAGGGTGGTGGAC-3', and DS, 5'-GTTGGGGGCCGAGTTGGGATAG-3'. The P-values were determined by three independent experiments and calculated by using Microsoft Excel t-test function. A $P < 0.05$ (in some cases $P < 0.1$) is considered to be statistically significant.

Co-immunoprecipitation (CoIP) and GST Pull-Down Assay

For each CoIP experiment, 1 mg nuclear extract from MEL or CB3 cells was used. The nuclear extracts were precleared by incubating with anti-rabbit immunoglobulin G (IgG) beads (eBioscience) for 30 min. The precipitation was achieved by incubating 2.5 μ g of appropriate antibody with the precleared cell nuclear extract for 2.5 hours on a spinning wheel at 4°C. The protein complexes were captured by incubating the antibody containing cell nuclear extracts with anti-rabbit IgG beads for 2 hours. Then the beads were collected by briefly spinning in a microcentrifuge and washed three times with RIPA buffer. All the incubations were performed at 4°C on a spinning wheel. The immunoprecipitated protein complexes were eluted with

Laemmli buffer (Bio-Rad) at 95°C for 10min, then loaded onto 4-20% Ready Gels (Bio-Rad), and analyzed by western blotting. The antibodies used for CoIP are the same as those used for western blotting experiments, except that Pol II (N-20) sc-899 (Santa Cruz Biotechnology) was used for the Co-IP experiments.

Full length or truncated human USF1 (hUSF1) cDNA was inserted into pGEX-5X-1 vector (Promega) as described by Huang et al. (243). The USF1-N construct was generated by digesting the pGEX-5X-1-USF1 plasmid that contains the full length USF1 with BlnI and XhoI, to delete the 202-310 amino acids of USF1. The USF1-LZ construct was generated by digesting the pGEX-5X-1-USF1 plasmid with NheI and SacI to remove the 262-291 amino acid residues of USF1. GST and GST-hUSF1 fusion proteins were expressed in and purified from *E.coli*. Equal amounts of the proteins were coupled to glutathione sepharose 4 fast flow beads (GE Healthcare, 17-5132-01) according to the manufacturer's instruction. Equal amounts of protein coupled beads were incubated with recombinant NF-E2 (55) in 300µl pull-down binding buffer (50mM Tris-HCl PH8.0, 2mM EDTA, 150mM NaCl, 0.1% NP-40, 0.1µg/µl BSA, 1mM DTT, 1mM PMSF and protease inhibitors cocktail, Roché) for 1 hour at 4°C. The beads were washed 3-5 times with 1mL pull-down binding buffer (without BSA), collected by briefly spinning in a microcentrifuge, boiled with Laemmli sample buffer (Bio-Rad), and loaded onto 4-20% Ready gels (Bio-Rad), followed by western blotting assay with antibody against NF-E2 (C-19) (Santa Cruz, sc-291).

RNA Isolation, Reverse Transcription, and Real-Time PCR

RNA was isolated by using the guanidine thiocyanate method (281) and reverse transcribed by using iScript cDNA synthesis kit (Bio-Rad) according to the manufacturer's protocol. Quantitative PCR (qPCR) was performed using MyiQ (Bio-Rad) and reactions were carried out using iQ SYBR green super mix (Bio-Rad). Real-Time PCR for ChIP samples and

cDNAs was performed and analyzed as described by Crusselle-Davis et al. (269), except that for cDNA samples, the PCR conditions described above were modified as follows: 95°C for 5min, 40 cycles of 94°C for 30s, 60°C for 30s, 72°C for 1min, final extension at 72°C for 5min. β -actin was used as endogenous control for RT-qPCR. The qPCR primers used for amplifying the murine β maj-globin gene were the same as previously described by Levings et al. (280). The RT primers for β -actin are: US, 5' - GTGGGCCGCTCTAGGCACCA-3', and DS, 5' - TGGCCTTAGGGTGCAGGGGG-3'.

***In vitro* Transfer/Dissociation Assay**

The *in vitro* Pol II transfer/dissociation analysis was performed as described by Vieira *et al.* (55), with minor modifications. Briefly, a plasmid containing the wild-type human β -globin LCR was linearized and immobilized on streptavidin-coated magnetic beads as described by Leach *et al.* (53). 1.5 μ g of the immobilized LCR was incubated with 500 μ g of MEL whole cell extracts for 30min at 30°C. 1 μ g of wild-type (pRS β) (103) or mutant human β -globin gene construct was added in experiments in which dissociation was assayed in the presence of DNA templates. The three β -globin gene mutants, pRS β INImut (INImut), pRS β +60E-boxmutb (+60E-boxmut), and pRS β MAREa (NF-E2mut1), were described previously by Leach *et al.* (103). The β -globin gene construct lacking the promoter (β -) was generated by deleting a ~1.2 Kbp region at the 5' end of the gene. After incubating beads-LCR-protein-complexes with β -globin gene constructs for 30 min at 30°C, the tubes were placed on a magnetic device and the supernatant was collected from each tube, and subjected to western blotting analysis. In some experiments, 60ng of recombinant NF-E2, A-USF, or BSA was added with or without the β -globin gene construct during the dissociation reaction. The dissociation assays were performed following completely random experimental design (i.e. all β -globin gene constructs and recombinant proteins were added to a random selected tube that contain LCR/protein complexes). The experiments were repeated at

least three times and the results were reproducible. To quantitate the efficiency of the transfer of Pol II from the LCR to the β -globin gene template, I modified the above described procedures as follows. After incubating the LCR coupled streptavidin beads with MEL nuclear extracts and removing unbound material, each sample was equally divided in half. One half of each sample was boiled with Laemmli sample buffer (Bio-Rad), loaded onto 4-20% Ready gel (Bio-Rad), and then analyzed by western blotting using an antibody against Pol II (clone CTD4H8, 05-623, Upstate). The other half of each sample was subjected to an *in vitro* transfer assay as described previously (55). The supernatants, plus or minus the β -globin template, were collected at the end of the assay and cross-linked with 0.5% formaldehyde at room temperature for 10min. Then, 0.125M glycine was added to each sample, incubating at room temperature for 5min to stop the cross-linking reactions. All samples were dialyzed against ChIP dilution buffer, and then subjected to immunoprecipitation (IP) analysis using the CIP assay described previously with antibodies against Pol II (clone CTD4H8 05-623, Upstate) and rabbit IgG (Bethyl Laboratories). Precipitation of the β -globin gene template was monitored by quantitative Real-Time PCR. The primers for human β -globin gene are: US, 5'-ATTGCATCAGTGTGGAAGTC-3', and DS, 5'-ATTGCCCTGAAAGAAAGAGATTAG-3'.

CHAPTER 3 REGULATION OF β -GLOBIN GENE EXPRESSION BY CIS- AND TRANS- REGULATORS

Introduction

The human β -globin gene is located on chromosome 11 and consists of five β -like globin genes (5'- ϵ -G γ -A γ - δ - β -3'), which are expressed in erythroid cells in a tissue and developmental stage-specific manner (Figure 1-1A). The ϵ -globin gene is expressed in the embryonic yolk sac blood islands; the two γ -globin genes (G γ and A γ) are expressed in the fetal liver; and the δ - and β -globin genes are expressed during adult bone marrow hematopoiesis (1). The mouse β -globin gene locus is located on chromosome 7. It is homologous to the human β -globin gene locus and consists of five β -like globin genes (5'- $\epsilon\gamma$ - β h1- ϕ - β maj- β min-3'). The $\epsilon\gamma$ - and β h1-globin genes are expressed at the embryonic stage, and the β maj- and β min-globin genes are expressed predominantly after birth.

The locus control region (LCR) located at about 50kb upstream of the human β -globin gene functions as a powerful enhancer to regulate expression of the β -like globin genes during development. Transgenic mice studies have shown that in the absence of the LCR, human γ - and β - globin expression levels are low relative to those of the endogenous mouse globin genes (57-59). However, when the LCR (containing HS1-HS5) is linked to the human β -globin gene, all resulting transgenic mice lines express the gene at a level comparable to the endogenous β -globin gene (60). In addition, the Hispanic thalassemia deletion (deletion of HS2-5 and ~20 kb upstream of this region) leads to an alternation of the general DNase I sensitivity of the β -globin locus (45). These results suggest that the LCR not only functions as an enhancer, but also carries chromatin domain opening activity. The LCR contains several DNaseI hypersensitive sites (HS), which provide multiple binding sites for a variety of transcription factors. Among all of the HS sites within the LCR, HS2 exhibits the most powerful enhancer activity (67). LCR HS2 contains

multiple DNA motifs to provide binding sites for a variety of transcription factors, including two tandemly repeated NF-E2 binding sites, one GATA-1 binding site, one EKLF binding site, and two E-box elements, which serve as binding sites for helix-loop-helix proteins such as Tal1, USF and TFII-I (49). LCR HS3 also has enhancer activity and one study has shown that it preferentially activates ϵ - and γ -globin gene expression (68). There is also one NF-E2 binding site and a tandem GATA-1 binding site in LCR HS3. Mutations within these NF-E2 and/or GATA-1 binding sites in either HS2 or HS3 all lead to a decrease in downstream globin gene expression and loss of or reduced DNase I hypersensitivity (219).

The ubiquitously expressed upstream stimulating factor (USF) has been shown to function as an activator for adult β -globin gene expression in mouse erythroleukemia cell lines (269). USF belongs to the basic-helix-loop-helix-leucine zipper (bHLH-LZ) family, which interacts with the E-box elements with high affinity (234). USF has two homologous family members - USF1 and USF2. These proteins have very similar C-termini, which contain the basic region, the helix-loop-helix, and the leucine zipper domain, and serve as a protein-protein and protein-DNA interaction domain. The N-terminus of USF1 and USF2 contains the transactivation domain and is less conserved between these proteins. Considering that USF1 and USF2 can interact with E-boxes as hetero- and homo-dimers, the difference in their N-termini may provide unique regulatory functions. In the β -globin locus, it has been shown that USF1 and USF2 interact with the +60 E-box element of the human β -globin gene promoter at the adult stage, whereas USF2 together with TFII-I interact with the initiator (Inr)/ E-box element of the β -globin gene promoter at the embryonic stage (103). In addition, when overexpressing a dominant-negative form of USF (named A-USF) in MEL cells to sequester USF from binding to the DNA, the adult β major globin gene expression level decreased. Complementarily, when overexpressing an ectopic USF1

in MEL cells, there was a dramatic increase in the β maj-globin gene expression (269). Moreover, USF1 interacts with histone acetyltransferase (HAT) p300 in MEL cells which express adult β -globin gene, but not in human K562 cells which express the embryonic ϵ - and fetal γ -globin genes (282). These observations suggest that USF functions as an activator for adult β -globin gene expression.

Another helix-loop-helix protein, TFII-I, is also involved in the regulation of β -globin gene expression. TFII-I is a basal transcription factor that cooperates with USF to bind to the initiator (Inr) /E-box elements (103). Both K562 (embryonic stage erythroid cells) and MEL (adult stage erythroid cells) cells transfected with dominant-negative TFII-I show increased levels of β -globin expression. Ectopically expressing TFII-I in MEL cells leads to a decrease in β -globin gene expression. Interestingly, in K562 cells, USF1 mainly interacts with TFII-I, but not with its homologous family member USF2; while in MEL cells, USF1 primarily interacts with USF2, which is accompanied by decreased interactions between USF1 and TFII-I (269). Together with the observation that TFII-I interacts with the Inr/E-box motif at the β -globin gene promoter in embryonic stage K562 cells (103), all of the facts described above imply that TFII-I functions as an inhibitor for adult β -globin gene expression. TFII-I has several isoforms, but currently, their exact functions are still unknown.

The association of RNA polymerase II (RNA Pol II) at the promoter of a gene is an essential marker of gene transcription. It has been shown that RNA Pol II associates with specific enhancer elements prior to transcription of the genes that the enhancer regulates (280). The enhancer recruited RNA Pol II may be transferred to the gene promoter. It has been shown that the transcription pre-initiation complex (PIC) is required for recruiting RNA Pol II to the gene promoter (162,163). The PIC component TFIIIB helps the correct positioning of RNA Pol II at

the gene promoter, and TFIID interacts with RNA Pol II C-terminal domain (CTD) to stabilize it on the gene (283). In the β -globin gene locus, it has been shown that RNA Pol II is recruited to the LCR HS sites *in vitro* and *in vivo* (53,78,284,285), as well as to the promoters of the globin genes.

Histone modifications serve as marks for accessible or condensed chromatin. Generally, acetylated histones indicate an open chromatin environment. Histone acetyltransferases (HATs) add acetyl groups to histone tails, rendering the chromatin more accessible. On the contrary, a histone deacetylase (HDAC) can remove acetyl groups from histone tails, rendering the chromatin less accessible (286). The methylated histone residues, such as H3K9, K27, and H4K20 are normally associated with heterochromatin, which leads to transcription repression; while the methylated residues H3K4, K36, and K79 have been shown to associate with euchromatin, which correlates with transcription activation (287). It has been found that TFII-I interacts with HDAC3, a histone deacetylase, at the adult β -globin gene promoter in K562 cells where the β -globin gene is repressed (269); while USF1 interacts with p300, a histone acetyltransferase, at the adult β -globin gene promoter in MEL cells (282), where the β -globin gene is actively transcribed.

Embryonic stem (ES) cells are derived from the inner cell mass (ICM) of day 5-8 blastocysts (288,289) or morula-stage embryos (290). ES cells are pluripotent cells that are capable of differentiating into cells of all lineages. ES cells are able to maintain a karyotypically stable prolonged self-renewal status and have the potential to be differentiated into cells of the three germ layers, both *in vitro* and *in vivo* (288,289,291). The *in vitro* differentiation capacity of the ES cells is limited, however, they can be induced *in vitro* to differentiate into hematopoietic cell lineages, and represent a powerful tool for investigating hematopoietic differentiation and

development (279). Numerous studies have attempted to understand the mechanisms of ES cell unique properties, however, except for the identification of the pluripotency-associated transcription factors OCT4, NANOG and SOX2 as markers for ES cells (292), not much is known about the mechanisms regulating pluripotency. Recent studies indicate that epigenetic regulatory mechanisms may regulate the self-renewal, pluripotency, and lineage-specific differentiation of ES cells (293). The genome of ES cells is mostly in an open and accessible chromatin configuration, but upon ES cell differentiation, a more compact heterochromatic structure forms in these cells (294).

Polycomb Group (PcG) proteins act as repressors of transcription. They were firstly purified from *Drosophila* embryos and found to be required for maintaining the inactive state of homeotic and other important regulators during development (295). There are two polycomb repressor complexes, PRC1 and PRC2. The mammalian PRC2 complex contains EZH2, EED and Suz12 (296). The EZH2 protein is a histone methyltransferase (HMTs) that adds methyl groups to lysine 27 of histone 3 (H3K27) and to a much lesser extent to H3K9 (296-298). It has been shown that EED is required for EZH2 mediated H3K27 methylation (299) and Suz12 is associated with developmentally important transcription factors, H3K27 trimethylation (H3K27me₃) and gene repression (300). The methylated H3K27 serves as a docking site for PRC1 (301). Therefore, a current model for PcG mediated gene silencing proposes that the PRC2 complex first adds H3K27 methylation, then the methylated H3K27 mark interacts with the PRC1 complex and recruits it to the gene locus, and finally the PRC1 complex further stabilizes the repressive chromatin structure by catalyzing mono-ubiquitinylation of histone H2A at lysine 119 (H2Aub1) and thereby impeding RNA Pol II elongation (302),(303). H3K27me₃ has been found at the promoters of silent-lineage-associated genes in mouse ES cells (304).

By using an *in vitro* embryonic stem (ES) cell differentiation system I have shown that during differentiation along the erythroid lineage, RNA Pol II is recruited to the β -globin gene LCR prior to its association with the downstream globin gene promoters. Also, along with ES cell differentiation into the erythroid lineage, the association of USF1 and USF2 to the LCR and β -globin gene promoter increased, whereas the association of TFII-I to these regions decreased when the erythroid lineage cells were transitioning to the adult stage. It is further shown here that TFII-I interacts with Suz12, a component of the PRC2 complex. The data support the hypothesis that the LCR serves as the primary site of transcription complex recruitment in the globin gene locus. The data further suggest that TFII-I contributes to the formation of an inaccessible chromatin structure around the β -globin gene promoter in undifferentiated and perhaps embryonic cells.

Results

Our laboratory has established a system by which mouse embryonic stem cells differentiate into globin gene expressing erythroid cells. The protocol was originally developed in the laboratory of Dr. Nakano and is based on the growth of mouse ES cells on a confluent layer of OP9 stromal cells in the presence of erythropoietin (Epo) (279). Stromal cell lines can support the proliferation and differentiation of hematopoietic cells. The OP9 cells used here lack expression of macrophage colony-stimulating factor (M-CSF) (305), thus avoiding macrophage proliferation and allowing efficient production of erythroid cells. Leukemia inhibitory factor (LIF) is required to maintain the undifferentiated state and pluripotency of ES cells. LIF is a member of interleukin-6 (IL-6) family of cytokines, and is known to bind its transmembrane receptor, LIFR, which heterodimerizes with the signal transduction receptor, gp130. The pluripotency of ES cells depends on the intracellular signal transduction pathway activated by the binding of LIF to its receptor, including the phosphorylation by the Janus family of tyrosine

kinase (JAKs) (306). Mouse embryonic fibroblasts (MEFs) secrete LIF, IL-6 and other factors known to maintain ES cell pluripotency. Erythropoietin (Epo) is a major growth factor for erythroid cells. Epo interacts with its receptor EpoR, a cell surface receptor expressed in erythroid, megakaryocytic, and mast cells, activating signal cascades responsible for cell proliferation, differentiation and survival of erythroid progenitors (212).

The ES cells were induced to differentiate along the erythroid lineage in the presence of Epo as described in the Material and Methods section. At defined time points, the cells were subjected to gene expression analysis using RT-PCR, and to protein-chromatin interaction analysis using ChIP assays.

It is shown that undifferentiated ES cells express neither embryonic $\epsilon\gamma$ - nor adult β maj-globin genes, while they do express the Rex-1 gene, a marker for early development (Figure 3-1). When induced to differentiate, the embryonic $\epsilon\gamma$ - and adult β maj-globin genes are sequentially activated in a developmental manner such that the $\epsilon\gamma$ -globin gene is expressed as early as Day 5 of induction, whereas there is no or low level of β maj-globin gene expression detectable at that time point. The β maj-globin is expressed at a higher level at Day 8 and becomes predominant at Day 12, coinciding with a gradually decreased expression level of the $\epsilon\gamma$ -globin gene at these later stages of differentiation. The expression levels of the Rex-1 gene, relative to that of the β -actin gene, at Day 12 is reduced compared to Day 5 and Day 8. The fact that expression of the Rex-1 gene is still detectable at Day 12 is probably due to unsynchronized ES cell differentiation, where a little portion of the cells are still in an undifferentiated stage even at Day 12, or that there is a fraction of differentiated cells that fail to silence the Rex-1 gene *in vitro*. Overall, the RT-PCR data show that the ES/OP9 *in vitro* differentiation system used in this study is a powerful tool for analyzing gene regulation during erythroid differentiation.

Next we examined the interactions of RNA Pol II, general transcription factors and histone modifications with the β -globin gene locus during erythroid lineage-specific ES cell differentiation (Figure 3-2). Since it has been shown that HS2 and HS3 carry most of the enhancer activity of the LCR, we examined the association of the transcription regulatory proteins with LCR HS2 and HS3. We found that histone modifications that mark permissive chromatin such as dimethylated H3K4 (H3K4me2) and acetylated histone H4 (AcH4) are already present at the LCR HS2 and HS3 core element regions in the undifferentiated cells (Day 0), whereas no such modifications are detectable at the promoters of downstream $\epsilon\gamma$ - and β maj-globin genes. During differentiation, these histone marks are detected at both of the promoters of $\epsilon\gamma$ - and β maj-globin concurring with the expression of these globin genes (Day 5 and Day 8). The observation that AcH4, but not as much H3K4me2, is detected at the β maj-globin gene promoter at Day 5 of differentiation when the gene is silenced or expressed at very low level, suggests that the chromatin structure surrounding the β maj-globin gene at this time point is already poised for activation. The above described histone modifications at the LCR at all differentiation stages are limited to the HS sites (HS2 and HS3) and are not found in regions located between the HS sites, such as the region between the HS2 and HS3 core elements (HS3/2 flank). This is consistent with observations from chromosome conformation capture (3C) experiments indicating that the LCR HS sites interact with each other during erythroid maturation (149).

Consistent with the distribution of the transcriptionally active chromatin marks, RNA Pol II and general transcription factor TBP firstly interact with the LCR HS2 and HS3 but not with the $\epsilon\gamma$ - and β maj-globin gene promoters in the undifferentiated ES cells. During differentiation, the association of RNA Pol II is only detected at the embryonic $\epsilon\gamma$ -globin gene promoter but not

at the adult β maj-globin gene promoter at Day 5. RNA Pol II and TBP are both detected at the promoters of $\epsilon\gamma$ - and β maj-globin genes at Day 12, coinciding with the expression pattern of these globin genes.

At every stage of differentiation there is a small fraction of OP9 and MEF cells present in the culture taken for RT-PCR or ChIP analysis. Thus, I performed ChIP assays on MEF and OP9 cells to examine the association of RNA Pol II and active histone marks with the globin gene locus (Figure 3-3). We observed that in both of the MEF and OP9 cells there is no RNA Pol II association with LCR HS2, $\epsilon\gamma$ - and β maj-globin gene promoters. As a positive control, I detected the interaction of RNA Pol II with the promoter of a housekeeping gene, glyceraldehyde 3-phosphate dehydrogenase (GAPDH) in these cells. Also, I used the human necdin gene, which is exclusively expressed in brain, as the negative control, and the data show that there is no RNA Pol II associated with the necdin gene promoter. However, I do detect low levels of the association of H3K4me2 throughout the β -globin locus in MEF and OP9 cells. Considering the significantly low number (less than 10%) of MEF and OP9 cells present in the undifferentiated or differentiated ES cell cultures, I believe that the H3K4me2 levels detected at the β -globin locus are mainly contributed by the ES cells.

Previous studies have shown that the helix-loop-helix transcription factors USF and TFII-I antagonistically regulate β -globin gene expression in K562 and MEL cells, indicating that USF functions as an activator, whereas TFII-I functions as an inhibitor for adult β -globin gene expression (269). Here, by using the ES/OP9 *in vitro* differentiation system to differentiate ES cells up to Day 8, I examined the interactions of USF and TFII-I with the β -globin LCR HS2 and the β maj-globin gene promoter throughout differentiation (Figure 3-4). I analyzed H3K4me2 levels as a positive control for the ChIP assays. I observed that this histone modification which

marks transcription competent chromatin is detectable at the LCR HS2 at all differentiation stages. The H3K4me2 association with the β maj-globin gene promoter is at relatively low levels compared to that of LCR HS2 in the undifferentiated and early differentiated (Day 0 and Day 5) cells. But upon activation of the adult β maj-globin gene at Day 8, the H3K4me2 mark is present at the β maj-globin gene promoter at high levels. At Day 5, when only the embryonic $\epsilon\gamma$ -globin gene is transcribed, USF1 only associates with the LCR HS2 but not with the adult β maj-globin gene promoter. At Day 8, when the adult β maj-globin gene is activated, USF1 associates with the LCR HS2 and with the β maj-globin gene promoter. Interestingly, USF2 interacts with the LCR and the β maj-globin gene promoter throughout differentiation, suggesting that USF1 and USF2 exert different functions during erythroid differentiation. Also, I observed that TFII-I only interacts with LCR HS2 and the β maj-globin gene promoter at Day 5, where the embryonic $\epsilon\gamma$ - but not the adult β maj-globin gene is expressed. I was not able to detect interactions between TFII-I and the globin locus at Day 0 or Day 8 cells, suggesting that TFII-I may be involved in preventing premature β -globin gene expression. The data support the hypothesis that USF is an activator for adult β -globin gene expression, while TFII-I is a repressor. Also, consistent with the results from previous *in vitro* binding assays (103), here I show that USF and TFII-I bind to the LCR and adult β -globin gene promoter in the context of intact cells.

By using affinity chromatography, it has been shown that the PRC2 complex associates with histone deacetyltransferases (HDAC) (296). Our lab previously showed that TFII-I and HDAC3 colocalize at the adult β -globin gene promoter in K562 cells, which represent embryonic-like erythroid cells (269). The protein interaction between TFII-I and HDAC3 was further confirmed by co-immunoprecipitation (CoIP) assays and this interaction only occurred in K562 cells, but not in MEL cells which express the adult β -globin gene upon dimethyl sulfoxide

(DMSO) induction. To extend these studies, I wished to examine interactions between TFII-I and components of the PRC2 complex. Thus, I performed CoIP assays by using whole cell extracts from K562 cells to and examined interactions between TFII-I and Suz12, a component of PRC2 (Figure 3-5). Indeed, when using antibodies against Suz12 to pull down its associated protein complex in K562 whole cell extracts, I did detect TFII-I in the immunoprecipate. I also used an unspecific antibody against IgG as a negative control and showed that there was no interaction between IgG and TFII-I.

Since the PRC2 complex acts as a repressor of gene transcription by methylating lysine 27 at histone 3 (H3K27) (296-298) and previous data from our laboratory show that TFII-I functions as a repressor of adult β -globin gene expression, I therefore examined the interactions of TFII-I, Suz12 and trimethylated H3K27 (H3K27me3) with the adult β -globin gene promoter during *in vitro* differentiation of erythroid cells (Figure 3-6). I found that after induction of erythroid differentiation, low levels of the adult β maj-globin gene was expressed in Day 5 cells. After 12 days of differentiation, there was more than a ~30 fold increase in the expression levels of the adult β maj-globin gene compared to Day 5 cells (Figure 3-6A, B). I observed that TFII-I, Suz12, and the repressive histone modification mark H3K27me3 are associated with the adult β maj-globin gene promoter in Day 5 cells but not in Day 12 cells (Figure 3-6C). By performing quantitative Real-Time PCR, I further show that the decreased association of Suz12 with β maj-globin gene promoter from Day 5 to Day 12 is significant (Figure 3-6D). In addition, as a positive control, I found that the H3K4me2 levels associated with LCR HS2 at Day 5 and Day 12 did not change. I observed that there was no TFII-I or Suz12 association with the negative control HS3/2 flanking region (Figure 3-7). As has been shown previously (Figure 3-2) this

region also does not interact with RNA Pol II, TBP, or permissive histone marks such as H3K4me2 and AcH4 during the *in vitro* differentiation of erythroid cells.

Discussion

ES cells are pluripotent and capable of differentiating into every somatic cell type. Therefore, they must have the ability to initiate transcription profiles for all of these different cell types and maintain their self-renewal properties at the same time. Various studies aimed at understanding the molecular mechanisms of pluripotency have been done, but our knowledge about this process is very limited. Recent studies suggest that epigenetic mechanisms contribute to the unique characteristics of ES cells. Epigenetic regulatory mechanisms include histone modifications, DNA methylation, ATP-dependent nucleosome remodeling, incorporation of histone variants, etc, through which the gene expression profiles are tightly and dynamically regulated.

Numerous studies have shown that epigenetic alterations play a prominent role during ES cell differentiation and mammalian development. Consistently, the reprogramming of somatic nuclei into pluripotent nuclei is associated with large scale epigenetic modifications (307). Indeed, it has been demonstrated that the chromatin structure of mouse ES cells is hyperdynamic, in which the core histone proteins loosely associate with DNA. These ES cells harbor a euchromatic chromatin environment, allowing them to be highly permissive for gene expression. On the contrary, upon development and differentiation, the genome structure becomes more condensed and heterochromatic, causing the loss of pluripotency (308).

The results of the present studies show that ES cell differentiation into erythroid cells *in vitro* is associated with changes in histone modification patterns that correlate with the transcription of erythroid specific genes (Figure 3-2). In undifferentiated ES cells (Day 0), there is already an open chromatin environment at the LCR, but not at the downstream globin gene

promoters, which is indicated by the association of active H3K4me2 and AcH4 marks at the LCR core HS sites. RNA Pol II and the general transcription factor TBP associate with the LCR but not with the globin gene promoters in undifferentiated cells (280). These observations suggest that as early as Day 0, the LCR is permissive for gene transcription. In the later stages of ES cell differentiation (Day 5 and Day 12), we observe that the active chromatin marks (especially H3K4me2) as well as RNA Pol II and TBP are present at the downstream $\epsilon\gamma$ - and β -globin gene promoters in a temporal manner coinciding with different developmental stages. It has been shown that there are intergenic transcripts that initiate within the LCR HS2 core element and proceed unidirectionally towards the globin genes (284). The findings of intergenic transcripts led investigators to propose that LCR recruited RNA Pol II tracks along the globin locus, thereby opening chromatin structure, and finally reaches the globin genes to generate globin mRNAs. Recent data from 3C experiments show interactions between the LCR and globin gene promoters in differentiated erythroid cells. Also, the LCR is relocated from within inaccessible chromatin territories (CT) to the surface of these heterochromatic regions during differentiation (309). Therefore, our observations that active histone modification marks and RNA Pol II as well as TBP associate with the LCR in undifferentiated ES cells suggests that the LCR is always associated with interchromosome compartments (IC), which is the space between CTs that is devoid of chromatin and enriched for RNA Pol II. The intergenic transcription detected in the LCR at early stages of differentiation could maintain the position of the LCR in the IC and if transcription proceed all the way to the globin genes it could reel the genes into the IC.

Previous studies in our laboratory showed that USF interacts with LCR HS2 and with the adult β -globin gene promoter and functions as activator of adult β -globin gene expression (269).

Overexpression of USF1 in mouse erythroleukemia (MEL) cells that only express the adult β -globin gene leads to an increase in adult β maj-globin gene expression, whereas introducing a dominant-negative form of USF1 (A-USF) in MEL cells cause a reduction in β maj-globin gene expression. The decrease of β maj-globin gene expression in A-USF transfected MEL cells is accompanied by a reduced association of RNA Pol II and histone acetyltransferase p300 at both LCR HS2 and the β maj-globin gene promoter. Felsenfeld's laboratory has shown that USF interacts with the coregulators p300, CBP, SET7/9 and PCAF and possibly recruits these factors to the cHS4 insulator in the chicken β -globin locus (64). Our laboratory has found that p300 and CBP interacts with LCR HS2 and the β maj-globin gene promoter in MEL cells (282). In human erythroleukemia K562 cells, which only express embryonic stage globin genes, p300 and CBP only interact with LCR HS2, but not with the globin gene promoters. Further data derived from CoIP assays show that USF1 only interacts with p300 in MEL cells, but not in K562 cells. The fact that histone acetyltransferase p300 can add acetyl groups to lysine residues of N-terminal histone tails suggests that USF1 recruits p300 to the adult β -globin gene promoter to facilitate chromatin opening and transcription activation. In K562 cells where the adult β -globin gene is silenced, the recruitment of p300 and CBP to LCR HS2 are possibly mediated by GATA-1 and NF-E2, both of them associated with CBP (49,199).

Our laboratory also previously identified TFII-I as a repressor of adult β -globin gene expression (269). By introducing ectopic TFII-I in K562 cells, there is a decrease in adult β -globin gene expression, whereas overexpressing a dominant-negative form of TFII-I (p70) in K562 cells leads to an increase in β -globin gene expression. Interestingly, TFII-I interacts with a histone deacetylase (HDAC) only in K562 cells, but not in MEL cells, and they co-localize at the adult β -globin gene promoter (269,282). HDACs are known to remove acetyl groups from

histone N-terminal tails to maintain a repressive heterochromatic environment. USF1 interacts mainly with TFII-I in K562 cells and primarily with USF2 in MEL cells. This is consistent with our previous *in vitro* data suggesting that TFII-I and USF2 associate with the adult β -globin gene initiator (Inr)/ E-box element in K562 cells; whereas USF1 and USF2 associate with the +60 E-box in MEL cells in which the gene is transcribed (53,103).

In the present study, I further investigated the mechanism of the repressor activity of TFII-I. Components of PRC2 are expressed at high levels in embryonic tissues and are required for early mammalian development. They occupy a set of developmental genes in ES cells and silence their expression to maintain the pluripotency of ES cells (304). The interaction between TFII-I and Suz12 (Figure 3-5) suggests that at least part of the repressive activity of TFII-I is mediated by the PRC2 complex. This is further supported by the observation that there is an enrichment of H3K27me3 at the adult β -globin promoter during early stages of erythroid *in vitro* differentiation. The PRC2 complex is known to methylate H3K27, which serves as a docking site for recruiting the PRC1 complex to further stabilize repressive chromatin structure by DNA methylation (303). I did not analyze the recruitment of the PRC1 complex to the globin gene locus and believe that PRC1 is likely not recruited because the β -globin promoter is not methylated in embryonic erythroid cells (280). Here I show that during ES cells differentiation into the erythroid lineage, Suz12, H3K27me3 as well as TFII-I interact with the LCR HS2 and with the β maj-globin gene promoter at early stages of differentiation (Day 5) to repress the adult β maj-globin gene expression. The PRC2 mediated repression is removed by an unknown mechanism but could involve increased activity of USF1 competing with TFII-I for heterodimerization with USF2 at Day 12 of differentiation, when adult β maj-globin gene expression becomes dominant in these cells (Figure 3-6). Our observation of TFII-I interacting

with both of HDAC3 and Suz12 is consistent with a previous report, indicating that the PRC2 complex co-purifies with HDAC proteins (296). Our data further suggest that TFII-I act as a repressor for adult β -globin gene expression by recruiting HDAC and PRC2 complexes leading to the formation of repressive chromatin structure at the adult β -globin gene.

The erythroleukemia cell lines used in the previous studies to analyze the regulation of globin genes do not accurately reflect the identities of erythroid cells that differentiate during mammalian development. Therefore, in this study, I examined the association of USF and TFII-I at the β -globin locus during mouse ES cell erythroid differentiation. The ES cell system will likely not recapitulate precisely the steps involved in erythropoiesis as they occur in a living animal but certainly represents an advantage over transformed cell lines (Figure 3-1). The ES cell differentiation experiment shows that when the adult β -globin gene is activated (Figure 3-4, Day 8), USF1 associates with and TFII-I dissociates from the adult β maj-globin gene promoter. In addition, at Day 5 when the embryonic stage globin genes are primarily expressed, I observe that both USF1 and TFII-I associate with LCR HS2 but that only TFII-I interacts with the β maj-globin gene promoter. These results further suggest that TFII-I act as repressor and USF as an activator of adult β -globin gene expression, and furthermore that the LCR recruits not only components involved in activating globin gene expression but also potential repressor activities.

Taken together, our present study is consistent with the hypothesis that the LCR recruits multiple transcription regulatory proteins (including RNA Pol II, histone modification enzymes, general transcription factors and transcription factors required for globin gene regulation) through protein-DNA and protein-protein interactions and organize them to form a transcriptional holocomplex prior to globin gene expression. We propose that the transcription

complexes assembled at the LCR are subsequently delivered to globin gene promoters by a looping mechanism.

In order to understand the mechanism of LCR mediated globin gene activation it will be important to identify transcription regulatory proteins that are required for the transfer of transcription complexes from the LCR to the β -globin gene promoter. Furthermore, the antagonistic function of HLH proteins USF and TFII-I with respect to globin gene regulation needs to be further examined. The identification of interacting proteins will be important to comprehend how these proteins function.

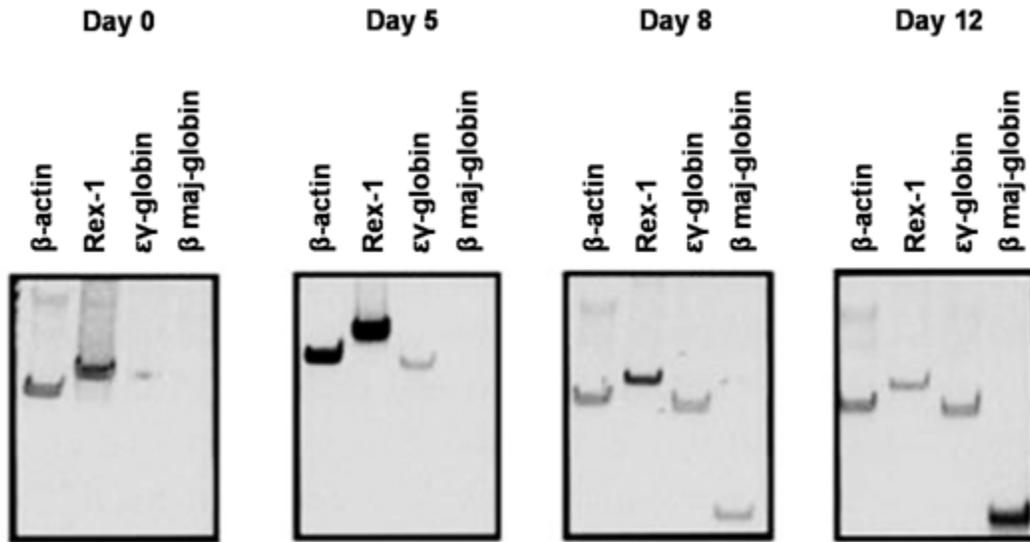


Figure 3-1. Sequential activation of globin gene transcription during *in vitro* erythroid differentiation of murine embryonic stem cells. PCR analysis of DNase I treated and reverse-transcribed total RNA extracted from differentiating embryonic stem cells at the indicated time points. All primer sets span introns, with the exception of Rex-1, and the size of each RT-PCR product is as follows: Rex-1, ~600 bp; β -actin, 480 bp; $\epsilon\gamma$ -globin, 400 bp; β maj-globin, 220 bp. None of the samples showed genomic DNA amplification (not shown). (This ES cell differentiation experiment and RT-PCR are done by Dr. Padraic P. Levings and repeated by me.)

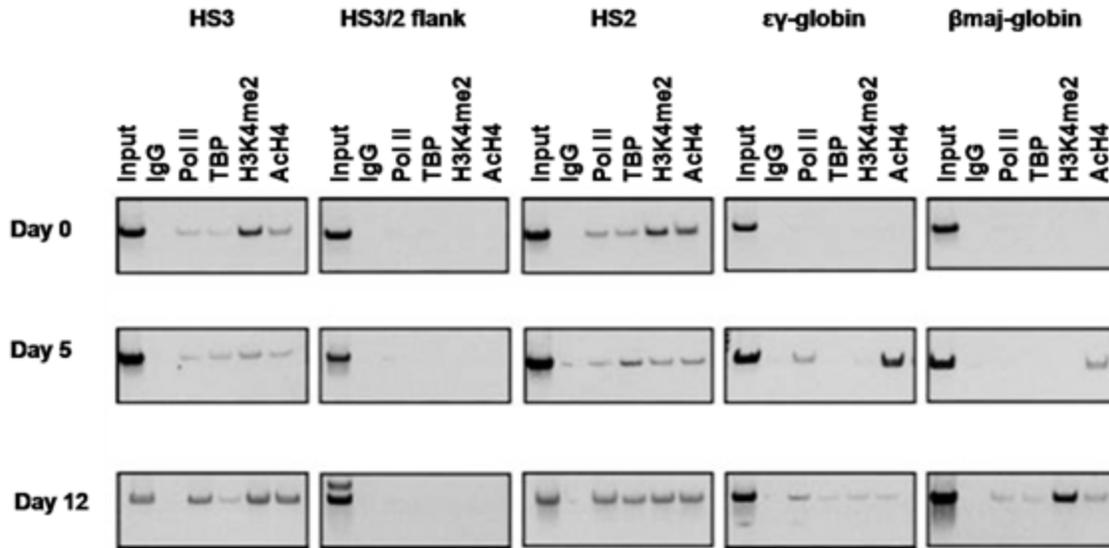


Figure 3-2. Interaction of transcription factors and RNA polymerase II with the β -globin locus during *in vitro* murine embryonic stem (ES) cells erythroid differentiation. Chromatin immunoprecipitation (ChIP) assays were performed on undifferentiated (Day 0) and differentiated (Day 5 and 12) ES cells. The cells were incubated in formaldehyde and the cross-linked chromatin was fragmented, isolated, and precipitated with antibodies specific for chicken anti-IgG (IgG) (as negative control), RNA polymerase II (Pol II), TATA binding protein (TBP), di-methylated histone H3 lysine 4 (H3K4me2), and acetylated histone H4 (AcH4). DNA purified from the precipitate was analyzed by PCR with primers corresponding to the LCR and the globin gene promoter regions in the murine β -globin locus as indicated. (This ES cell differentiation experiment and ChIP analysis were performed by Dr. Padraic P. Levings and the ChIP analyses of Day 0 cells are repeated by me.)

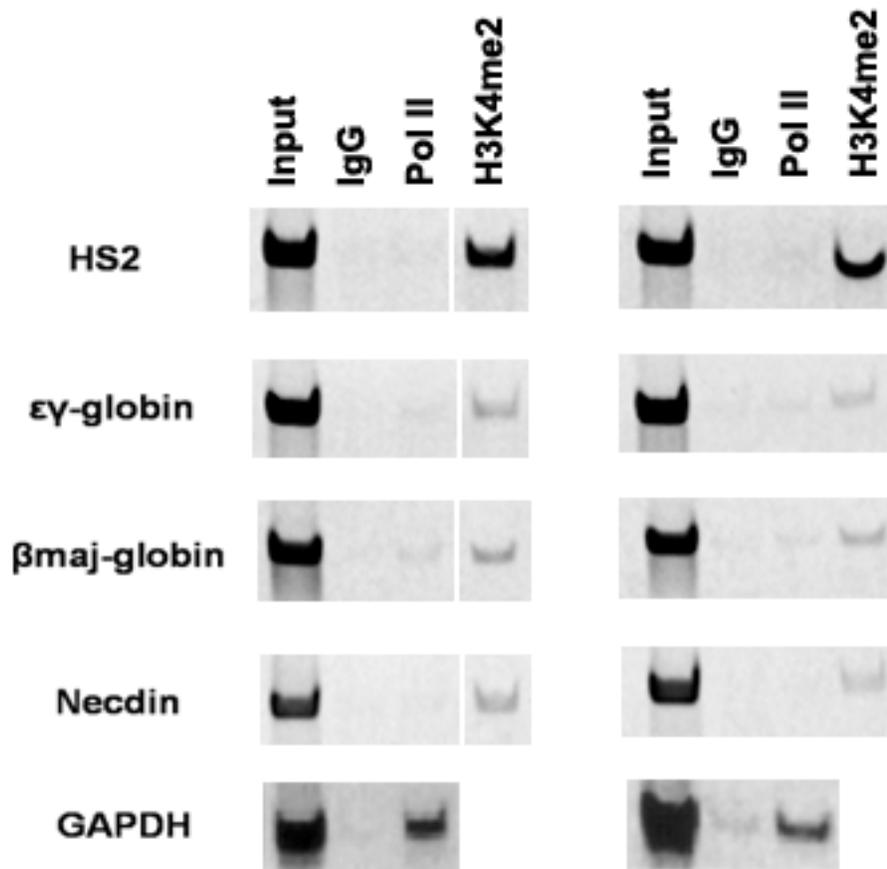


Figure 3-3. RNA polymerase II (RNA Pol II) is not recruited to the β -globin gene locus in MEF and OP9 cells. ChIP analysis of the association of RNA Pol II and H3K4me2 with the murine β -globin gene locus and GAPDH control gene was performed using Pol II, H3K4me2, and IgG-specific antibodies. The precipitated DNA was analyzed by PCR using sets of primers specific for LCR element HS2 and promoters for $\epsilon\gamma$ - and β major-globin gene, as well as Necdin and GAPDH. PCR amplification products were run on an acrylamide gel and stained with SYBR green.

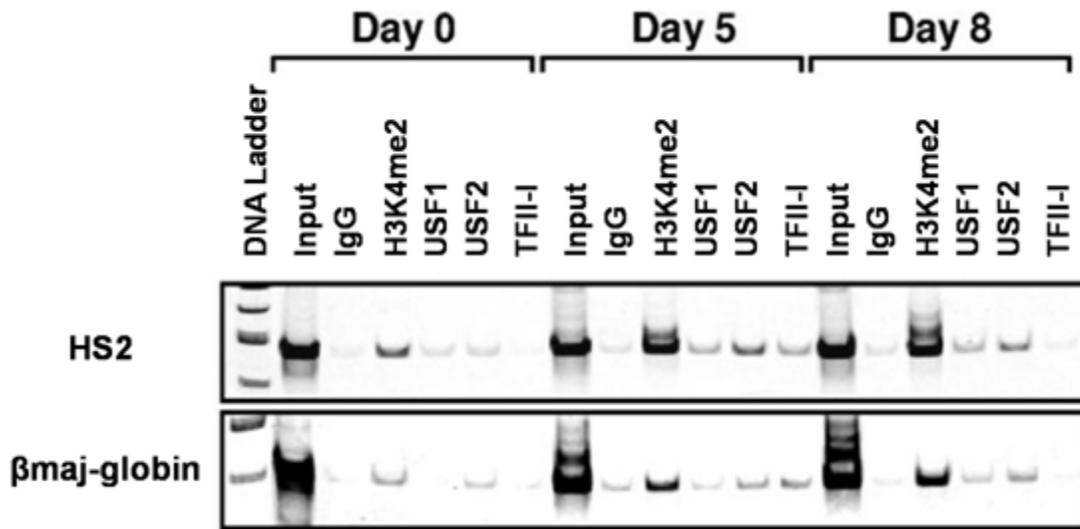


Figure 3-4. Interaction of USF and TFII-I with the β -globin gene locus during erythroid differentiation of murine embryonic stem cells. Murine embryonic stem cells were cultured and induced to differentiate as previously described by Levings et al. (280). At Day 5, after the addition of erythropoietin (Epo), cells were collected and subjected to ChIP analysis using antibodies against dimethylated histone H3K4 (H3K4me2), USF1, USF2, and TFII-I. IgG antibodies were used as negative controls in these experiments. The DNA was isolated from the precipitate and analyzed by PCR using sets of primers specific for LCR element HS2 and the β maj-globin gene promoter.

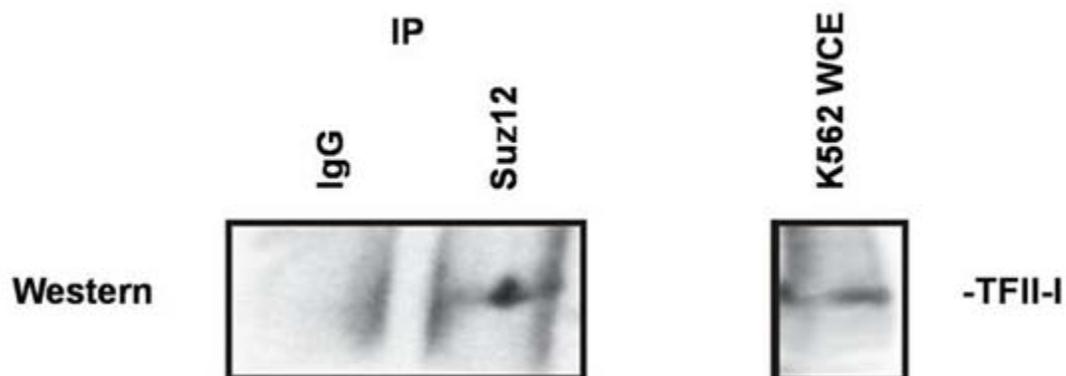


Figure 3-5. Interaction of Suz12 with TFII-I in K562 cells. K562 whole cell extracts (WCE) were precleared with anti-(rabbit) IgG-beads and precipitated with 2.5 μ g of anti-Suz12 or anti-IgG (as negative control) antibodies. The complexes were captured by incubation with anti-(rabbit IgG) beads. Complexes were eluted off the beads with Laemmli buffer by incubation at 95 °C for 10 min and loaded onto a 10% Ready gel (Bio-Rad). The membrane was probed with anti-TFII-I antibody. The lane labeled K562 WCE represents a regular western blot for TFII-I with protein extract from K562 cells (as positive control).

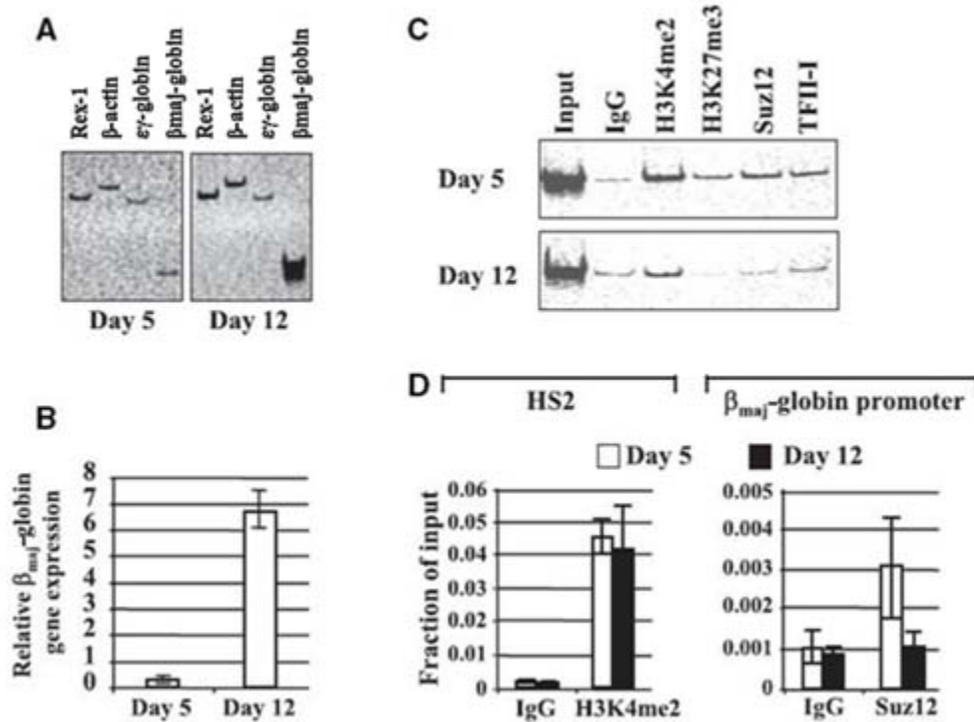


Figure 3-6. The interaction of Suz12 with the β -globin gene promoter decreases with increased β -globin gene expression during erythroid differentiation of mouse embryonic stem cells. (A) RT-PCR analysis of $\epsilon\gamma$ - and β_{maj} -globin gene expression during erythroid differentiation of mouse embryonic stem cells. RNA was isolated from the embryonic stem cells incubated in the presence of Epo for 5 or 12 days, as indicated. The RNA was reverse transcribed and subjected to RT-PCR using primers specific for the control genes Rex-1 and β -actin as well as the embryonic $\epsilon\gamma$ - and adult β_{maj} -globin genes. (B) Quantitative RT-PCR analysis of β_{maj} -globin gene expression at Day 5 and 12 of erythroid differentiation of mouse embryonic stem cells. RNA was isolated from the cells at the indicated time points after addition of Epo and subjected to quantitative RT-PCR analysis using primers specific for the β_{maj} -globin gene. (C) Analysis of modified histones, Suz12 and TFII-I interactions with the β_{maj} -globin gene promoter at Days 5 and 12 of erythroid differentiation. Cells were collected at the indicated time points and subjected to ChIP analysis using antibodies specific for histone H3 dimethylated at K4 (H3K4me2), histone H3 trimethylated at K27 (H3K27me3), Suz12, TFII-I and the negative control IgG. (D) Quantitative analysis of Suz12 interactions with the β_{maj} -globin gene promoter at Day 5 and 12 of erythroid differentiation. Cells were taken at the indicated time points and subjected to ChIP with the indicated antibodies. The precipitated DNA was subjected to quantitative PCR using primers specific for the β_{maj} -globin gene. In the left panel, chromatin was precipitated with antibodies specific for histone H3 dimethylated at K4 (H3K4me2) and analyzed by quantitative PCR using primers specific for LCR element HS2.

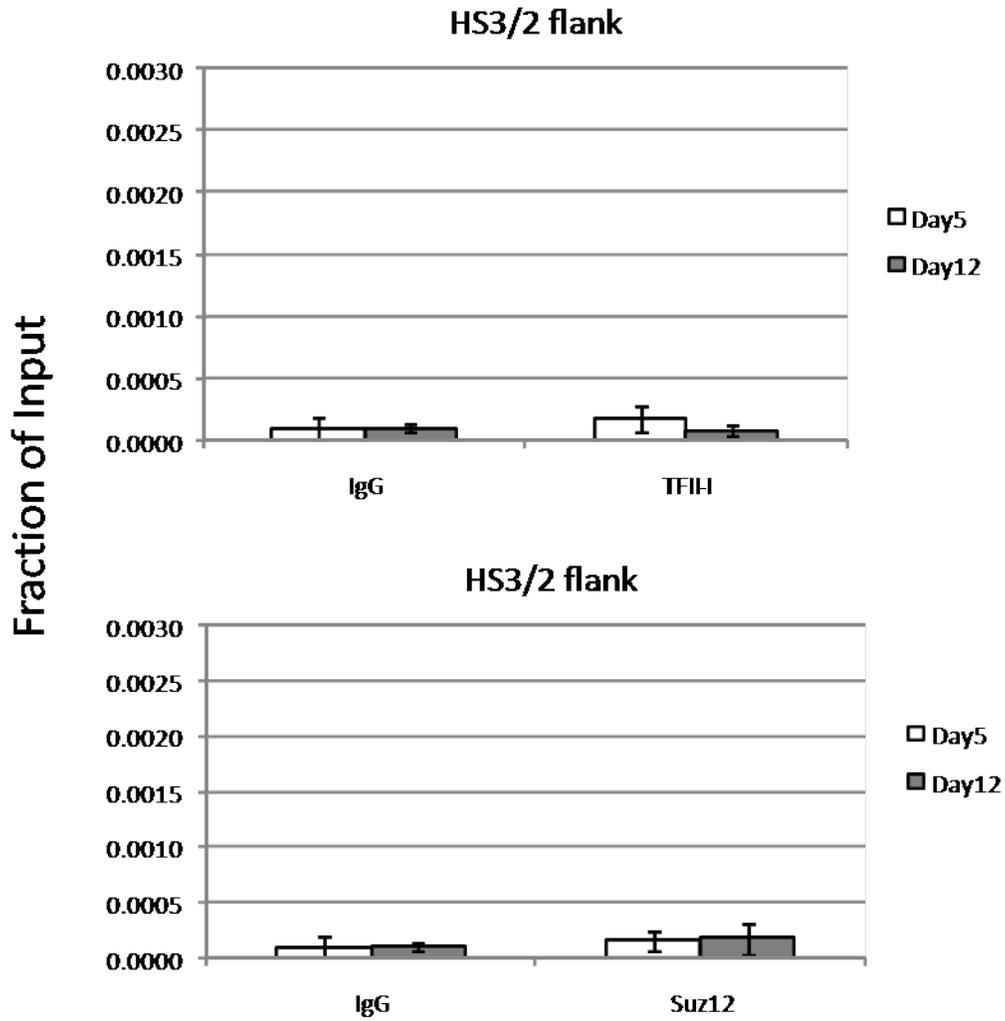


Figure 3-7. TFII-I and Suz12 do not interact with the LCR HS3/2 flanking region. Quantitative analysis of TFII-I and Suz12 interactions with the LCR HS3/2 flanking region at Days 5 and 12 of erythroid differentiation of murine embryonic stem cells. Cells were taken at the indicated time points and subjected to ChIP with the indicated antibodies. The precipitated DNA was subjected to quantitative PCR using primers specific for the LCR HS3/2 flanking region (HS3/2 flank).

CHAPTER 4

NF-E2 AND USF COOPERATE TO REGULATE RNA POLYMERASE II IN THE β -GLOBIN GENE LOCUS

Introduction

As a model system of tissue- and stage-specific gene regulation during development, the β -globin gene locus has been extensively studied in the past three decades. Although the precise model of β -globin gene regulation is still controversial, it is well accepted that the β -globin gene locus is regulated by proximal and distal DNA elements as well as by erythroid-specific and ubiquitously expressed transcription factors that act through the regulatory DNA elements.

The most powerful distal regulatory element of the β -globin gene locus is the locus control region (LCR). Results from studies with transgenic mice suggest that the LCR carries strong enhancer and chromatin domain-opening activities, and it regulates downstream β -like globin gene expression in a copy-number dependent manner during development. Because of these intriguing properties of the LCR, it has been studied intensively during past twenty years. The human β -globin LCR contains five DNase I hypersensitive sites, HS1-HS5. Among these, HS2 and HS3 feature the majority of the known activities of the LCR (310). HS2 is considered to be the most powerful enhancer element of the LCR in the regulation of β -globin gene expression. By itself, HS2 can activate human β -globin gene expression at high levels in yeast artificial chromosome (β YAC) transgenic mice (221,311). HS3 plays an important role in the embryonic ϵ - and γ -globin gene expression and it cooperates with HS2 for long-range enhancer activity (312,313). LCR HS2 and HS3 contain multiple DNA motifs which serve as binding sites for a variety of erythroid-specific and ubiquitously expressed transcription factors. HS2 contains a tandem MARE (Maf recognition element) sequence which serves as a binding site for NF-E2, a GATA-1 binding site, an EKLF binding site, two E-Box elements which recruit helix-loop-helix proteins such as USF, Tal1, and TFII-I, and a YY1 binding site (Figure 4-1A). In HS3, there is

one NF-E2 binding site, two GATA-1 binding sites and an EKLF binding site. It has been shown that these transcription factors are associated with the LCR *in vivo*, and single and combinatorial mutations of these binding sites show decreased β -globin gene expression as well as DNase I resistance. Therefore, it is suggested that the strong regulatory activity of the LCR is attained by the association of multiple transcription regulatory proteins to its HS sites.

The most important proximal regulatory element of the β -globin gene is its promoter. The upstream promoter of the human β -globin gene contains a non-canonical TATA-box (CATAAA), a CACCC-box (EKLF binding site), and multiple GATA-1 binding sites (314). It has been shown that the non-canonical TATA-box is essential for β -globin gene transcription. Furthermore, EKLF and GATA-1 have been shown to interact with the β -globin gene promoter *in vivo* (47,103). The downstream promoter region contains an initiator (Inr) element, a MARE-like sequence, and three E-Box elements located at +20, at +60, and immediately downstream of the initiation site overlapping the initiator (Figure 4-1A) (103). The E-box at +20 is not conserved between mice and human. The combined initiator E-box interacts with TFII-I and USF and the +60 E-box interacts with USF1 and USF2. Also, the association of NF-E2 with the MARE-like sequence from the human β -globin downstream promoter region has been observed *in vitro*.

Since the β -globin gene is exclusively expressed in erythroid cells, it is reasonable to speculate that β -globin gene expression is regulated by both erythroid-specific and ubiquitously expressed transcription factors. To date, the most well-studied erythroid-specific transcription factors involved in β -globin gene regulation are NF-E2, GATA-1 and EKLF.

NF-E2 belongs to the basic-leucine zipper (bZip) family of transcription factors. It consists of two subunits, p45, an tissue-specific protein containing a DNA binding and a transactivation

domain, and a ubiquitously expressed small Maf protein, commonly p18/MafK, which contains a DNA binding domain but lacks a transactivation domain. The transactivation domain of p45 interacts with a variety of proteins to regulate globin gene expression. These proteins include histone acetyltransferase CBP, histone methyltransferases MLL2 and G9a, and components of the basal transcription machinery, such as TAFII130 (226). The small p18 Maf protein dimerizes with NF-E2 p45 through their bZip domains and helps to stabilize the association of NF-E2 with its cognate MARE sequence. Based on studies using mouse erythroleukemia (MEL) cell lines lacking the p45 subunit, NF-E2 p45 is thought to be required for α - and β -globin gene expression, hyperacetylation of histones and the association of RNA Pol II with the β -globin gene promoter. However, NF-E2 p45 deficient mice display normal hematopoiesis and express the adult β -globin gene as high as ~70% of that of the wild type. The mechanisms that apparently compensate for the loss of p45 with regard to β -globin gene regulation *in vivo* are currently unknown, but could involve NF-E2 related proteins expressed in erythroid cells. Additionally, p18/MafK deficient mice also display a normal phenotype, which could be due to the functional compensation from other small Maf proteins, such as MafG (315).

GATA-1 is required for normal erythropoiesis. It has been shown that in a mouse erythroid cell line deficient of GATA-1, the recruitment of RNA Pol II to LCR HS1-HS3 and to β -globin gene promoter is abolished. As a result, GATA-1 deficiency leads to embryonic lethality due to severe anemia. EKLF functions as an adult switching factor, essential for adult β -globin gene expression in mice. EKLF null mice express embryonic globin genes normally, but they fail to express the adult β -globin gene and die *in utero* at around day 15 (316).

Besides the general transcription factors, ubiquitously expressed transcription factors involved in β -globin gene regulation include the helix-loop-helix proteins USF and TFII-I.

Previously, based on studies using erythroid cells lines and *in vitro* erythroid lineage-specific ES cell differentiation, we have shown that USF functions as an activator while TFII-I functions as repressor of adult β -globin gene expression. Moreover, our laboratory recently successfully generated transgenic mice expressing an erythroid-specific dominant-negative form of USF (A-USF) (317). The founder A-USF transgenic mice is a male, but all of its transgenic male offsprings are embryonic lethal. The male transgenic embryos were pale and reabsorbed at 14.5 days postcoitum (dpc). Thus it is likely that the transgene is integrated into the X chromosome and that the expression of A-USF in all erythroid cells affects survival. The variations in the phenotypes of female transgenic mice are probably due to the difference in silencing the transgene. Therefore, adult female transgenic mice were treated with phenylhydrazine to induce hemolytic anemia. We detected significantly reduced USF and RNA Pol II associations with the adult β major-globin gene promoter in the spleens of A-USF transgenic mice compared to the wild type littermates. The observations obtained from the A-USF transgenic mice further support that USF acts as an activator of adult β -globin gene expression.

Recent studies using the 3C technique has shown that, in differentiated mouse erythroid cells, the β -globin locus forms an 'active chromatin hub' (ACH) in which the LCR interacts with the adult β -globin gene promoter (149). Since NF-E2, GATA-1, EKLF, USF, and TFII-I have all been shown to interact with both LCR HS2 and the β -globin gene promoter *in vivo*, it would be interesting to study whether they can function as a bridge to bring the LCR and β -globin gene promoter together. In fact, it has been shown that GATA-1 and EKLF are both required for ACH formation (156,157). Although it is unknown whether NF-E2 and USF function in β -globin ACH formation, they may function in the context of the loop.

RNA Pol II has been shown to associate with LCR HS1-HS4 and with the β -globin gene promoter *in vivo* (47,280). In our previous *in vitro* ES cell differentiation system, we showed that RNA Pol II associates only with the LCR in undifferentiated ES cells, prior to the expression of the downstream globin gene. In addition, based on studies using an *in vitro* transfer system, our laboratory's previous data also suggest that the LCR is able to deliver its associated RNA Pol II to an adult β -globin gene and the presence of additional NF-E2 helps to promote the transfer of RNA Pol II (55). Moreover, in erythroid cells lacking endogenous NF-E2 p45, the association of RNA Pol II to the β -globin gene promoter is abolished, while its association with the LCR is remains relatively unperturbed. However, in the absence of functional USF, the interactions of RNA Pol II with both the LCR and the β -globin gene promoter are reduced. Therefore, it is important to investigate the mechanisms of how RNA Pol II and other components of the transcription apparatus are transferred from the LCR to the β -globin gene promoter.

Based on studies using mouse erythroleukemia cells, I demonstrate that NF-E2 p45 is essential for the high level association of RNA Pol II and various transcription regulatory proteins to the LCR and the adult β -globin gene promoter in differentiated erythroid cells. Interestingly, the association of CTD phosphorylated RNA Pol II to the β -globin gene locus requires the presence of NF-E2 p45. I also found that NF-E2, USF and RNA Pol II interact with each other in differentiated erythroid cells. Moreover, by using an *in vitro* transfer/dissociation assay, I show that NF-E2 facilitates the transfer of RNA Pol II and several transcription factors previously associated with the LCR to the adult β -globin gene promoter and that the transfer of RNA Pol II requires the presence of the +60 E-box at the promoter. I further show that USF dissociates RNA Pol II from the LCR in the absence of a β -globin gene promoter. Since USF directly interacts with NF-E2, I propose a model in which the association and activity of RNA

Pol II and transcription regulatory factors at the β -globin locus is cooperatively regulated by two transcription factors, tissue-specific NF-E2 and ubiquitously expressed USF.

Results

In the course of our studies I used three different cell lines: MEL, CB3 and CB3/NF-E2. MEL cells are murine erythroleukemia cells that are arrested at a pro-erythroblast stage and can be differentiated by a variety of chemical inducers including dimethyl sulfoxide (DMSO). DMSO mediated induction of MEL cell differentiation has been widely used to study gene regulatory mechanisms in the globin gene loci (49,52). Figure 4-1B shows that incubation of MEL cells for 3 days with 1.5% DMSO leads to a more than 50-fold increase in expression of the adult β maj-globin gene. The second cell line, CB3, is also derived from MEL cells but lacks expression of NF-E2 (p45), due to viral insertion into the p45 gene locus (229). CB3 cells fail to express the adult β -globin gene upon exposure to DMSO (Figure 4-1D). The third cell line that was used was the CB3/NF-E2 cell line. These cells represent CB3 cells stably transfected with an expression construct for p45, the large subunit of NF-E2. These cells express relatively high levels of the adult β maj-globin gene in the absence of DMSO induction (Figure 4-1F) but fail to increase globin gene expression in response to DMSO (data not shown). I compared the expression of several proteins previously implicated in globin gene regulation in uninduced and DMSO induced MEL (Figure 4-1C) and CB3 (Figure 4-1E) cells by western blotting analysis. The data demonstrate that the expression of transcription factors USF1, USF2, and NF-E2 (p45) increases during differentiation of MEL cells. In contrast, expression of RNA Pol II, TFIIB, MafK, CBP, and GAPDH remain similar between uninduced and induced MEL cells. The situation in CB3 cells is somewhat different in that these cells do not express NF-E2 (p45), as expected, and do not reveal an increase in USF expression upon exposure to DMSO. In contrast to CB3 cells, CB3/NF-E2 cells express NF-E2 (p45) (Figure 4-1G), however, the USF protein

levels are difficult to compare between CB3 and CB3/NF-E2 cells due to the difference in intensities of the bands corresponding to the internal control GAPDH.

I next examined the interaction of the above-mentioned proteins with the globin gene locus in uninduced and DMSO induced MEL and CB3 cells (Figures 4-2 and 4-3). I focused the attention on two elements that have previously been shown to be the major sites of interactions for both NF-E2 and USF, namely LCR element HS2 and the adult β maj-globin gene promoter (269,318). Both of these elements harbor multiple E-box motifs and HS2 also contains two consensus MARE elements. NF-E2 also interacts with the adult β -major globin gene promoter although it lacks a consensus MARE sequence (104,318). As negative controls, I also examined the interaction of the proteins with either a DNA element located in between HS2 and HS3 (HS3/2 flanking region) or a GAPDH 3' untranslated region (UTR). The proteins examined here interact significantly less efficiently with the negative control region in uninduced or induced MEL or CB3 cells (55,269,280,282) (Figures 4-7 and 4-8). Figure 4-7 shows that NF-E2 (p45), MafK, USF1, USF2, CBP and TFIIB do not interact with the negative control HS3/2 flanking region in DMSO induced or uninduced MEL, CB3 and CB3/NF-E2 cells. Some of the values for the IgG ChIP were zero after quantitative PCR with DNA primers against HS3/2 flanking region, thus the ChIP data presented in Figure 4-7 and 4-8 are shown as fraction of input.

Figure 4-2A demonstrates that NF-E2 (p45), MafK, USF1, USF2, CBP, and TFIIB are already associated with LCR element HS2 in uninduced MEL cells. Only USF2, MafK, and TFIIB are significantly enriched at the adult β maj-globin gene promoter in uninduced MEL cells, while all the other factors are not enriched. Upon induction of MEL cell differentiation there is a 2 to 5-fold increase in the association of all of the factors with LCR element HS2 and a 5 to 10-fold increase in the interaction of all factors with the adult β -globin gene promoter. Only upon

induction of differentiation do NF-E2 (p45), USF1, and CBP reveal significant enrichment at the adult β -globin gene promoter. The situation is very different in CB3 cells (Figure 4-2B). The proteins MafK, USF1, USF2, CBP, and TFIIB reveal significant enrichment at LCR element HS2 and at the adult β maj-globin gene promoter, but at lower levels compared to MEL cells. None of the protein/chromatin interactions increase upon exposure of CB3 cells to DMSO. We next analyzed the interaction of NF-E2 and USF in CB3/NF-E2 cells, in which expression of p45 has been restored by stable transfection with a p45 expressing DNA construct. The data demonstrate that NF-E2 binds efficiently at LCR HS2 and at the adult β maj-globin gene promoter in CB3 cells expressing p45 (Figure 4-2C). Interestingly, the level of NF-E2 binding is higher than in uninduced MEL cells. Reexpression of p45 did not lead to dramatically increased interactions of USF1 and USF2 with LCR HS2 and the globin promoter.

Previous work by Johnson et al. demonstrated that NF-E2 is required for the recruitment of RNA Pol II to the adult β -globin gene promoter but not to LCR element HS2 (49). RNA Pol II is recruited to DNA in its unphosphorylated form, is first phosphorylated at Ser-5 in the CTD during transcription initiation, and subsequently phosphorylated at Ser-2 in the CTD to allow efficient elongation (319). I examined the interaction of total RNA Pol II (Pol II/CTD), unphosphorylated RNA Pol II (Pol II/P-), Ser-5 phosphorylated RNA Pol II (Pol II/S5P), and Ser-2 phosphorylated RNA Pol II (Pol II/S2P) with the globin gene locus in uninduced and DMSO induced MEL and CB3 cells (Figure 4-3A, B). The results demonstrate that RNA Pol II was recruited to LCR element HS2 but not as efficiently to the adult β -globin gene promoter in uninduced MEL cells (Figure 4-3A). Interestingly, there was no Ser-5 phosphorylated RNA Pol II in LCR element HS2 and the β maj-globin gene promoter in uninduced MEL cells. A positive control experiment demonstrated that Ser-5 phosphorylated RNA Pol II was efficiently recruited

to the GAPDH gene in uninduced MEL cells (Figure 4-3D). I observed a significant enrichment for the Ser-2 phosphorylated form of Pol II at LCR element HS2 but not at the β maj-globin gene promoter in uninduced cells. The presence of Ser-2 phosphorylated RNA Pol II indicates ongoing intergenic transcription within the LCR in uninduced MEL cells. After induction of differentiation, I detected a strong increase in the association of RNA Pol II, including unphosphorylated RNA Pol II as well as Ser-5 and Ser-2 phosphorylated forms of RNA Pol II, with both LCR HS2 and with the adult β maj-globin gene promoter (Figure 4-3A). I found that the Ser-5 phosphorylated form of RNA Pol II significantly associated with LCR HS2 but not with the β maj-globin gene promoter in MEL cells that were induced by 1.5% DMSO for only 24 hrs. This result suggests that elongation competent RNA Pol II complexes are first assembled at the LCR. In CB3 cells low but significant levels of Pol II were associated with LCR element HS2 but not with the adult β -globin gene promoter (Figure 4-3B), consistent with previous observations made by Johnson et al. (49). There was no increase in the association of RNA Pol II with the globin locus associated elements upon incubation of CB3 cells with DMSO. Importantly, in contrast to MEL cells, there were no Ser-5 or Ser-2 phosphorylated forms of RNA Pol II were detectable at LCR element HS2 and at the adult β -globin gene promoter in CB3 cells either before or after incubation with DMSO (Figure 4-3B), demonstrating that NFE2 (p45) is required for the phosphorylation of RNA Pol II in the globin gene locus. Similar to what I observed at HS2, RNA Pol II is also recruited to HS3 and phosphorylation of RNA Pol II at HS3 is also dependent on NF-E2 (p45) and erythroid differentiation (data not shown). I next addressed the question of whether re-expression of p45 in NF-E2 deficient CB3 cells would restore the association and phosphorylation of RNA Pol II at the β -globin gene locus. Re-expression of NF-E2 (p45) led to 4-fold increase in the association of Pol II with LCR HS2 when compared to CB3

cells (Figure 4-3C). The levels of Pol II bound at HS2 in these cells were comparable to those detected in uninduced MEL cells. However, the association of RNA Pol II with the adult β maj-globin gene promoter was increased about 40-fold compared to CB3 cells and was much higher than in uninduced MEL cells. These data show that NF-E2 is important for the efficient recruitment of RNA Pol II to the β maj-globin gene. I did not observe an increase in Ser-5 phosphorylated RNA Pol II at the LCR in CB3 cells expressing p45. In contrast there was a significant increase of both Ser-5 and Ser-2 phosphorylated RNA Pol II at the adult β -globin gene promoter.

I also examined the interaction of the various forms of RNA Pol II with the negative control regions, HS3/2 flanking region or GAPDH 3' UTR (Figure 4-8). The interactions of total RNA Pol II (Pol II CTD) and unphosphorylated RNA Pol II (Pol II/P-) with the HS3/2 flanking region were undetectable in CB3 cells with or without DMSO induction (Figure 4-8A). Since there was no recruitments of Ser-5 or Ser-2 phosphorylated RNA Pol II to the LCR HS2 or adult β maj-globin gene promoter in CB3 cells, the interactions of these forms of RNA Pol II with HS3/2 flanking region were not examined. In uninduced MEL cells, there was a low level of total RNA Pol II (Pol II CTD) but no unphosphorylated RNA Pol II (Pol II/P-) associated with HS3/2 flanking region, whereas in MEL cells induced by DMSO for 3days, there were increased levels of the association of total and unphosphorylated RNA Pol II to the HS3/2 flanking region (Figure 4-8A). However, the interactions of these forms of RNA Pol II were at 10-20 times lower levels compared to the interactions with LCR HS2 and the β maj-globin gene promoter in DMSO induced MEL cells. Even in the uninduced MEL cells, the association of total RNA Pol II to both of LCR HS2 and the β maj-globin gene promoter were significantly higher compared to levels detected at the HS3/2 flanking region.

Figure 4-8B shows that Ser-5 phosphorylated RNA Pol II (Pol II/S5P) was not associated with the HS3/2 flanking region or with the GAPDH 3' UTR in MEL cells induced by 1.5% DMSO for 1 or 3 days. Since Ser-5 phosphorylated RNA Pol II is only present during transcription initiation and early elongation, the observation that it was undetectable at the GAPDH 3' UTR confirmed that the RNA Pol II/S5P antibody I used here is specific and does not detect other forms of RNA Pol II. The interaction of Ser-5 phosphorylated RNA Pol II with the HS3/2 flanking region in uninduced MEL cells was not examined, because Ser-5 phosphorylated RNA Pol II was undetectable at both LCR HS2 and the β maj-globin gene promoter in MEL cells. Ser-2 phosphorylated RNA Pol II (Pol II/S2P) was undetectable at the HS3/2 flanking region in MEL cells incubated with or without DMSO. There was no association of all forms of RNA Pol II to the HS3/2 flanking region in CB3/NF-E2 cells (Figure 4-8C).

Because our data and those previously published by Johnson et al. suggest that USF and NF-E2 are both required for the recruitment of RNA Pol II to the adult β -globin gene promoter (49,269), I examined whether these proteins interact in erythroid cells (Figure 4-4). In co-immunoprecipitation experiments I found that USF1 and USF2 interact with NFE2 (p45) in uninduced MEL cells with comparable efficiency. After induction of differentiation, there appeared to be an increase in interaction between USF2 and NF-E2, relative to interactions between USF1 and NF-E2 (Figure 4-4). I performed reciprocal experiments in which I immunoprecipitated with a NF-E2 p45 antibody and performed the western blotting with USF antibodies and the results confirmed the interactions between USF and NF-E2. I also observed interactions of USF2 with USF1, NF-E2, and RNA Pol II in DMSO induced MEL cells. There was a reproducible but weak signal for RNA Pol II in USF2 precipitated material from uninduced cells (compare the signal in the lane labeled USF2 with the signal in the lane labeled

IgG), suggesting that USF2 but not USF1 interacts with RNA Pol II in undifferentiated MEL cells. The more efficient coimmunoprecipitation of USF with RNA Pol II in differentiated cells could be due to the fact that there is an increase in expression of USF during differentiation (320), or that there is an additional activity induced upon induction that mediates interactions between USF and RNA Pol II. This activity could in fact be the increased expression of NF-E2. It is also interesting to note that in undifferentiated MEL cells USF1 homodimers may be more abundant than USF1/USF2 heterodimers. This does not appear to be the case in differentiated MEL cells, in which we detected efficient interaction of USF2 with both USF1 and NF-E2 (p45). Furthermore, USF2 but not USF1 interacts efficiently with the co-activator CBP in MEL cells (Figure 4-4). TFIIB precipitates together with USF2 in uninduced and induced MEL cells. In contrast to MEL cells I did not detect interactions between USF and RNA Pol II in CB3 cells although the interaction between USF2 and CBP are detectable (Figure 4-4). I did not detect USF, NF-E2, or CBP in samples precipitated with antibodies against RNA Pol II. This is likely because only a small fraction of nuclear RNA Pol II will associate with these proteins at any given time. All of the protein/protein interaction data here have been reproduced.

NF-E2 and USF could be part of a large protein complex and thus the interaction could either be direct or mediated by other proteins. I carried out GST-pull down assays using GST-tagged recombinant USF1 and recombinant his-tagged NF-E2 (p45 tethered to mafG). The results demonstrate that USF1 directly interacts with NF-E2. Since both proteins contain a leucine-zipper interaction domain, I analyzed the interaction of different mutants of USF1 with NF-E2. The USF1 mutants contain either the truncated C-terminal immediately downstream of USR (USF specific region) and the basic region, or with further downstream deletion of the helix-loop-helix domain as indicated in Figure 4-4B and C. I detected efficient interactions only

between full length USF1 and NF-E2 demonstrating that the N-terminus of USF1 is important for the interaction with NF-E2. I next analyzed USF1 N-terminal mutants and results demonstrate that the C-terminus is also important for the interaction between USF1 and NF-E2. However, in contrast to the C-terminal mutants, deletion of the leucine zipper or the entire N-terminus still allowed interactions between the two proteins albeit with much lower efficiency. This result demonstrates that the USF1 leucine zipper domain is not the most critical determinant among all the other USF1 domains for the specific interaction between USF1 and NF-E2.

I next examined the mechanism(s) by which NF-E2 and USF could regulate LCR mediated recruitment of RNA Pol II to the adult β -globin gene promoter using a protein dissociation/transfer assay (Figure 4-5A). Our laboratory previously established a method that allows the analysis of Pol II transfer from the LCR to the adult β -globin gene promoter (55). A plasmid containing all 5 HS sites from the human β -globin LCR is linearized, biotinylated, and immobilized on streptavidin coated magnetic beads. The immobilized LCR is incubated with protein extracts from MEL cells. Studies from our laboratory previously demonstrated that under the applied conditions RNA Pol II is recruited to LCR elements HS2 and HS3 (53). After removing all material not bound to the LCR, the immobilized protein/DNA complex is incubated with DNA templates containing the adult β -globin promoter or mutants. In previous experiments from our laboratory, the globin gene template along with associated proteins from the immobilized LCR complex is removed and immunoprecipitation is performed with RNA Pol II specific antibodies followed by PCR using β -globin gene specific DNA primers (55). In the initial experiments described here I removed all material from the LCR complex after incubation with or without β -globin promoter containing templates and subjected the material to western blotting analysis using NF-E2, USF, or RNA Pol II specific antibodies as shown in Figure 4-5B.

In most of these experiments I used β -globin gene templates either containing the β -globin gene promoter (β^+) or lacking the β -globin gene promoter (β^-). This experimental setup allows me to monitor β -globin promoter dependent dissociation of proteins from the LCR. The data demonstrate that there is an increase in the dissociation of USF1, USF2, NF-E2 (p45) and RNA Pol II from the LCR in the presence of the adult β -globin gene promoter (Figure 4-5B), suggesting that the promoter efficiently competes for the binding of these proteins. I observed a reproducible decrease in the dissociation of RNA Pol II from the immobilized LCR if β -globin templates are used that carry mutations in an E-box located 60bp downstream of the transcription initiation sites (+60Eboxmut, Figure 4-5C). Our laboratory previously demonstrated that this E-box interacts with USF and is required for the efficient *in vitro* transcription of the β -globin gene (103). A mutation of a partial MARE sequence (NF-E2mut) in the downstream promoter region did not affect the increase in β -globin promoter mediated dissociation of RNA Pol II from the LCR. Expression of a dominant negative mutant of USF (A-USF) in MEL cells reduces recruitment of RNA Pol II to LCR HS2 and the adult β -globin gene promoter (269). Interestingly, the addition of A-USF increases the dissociation of Pol II from the LCR even in the absence of a β -globin gene promoter, while BSA, AAV Rep protein, or NF-E2 were all unable to do so (Figure 4-5D). Previous studies from our laboratory have shown that NF-E2 increases the efficiency of transfer of RNA Pol II from the LCR to the adult β -globin gene promoter (55). I show here that NF-E2 only facilitates dissociation of RNA Pol II from the LCR in the presence of the adult β -globin gene promoter (Figure 4-5D). The fact that the USF-binding site is required for efficient dissociation/transfer suggests that USF may be required to stabilize NF-E2 binding at the promoter, which lacks a consensus MARE sequence.

I next verified some of the results from the *in vitro* dissociation experiment using a quantitative assay in which I performed the incubation experiments as described above but examined the transfer of RNA Pol II from the LCR to the β -globin gene promoter using immunoprecipitation (IP) with RNA Pol II antibodies followed by quantitative PCR using primers that amplify a fragment from the β -globin gene (β^+) and the mutant β -globin gene lacking the promoter (β^-) (Figure 4-5E). The data show that RNA Pol II is transferred from the LCR to the β -globin gene in a promoter dependent manner and that NF-E2, but not A-USF, facilitates the transfer of RNA Pol II to the β -globin gene. To control for the initial amount of Pol II recruited to the LCR, I performed western blotting experiments. The data demonstrate that all samples analyzed quantitatively had about the same amount of Pol II recruited to the LCR before dissociation/transfer was analyzed. The fact that A-USF did not increase recruitment of RNA Pol II to the β -globin gene promoter even it dissociates USF and RNA Pol II from the LCR, is likely due to that A-USF associated USF is unable to bind to the +60 E-box at the β -globin promoter to facilitate the recruitment of RNA Pol II to the promoter.

Discussion

Previous studies have shown that β -globin LCR HS sites recruit transcription complexes and that the LCR is required for the association of the β -globin gene locus with transcription factories (55,274,284,285). Furthermore, long intergenic noncoding transcripts originate from within or upstream of the LCR and are detectable throughout the globin gene locus in a developmental stage-specific manner (321). During the differentiation of erythroid cells, transcription complexes and other activities first associate with the LCR before they are detectable at the globin gene promoters (280,322,323). These data suggest that the LCR is the primary attachment site for the recruitment of transcription complexes and that these complexes are delivered to the globin gene locus by a tracking, linking, or looping mechanism (324). Not

consistent with this model, however, is the observation that even in the absence of the LCR RNA Pol II is efficiently recruited to the adult β -globin gene promoter, although it exhibits defects in elongation of transcription (75). Sawado et al. discussed the possibility that the LCR provides activities for the efficient elongation of RNA Pol II at the β -globin gene promoter (75). If RNA Pol II is first recruited to the LCR, another possibility is that the elongation competent transcription complexes are assembled at the LCR and transferred to the high affinity globin gene promoters.

To further elucidate how the LCR and interacting proteins mediate recruitment and activity of RNA Pol II in the β -globin gene locus I analyzed MEL cells expressing or not expressing NF-E2 (p45). NF-E2 (p45) has previously been shown to be required for the recruitment of RNA Pol II to the adult β -globin gene promoter but not to the LCR. I demonstrate here that RNA Pol II is recruited to LCR HS2 but not to the β -globin gene promoter in undifferentiated MEL cells (Figure 4-3). Recruitment of RNA Pol II to LCR HS2 is inefficient in the absence of NF-E2 (p45). Perhaps more importantly, however, is the observation that in the absence of NF-E2 (p45) there is no phosphorylated RNA Pol II detectable at LCR HS2 or the β maj-globin gene promoter. This result suggests that NF-E2 is not only important for the efficient recruitment of RNA Pol II to the globin gene locus but plays a role in converting RNA Pol II into an elongation competent form, as has been discussed previously by Sawado et al. (75,318). Re-expression of p45 in NF-E2 deficient CB3 cells caused a strong increase in the recruitment of RNA Pol II with the β maj-globin gene promoter. Expression of p45 also led to the phosphorylation of RNA Pol II at the promoter but not at LCR HS2. This result is interesting and somewhat contrasts the findings in MEL cells showing that Ser5-phosphorylated RNA Pol II is first detectable at the LCR during DMSO induced differentiation. Expression of relative high levels of p45 in uninduced CB3 cells

likely causes the local remodeling of the chromatin structure at the β maj-globin gene promoter, consistent with findings from the Brandt and Bresnick laboratories (276,277). According to the transfer model the open chromatin structure at the promoter would lead to the efficient transfer of Ser5-phosphorylated RNA Pol II to the promoter. Alternatively, the NF-E2 (p45) induced opening of the chromatin structure at the adult globin gene promoter may bypass the need for LCR mediated RNA Pol II recruitment.

I detected low levels of RNA Pol II binding at the adult β -globin promoter in uninduced MEL cells, which possibly attribute to the presence of cells that spontaneously differentiated in the absence of DMSO. Alternatively, low levels of RNA Pol II could be recruited to the β maj-globin gene in an LCR dependent or independent manner in undifferentiated cells. I observed that, along with RNA Pol II, USF2, MafK, and TFIIB are already associated with LCR HS2 and with the adult β -globin gene promoter in undifferentiated MEL cells, while USF1, CBP and NF-E2 only are only associated with HS2 in undifferentiated cells. Using Co-IP I found that after differentiation of MEL cells there is an increased interaction of USF2 with NF-E2 (p45), USF1, and RNA Pol II (Figure 4-4). This is accompanied by an increased recruitment of all of these proteins to the globin gene locus. Previous studies had already shown that the NF-E2 activity increases during differentiation of erythroid cells (325). My data suggest that increased expression of USF and NF-E2 facilitate the formation of large protein complexes that regulate RNA Pol II recruitment and activity in the β -globin gene locus.

Based on my data, I propose that partial elongation incompetent transcription complexes are first assembled at the LCR in undifferentiated MEL cells. This is in part mediated by USF2 and TFIIB. Previous data from our laboratory showing that expression of a dominant negative mutant of USF (A-USF) in undifferentiated MEL cells reduces the recruitment of RNA Pol II to

LCR HS2 (269) are consistent with the hypothesis that USF participates in the recruitment of RNA Pol II to the LCR in undifferentiated MEL cells. During differentiation, an increase in expression of NF-E2 (p45), USF, and other proteins leads to the efficient recruitment of additional activities to the LCR, including those that convert RNA Pol II from a transcriptionally inert to a transcriptionally competent form. The assembly of elongation competent transcription complexes is indispensable of the presence of NF-E2 and is also accompanied by a conformational change in the globin gene locus that brings the adult β -globin gene in close proximity to the LCR (Figure.4-6) (149). The presence of high affinity basal promoter elements in the adult β -globin gene promoter facilitates the transfer of elongation competent transcription complexes to the promoter. The transfer is mediated at least in part by NF-E2. However, the association of NF-E2 with the β -globin gene promoter is likely stabilized through its interaction with USF and the presence of USF binding sites. The *in vitro* RNA Pol II transfer experiments revealed that NF-E2 is able to dissociate RNA Pol II from the LCR and that this process requires the presence of a β -globin promoter template. The transfer to the promoter also required a USF binding sites. These data are consistent with our findings from MEL cells and further demonstrate that USF and NF-E2 cooperate to mediate the transfer/recruitment of RNA Pol II to the adult β -globin promoter. The function of NF-E2 *in vivo* is likely more complex and involves the remodeling of chromatin structure at the adult β -globin gene promoter, which would further facilitate recruitment of the transcription complex.

The maintenance of proximity between the LCR and promoter guarantees the continued loading of elongation competent transcription complexes to the adult β -globin gene promoter (Figure 4-6). It is also possible, however, that the conformational change bringing the β -globin gene into close proximity to the LCR precedes the assembly of active transcription complexes.

Our data do not clearly distinguish between these two possibilities. The only piece of evidence arguing for LCR mediated assembly of elongation competent transcription complexes is the observation that Ser-5 phosphorylated RNA Pol II is first detectable at LCR HS2 during the differentiation of MEL cells. In contrast to other hematopoietic specific transcription factors, like GATA-1, Fog-1, EKLF, and NL1, NF-E2 (p45) is not critical for mediating proximity between the LCR and the adult β -globin gene or for the formation of an active chromatin hub (ACH) (156,157,228,326). However, my data suggest that NF-E2 and USF function in the context of the ACH and perhaps mediate the assembly of elongation competent transcription complexes at the LCR and at the adult β -globin gene promoter. Dr. Suming Huang at UF has characterized proteins that associate with USF1 in HeLa cells (Huang et al., unpublished data). Interestingly, three of these proteins are implicated in the recruitment and activity of RNA Pol II. These proteins are TBP associated factors (TAFs) 4 and 6, as well as the elongation factor EF1A2. In this respect it is interesting to note that USF1 only associates with RNA Pol II and the β -globin gene locus after differentiation of erythroid cells.

Many proteins contribute to expression of the globin gene locus and NF-E2 and USF function within a cascade of events that regulate accessibility and location of the globin genes in the nucleus (327). It will be increasingly important to determine how the different transcription factors function together to mediate extremely high-level transcription of the β -like globin genes during erythroid differentiation and development. Another protein that acts at the adult β -globin gene is EKLF, which recruits chromatin remodeling complexes to the promoter but also contacts components of the transcription initiation complex (233,328). Future studies will address if and how EKLF communicates with USF and NF-E2 in regulating expression of the adult β -globin gene.

Figure 4-1. DMSO mediated increase in USF, NF-E2 (p45), and β -globin expression in MEL but not CB3 cells. (A) Schematic of cis-acting elements within LCR HS2 and sequence alignment of the β -globin downstream promoter. The LCR HS2 contains a tandem NF-E2 binding site, E-box elements (E), CACC motif (EKLF binding site) and GATA-1 and YY1 binding sites as indicated. The adult β -globin downstream promoter region contains three E-box elements and MARE/AP1-like element in human (H), mouse (M) and rabbit (R) as indicated. The E-box elements overlapping the initiator and at +60 downstream of the transcription start site (+1) are conserved in all three species, whereas the one located at +20 is only present in the human and rabbit genes. (B) Quantitative RT-PCR analysis of β -globin gene expression in MEL cells incubated with or without 1.5% DMSO for 3 days. RNA was isolated from MEL cells, reverse transcribed, and subjected to qPCR using primers specific for the adult β maj-globin gene. The results are shown as the relative expression normalized to transcription of the β -actin gene. The error bars represent the standard error of the mean (SEM) from three independent experiments. (C) Western blotting analysis of NF-E2 (p45), USF1, USF2, Pol II, TFIIB, GAPDH, MafK, and CBP in MEL cells incubated with or without 1.5% DMSO for 3 days. 60 μ g of protein from whole cell extracts was electrophoresed in 4-20% Ready Gels (Bio-Rad), transferred to nitrocellulose membranes and incubated with antibodies as indicated. (D) Quantitative RT-PCR analysis of β -globin gene expression in CB3 cells incubated with or without 1.5% DMSO for 3 days. RNA was isolated and analyzed as described in B. (E) Western blotting analysis of NF-E2 (p45), USF1, USF2, Pol II, TFIIB, GAPDH, MafK, and CBP in CB3 cells incubated with or without 1.5% DMSO for 3 days. Proteins were processed and analyzed as described in C. (F) Quantitative RT-PCR analysis of β -globin gene expression in CB3 and CB3/NF-E2 cells. RNA was processed and analyzed as described in A. (G) Western blotting analysis of NF-E2/p45, USF1, USF2, and GAPDH expression in CB3 and CB3/NF-E2 cells. Proteins were processed and analyzed as described in C. (Figure (A) is adapted from Johnson *et al.* (49) and Leach *et al.* (103).)

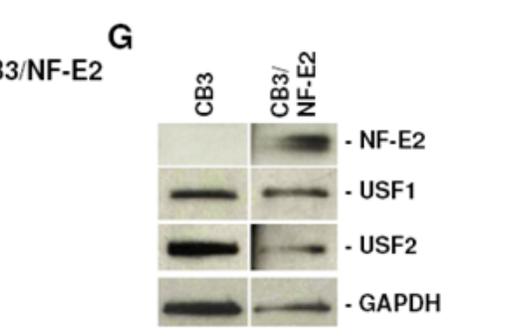
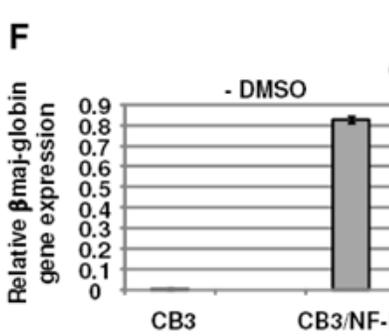
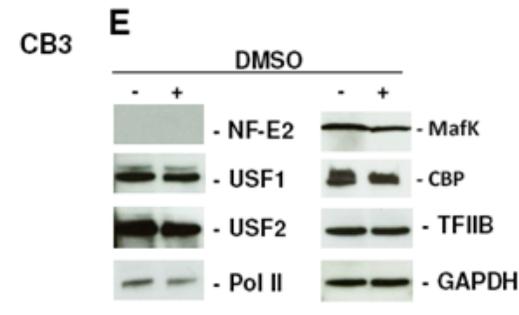
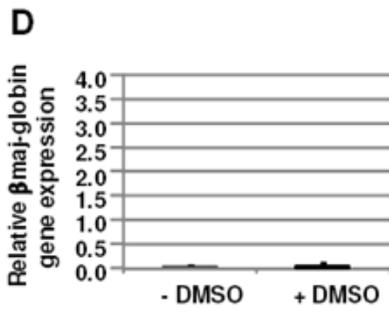
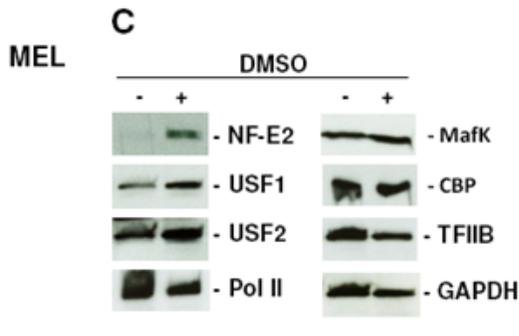
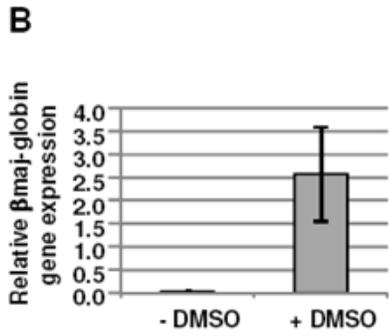
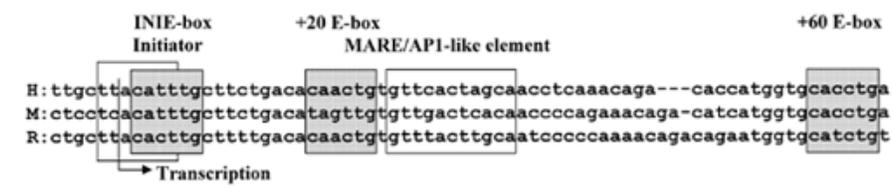
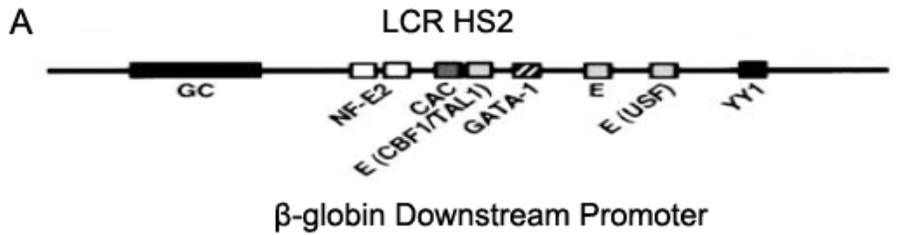


Figure 4-2. Lack of NF-E2 (p45) reduces the assembly of protein complexes at LCR HS2 and at the adult β maj-globin gene promoter. (A) ChIP analysis of protein chromatin interactions in LCR HS2 and the adult β maj-globin gene promoter in MEL cells incubated with or without 1.5% DMSO for 3 days. After crosslinking MEL cells with 1% formaldehyde, chromatin was isolated, fragmented by sonication, and subjected to immunoprecipitation with antibodies against NF-E2 (p45), MafK, USF1, USF2, CBP, and TFIIB. Reactions with the IgG antibody served as a negative control. The DNA was purified from the precipitate and subjected to qPCR using primers specific for LCR HS2 and the adult β maj-globin gene promoter as indicated. Error bars represent SEMs of three independent experiments. (B) ChIP analysis of protein chromatin interactions in LCR HS2 and the adult β maj-globin gene promoter in CB3 cells incubated with or without 1.5% DMSO for 3 days. DNA was isolated from immunoprecipitated material and analyzed as described in A. (C) Comparative ChIP analysis of protein chromatin interactions in CB3, MEL, and CB3/NF-E2 cells. Crosslinked chromatin was precipitated with IgG or antibodies against NF-E2, USF1, or USF2 and DNA was analyzed as described in panel A. In A and B, * indicate p-values of <0.05, Δ indicate p-values of <0.1 between DMSO induced versus uninduced samples; ** indicate p-values of <0.05, $\Delta\Delta$ indicate p-values of <0.1 between specific antibody versus IgG. In C, * indicate p-values of <0.05, Δ indicate p-values of 0.1 between MEL or CB3/NF-E2 cells versus CB3 cells; ** indicate p-values of <0.05, $\Delta\Delta$ indicate p-values of <0.1 between CB3/NF-E2 cells versus MEL cells; *** indicate p-values of <0.05 between CB3/NF-E2 cells versus MEL and CB3 cells.

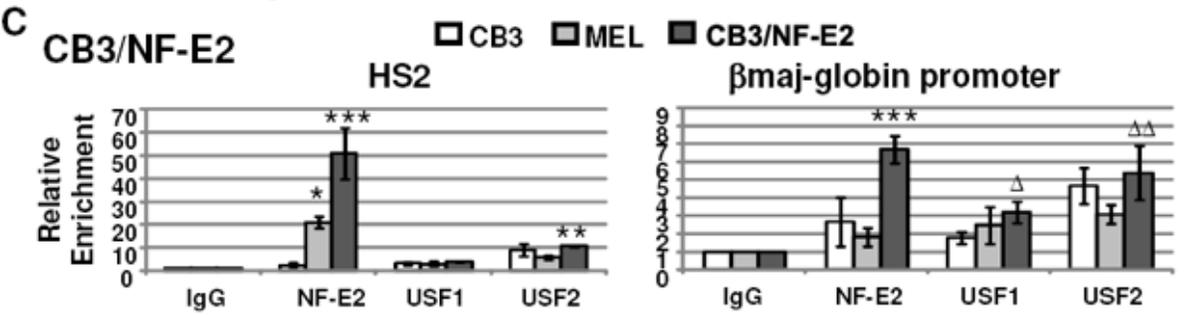
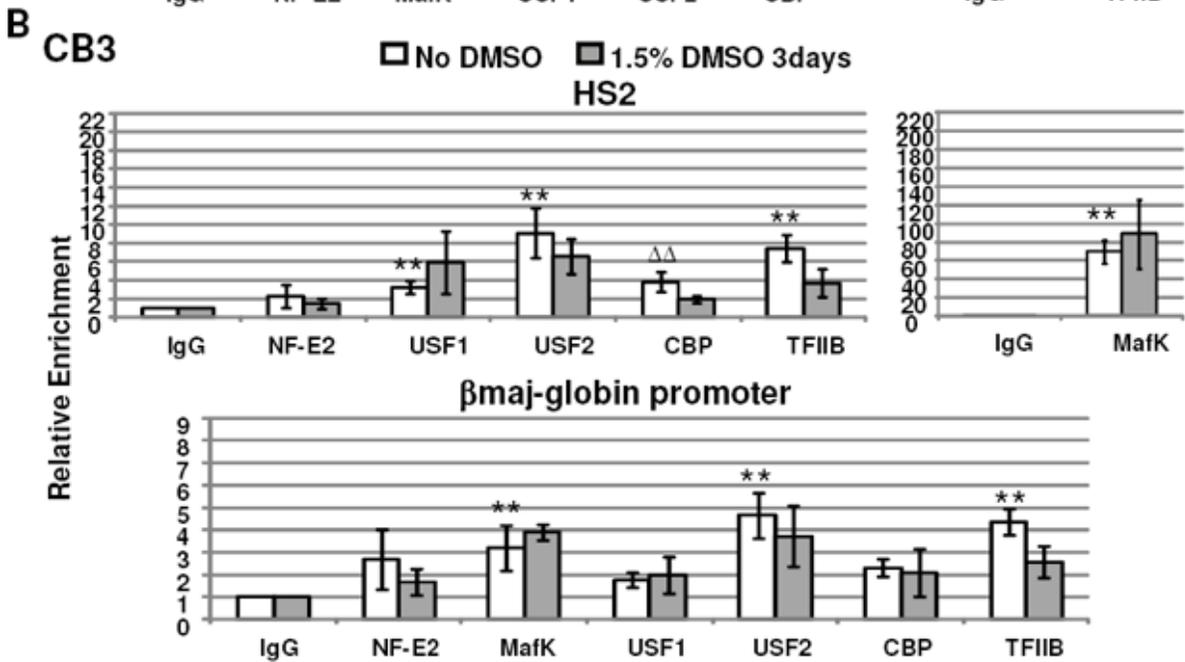
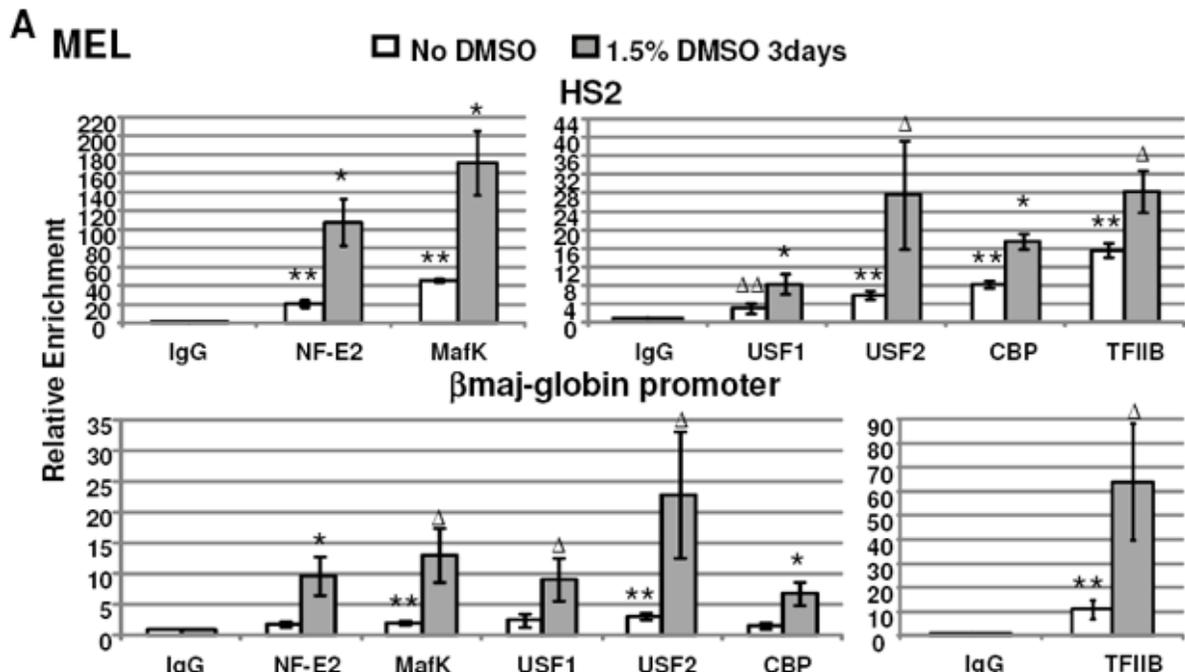


Figure 4-3. Efficient recruitment of Pol II and CTD serine 5 and serine 2 phosphorylation at the β -globin gene locus requires NF-E2. (A) ChIP analysis of Pol II interactions in the β -globin gene locus in MEL cells incubated with or without 1.5% DMSO for 1 or 3 days. Chromatin was isolated from crosslinked MEL cells, fragmented by sonication, and immunoprecipitated with antibodies specific for the Pol II CTD (Pol II/CTD), for unphosphorylated Pol II (Pol II/P-), or for Pol II phosphorylated at serine 5 (Pol II/S5P) or serine 2 (Pol II/S2P) of the CTD. IgG or IgM antibodies were used in these experiments as negative controls. DNA was isolated from the precipitates and subjected to qPCR with DNA primers specific for LCR HS2 or the adult β maj-globin gene promoter as indicated. Error bars represent SEMs from three independent experiments. (B) ChIP analysis of Pol II interactions in the β -globin gene locus in CB3 cells incubated with or without 1.5% DMSO for 3 days. Chromatin precipitation and DNA analysis by qPCR was performed as described in panel A. (C) Comparative ChIP analysis of Pol II interactions in CB3, MEL and CB3/NF-E2 cells. Chromatin precipitation and DNA analysis by qPCR was performed as described in panel A. (D) ChIP analysis of Ser-5 phosphorylated Pol II at the GAPDH promoter in undifferentiated and DMSO induced MEL and CB3 cells (as indicated). Cells were grown in the absence or presence of DMSO (1.5% for 3 days). DNA was isolated from immunoprecipitated material and analyzed by qPCR as described in A. In A, B and D, * indicate p-values of <0.05 , Δ indicate p-values of <0.1 between DMSO induced versus uninduced samples; ** indicate p-values of <0.05 between specific antibody versus IgG. In C, * indicate p-values of <0.05 between MEL or CB3/NF-E2 cells versus CB3 cells; ** indicate p-values of <0.05 , $\Delta\Delta$ indicate p-values of <0.1 between CB3/NF-E2 cells versus MEL and CB3 cells.

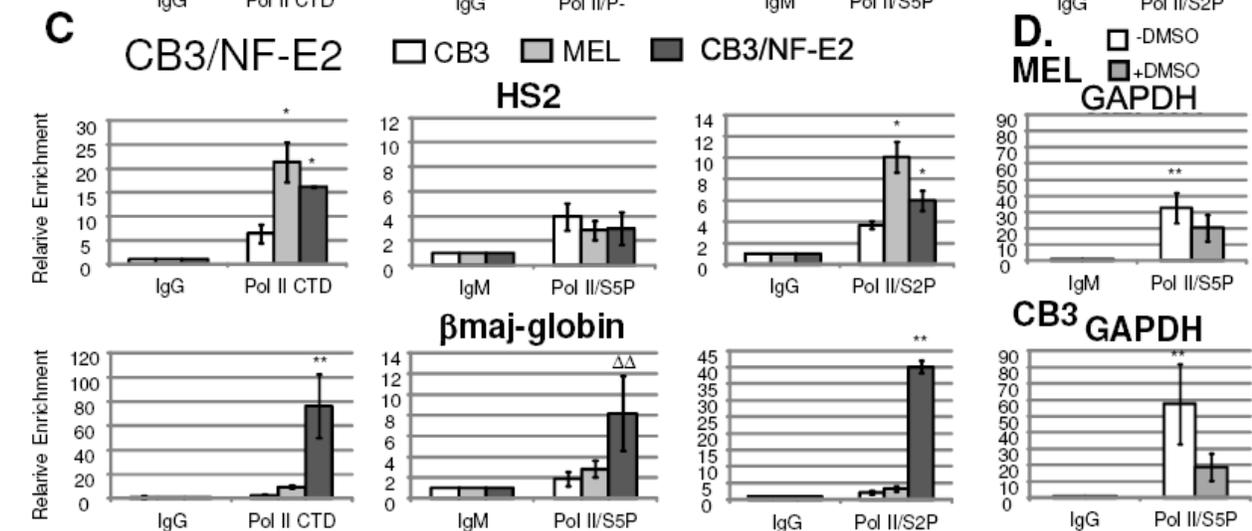
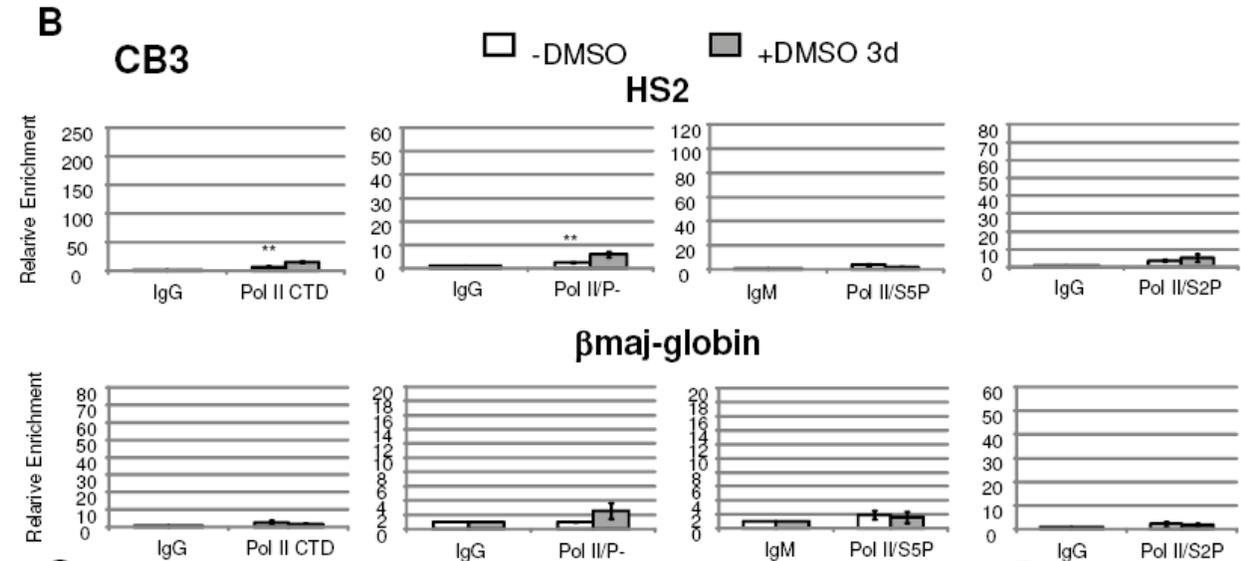
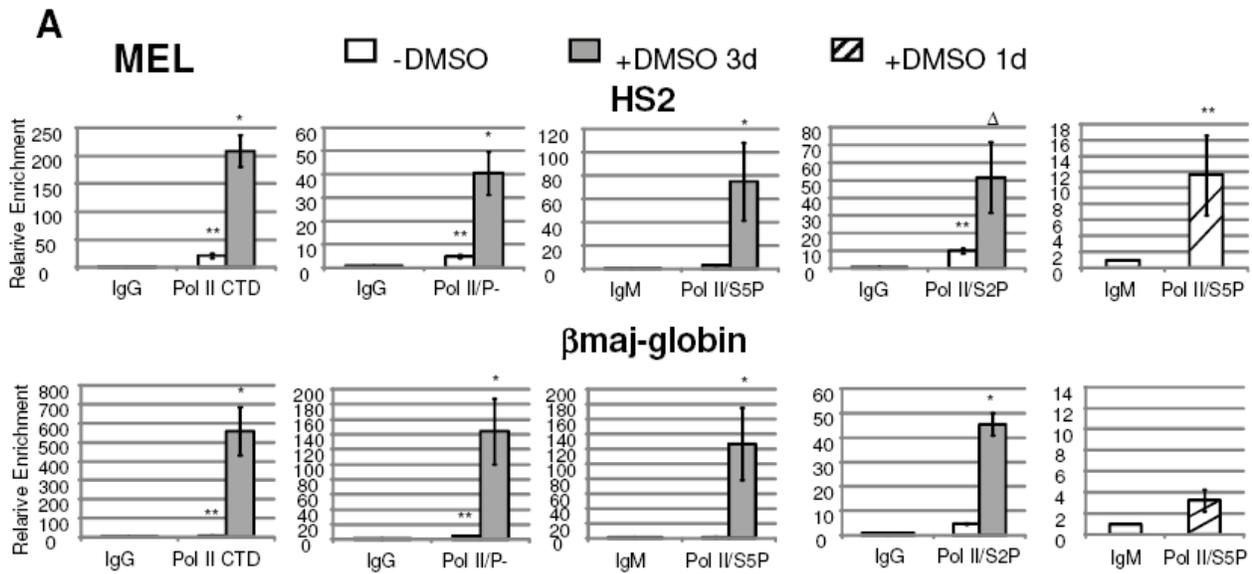


Figure 4-4. Interactions of USF1 and USF2 with NF-E2 (p45) and Pol II during erythroid differentiation of MEL and CB3 cells. (A) Co-immunoprecipitation experiments were performed by first subjecting nuclear extracts from MEL or CB3 cells incubated with or without 1.5% DMSO for 3 days to immunoprecipitation with antibodies specific for IgG, Pol II (N-20), USF1, USF2, and TFIIB. The immunoprecipitated material was electrophoresed using 4-20% Ready Gels (Bio-Rad) and transferred to nitrocellulose membranes. The nitrocellulose membranes were incubated with antibodies against Pol II, NF-E2, USF1, USF2, and CBP, as indicated, and subjected to ECL plus chemiluminescence (Amersham). (B) Generation and expression of USF1/GST fusion proteins in *E.coli*. cDNA encoding full length or truncated USF1 were ligated into the pGEX-5X-1 vector. Fusion proteins were expressed in and purified from *E.coli* and analyzed by SDS-PAGE. The following USF1 derived proteins were purified; USF1, fulllength USF1 protein; USF-M1, deletion of the N-terminal half; USF1-M2, deletion of the N-terminus and the basic region; M3, deletion of the N-terminus, the basic region and the helix-loop-helix (HLH) domain; USF1-LZ, deletion of the leucine zipper; USF1-N, deletion of the C-terminus as well as the basic region, the HLH domain, and the LZ domain. (C) Interaction of USF1 with NF-E2. Equal amounts of GST-tagged wild type and mutant USF1 fusion proteins were incubated with His-tagged NF-E2. After washing, proteins were eluted from the GST-beads, electrophoresed using SDS-PAGE and subjected to western-blotting analysis using an antibody specific for NF-E2 (p45). Fast Green staining was performed to monitor the equal loading amount of all GST-fusion proteins prior to the antibody incubation with the membrane.

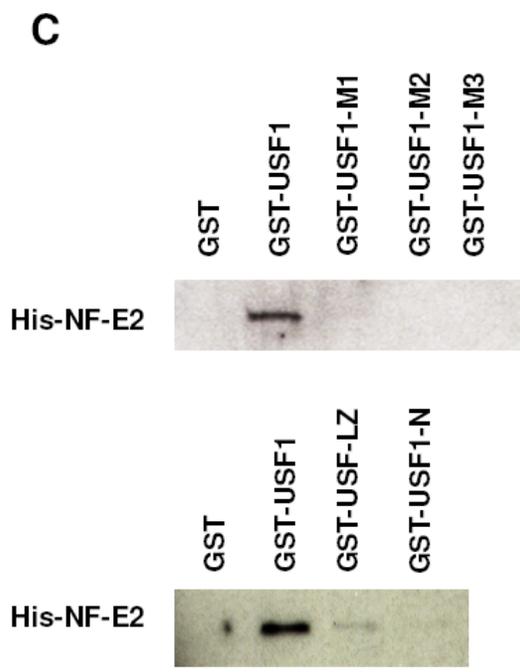
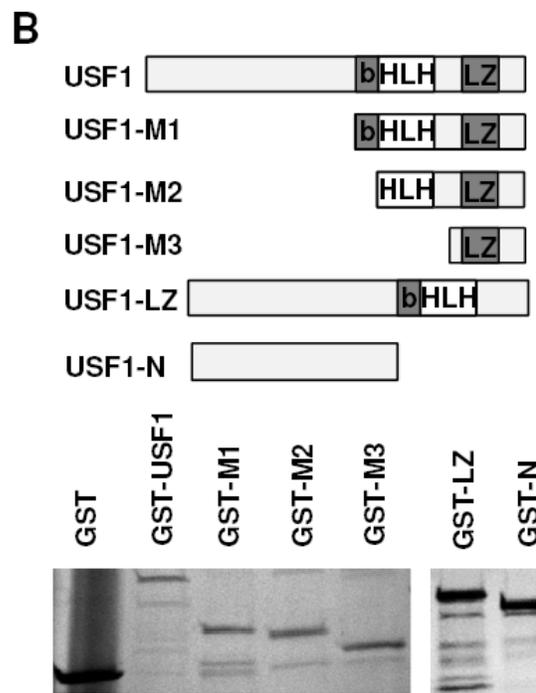
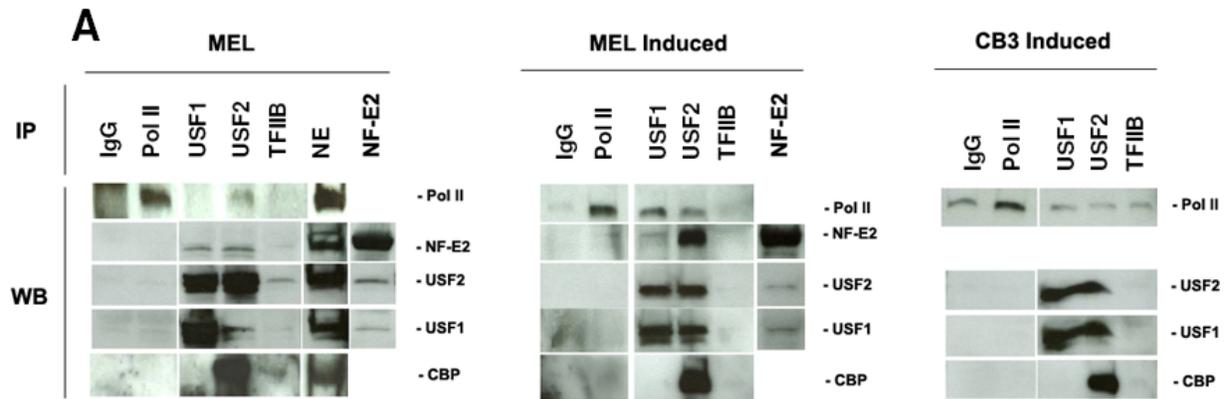
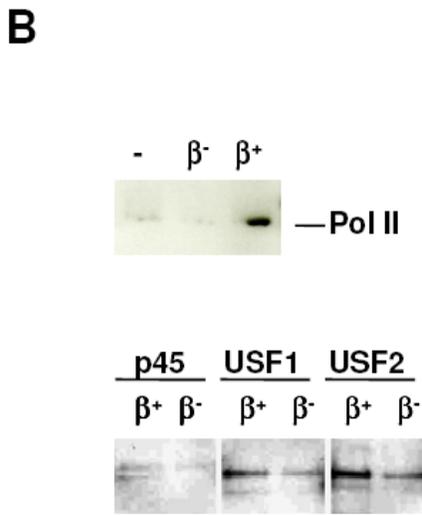
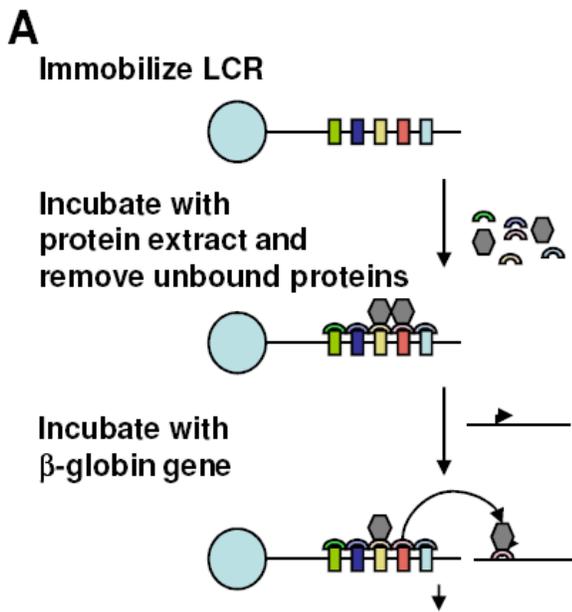
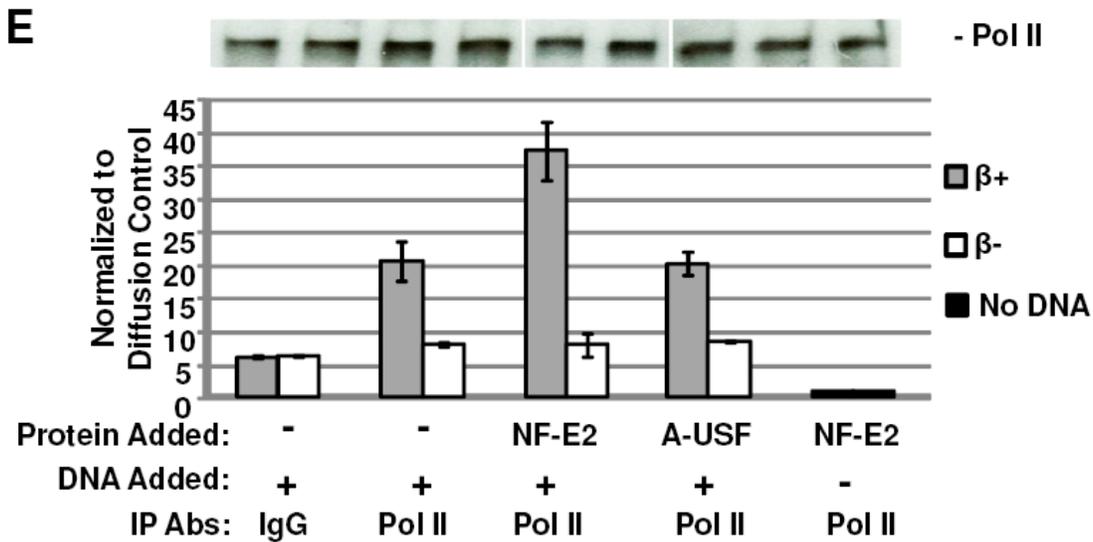
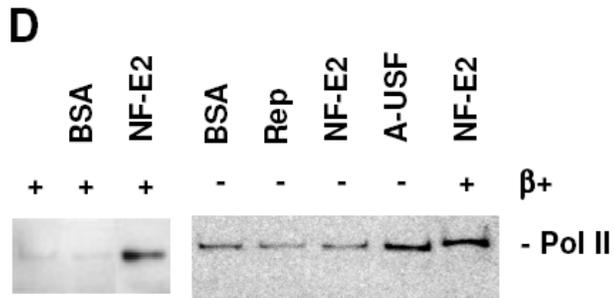
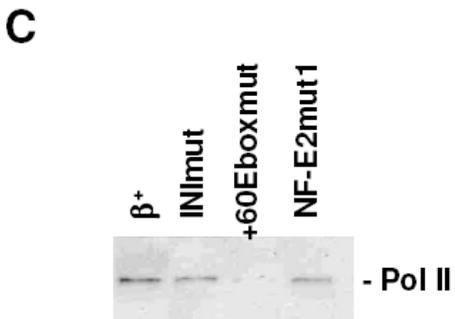


Figure 4-5. USF and NF-E2 regulate the recruitment and dissociation of RNA Pol II to and from immobilized LCR templates. (A) Scheme of the experimental strategy. A linearized and biotinylated plasmid containing the human β -globin LCR was immobilized on streptavidin coated magnetic beads as described by Vieira et al. (55). The LCR was then incubated with whole cell extracts from MEL cells. Unbound material was removed and the LCR/protein complex was washed several times and incubated with different DNA templates in the presence or absence of recombinant NF-E2 (p45 tethered to MafG, 16, 37) or A-USF (dominant negative form of USF). Proteins that dissociate from the LCR after the incubation step were analyzed using western blotting analysis. Transfer of proteins to the β -globin gene promoter was analyzed by immunoprecipitation (IP) followed by quantitative PCR. (B) β -globin promoter mediated dissociation of RNA Pol II, NF-E2 (p45), USF1, and USF2, from the LCR in the presence of a plasmid containing the β -globin gene with (β^+) or without (β^-) its promoter. Proteins and DNA were removed from the immobilized LCR after incubation for 30 min at 37°C. (C) Analysis of the effect of β -globin promoter mutations on the dissociation of RNA Pol II. DNA plasmids containing the wild-type β -globin promoter (β^+) or the promoter with mutations in the initiator (INI_{mut}), the +60 E-box (+60Ebox_{mut}), or the partial MARE sequence (NF-E2_{mut}) were incubated for 30 min with the immobilized LCR/protein complex in the presence of recombinant NF-E2. Dissociated proteins were removed and analyzed by western blotting experiments using an antibody specific for RNA Pol II. (D) Effect of NF-E2 and A-USF on the dissociation of RNA Pol II from the immobilized LCR. The immobilized LCR/protein complex was incubated for 30 min with BSA, AAV Rep (Rep), NF-E2, or A-USF in the absence or presence of a plasmid containing the wild-type β -globin gene promoter (β^+). Dissociated proteins were removed and analyzed by western blotting experiments using an antibody specific for RNA Pol II. (E) Quantitative PCR analysis of RNA Pol II transfer from immobilized LCR templates to the β -globin gene promoter. Immobilized LCR/protein complex was incubated with or without a plasmid containing the β -globin gene with (β^+) or without (β^-) its promoter region for 30 min at 37°C. Unbound material was subjected to immunoprecipitation using IgG or RNA Pol II specific antibodies. The DNA was isolated from the precipitate and subjected to quantitative Real-Time PCR using primers specific for the β -globin gene. The experiment has been repeated and the error bars represent the SEM. An aliquot was taken from the immobilized LCR/protein complex in each transfer/dissociation assay and analyzed by western blotting using a RNA Pol II specific antibody (shown above the graph).



Western Blot: Pol II
 IP: with Pol II antibodies followed by PCR



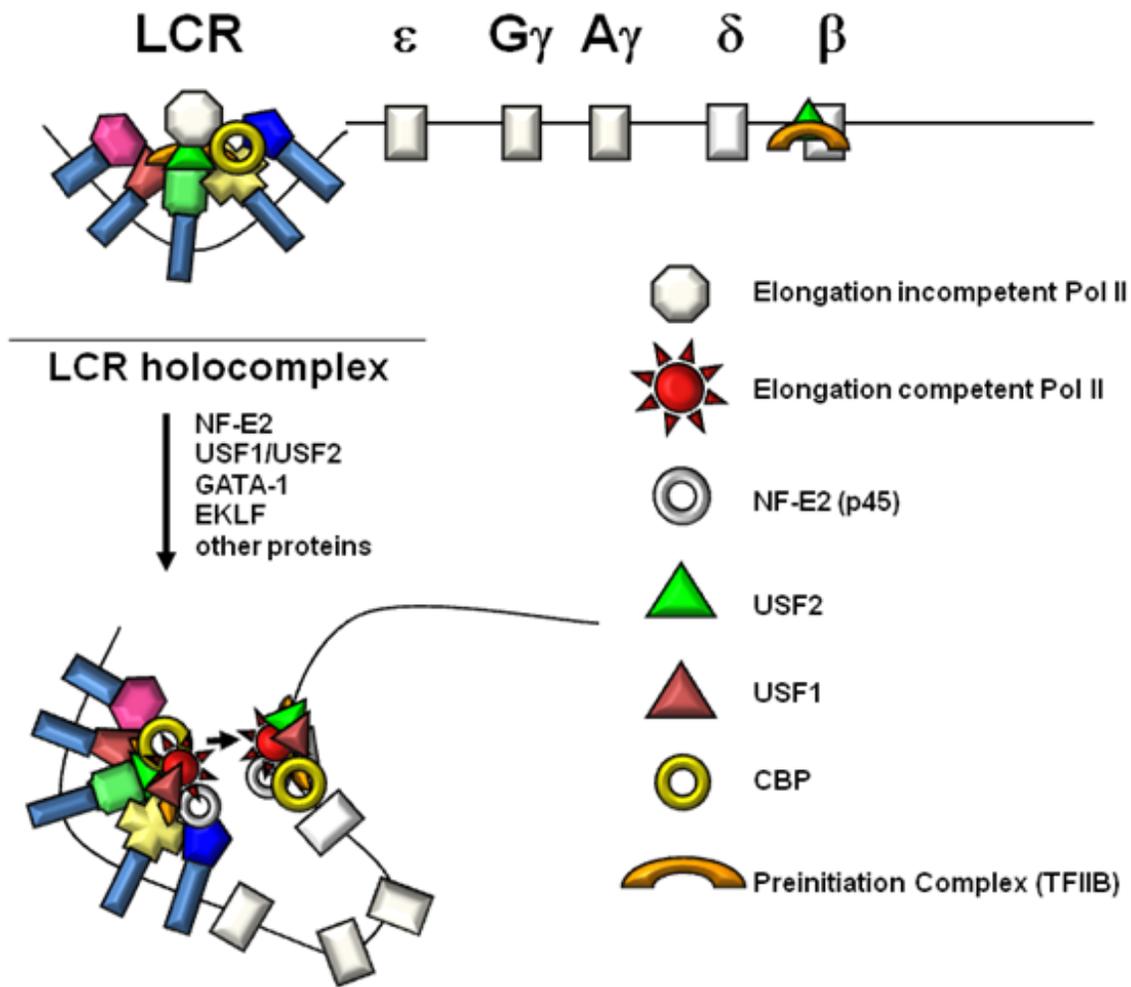


Figure 4-6. Model of NF-E2 and USF mediated assembly and transfer of elongation competent transcription complexes in the β -globin gene locus. Incomplete elongation incompetent Pol II transcription complexes are first recruited to the LCR. This is mediated in part by USF2 and its associated co-factor CBP. After erythroid differentiation, expression of NF-E2 (p45) and USF increases and these proteins increasingly associate with the LCR. This leads to the recruitment or assembly of transcriptionally competent Pol II complexes, and phosphorylation of the Pol II CTD. The differentiation of erythroid cells is also accompanied by a conformational change in the globin locus that juxtaposes the adult-globin gene with the LCR. This facilitates the transfer of elongation competent transcription complexes from the LCR to the adult-globin gene.

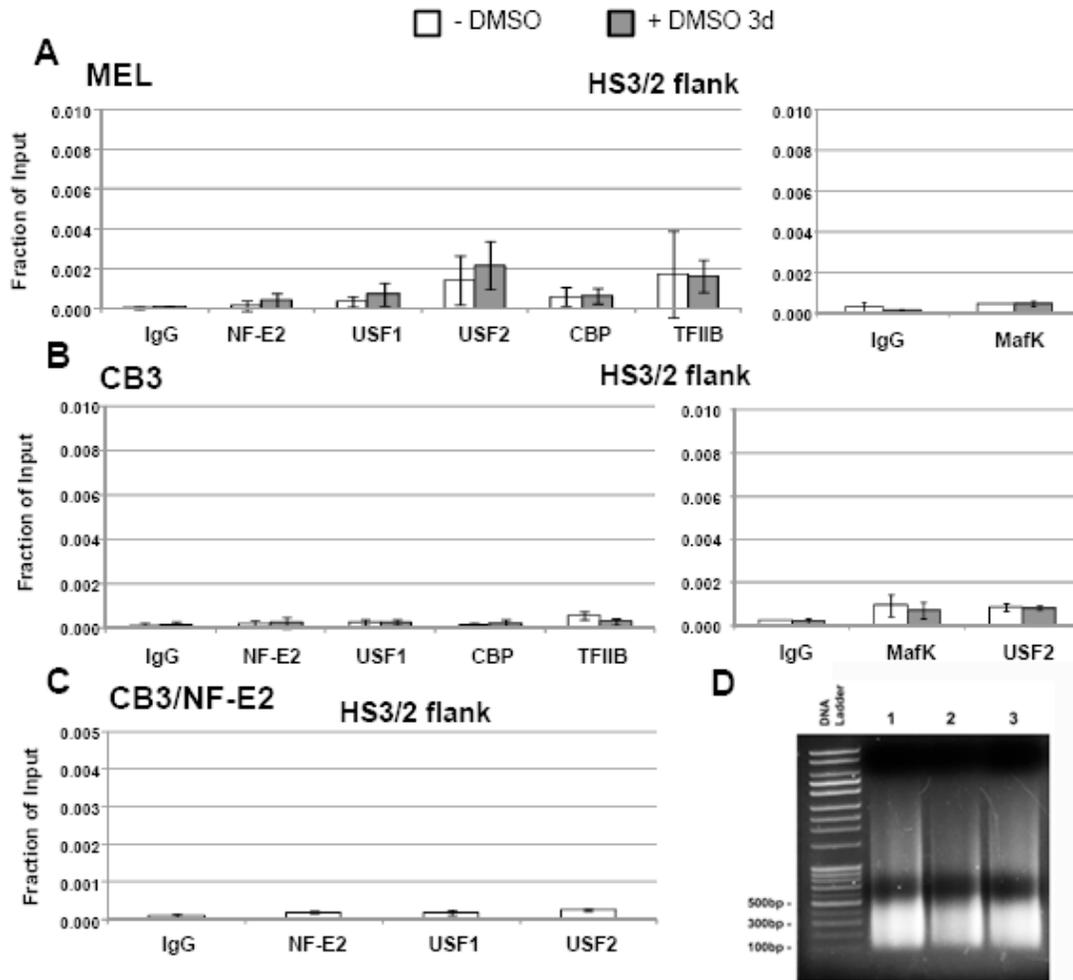


Figure 4-7. Interaction of protein complexes with the LCR HS3/2 flanking region in MEL, CB3 and CB3/NF-E2 cells. (A) ChIP analysis of protein chromatin interactions in the LCR HS3/2 flanking region (HS3/2 flank) in MEL cells incubated with or without 1.5% DMSO for 3 days. After crosslinking MEL cells with 1% formaldehyde, chromatin was isolated, fragmented by sonication, and subjected to immunoprecipitation with antibodies against NF-E2 (p45), MafK, USF1, USF2, CBP, and TFIIB. Reactions with the IgG antibody served as a negative control. The DNA was purified from the precipitate and subjected to qPCR using primers specific for LCR HS3/2 flanking region as indicated. Error bars represent standard deviations of three independent experiments. (B) ChIP analysis of protein chromatin interactions in the LCR HS3/2 flanking region in CB3 cells incubated with or without 1.5% DMSO for 3 days. DNA was isolated from immunoprecipitated material and analyzed as described in A. (C) ChIP analysis of protein chromatin interactions in CB3/NF-E2 cells. Crosslinked chromatin was precipitated with IgG or antibodies against NF-E2 (p45), USF1, or USF2 and DNA was analyzed as described in panel A. (D) Example of the extent of DNA fragmentation after sonication of crosslinked chromatin from MEL cells. After sonication, the DNA was electrophored in an agarose gel. The resulting DNA fragments are generally between 100bp and 500bp.

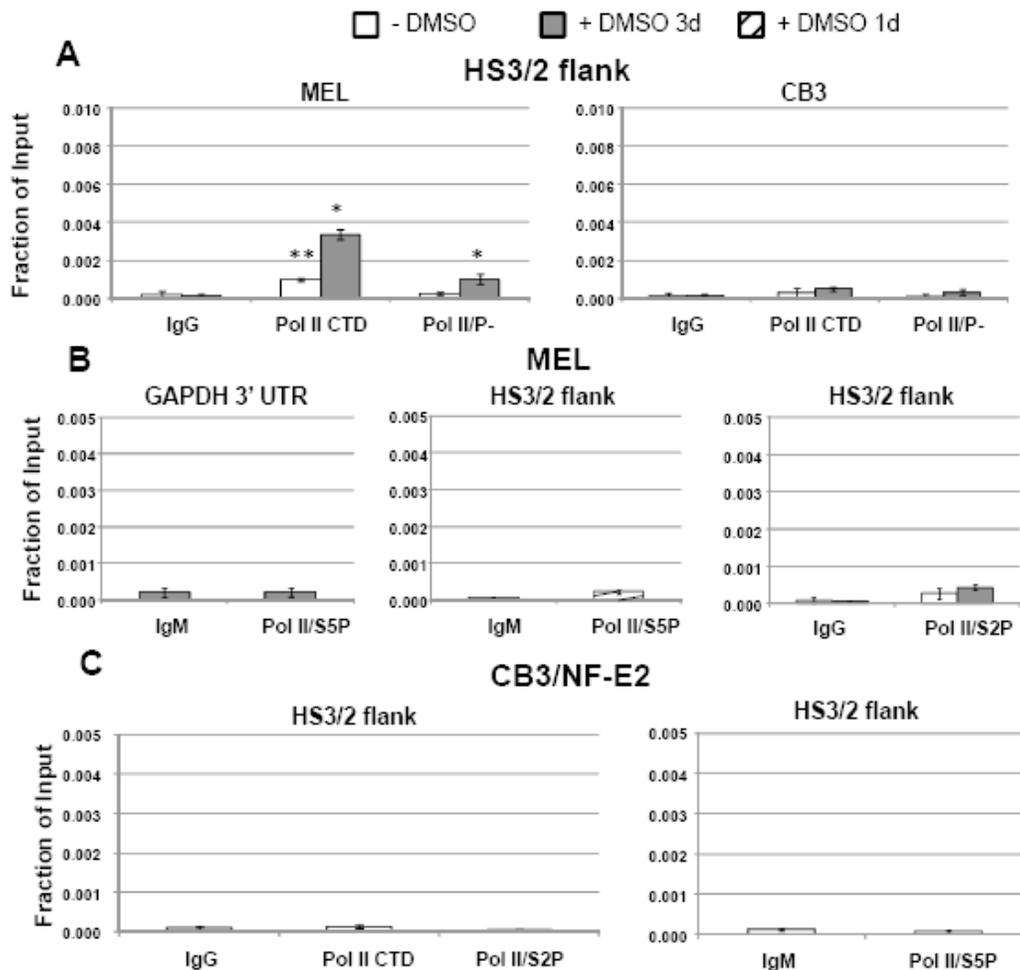


Figure 4-8. Interaction of RNA Pol II with the LCR HS3/2 flanking region or the GAPDH 3' untranslated region (UTR) in MEL, CB3 and CB3/NF-E2 cells. (A, B) ChIP analysis of RNA Pol II interactions with the HS3/2 flanking region (HS3/2 flank) or the GAPDH 3' UTR in MEL and CB3 cells incubated with or without 1.5% DMSO for 1 or 3 days. Chromatin was isolated from crosslinked cells, fragmented by sonication, and immunoprecipitated with antibodies specific for the RNA Pol II CTD (Pol II/CTD), for unphosphorylated RNA Pol II (Pol II/P-), or for RNA Pol II phosphorylated at serine 5 (Pol II/S5P) or serine 2 (Pol II/S2P) of the CTD. IgG or IgM antibodies were used in these experiments as negative controls. DNA was isolated from the precipitates and subjected to qPCR with DNA primers specific for the LCR HS3/2 flanking region or the GAPDH 3' UTR as indicated. Error bars represent standard deviations from three independent experiments. (C) ChIP analysis of Pol II interactions with the LCR HS3/2 flanking region in CB3/NF-E2 cells. Chromatin precipitation and DNA analysis by qPCR was performed as described in panels A and B. In A, * indicate p-values of <0.05 between DMSO induced versus uninduced samples; ** indicate p-values of <0.05 between specific antibody versus IgG.

CHAPTER 5 CONCLUSIONS AND FUTURE DIRECTIONS

***In vitro* ES Cell Erythroid Differentiation**

Overview

We have used an *in vitro* assay to induce murine ES cells to differentiate into the erythroid lineage and recapitulated the expression of β -like globin genes in a developmental manner. By using this system, we have observed that in the undifferentiated ES cells, there is no globin gene expression. The embryonic $\epsilon\gamma$ -globin genes are activated first as early as 5 days after differentiation and reaches their highest expression level at around Day 8. The adult β maj-globin gene is first observed at low levels at Day 8 and then it is up-regulated upon the initiation of definitive erythropoiesis (Day 10-12). At the mean time, the expression of embryonic $\epsilon\gamma$ -globin genes decreased. These results suggest that the ES cell differentiation system we used is a powerful tool for investigating hematopoietic development and differentiation.

However, from my data, it seems that the ES cells are not uniformly differentiated, probably due to their different initial states. Thus the cells that I collected at each time point are a mixture of cells at different differentiation stages. Even though the RT-PCR results show that the majority of the cells are at the same developmental stage, it still needs to be concerned with the variation generated by the small population of cells that are in a different developmental stage.

To eliminate this variation, one possibility is that we can use stem cell surface markers such as Oct4, Nanog and/or Sox2 to pre-sort the ES cells, followed by cell synchronization to arrest all the cells at the same stage before differentiation. However, due to the unique properties of ES cells, the effects of cell cycle arresting chemicals on the ES cell pluripotency are still under investigation (329). It has been shown that nocodazole (a drug often used to arrest cells at G2/M phase)-synchronized mouse ES cells show specific downregulation of cyclin-dependent kinase

(Cdk) 2, that establishes a somatic cell-like cell cycle in mouse ES cells and induces expression of differentiation markers (330). Alternatively, the differentiated cells collected at each time point can be sorted based on their surface markers before subject to further analysis. CD117, ERY-1 and Ter-119 are the most widely used cell surface markers to identify erythroid cells at different developmental stages. CD117 marks the most immature hematopoietic cells; ERY-1 is an intermediate mature marker; Ter-119 marks the mature erythroid cells (274). By sorting the cells with a combination of these three cell surface selection markers, we can obtain hematopoietic cells of the same developmental stage from our ES cells differentiation system.

The Antagonistic Role of USF and TFII-I in β -Globin Gene Regulation

By using our ES cells differentiation assay, I observed that TFII-I is detectable at HS2 and the adult β maj-globin gene promoter at Day 5, when embryonic $\epsilon\gamma$ -globin genes are expressed, while USF1 is found only at HS2 at this stage. At Day 8, USF1 interacts with both HS2 and the β maj-globin gene promoter, while TFII-I is no longer detectable at these regions. Previous studies in our laboratory have shown that, by overexpressing a dominant-negative form of USF (A-USF) or TFII-I (p70) in erythroid cells, USF activates and TFII-I represses adult β -globin gene expression (269). The data presented here further support this antagonistic regulatory role of USF and TFII-I.

However, even though the ES cell differentiation system has advantages over the erythroleukemia cell lines that are normally used for studying β -globin gene regulation, it cannot accurately reflect the identities of erythroid cells that differentiate during mammalian development. Hence it would be important to investigate the functions of USF and TFII-I during the differentiation and development of hematopoietic cells in animal models. Since USF or TFII-I knockout mice are embryonic lethal, our laboratory proceeded to create USF knockdown mice by expressing the dominant-negative form of USF in the mice. By placing the A-USF cDNA

under the control of the LCR and β -globin gene promoter and protecting them from position effects by a flanking pair of chicken HS4 insulators, these transgenes should be expressed exclusively in erythroid cells of the transgenic mice. Our laboratory has analyzed the A-USF transgenic mice and found that, although the heterozygous mice likely have the transgene integrated into X chromosome, when subjected to hemolytic anemia induction by phenylhydrazine (PHZ), there is a decrease in adult β maj-globin gene expression, accompanied by a reduction of RNA Pol II at the β maj-globin gene promoter (317). The same strategy could be pursued to analyze the function of TFII-I during erythroid development. For dominant negative TFII-I (p70) transgenic mice, we would expect to see effects on increased β maj-globin gene expression levels early in erythroid development. Currently, no p70 transgenic mice have been generated.

Recruitment of Histone Modification Enzymes to the β -Globin Locus by TFII-I

Covalent histone modifications, including methylation, acetylation, phosphorylation, ubiquitination, sumoylation and ADP-ribosylation, take place at the long, flexible N-terminal tails of the core histones. These modifications can alter chromatin structure by changing the histone-DNA and histone-histone contacts in order to create transcriptionally permissive or repressive chromatin environments around the genes (331). The enzymes that modify the histone tails are normally recruited to the gene by transcription factors.

In this study, I investigated whether TFII-I can recruit histone modifying enzymes to the β -globin gene locus. Previous studies have shown that TFII-I interacts with HDAC3, a histone deacetyltransferase, which can remove acetyl groups from the lysine residues at the histone tails and promote chromatin condensation to aid in transcriptional repression (332). Here I show that TFII-I also interacts with Suz12, a component of the PRC2 complex, in both embryonic and adult erythroid cells. The PRC2 complex contains another two main components, EED and the

histone methyltransferase, EZH2. EZH2 preferentially methylates H3K27, and EED is required for this process (296,299). I observed that Suz12 and the repressive chromatin mark, H3K27me₃, are detected at the silenced β major-globin gene promoter *in vitro* in Day 5 differentiated ES cells and that these associations disappear after 12 days of differentiation. To further prove that the recruitment of Suz12 by TFII-I to the β -globin locus does exert transcriptional repression, it would be essential to examine if EZH2 and EED are associated with the adult β -globin gene promoter. Also, since the repressive function of PRC2 on gene transcription is stabilized by recruitment of PRC1 to the H3K27 methylation site, it would be important to test if PRC1 is present at the β -globin gene promoter.

TFII-I is composed of six direct reiterated I-repeats, R1-R6, with each of them consisting of a putative helix-loop-helix (HLH) motif to interact with other proteins (262). Therefore, it would be interesting to determine if TFII-I interacts with both HDAC3 and Suz12 at the same time to exert strong repression, or if there is competition between HDAC3 and Suz12. TFII-I R5 and R6 domains are the transactivation domains. It has been reported that HDAC3 binds to the R3 and R4 domains of TFII-I (333). Therefore, it would be interesting to study if Suz12 binds to R3 and R4 as well or if it binds to other reiterated I-repeats. This can be done by expressing recombinant TFII-I mutants and Suz12 in *E.coli* and performing pull-down assays to see if they interact.

The Cooperative Role of NF-E2 and USF in β -Globin Gene Regulation

***In Vitro* Transfer/Dissociation Assay**

From the *in vitro* transfer/dissociation assay, I show that NF-E2 can dissociate RNA Pol II from the LCR in the presence of the wild type β -globin gene, while a dominant-negative form of USF (A-USF) can dissociate RNA Pol II from the LCR in the absence of the β -globin gene. These results suggest that USF is required for the recruitment of RNA Pol II to the LCR, while

NF-E2 is essential for the transfer of RNA Pol II from the LCR to the β -globin gene promoter. I also observed that NF-E2 facilitated RNA Pol II transfer from the LCR to the β -globin gene requires the presence of the +60 E-box element at the promoter. Our laboratory's previous data show that, in the adult erythroid environment (MEL cells extract) USF1 and USF2 bind to the +60 E-box element. Since NF-E2 is required for RNA Pol II transfer from the LCR to the β -globin promoter and NF-E2 interacts with USF, these data suggest that the transcription holocomplex including NF-E2, USF and RNA Pol II form on the LCR first, and then USF binds to the +60 E-box at the β -globin gene promoter. By directly interacting with NF-E2, USF pulls NF-E2 from the LCR to the β -globin gene promoter. Since NF-E2 also interacts with RNA Pol II, this process brings RNA Pol II to the β -globin promoter as well. This hypothesis is consistent with my ChIP and CoIP data shown in Chapter 4 in this thesis. It is also consistent with the results from a primer extension *in vitro* transcription experiment showing that the +60 E-box mutated β -globin gene is not transcribed. Due to the weak binding of NF-E2 to the MARE-like sequence in the β -globin gene promoter, it is reasonable to speculate that other interactions may exist to stabilize NF-E2 at the β -globin promoter (103). To examine if USF1 can stabilize NF-E2 at the β -globin promoter thus to promote β -globin gene expression, I could link a biotinylated wild-type or +60 E-box mutated β -globin gene to the streptavidin beads and incubate the beads-DNA construct with MEL nuclear extracts, then after removing unbound proteins, we can perform immunoblot analysis on the remaining bound proteins to detect NF-E2 and RNA Pol II. I would expect to see reduced NF-E2 and RNA Pol II protein in the samples incubated with the +60 E-box mutated β -globin gene.

The Different Role of USF1 and USF2 in β -Globin Gene Regulation

USF1 and USF2 can form homo- or heterodimers to regulate gene expression. In transgenic mouse studies, USF1 knockout mice only show slight behavioral abnormalities, while

USF2 knockout mice display obvious growth defects. It was shown that in USF1 knockout mice, there is an increased level of USF2, whereas in USF2 knockout mice, there is a reduced level of USF1 (234). These data suggest that the functions of USF1 and USF2 may not be completely redundant. Figure 3-4, 4-2, 4-4 (ChIP data from MEL/CB3 cells and ES differentiation, together with CoIP in MEL cells) and ChIP data show that USF2 and TFII-I associate with adult β -globin gene promoter in embryonic stage K562 cells, whereas USF1 and USF2 associate with the same promoter in adult environment MEL cells (103). These data demonstrate that USF1 and USF2 may exhibit distinct functions in the regulation of the β -globin locus. It is possible that USF1 functions as a β -globin gene activator, whereas USF2 functions as a dimerization partner which remains associated with the β -globin gene promoter during development. When the adult β -globin gene is activated, USF2 dimerizes with USF1 and binds to the +60 E-box of the β -globin gene promoter; when the β -globin gene is silenced, USF2 dimerizes with TFII-I and binds to the Inr/E-box of the promoter. In the undifferentiated mouse ES cells, we observe that USF2 is detectable at both LCR HS2 and the adult β maj-globin gene promoter, which further supports our hypothesis that the function of USF2 is like p18/MafK, which can change its dimerization partners during development to repress or activate adult β -globin gene.

NF-E2 p45 Knockout Mice vs. NF-E2 p45 Null Murine Erythroid Cells

From the studies utilizing NF-E2 p45 null CB3 cells, NF-E2 p45 is thought to be required for β -globin gene expression, recruitment of RNA Pol II to the β -globin promoter, and hyperacetylation of histones (49). However, NF-E2 p45 deficient mice show normal erythropoiesis and express the β -globin gene at ~70% of the levels observed in wild type mice (228). This difference was thought to be due to the presence of NF-E2 related factors (Nrf), but in the NF-E2 p45 knockout mice, although there is an increased binding of Nrf-2 at the LCR HS2, there is no change in the level of Nrf-2 association at the adult β maj-globin gene promoter.

Also, the association of p18/MafK is reduced at the LCR HS2 and almost lost at the β maj-globin gene promoter. In addition, mice lacking both NF-E2 p45 and Nrf-2 or Nrf-3 fail to display an erythroid defective phenotype more severe than NF-E2 p45 knockout mice. The active chromatin hub (ACH) in NF-E2 p45 knockout mice still forms normally, which suggests that NF-E2 is not required for the ACH formation. Therefore, further experiments such as generating NF-E2/Nrf2/Nrf3 triple knockout mice need to be done to investigate whether functional compensation by Nrfs exists in the NF-E2 knockout mice.

I suspect the striking differences between NF-E2 p45 null cells and null mice are due to failed ACH formation in CB3 cells which lack the endogenous NF-E2. While it is known that GATA-1 and EKLF are required for ACH formation, the levels of GATA-1 and EKLF in CB3 cells are not altered as compared to wild type MEL cells. However, in the CB3 cells, the absence of NF-E2 p45 causes the ~4-fold reduction of GATA-1 associated to the β maj-globin gene promoter and this reduction can be rescued by overexpressing NF-E2 p45 in the CB3 cells (47). This observation leads us to suspect that the lack of NF-E2 p45 in CB3 cells almost abolished GATA-1 binding at the β maj-globin gene promoter. This may disrupt the ACH formation in CB3 cells, thus abolishing RNA Pol II association at the promoter due to the lack of NF-E2 p45. In the NF-E2 p45 knockout mice, RNA Pol II can still bind to the β maj-globin gene promoter, which can be explained under the assumption that ACH formation is unperturbed in these mice. The transfer of RNA Pol II to the β -globin promoter may not necessarily require NF-E2 p45. Perhaps some transcription factors which interact with RNA Pol II (such as USF) are automatically transferred from the LCR to the promoter together with RNA Pol II due to the higher affinities of the transcription factors to the promoter than to the LCR and the closeness between the LCR and the promoter. However, this cannot be achieved in CB3 cells because

there is possibly no ACH formation. Thus the assistance of NF-E2 is required for RNA Pol II loading to the promoter. The difference is also possibly in part due to the somewhat artificial nature of CB3 cells, which are transformed cells that likely have very different expression profiles and chromatin structure.

My data from the NF-E2 p45 transfected CB3 (CB3/NF-E2) cells are consistent with this hypothesis. I observe that, when overexpressing NF-E2 p45 in CB3 cells, there is an increase of β -globin gene expression; although the level of expression is only 20% of that found in DMSO induced MEL cells. I also induced CB3/NF-E2 cells with DMSO, however, β -globin gene expression was not increased (data not shown). Thus we suspect that the low level expression of β -globin gene expression in CB3/NF-E2 cells is due to insufficient RNA Pol II loading onto the promoter directly by NF-E2 without the assistance of the LCR (since there is possibly no ACH in CB3 cells). My ChIP data show that there is no Serine 5 phosphorylated RNA Pol II associated with the LCR in CB3/NF-E2 cells even when the β -globin gene is expressed, further confirming this hypothesis.

Thus, I reason that the difference of β -globin gene expression between NF-E2 p45 null mice and CB3 cells could be related to the presence or absence of the ACH. Further experiments need to be performed to examine if the ACH exist in CB3 cells or not. It is possible that NF-E2 p45 and USF are not required for the ACH formation, but they function in the context of the loop. However, the possible redundancy between NF-E2 and Nrfs cannot be ruled out.

The Model of the LCR Regulated β -Globin Gene Expression

Based on the previous and current data from our laboratory and others, we propose a model of LCR regulated β -globin gene expression as follows:

Upon erythroid cell maturation, the LCR is relocated from within inaccessible chromatin territories (CT) to the surface of these territories (309) which are enriched for transcription

complexes in regions known as transcription factories (334). The association of the LCR with the transcription factories is mediated by USF and GATA-1, and synthesis of the unidirectional intergenic transcripts from the LCR to the downstream β -globin gene (284) could reel the β -globin gene to the surface of CTs. If the LCR is fixed at the transcription factories, the β -globin gene promoter can be reeled to the LCR by continuous transcription of the intervening intergenic region followed by association of the promoter with a transcription factory. Once the β -globin gene promoter is in close proximity to the LCR, GATA-1 and EKLF recruit the promoter to the LCR, while USF specifically binds to the +60 E-box of the promoter and loads NF-E2, together with RNA Pol II, onto the promoter. At the beginning of this process, unphosphorylated RNA Pol II is present in the LCR. Upon the induction of erythroid maturation, NF-E2 mediates RNA Pol II phosphorylation at the LCR and helps to deliver this transcriptionally competent holocomplex from the LCR to the promoter with the assistance of USF, thus activating β -globin gene expression.

LIST OF REFERENCES

1. Stamatoyannopoulos G., N.A. (1994) Hemoglobin switching. *The molecular basis of blood diseases*. W. B. Saunders, Philadelphia, PA., 107-155.
2. Weatherall, D.J. (2001) Phenotype-genotype relationships in monogenic disease: lessons from the thalassaemias. *Nat Rev Genet*, **2**, 245-255.
3. Angastiniotis, M. and Modell, B. (1998) Global epidemiology of hemoglobin disorders. *Ann N Y Acad Sci*, **850**, 251-269.
4. Roth, E.F., Jr., Shear, H.L., Costantini, F., Tanowitz, H.B. and Nagel, R.L. (1988) Malaria in beta-thalassemic mice and the effects of the transgenic human beta-globin gene and splenectomy. *J Lab Clin Med*, **111**, 35-41.
5. Weatherall, D.J., Clegg J. B. (2001) The thalassaemia syndromes. *Malden, Massachusetts: Blackwell Science*.
6. Pauling, L., Itano, H.A. and et al. (1949) Sickle cell anemia, a molecular disease. *Science*, **109**, 443.
7. Steinberg, M.H., Forget, B.G., Higgs, D.R., Nagel, R.L. (2001) Disorders of hemoglobin: genetics, pathophysiology and clinical management. *Cambridge: Cambridge University Press*.
8. Urbinati, F., Madigan, C. and Malik, P. (2006) Pathophysiology and therapy for haemoglobinopathies. Part II: thalassaemias. *Expert Rev Mol Med*, **8**, 1-26.
9. Weatherall, D.J., Clegg, J.B. (1981) The thalassaemia syndrome. *Oxford: Blackwell Scientific*.
10. Cooley, T.B., Lee, P. (1925) A series of cases of splenomegaly in children with anemia and peculiar bone changes. *Trans Am Pediatr Soc.*, **37**, 29.
11. Jensen, C.E., Tuck, S.M., Agnew, J.E., Koneru, S., Morris, R.W., Yardumian, A., Prescott, E., Hoffbrand, A.V. and Wonke, B. (1998) High prevalence of low bone mass in thalassaemia major. *Br J Haematol*, **103**, 911-915.
12. Otsuka, S., Maruyama, H. and Listowsky, I. (1981) Structure, assembly, conformation, and immunological properties of the two subunit classes of ferritin. *Biochemistry*, **20**, 5226-5232.
13. Gutteridge, J.M., Rowley, D.A. and Halliwell, B. (1981) Superoxide-dependent formation of hydroxyl radicals in the presence of iron salts. Detection of 'free' iron in biological systems by using bleomycin-dependent degradation of DNA. *Biochem J*, **199**, 263-265.

14. Bacon, B.R. and Britton, R.S. (1990) The pathology of hepatic iron overload: a free radical--mediated process? *Hepatology*, **11**, 127-137.
15. Bonkovsky, H.L. (1991) Iron and the liver. *Am J Med Sci*, **301**, 32-43.
16. Buja, L.M. and Roberts, W.C. (1971) Iron in the heart. Etiology and clinical significance. *Am J Med*, **51**, 209-221.
17. Giardina, P.J. and Grady, R.W. (2001) Chelation therapy in beta-thalassemia: an optimistic update. *Semin Hematol*, **38**, 360-366.
18. Ley, T.J., Griffith, P. and Nienhuis, A.W. (1982) Transfusion haemosiderosis and chelation therapy. *Clin Haematol*, **11**, 437-464.
19. Callender, S.T. and Weatherall, D.J. (1980) Iron chelation with oral desferrioxamine. *Lancet*, **2**, 689.
20. Giardini, C. and Lucarelli, G. (1994) Bone marrow transplantation in the treatment of thalassemia. *Curr Opin Hematol*, **1**, 170-176.
21. Lucarelli, G., Clift, R.A., Galimberti, M., Angelucci, E., Giardini, C., Baronciani, D., Polchi, P., Andreani, M., Gaziev, D., Erer, B. *et al.* (1999) Bone marrow transplantation in adult thalassemic patients. *Blood*, **93**, 1164-1167.
22. Sadelain, M., Boulad, F., Galanello, R., Giardina, P., Locatelli, F., Maggio, A., Rivella, S., Riviere, I. and Tisdale, J. (2007) Therapeutic options for patients with severe beta-thalassemia: the need for globin gene therapy. *Hum Gene Ther*, **18**, 1-9.
23. Dzierzak, E.A., Papayannopoulou, T. and Mulligan, R.C. (1988) Lineage-specific expression of a human beta-globin gene in murine bone marrow transplant recipients reconstituted with retrovirus-transduced stem cells. *Nature*, **331**, 35-41.
24. Bender, M.A., Gelinas, R.E. and Miller, A.D. (1989) A majority of mice show long-term expression of a human beta-globin gene after retrovirus transfer into hematopoietic stem cells. *Mol Cell Biol*, **9**, 1426-1434.
25. Sadelain, M., Wang, C.H., Antoniou, M., Grosveld, F. and Mulligan, R.C. (1995) Generation of a high-titer retroviral vector capable of expressing high levels of the human beta-globin gene. *Proc Natl Acad Sci U S A*, **92**, 6728-6732.
26. May, C., Rivella, S., Callegari, J., Heller, G., Gaensler, K.M., Luzzatto, L. and Sadelain, M. (2000) Therapeutic haemoglobin synthesis in beta-thalassaemic mice expressing lentivirus-encoded human beta-globin. *Nature*, **406**, 82-86.
27. Kohn, D.B., Sadelain, M. and Glorioso, J.C. (2003) Occurrence of leukaemia following gene therapy of X-linked SCID. *Nat Rev Cancer*, **3**, 477-488.

28. Emery, D.W., Yannaki, E., Tubb, J., Nishino, T., Li, Q. and Stamatoyannopoulos, G. (2002) Development of virus vectors for gene therapy of beta chain hemoglobinopathies: flanking with a chromatin insulator reduces gamma-globin gene silencing in vivo. *Blood*, **100**, 2012-2019.
29. Pawliuk, R., Westerman, K.A., Fabry, M.E., Payen, E., Tighe, R., Bouhassira, E.E., Acharya, S.A., Ellis, J., London, I.M., Eaves, C.J. *et al.* (2001) Correction of sickle cell disease in transgenic mouse models by gene therapy. *Science*, **294**, 2368-2371.
30. Levasseur, D.N., Ryan, T.M., Pawlik, K.M. and Townes, T.M. (2003) Correction of a mouse model of sickle cell disease: lentiviral/antisickling beta-globin gene transduction of unmobilized, purified hematopoietic stem cells. *Blood*, **102**, 4312-4319.
31. Puthenveetil, G., Scholes, J., Carbonell, D., Qureshi, N., Xia, P., Zeng, L., Li, S., Yu, Y., Hiti, A.L., Yee, J.K. *et al.* (2004) Successful correction of the human beta-thalassemia major phenotype using a lentiviral vector. *Blood*, **104**, 3445-3453.
32. Bank, A., Dorazio, R., Leboulch, P. (2005) A phase I/II clinical trial of β -globin gene therapy for betathalassemia. *Ann N Y Acad Sci.*, **1054**, 308-316.
33. Bollekens, J.A. and Forget, B.G. (1991) Delta beta thalassemia and hereditary persistence of fetal hemoglobin. *Hematol Oncol Clin North Am*, **5**, 399-422.
34. Platt, O.S., Orkin, S.H., Dover, G., Beardsley, G.P., Miller, B. and Nathan, D.G. (1984) Hydroxyurea increases fetal hemoglobin production in sickle cell anemia. *Trans Assoc Am Physicians*, **97**, 268-274.
35. Charache, S., Dover, G.J., Moore, R.D., Eckert, S., Ballas, S.K., Koshy, M., Milner, P.F., Orringer, E.P., Phillips, G., Jr., Platt, O.S. *et al.* (1992) Hydroxyurea: effects on hemoglobin F production in patients with sickle cell anemia. *Blood*, **79**, 2555-2565.
36. Charache, S., Terrin, M.L., Moore, R.D., Dover, G.J., Barton, F.B., Eckert, S.V., McMahan, R.P. and Bonds, D.R. (1995) Effect of hydroxyurea on the frequency of painful crises in sickle cell anemia. Investigators of the Multicenter Study of Hydroxyurea in Sickle Cell Anemia. *N Engl J Med*, **332**, 1317-1322.
37. Platt, O.S. (2008) Hydroxyurea for the treatment of sickle cell anemia. *N Engl J Med*, **358**, 1362-1369.
38. Ley, T.J., DeSimone, J., Anagnou, N.P., Keller, G.H., Humphries, R.K., Turner, P.H., Young, N.S., Keller, P. and Nienhuis, A. W. (1982) 5-azacytidine selectively increases gamma-globin synthesis in a patient with beta⁺ thalassemia. *N Engl J Med*, **307**, 1469-1475.
39. Torrealba-de Ron, A.T., Papayannopoulou, T., Knapp, M.S., Fu, M.F., Knitter, G. and Stamatoyannopoulos, G. (1984) Perturbations in the erythroid marrow progenitor cell pools may play a role in the augmentation of HbF by 5-azacytidine. *Blood*, **63**, 201-210.

40. Carr, B.I., Rahbar, S., Asmeron, Y., Riggs, A. and Winberg, C.D. (1988) Carcinogenicity and haemoglobin synthesis induction by cytidine analogues. *Br J Cancer*, **57**, 395-402.
41. Perrine, S.P., Greene, M.F. and Faller, D.V. (1985) Delay in the fetal globin switch in infants of diabetic mothers. *N Engl J Med*, **312**, 334-338.
42. Perrine, S.P., Ginder, G.D., Faller, D.V., Dover, G.H., Ikuta, T., Witkowska, H.E., Cai, S.P., Vichinsky, E.P. and Olivieri, N.F. (1993) A short-term trial of butyrate to stimulate fetal-globin-gene expression in the beta-globin disorders. *N Engl J Med*, **328**, 81-86.
43. Sealy, L. and Chalkley, R. (1978) The effect of sodium butyrate on histone modification. *Cell*, **14**, 115-121.
44. 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.
45. 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.
46. Stamatoyannopoulos, G.G., F. (2001) Hemoglobin switching. In: *Molecular Basis of Blood Disorders, 3rd edn.* (ed. by G. Stamatoyannopoulos, P.W. Majerus, R.M. Perlmutter & H. Varmus), pp135-182. WB Saunders, Philadelphia, PA.
47. Johnson, K.D., Grass, J.A., Boyer, M.E., Kiekhäfer, 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.
48. Horak, C.E., Mahajan, M.C., Luscombe, N.M., Gerstein, M., Weissman, S.M. and Snyder, M. (2002) GATA-1 binding sites mapped in the beta-globin locus by using mammalian ChIP-chip analysis. *Proc Natl Acad Sci U S A*, **99**, 2924-2929.
49. 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.
50. Elnitski, L., Miller, W. and Hardison, R. (1997) Conserved E boxes function as part of the enhancer in hypersensitive site 2 of the beta-globin locus control region. Role of basic helix-loop-helix proteins. *J Biol Chem*, **272**, 369-378.
51. Bieker, J.J. and Southwood, C.M. (1995) The erythroid Kruppel-like factor transactivation domain is a critical component for cell-specific inducibility of a beta-globin promoter. *Mol Cell Biol*, **15**, 852-860.

52. 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.
53. 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.
54. Mahajan, M.C., Narlikar, G.J., Boyapaty, G., Kingston, R.E. and Weissman, S.M. (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 U S A*, **102**, 15012-15017.
55. 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*, **279**, 50350-50357.
56. Letting, D.L., Rakowski, C., Weiss, M.J. and Blobel, G.A. (2003) Formation of a tissue-specific histone acetylation pattern by the hematopoietic transcription factor GATA-1. *Mol Cell Biol*, **23**, 1334-1340.
57. 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.
58. Epner, E., Reik, A., Cimborra, D., Telling, A., Bender, M.A., Fiering, S., Enver, T., Martin, D.I., Kennedy, M., Keller, G. *et al.* (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.
59. 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.
60. 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.
61. 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.
62. Enver, T., Ebens, A.J., Forrester, W.C. and Stamatoyannopoulos, G. (1989) The human beta-globin locus activation region alters the developmental fate of a human fetal globin gene in transgenic mice. *Proc Natl Acad Sci U S A*, **86**, 7033-7037.

63. Grosveld, F., Antoniou, M., Berry, M., De Boer, E., Dillon, N., Ellis, J., Fraser, P., Hanscombe, O., Hurst, J., Imam, A. *et al.* (1993) The regulation of human globin gene switching. *Philos Trans R Soc Lond B Biol Sci*, **339**, 183-191.
64. West, A.G., Huang, S., Gaszner, M., Litt, M.D. and Felsenfeld, G. (2004) Recruitment of histone modifications by USF proteins at a vertebrate barrier element. *Mol Cell*, **16**, 453-463.
65. Pikaart, M.J., Recillas-Targa, F. and Felsenfeld, G. (1998) Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. *Genes Dev*, **12**, 2852-2862.
66. Chung, J.H., Whiteley, M. and Felsenfeld, G. (1993) A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. *Cell*, **74**, 505-514.
67. Bungert, J., Tanimoto, K., Patel, S., Liu, Q., Fear, M. and Engel, J.D. (1999) Hypersensitive site 2 specifies a unique function within the human beta-globin locus control region to stimulate globin gene transcription. *Mol Cell Biol*, **19**, 3062-3072.
68. 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.
69. Fraser, P. and Grosveld, F. (1998) Locus control regions, chromatin activation and transcription. *Curr Opin Cell Biol*, **10**, 361-365.
70. Fedosyuk, H. and Peterson, K.R. (2007) Deletion of the human beta-globin LCR 5'HS4 or 5'HS1 differentially affects beta-like globin gene expression in beta-YAC transgenic mice. *Blood Cells Mol Dis*, **39**, 44-55.
71. 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.
72. Navas, P.A., Peterson, K.R., Li, Q., McArthur, M. and Stamatoyannopoulos, G. (2001) The 5'HS4 core element of the human beta-globin locus control region is required for high-level globin gene expression in definitive but not in primitive erythropoiesis. *J Mol Biol*, **312**, 17-26.
73. Schubeler, D., Francastel, C., Cimbara, 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.
74. 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.

75. 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.
76. Bender, M.A., Bulger, M., Close, J. and Groudine, M. (2000) 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.
77. 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.
78. Routledge, S.J. and Proudfoot, N.J. (2002) Definition of transcriptional promoters in the human beta globin locus control region. *J Mol Biol*, **323**, 601-611.
79. 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.
80. Wobbe, C.R. and Struhl, K. (1990) Yeast and human TATA-binding proteins have nearly identical DNA sequence requirements for transcription in vitro. *Mol Cell Biol*, **10**, 3859-3867.
81. Nienhuis, A.W., Anagnou, N.P. and Ley, T.J. (1984) Advances in thalassemia research. *Blood*, **63**, 738-758.
82. Dierks, P., van Ooyen, A., Cochran, M.D., Dobkin, C., Reiser, J. and Weissmann, C. (1983) Three regions upstream from the cap site are required for efficient and accurate transcription of the rabbit beta-globin gene in mouse 3T6 cells. *Cell*, **32**, 695-706.
83. Cowie, A. and Myers, R.M. (1988) DNA sequences involved in transcriptional regulation of the mouse beta-globin promoter in murine erythroleukemia cells. *Mol Cell Biol*, **8**, 3122-3128.
84. Antoniou, M. and Grosveld, F. (1990) beta-globin dominant control region interacts differently with distal and proximal promoter elements. *Genes Dev*, **4**, 1007-1013.
85. Stewart, J.J. and Stargell, L.A. (2001) The stability of the TFIIA-TBP-DNA complex is dependent on the sequence of the TATAAA element. *J Biol Chem*, **276**, 30078-30084.
86. 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.
87. deBoer, E., Antoniou, M., Mignotte, V., Wall, L. and Grosveld, F. (1988) The human beta-globin promoter; nuclear protein factors and erythroid specific induction of transcription. *Embo J*, **7**, 4203-4212.

88. Delvoe, N.L., Destroismaisons, N.M. and Wall, L.A. (1993) Activation of the beta-globin promoter by the locus control region correlates with binding of a novel factor to the CAAT box in murine erythroleukemia cells but not in K562 cells. *Mol Cell Biol*, **13**, 6969-6983.
89. Wall, L., Destroismaisons, N., Delvoe, N. and Guy, L.G. (1996) CAAT/enhancer-binding proteins are involved in beta-globin gene expression and are differentially expressed in murine erythroleukemia and K562 cells. *J Biol Chem*, **271**, 16477-16484.
90. Zafarana, G., Rottier, R., Grosveld, F. and Philipsen, S. (2000) Erythroid overexpression of C/EBPgamma in transgenic mice affects gamma-globin expression and fetal liver erythropoiesis. *Embo J*, **19**, 5856-5863.
91. Gordon, C.T., Fox, V.J., Najdovska, S. and Perkins, A.C. (2005) C/EBPdelta and C/EBPgamma bind the CCAAT-box in the human beta-globin promoter and modulate the activity of the CACC-box binding protein, EKLF. *Biochim Biophys Acta*, **1729**, 74-80.
92. 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.
93. Sinha, S., Maity, S.N., Lu, J. and de Crombrughe, B. (1995) Recombinant rat CBF-C, the third subunit of CBF/NFY, allows formation of a protein-DNA complex with CBF-A and CBF-B and with yeast HAP2 and HAP3. *Proc Natl Acad Sci U S A*, **92**, 1624-1628.
94. Lewis, B.A. and Orkin, S.H. (1995) A functional initiator element in the human beta-globin promoter. *J Biol Chem*, **270**, 28139-28144.
95. Roy, A.L., Meisterernst, M., Pognonec, P. and Roeder, R.G. (1991) Cooperative interaction of an initiator-binding transcription initiation factor and the helix-loop-helix activator USF. *Nature*, **354**, 245-248.
96. Li, Y., Flanagan, P.M., Tschochner, H. and Kornberg, R.D. (1994) RNA polymerase II initiation factor interactions and transcription start site selection. *Science*, **263**, 805-807.
97. Smale, S.T. and Baltimore, D. (1989) The "initiator" as a transcription control element. *Cell*, **57**, 103-113.
98. Zenzie-Gregory, B., O'Shea-Greenfield, A. and Smale, S.T. (1992) Similar mechanisms for transcription initiation mediated through a TATA box or an initiator element. *J Biol Chem*, **267**, 2823-2830.
99. Lewis, B.A., Kim, T.K. and Orkin, S.H. (2000) A downstream element in the human beta-globin promoter: evidence of extended sequence-specific transcription factor IID contacts. *Proc Natl Acad Sci U S A*, **97**, 7172-7177.

100. Seto, E., Shi, Y. and Shenk, T. (1991) YY1 is an initiator sequence-binding protein that directs and activates transcription in vitro. *Nature*, **354**, 241-245.
101. Roy, A.L., Malik, S., Meisterernst, M. and Roeder, R.G. (1993) An alternative pathway for transcription initiation involving TFII-I. *Nature*, **365**, 355-359.
102. Weis, L. and Reinberg, D. (1992) Transcription by RNA polymerase II: initiator-directed formation of transcription-competent complexes. *Faseb J*, **6**, 3300-3309.
103. 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.
104. Kang, S.H., Vieira, K. and Bungert, J. (2002) Combining chromatin immunoprecipitation and DNA footprinting: a novel method to analyze protein-DNA interactions in vivo. *Nucleic Acids Res*, **30**, e44.
105. Donovan-Peluso, M., Acuto, S., O'Neill, D., Hom, A., Maggio, A. and Bank, A. (1991) The beta globin 3' enhancer element confers regulated expression on the human gamma globin gene in the human embryonic-fetal erythroleukemia cell line K562. *Blood*, **77**, 855-860.
106. Trudel, M. and Costantini, F. (1987) A 3' enhancer contributes to the stage-specific expression of the human beta-globin gene. *Genes Dev*, **1**, 954-961.
107. 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.
108. Stamatoyannopoulos G, G.F. (2001) Hemoglobin switching. In: *Stamatoyannopoulos G, Perlmutter R, Majerus W, Varmus H (eds) The Molecular Basis of Blood Diseases. Saunders, Philadelphia*, pp135-165.
109. Weatherall D, C.J. (1981) The Thalassemia Syndromes. *Oxford, UK Blackwell Scientific*.
110. Feingold, E.A. and Forget, B.G. (1989) The breakpoint of a large deletion causing hereditary persistence of fetal hemoglobin occurs within an erythroid DNA domain remote from the beta-globin gene cluster. *Blood*, **74**, 2178-2186.
111. Stamatoyannopoulos, G. (2005) Control of globin gene expression during development and erythroid differentiation. *Exp Hematol*, **33**, 259-271.
112. Kingsley, P.D., Malik, J., Emerson, R.L., Bushnell, T.P., McGrath, K.E., Bloedorn, L.A., Bulger, M. and Palis, J. (2006) "Maturation" globin switching in primary primitive erythroid cells. *Blood*, **107**, 1665-1672.

113. Gaensler, K.M., Kitamura, M. and Kan, Y.W. (1993) Germ-line transmission and developmental regulation of a 150-kb yeast artificial chromosome containing the human beta-globin locus in transgenic mice. *Proc Natl Acad Sci U S A*, **90**, 11381-11385.
114. Behringer, R.R., Ryan, T.M., Palmiter, R.D., Brinster, R.L. and Townes, T.M. (1990) Human gamma- to beta-globin gene switching in transgenic mice. *Genes Dev*, **4**, 380-389.
115. 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.
116. Raich, N., Enver, T., Nakamoto, B., Josephson, B., Papayannopoulou, T. and Stamatoyannopoulos, G. (1990) Autonomous developmental control of human embryonic globin gene switching in transgenic mice. *Science*, **250**, 1147-1149.
117. 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.
118. Li, Q., Blau, C.A., Clegg, C.H., Rohde, A. and Stamatoyannopoulos, G. (1998) Multiple epsilon-promoter elements participate in the developmental control of epsilon-globin genes in transgenic mice. *J Biol Chem*, **273**, 17361-17367.
119. Raich, N., Clegg, C.H., Grofti, J., Romeo, P.H. and Stamatoyannopoulos, G. (1995) GATA1 and YY1 are developmental repressors of the human epsilon-globin gene. *Embo J*, **14**, 801-809.
120. Filipe, A., Li, Q., Deveaux, S., Godin, I., Romeo, P.H., Stamatoyannopoulos, G. and Mignotte, V. (1999) Regulation of embryonic/fetal globin genes by nuclear hormone receptors: a novel perspective on hemoglobin switching. *Embo J*, **18**, 687-697.
121. 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.
122. 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.
123. Dillon, N. and Grosveld, F. (1991) Human gamma-globin genes silenced independently of other genes in the beta-globin locus. *Nature*, **350**, 252-254.
124. Stamatoyannopoulos, G., Josephson, B., Zhang, J.W. and Li, Q. (1993) Developmental regulation of human gamma-globin genes in transgenic mice. *Mol Cell Biol*, **13**, 7636-7644.

125. Harju, S., Navas, P.A., Stamatoyannopoulos, G. and Peterson, K.R. (2003) Role of gamma-globin gene silencing and chromatin sub-domain in globin gene switching. *Blood Cells Mol Dis*, **331**, 145-146.
126. Luo, H.Y., Mang, D., Patrinos, G.P. et al. (2004) A mutation in a GATA-1 binding Site 5' to the G γ -globin gene (nt -567, T>G) may be associated with increased levels of fetal hemoglobin. *Blood (ASH Annual Meeting Abstracts)*, **104**, Abstract 500.
127. Li, Q., Fang, X., Han, H. and Stamatoyannopoulos, G. (2004) The minimal promoter plays a major role in silencing of the galago gamma-globin gene in adult erythropoiesis. *Proc Natl Acad Sci U S A*, **101**, 8096-8101.
128. Asano, H. and Stamatoyannopoulos, G. (1998) Activation of beta-globin promoter by erythroid Kruppel-like factor. *Mol Cell Biol*, **18**, 102-109.
129. Donze, D., Townes, T.M. and Bieker, J.J. (1995) Role of erythroid Kruppel-like factor in human gamma- to beta-globin gene switching. *J Biol Chem*, **270**, 1955-1959.
130. Bieker, J.J. (1996) Isolation, genomic structure, and expression of human erythroid Kruppel-like factor (EKLF). *DNA Cell Biol*, **15**, 347-352.
131. 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.
132. Asano, H., Li, X.S. and Stamatoyannopoulos, G. (1999) FKLF, a novel Kruppel-like factor that activates human embryonic and fetal beta-like globin genes. *Mol Cell Biol*, **19**, 3571-3579.
133. Asano, H., Li, X.S. and Stamatoyannopoulos, G. (2000) FKLF-2: a novel Kruppel-like transcriptional factor that activates globin and other erythroid lineage genes. *Blood*, **95**, 3578-3584.
134. Emery, D.W., Gavriilidis, G., Asano, H. and Stamatoyannopoulos, G. (2007) The transcription factor KLF11 can induce gamma-globin gene expression in the setting of in vivo adult erythropoiesis. *J Cell Biochem*, **100**, 1045-1055.
135. Song, C.Z., Gavriilidis, G., Asano, H. and Stamatoyannopoulos, G. (2005) Functional study of transcription factor KLF11 by targeted gene inactivation. *Blood Cells Mol Dis*, **34**, 53-59.
136. Zhou, M., McPherson, L., Feng, D., Song, A., Dong, C., Lyu, S.C., Zhou, L., Shi, X., Ahn, Y.T., Wang, D. et al. (2007) Kruppel-like transcription factor 13 regulates T lymphocyte survival in vivo. *J Immunol*, **178**, 5496-5504.
137. Ristaldi, M.S., Drabek, D., Gribnau, J., Poddie, D., Yannoutsos, N., Cao, A., Grosveld, F. and Imam, A.M. (2001) The role of the -50 region of the human gamma-globin gene in switching. *Embo J*, **20**, 5242-5249.

138. Zhou, W., Clouston, D.R., Wang, X., Cerruti, L., Cunningham, J.M. and Jane, S.M. (2000) Induction of human fetal globin gene expression by a novel erythroid factor, NF-E4. *Mol Cell Biol*, **20**, 7662-7672.
139. Ramamurthy, L., Barbour, V., Tuckfield, A., Clouston, D.R., Topham, D., Cunningham, J.M. and Jane, S.M. (2001) Targeted disruption of the CP2 gene, a member of the NTF family of transcription factors. *J Biol Chem*, **276**, 7836-7842.
140. Ritchie, H.H., Wang, L.H., Tsai, S., O'Malley, B.W. and Tsai, M.J. (1990) COUP-TF gene: a structure unique for the steroid/thyroid receptor superfamily. *Nucleic Acids Res*, **18**, 6857-6862.
141. 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.
142. Ottolenghi, S., Mantovani, R., Nicolis, S., Ronchi, A. and Giglioni, B. (1989) DNA sequences regulating human globin gene transcription in nondeletional hereditary persistence of fetal hemoglobin. *Hemoglobin*, **13**, 523-541.
143. Vitale, M., Di Marzo, R., Calzolari, R., Acuto, S., O'Neill, D., Bank, A. and Maggio, A. (1994) Evidence for a globin promoter-specific silencer element located upstream of the human delta-globin gene. *Biochem Biophys Res Commun*, **204**, 413-418.
144. Acuto, S., Urzi, G., Schimmenti, S., Maggio, A., O'Neill, D. and Bank, A. (1996) An element upstream from the human delta-globin-encoding gene specifically enhances beta-globin reporter gene expression in murine erythroleukemia cells. *Gene*, **168**, 237-241.
145. O'Neill, D., Yang, J., Erdjument-Bromage, H., Bornschlegel, K., Tempst, P. and Bank, A. (1999) Tissue-specific and developmental stage-specific DNA binding by a mammalian SWI/SNF complex associated with human fetal-to-adult globin gene switching. *Proc Natl Acad Sci U S A*, **96**, 349-354.
146. O'Neill, D., Bornschlegel, K., Flamm, M., Castle, M. and Bank, A. (1991) A DNA-binding factor in adult hematopoietic cells interacts with a pyrimidine-rich domain upstream from the human delta-globin gene. *Proc Natl Acad Sci U S A*, **88**, 8953-8957.
147. Georgopoulos, K. (1997) Transcription factors required for lymphoid lineage commitment. *Curr Opin Immunol*, **9**, 222-227.
148. 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.
149. 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.

150. Travers, A. (1999) Chromatin modification by DNA tracking. *Proc Natl Acad Sci U S A*, **96**, 13634-13637.
151. Dorsett, D. (1999) Distant liaisons: long-range enhancer-promoter interactions in *Drosophila*. *Curr Opin Genet Dev*, **9**, 505-514.
152. Bulger, M. and Groudine, M. (1999) Looping versus linking: toward a model for long-distance gene activation. *Genes Dev*, **13**, 2465-2477.
153. Blackwood, E.M. and Kadonaga, J.T. (1998) Going the distance: a current view of enhancer action. *Science*, **281**, 60-63.
154. 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.
155. Splinter, E., Heath, H., Kooren, J., Palstra, R.J., Klous, P., Grosveld, F., Galjart, N. and de Laat, W. (2006) CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev*, **20**, 2349-2354.
156. 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.
157. Vakoc, C.R., Letting, D.L., Gheldof, N., Sawado, T., Bender, M.A., Groudine, M., Weiss, M.J., Dekker, J. and Blobel, G.A. (2005) Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. *Mol Cell*, **17**, 453-462.
158. Dahmus, M.E. (1994) The role of multisite phosphorylation in the regulation of RNA polymerase II activity. *Prog Nucleic Acid Res Mol Biol*, **48**, 143-179.
159. Hampsey, M. and Reinberg, D. (2003) Tails of intrigue: phosphorylation of RNA polymerase II mediates histone methylation. *Cell*, **113**, 429-432.
160. Maniatis, T. and Reed, R. (2002) An extensive network of coupling among gene expression machines. *Nature*, **416**, 499-506.
161. Dahmus, M.E. (1996) Reversible phosphorylation of the C-terminal domain of RNA polymerase II. *J Biol Chem*, **271**, 19009-19012.
162. Conaway, R.C. and Conaway, J.W. (1993) General initiation factors for RNA polymerase II. *Annu Rev Biochem*, **62**, 161-190.
163. Zawel, L. and Reinberg, D. (1995) Common themes in assembly and function of eukaryotic transcription complexes. *Annu Rev Biochem*, **64**, 533-561.
164. Akoulitchev, S., Makela, T.P., Weinberg, R.A. and Reinberg, D. (1995) Requirement for TFIIF kinase activity in transcription by RNA polymerase II. *Nature*, **377**, 557-560.

165. Rodriguez, C.R., Cho, E.J., Keogh, M.C., Moore, C.L., Greenleaf, A.L. and Buratowski, S. (2000) Kin28, the TFIIF-associated carboxy-terminal domain kinase, facilitates the recruitment of mRNA processing machinery to RNA polymerase II. *Mol Cell Biol*, **20**, 104-112.
166. Orphanides, G. and Reinberg, D. (2002) A unified theory of gene expression. *Cell*, **108**, 439-451.
167. Sawadogo, M. and Sentenac, A. (1990) RNA polymerase B (II) and general transcription factors. *Annu Rev Biochem*, **59**, 711-754.
168. Li, J., Moazed, D. and Gygi, S.P. (2002) Association of the histone methyltransferase Set2 with RNA polymerase II plays a role in transcription elongation. *J Biol Chem*, **277**, 49383-49388.
169. Li, B., Howe, L., Anderson, S., Yates, J.R., 3rd and Workman, J.L. (2003) The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J Biol Chem*, **278**, 8897-8903.
170. Krogan, N.J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D.P., Beattie, B.K., Emili, A., Boone, C. *et al.* (2003) Methylation of histone H3 by Set2 in *Saccharomyces cerevisiae* is linked to transcriptional elongation by RNA polymerase II. *Mol Cell Biol*, **23**, 4207-4218.
171. Chambers, R.S. and Dahmus, M.E. (1994) Purification and characterization of a phosphatase from HeLa cells which dephosphorylates the C-terminal domain of RNA polymerase II. *J Biol Chem*, **269**, 26243-26248.
172. Chambers, R.S., Wang, B.Q., Burton, Z.F. and Dahmus, M.E. (1995) The activity of COOH-terminal domain phosphatase is regulated by a docking site on RNA polymerase II and by the general transcription factors IIF and IIB. *J Biol Chem*, **270**, 14962-14969.
173. Ko, L.J. and Engel, J.D. (1993) DNA-binding specificities of the GATA transcription factor family. *Mol Cell Biol*, **13**, 4011-4022.
174. Martin, D.I., Zon, L.I., Mutter, G. and Orkin, S.H. (1990) Expression of an erythroid transcription factor in megakaryocytic and mast cell lineages. *Nature*, **344**, 444-447.
175. Merika, M. and Orkin, S.H. (1993) DNA-binding specificity of GATA family transcription factors. *Mol Cell Biol*, **13**, 3999-4010.
176. Yamamoto, M., Ko, L.J., Leonard, M.W., Beug, H., Orkin, S.H. and Engel, J.D. (1990) Activity and tissue-specific expression of the transcription factor NF-E1 multigene family. *Genes Dev*, **4**, 1650-1662.
177. Whyatt, D.J., deBoer, E. and Grosveld, F. (1993) The two zinc finger-like domains of GATA-1 have different DNA binding specificities. *Embo J*, **12**, 4993-5005.

178. 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.
179. Lee, E.Y., Chang, C.Y., Hu, N., Wang, Y.C., Lai, C.C., Herrup, K., Lee, W.H. and Bradley, A. (1992) Mice deficient for Rb are nonviable and show defects in neurogenesis and haematopoiesis. *Nature*, **359**, 288-294.
180. 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.
181. Ponka, P. (1997) Tissue-specific regulation of iron metabolism and heme synthesis: distinct control mechanisms in erythroid cells. *Blood*, **89**, 1-25.
182. Zon, L.I., Mather, C., Burgess, S., Bolce, M.E., Harland, R.M. and Orkin, S.H. (1991) Expression of GATA-binding proteins during embryonic development in *Xenopus laevis*. *Proc Natl Acad Sci U S A*, **88**, 10642-10646.
183. Ito, E., Toki, T., Ishihara, H., Ohtani, H., Gu, L., Yokoyama, M., Engel, J.D. and Yamamoto, M. (1993) Erythroid transcription factor GATA-1 is abundantly transcribed in mouse testis. *Nature*, **362**, 466-468.
184. Partington, G.A. and Patient, R.K. (1999) Phosphorylation of GATA-1 increases its DNA-binding affinity and is correlated with induction of human K562 erythroleukaemia cells. *Nucleic Acids Res*, **27**, 1168-1175.
185. Warren, A.J., Colledge, W.H., Carlton, M.B., Evans, M.J., Smith, A.J. and Rabbitts, T.H. (1994) The oncogenic cysteine-rich LIM domain protein rbtn2 is essential for erythroid development. *Cell*, **78**, 45-57.
186. 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.
187. Marin, M., Karis, A., Visser, P., Grosveld, F. and Philipsen, S. (1997) Transcription factor Sp1 is essential for early embryonic development but dispensable for cell growth and differentiation. *Cell*, **89**, 619-628.
188. Fox, A.H., Kowalski, K., King, G.F., Mackay, J.P. and Crossley, M. (1998) Key residues characteristic of GATA N-fingers are recognized by FOG. *J Biol Chem*, **273**, 33595-33603.
189. Gregory, R.C., Taxman, D.J., Seshasayee, D., Kensinger, M.H., Bieker, J.J. and Wojchowski, D.M. (1996) Functional interaction of GATA1 with erythroid Kruppel-like factor and Sp1 at defined erythroid promoters. *Blood*, **87**, 1793-1801.
190. Lipinski, M.M. and Jacks, T. (1999) The retinoblastoma gene family in differentiation and development. *Oncogene*, **18**, 7873-7882.

191. Tsai, S.F., Strauss, E. and Orkin, S.H. (1991) Functional analysis and in vivo footprinting implicate the erythroid transcription factor GATA-1 as a positive regulator of its own promoter. *Genes Dev*, **5**, 919-931.
192. Tsai, S.F., Martin, D.I., Zon, L.I., D'Andrea, A.D., Wong, G.G. and Orkin, S.H. (1989) Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. *Nature*, **339**, 446-451.
193. Fox, A.H., Liew, C., Holmes, M., Kowalski, K., Mackay, J. and Crossley, M. (1999) Transcriptional cofactors of the FOG family interact with GATA proteins by means of multiple zinc fingers. *Embo J*, **18**, 2812-2822.
194. Weiss, M.J. and Orkin, S.H. (1995) Transcription factor GATA-1 permits survival and maturation of erythroid precursors by preventing apoptosis. *Proc Natl Acad Sci U S A*, **92**, 9623-9627.
195. Crossley, M., Merika, M. and Orkin, S.H. (1995) Self-association of the erythroid transcription factor GATA-1 mediated by its zinc finger domains. *Mol Cell Biol*, **15**, 2448-2456.
196. Hung, H.L., Lau, J., Kim, A.Y., Weiss, M.J. and Blobel, G.A. (1999) CREB-Binding protein acetylates hematopoietic transcription factor GATA-1 at functionally important sites. *Mol Cell Biol*, **19**, 3496-3505.
197. Boyes, J., Byfield, P., Nakatani, Y. and Ogryzko, V. (1998) Regulation of activity of the transcription factor GATA-1 by acetylation. *Nature*, **396**, 594-598.
198. Barrett, D.M., Gustafson, K.S., Wang, J., Wang, S.Z. and Ginder, G.D. (2004) A GATA factor mediates cell type-restricted induction of HLA-E gene transcription by gamma interferon. *Mol Cell Biol*, **24**, 6194-6204.
199. 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.
200. Crossley, M. and Orkin, S.H. (1994) Phosphorylation of the erythroid transcription factor GATA-1. *J Biol Chem*, **269**, 16589-16596.
201. Collavin, L., Gostissa, M., Avolio, F., Secco, P., Ronchi, A., Santoro, C. and Del Sal, G. (2004) Modification of the erythroid transcription factor GATA-1 by SUMO-1. *Proc Natl Acad Sci U S A*, **101**, 8870-8875.
202. Osada, H., Grutz, G., Axelson, H., Forster, A. and Rabbitts, T.H. (1995) Association of erythroid transcription factors: complexes involving the LIM protein RBTN2 and the zinc-finger protein GATA1. *Proc Natl Acad Sci U S A*, **92**, 9585-9589.

203. Wadman, I., Li, J., Bash, R.O., Forster, A., Osada, H., Rabbitts, T.H. and Baer, R. (1994) Specific in vivo association between the bHLH and LIM proteins implicated in human T cell leukemia. *Embo J*, **13**, 4831-4839.
204. Evans, T., Reitman, M. and Felsenfeld, G. (1988) An erythrocyte-specific DNA-binding factor recognizes a regulatory sequence common to all chicken globin genes. *Proc Natl Acad Sci U S A*, **85**, 5976-5980.
205. Weiss, M.J., Yu, C. and Orkin, S.H. (1997) Erythroid-cell-specific properties of transcription factor GATA-1 revealed by phenotypic rescue of a gene-targeted cell line. *Mol Cell Biol*, **17**, 1642-1651.
206. Pevny, L., Lin, C.S., D'Agati, V., Simon, M.C., Orkin, S.H. and Costantini, F. (1995) Development of hematopoietic cells lacking transcription factor GATA-1. *Development*, **121**, 163-172.
207. Shimizu, R., Takahashi, S., Ohneda, K., Engel, J.D. and Yamamoto, M. (2001) In vivo requirements for GATA-1 functional domains during primitive and definitive erythropoiesis. *Embo J*, **20**, 5250-5260.
208. Whyatt, D., Lindeboom, F., Karis, A., Ferreira, R., Milot, E., Hendriks, R., de Bruijn, M., Langeveld, A., Gribnau, J., Grosveld, F. *et al.* (2000) An intrinsic but cell-nonautonomous defect in GATA-1-overexpressing mouse erythroid cells. *Nature*, **406**, 519-524.
209. Kulesa, H., Frampton, J. and Graf, T. (1995) GATA-1 reprograms avian myelomonocytic cell lines into eosinophils, thromboblats, and erythroblats. *Genes Dev*, **9**, 1250-1262.
210. Hirasawa, R., Shimizu, R., Takahashi, S., Osawa, M., Takayanagi, S., Kato, Y., Onodera, M., Minegishi, N., Yamamoto, M., Fukao, K. *et al.* (2002) Essential and instructive roles of GATA factors in eosinophil development. *J Exp Med*, **195**, 1379-1386.
211. Chiba, T., Nagata, Y., Kishi, A., Sakamaki, K., Miyajima, A., Yamamoto, M., Engel, J.D. and Todokoro, K. (1993) Induction of erythroid-specific gene expression in lymphoid cells. *Proc Natl Acad Sci U S A*, **90**, 11593-11597.
212. Lacombe, C. and Mayeux, P. (1999) The molecular biology of erythropoietin. *Nephrol Dial Transplant*, **14 Suppl 2**, 22-28.
213. Gregory, T., Yu, C., Ma, A., Orkin, S.H., Blobel, G.A. and Weiss, M.J. (1999) GATA-1 and erythropoietin cooperate to promote erythroid cell survival by regulating bcl-xL expression. *Blood*, **94**, 87-96.
214. Kotkow, K.J. and Orkin, S.H. (1995) Dependence of globin gene expression in mouse erythroleukemia cells on the NF-E2 heterodimer. *Mol Cell Biol*, **15**, 4640-4647.

215. Andrews, N.C., Erdjument-Bromage, H., Davidson, M.B., Tempst, P. and Orkin, S.H. (1993) Erythroid transcription factor NF-E2 is a haematopoietic-specific basic-leucine zipper protein. *Nature*, **362**, 722-728.
216. Andrews, N.C., Kotkow, K.J., Ney, P.A., Erdjument-Bromage, H., Tempst, P. and Orkin, S.H. (1993) 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.
217. Blank, V. and Andrews, N.C. (1997) The Maf transcription factors: regulators of differentiation. *Trends Biochem Sci*, **22**, 437-441.
218. Stamatoyannopoulos, J.A., Goodwin, A., Joyce, T. and Lowrey, C.H. (1995) NF-E2 and GATA binding motifs are required for the formation of DNase I hypersensitive site 4 of the human beta-globin locus control region. *Embo J*, **14**, 106-116.
219. Pomerantz, O., Goodwin, A.J., Joyce, T. and Lowrey, C.H. (1998) Conserved elements containing NF-E2 and tandem GATA binding sites are required for erythroid-specific chromatin structure reorganization within the human beta-globin locus control region. *Nucleic Acids Res*, **26**, 5684-5691.
220. Forsberg, E.C., Downs, K.M. and Bresnick, E.H. (2000) Direct interaction of NF-E2 with hypersensitive site 2 of the beta-globin locus control region in living cells. *Blood*, **96**, 334-339.
221. Caterina, J.J., Ryan, T.M., Pawlik, K.M., Palmiter, R.D., Brinster, R.L., Behringer, R.R. and Townes, T.M. (1991) Human beta-globin locus control region: analysis of the 5' DNase I hypersensitive site HS 2 in transgenic mice. *Proc Natl Acad Sci U S A*, **88**, 1626-1630.
222. 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.
223. Ley, T.J., Hug, B., Fiering, S., Epner, E., Bender, M.A. and Groudine, M. (1998) Reduced beta-globin gene expression in adult mice containing deletions of locus control region 5' HS-2 or 5' HS-3. *Ann N Y Acad Sci*, **850**, 45-53.
224. 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.
225. Cheng, X., Reginato, M.J., Andrews, N.C. and Lazar, M.A. (1997) The transcriptional integrator CREB-binding protein mediates positive cross talk between nuclear hormone receptors and the hematopoietic bZip protein p45/NF-E2. *Mol Cell Biol*, **17**, 1407-1416.
226. Amrolia, P.J., Ramamurthy, L., Saluja, D., Tanese, N., Jane, S.M. and Cunningham, J.M. (1997) The activation domain of the enhancer binding protein p45NF-E2 interacts with

- TAFII130 and mediates long-range activation of the alpha- and beta-globin gene loci in an erythroid cell line. *Proc Natl Acad Sci U S A*, **94**, 10051-10056.
227. 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.
 228. Kooren, J., Palstra, R.J., Klous, P., Splinter, E., von Lindern, M., Grosveld, F. and de Laat, W. (2007) Beta-globin active chromatin Hub formation in differentiating erythroid cells and in p45 NF-E2 knock-out mice. *J Biol Chem*, **282**, 16544-16552.
 229. 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.
 230. Kotkow, K.J. and Orkin, S.H. (1996) Complexity of the erythroid transcription factor NF-E2 as revealed by gene targeting of the mouse p18 NF-E2 locus. *Proc Natl Acad Sci U S A*, **93**, 3514-3518.
 231. 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.
 232. 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.
 233. 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.
 234. Corre, S. and Galibert, M.D. (2005) Upstream stimulating factors: highly versatile stress-responsive transcription factors. *Pigment Cell Res*, **18**, 337-348.
 235. Sha, M., Ferre-D'Amare, A.R., Burley, S.K. and Goss, D.J. (1995) Anti-cooperative biphasic equilibrium binding of transcription factor upstream stimulatory factor to its cognate DNA monitored by protein fluorescence changes. *J Biol Chem*, **270**, 19325-19329.
 236. Groenen, P.M., Garcia, E., Debeer, P., Devriendt, K., Fryns, J.P. and Van de Ven, W.J. (1996) Structure, sequence, and chromosome 19 localization of human USF2 and its rearrangement in a patient with multicystic renal dysplasia. *Genomics*, **38**, 141-148.
 237. Luo, X. and Sawadogo, M. (1996) Functional domains of the transcription factor USF2: atypical nuclear localization signals and context-dependent transcriptional activation domains. *Mol Cell Biol*, **16**, 1367-1375.

238. Qyang, Y., Luo, X., Lu, T., Ismail, P.M., Krylov, D., Vinson, C. and Sawadogo, M. (1999) Cell-type-dependent activity of the ubiquitous transcription factor USF in cellular proliferation and transcriptional activation. *Mol Cell Biol*, **19**, 1508-1517.
239. Ferre-D'Amare, A.R., Pognonec, P., Roeder, R.G. and Burley, S.K. (1994) Structure and function of the b/HLH/Z domain of USF. *Embo J*, **13**, 180-189.
240. Andrews, G.K., Lee, D.K., Ravindra, R., Lichtlen, P., Sirito, M., Sawadogo, M. and Schaffner, W. (2001) The transcription factors MTF-1 and USF1 cooperate to regulate mouse metallothionein-I expression in response to the essential metal zinc in visceral endoderm cells during early development. *Embo J*, **20**, 1114-1122.
241. Ge, Y., Jensen, T.L., Matherly, L.H. and Taub, J.W. (2003) Physical and functional interactions between USF and Sp1 proteins regulate human deoxycytidine kinase promoter activity. *J Biol Chem*, **278**, 49901-49910.
242. Liu, M., Whetstone, J.R., Payton, S.G., Ge, Y., Flatley, R.M. and Matherly, L.H. (2004) Roles of USF, Ikaros and Sp proteins in the transcriptional regulation of the human reduced folate carrier B promoter. *Biochem J*, **383**, 249-257.
243. Huang, S., Li, X., Yusufzai, T.M., Qiu, Y. and Felsenfeld, G. (2007) USF1 recruits histone modification complexes and is critical for maintenance of a chromatin barrier. *Mol Cell Biol*, **27**, 7991-8002.
244. Han, S.Y., Kim, J.C., Suh, J.M. and Chung, I.K. (2001) Cell type-dependent regulation of human DNA topoisomerase III alpha gene expression by upstream stimulatory factor 2. *FEBS Lett*, **505**, 57-62.
245. Bungert, J., Kober, I., During, F. and Seifart, K.H. (1992) Transcription factor eUSF is an essential component of isolated transcription complexes on the duck histone H5 gene and it mediates the interaction of TFIID with a TATA-deficient promoter. *J Mol Biol*, **223**, 885-898.
246. Chiang, C.M. and Roeder, R.G. (1995) Cloning of an intrinsic human TFIID subunit that interacts with multiple transcriptional activators. *Science*, **267**, 531-536.
247. Meisterernst, M., Horikoshi, M. and Roeder, R.G. (1990) Recombinant yeast TFIID, a general transcription factor, mediates activation by the gene-specific factor USF in a chromatin assembly assay. *Proc Natl Acad Sci U S A*, **87**, 9153-9157.
248. Sawadogo, M. and Roeder, R.G. (1985) Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. *Cell*, **43**, 165-175.
249. Sawadogo, M. (1988) Multiple forms of the human gene-specific transcription factor USF. II. DNA binding properties and transcriptional activity of the purified HeLa USF. *J Biol Chem*, **263**, 11994-12001.

250. Du, H., Roy, A.L. and Roeder, R.G. (1993) Human transcription factor USF stimulates transcription through the initiator elements of the HIV-1 and the Ad-ML promoters. *Embo J*, **12**, 501-511.
251. Gregor, P.D., Sawadogo, M. and Roeder, R.G. (1990) The adenovirus major late transcription factor USF is a member of the helix-loop-helix group of regulatory proteins and binds to DNA as a dimer. *Genes Dev*, **4**, 1730-1740.
252. Sirito, M., Walker, S., Lin, Q., Kozlowski, M.T., Klein, W.H. and Sawadogo, M. (1992) Members of the USF family of helix-loop-helix proteins bind DNA as homo- as well as heterodimers. *Gene Expr*, **2**, 231-240.
253. Sirito, M., Lin, Q., Maity, T. and Sawadogo, M. (1994) Ubiquitous expression of the 43- and 44-kDa forms of transcription factor USF in mammalian cells. *Nucleic Acids Res*, **22**, 427-433.
254. Viollet, B., Lefrancois-Martinez, A.M., Henrion, A., Kahn, A., Raymondjean, M. and Martinez, A. (1996) Immunochemical characterization and transacting properties of upstream stimulatory factor isoforms. *J Biol Chem*, **271**, 1405-1415.
255. Kirschbaum, B.J., Pognonec, P. and Roeder, R.G. (1992) Definition of the transcriptional activation domain of recombinant 43-kilodalton USF. *Mol Cell Biol*, **12**, 5094-5101.
256. Sirito, M., Lin, Q., Deng, J.M., Behringer, R.R. and Sawadogo, M. (1998) Overlapping roles and asymmetrical cross-regulation of the USF proteins in mice. *Proc Natl Acad Sci U S A*, **95**, 3758-3763.
257. Casado, M., Vallet, V.S., Kahn, A. and Vaulont, S. (1999) Essential role in vivo of upstream stimulatory factors for a normal dietary response of the fatty acid synthase gene in the liver. *J Biol Chem*, **274**, 2009-2013.
258. Roy, A.L., Carruthers, C., Gutjahr, T. and Roeder, R.G. (1993) Direct role for Myc in transcription initiation mediated by interactions with TFII-I. *Nature*, **365**, 359-361.
259. Novina, C.D., Cheriyaath, V. and Roy, A.L. (1998) Regulation of TFII-I activity by phosphorylation. *J Biol Chem*, **273**, 33443-33448.
260. Yang, W. and Desiderio, S. (1997) BAP-135, a target for Bruton's tyrosine kinase in response to B cell receptor engagement. *Proc Natl Acad Sci U S A*, **94**, 604-609.
261. Cheriyaath, V., Novina, C.D. and Roy, A.L. (1998) TFII-I regulates Vbeta promoter activity through an initiator element. *Mol Cell Biol*, **18**, 4444-4454.
262. Roy, A.L., Du, H., Gregor, P.D., Novina, C.D., Martinez, E. and Roeder, R.G. (1997) Cloning of an inr- and E-box-binding protein, TFII-I, that interacts physically and functionally with USF1. *Embo J*, **16**, 7091-7104.

263. Cheriya, V. and Roy, A.L. (2000) Alternatively spliced isoforms of TFII-I. Complex formation, nuclear translocation, and differential gene regulation. *J Biol Chem*, **275**, 26300-26308.
264. Wang, Y.K., Perez-Jurado, L.A. and Francke, U. (1998) A mouse single-copy gene, Gtf2i, the homolog of human GTF2I, that is duplicated in the Williams-Beuren syndrome deletion region. *Genomics*, **48**, 163-170.
265. Perez Jurado, L.A., Wang, Y.K., Peoples, R., Coloma, A., Cruces, J. and Francke, U. (1998) A duplicated gene in the breakpoint regions of the 7q11.23 Williams-Beuren syndrome deletion encodes the initiator binding protein TFII-I and BAP-135, a phosphorylation target of BTK. *Hum Mol Genet*, **7**, 325-334.
266. Parker, R., Phan, T., Baumeister, P., Roy, B., Cheriya, V., Roy, A.L. and Lee, A.S. (2001) Identification of TFII-I as the endoplasmic reticulum stress response element binding factor ERSF: its autoregulation by stress and interaction with ATF6. *Mol Cell Biol*, **21**, 3220-3233.
267. Novina, C.D., Kumar, S., Bajpai, U., Cheriya, V., Zhang, K., Pillai, S., Wortis, H.H. and Roy, A.L. (1999) Regulation of nuclear localization and transcriptional activity of TFII-I by Bruton's tyrosine kinase. *Mol Cell Biol*, **19**, 5014-5024.
268. Montano, M.A., Kripke, K., Norina, C.D., Achacoso, P., Herzenberg, L.A., Roy, A.L. and Nolan, G.P. (1996) NF-kappa B homodimer binding within the HIV-1 initiator region and interactions with TFII-I. *Proc Natl Acad Sci U S A*, **93**, 12376-12381.
269. Crusselle-Davis, V.J., Vieira, K.F., Zhou, Z., Anantharaman, A. and Bungert, J. (2006) Antagonistic regulation of beta-globin gene expression by helix-loop-helix proteins USF and TFII-I. *Mol Cell Biol*, **26**, 6832-6843.
270. Bank, A., O'Neill, D., Lopez, R., Pulte, D., Ward, M., Mantha, S. and 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 N Y Acad Sci*, **1054**, 48-54.
271. Mahajan, M.C. and Weissman, S.M. (2002) DNA-dependent adenosine triphosphatase (helicase-like transcription factor) activates beta-globin transcription in K562 cells. *Blood*, **99**, 348-356.
272. 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.
273. Kosak, S.T. and Groudine, M. (2004) Form follows function: The genomic organization of cellular differentiation. *Genes Dev*, **18**, 1371-1384.
274. Ragozy, T., Bender, M.A., Telling, A., Byron, R. and Groudine, M. (2006) The locus control region is required for association of the murine beta-globin locus with engaged transcription factories during erythroid maturation. *Genes Dev*, **20**, 1447-1457.

275. Hung, H.L., Kim, A.Y., Hong, W., Rakowski, C. and Blobel, G.A. (2001) Stimulation of NF-E2 DNA binding by CREB-binding protein (CBP)-mediated acetylation. *J Biol Chem*, **276**, 10715-10721.
276. Chaturvedi, C.P., Hosey, A.M., Pali, C., Perez-Iratxeta, C., Nakatani, Y., Ranish, J.A., Dilworth, F.J. and Brand, M. (2009) Dual role for the methyltransferase G9a in the maintenance of beta-globin gene transcription in adult erythroid cells. *Proc Natl Acad Sci U S A*, **106**, 18303-18308.
277. Demers, C., Chaturvedi, C.P., Ranish, J.A., Juban, G., Lai, P., Morle, F., Aebersold, R., Dilworth, F.J., Groudine, M. and Brand, M. (2007) Activator-mediated recruitment of the MLL2 methyltransferase complex to the beta-globin locus. *Mol Cell*, **27**, 573-584.
278. Caraveo, G., van Rossum, D.B., Patterson, R.L., Snyder, S.H. and Desiderio, S. (2006) Action of TFII-I outside the nucleus as an inhibitor of agonist-induced calcium entry. *Science*, **314**, 122-125.
279. Kitajima, K., Tanaka, M., Zheng, J., Sakai-Ogawa, E. and Nakano, T. (2003) In vitro differentiation of mouse embryonic stem cells to hematopoietic cells on an OP9 stromal cell monolayer. *Methods Enzymol*, **365**, 72-83.
280. Levings, P.P., Zhou, Z., Vieira, K.F., Crusselle-Davis, V.J. and Bungert, J. (2006) Recruitment of transcription complexes to the beta-globin locus control region and transcription of hypersensitive site 3 prior to erythroid differentiation of murine embryonic stem cells. *Febs J*, **273**, 746-755.
281. Chomeczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem*, **162**, 156-159.
282. Crusselle-Davis, V.J., Zhou, Z., Anantharaman, A., Moghimi, B., Dodev, T., Huang, S. and Bungert, J. (2007) Recruitment of coregulator complexes to the beta-globin gene locus by TFII-I and upstream stimulatory factor. *Febs J*, **274**, 6065-6073.
283. Usheva, A., Maldonado, E., Goldring, A., Lu, H., Houbavi, C., Reinberg, D. and Aloni, Y. (1992) Specific interaction between the nonphosphorylated form of RNA polymerase II and the TATA-binding protein. *Cell*, **69**, 871-881.
284. 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.
285. 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.
286. Kouzarides, T. (2007) Chromatin modifications and their function. *Cell*, **128**, 693-705.
287. Strahl, B.D. and Allis, C.D. (2000) The language of covalent histone modifications. *Nature*, **403**, 41-45.

288. Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S. and Jones, J.M. (1998) Embryonic stem cell lines derived from human blastocysts. *Science*, **282**, 1145-1147.
289. Stojkovic, M., Lako, M., Stojkovic, P., Stewart, R., Przyborski, S., Armstrong, L., Evans, J., Herbert, M., Hyslop, L., Ahmad, S. *et al.* (2004) Derivation of human embryonic stem cells from day-8 blastocysts recovered after three-step in vitro culture. *Stem Cells*, **22**, 790-797.
290. Strelchenko, N., Verlinsky, O., Kukhareno, V. and Verlinsky, Y. (2004) Morula-derived human embryonic stem cells. *Reprod Biomed Online*, **9**, 623-629.
291. Evans, M.J. and Kaufman, M.H. (1981) Establishment in culture of pluripotential cells from mouse embryos. *Nature*, **292**, 154-156.
292. Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G. *et al.* (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell*, **122**, 947-956.
293. Meshorer, E. and Misteli, T. (2006) Chromatin in pluripotent embryonic stem cells and differentiation. *Nat Rev Mol Cell Biol*, **7**, 540-546.
294. Arney, K.L. and Fisher, A.G. (2004) Epigenetic aspects of differentiation. *J Cell Sci*, **117**, 4355-4363.
295. Ringrose, L. and Paro, R. (2004) Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu Rev Genet*, **38**, 413-443.
296. Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P. and Reinberg, D. (2002) Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev*, **16**, 2893-2905.
297. Muller, J., Hart, C.M., Francis, N.J., Vargas, M.L., Sengupta, A., Wild, B., Miller, E.L., O'Connor, M.B., Kingston, R.E. and Simon, J.A. (2002) Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell*, **111**, 197-208.
298. Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S. and Zhang, Y. (2002) Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science*, **298**, 1039-1043.
299. Kirmizis, A., Bartley, S.M., Kuzmichev, A., Margueron, R., Reinberg, D., Green, R. and Farnham, P.J. (2004) Silencing of human polycomb target genes is associated with methylation of histone H3 Lys 27. *Genes Dev*, **18**, 1592-1605.
300. Lee, T.I., Jenner, R.G., Boyer, L.A., Guenther, M.G., Levine, S.S., Kumar, R.M., Chevalier, B., Johnstone, S.E., Cole, M.F., Isono, K. *et al.* (2006) Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell*, **125**, 301-313.

301. Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A. and Pirrotta, V. (2002) Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell*, **111**, 185-196.
302. Zhou, W., Zhu, P., Wang, J., Pascual, G., Ohgi, K.A., Lozach, J., Glass, C.K. and Rosenfeld, M.G. (2008) Histone H2A monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation. *Mol Cell*, **29**, 69-80.
303. Bantignies, F. and Cavalli, G. (2006) Cellular memory and dynamic regulation of polycomb group proteins. *Curr Opin Cell Biol*, **18**, 275-283.
304. Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K. *et al.* (2006) Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature*, **441**, 349-353.
305. Kodama, H., Nose, M., Niida, S., Nishikawa, S. and Nishikawa, S. (1994) Involvement of the c-kit receptor in the adhesion of hematopoietic stem cells to stromal cells. *Exp Hematol*, **22**, 979-984.
306. Heinrich, P.C., Behrmann, I., Haan, S., Hermanns, H.M., Muller-Newen, G. and Schaper, F. (2003) Principles of interleukin (IL)-6-type cytokine signalling and its regulation. *Biochem J*, **374**, 1-20.
307. Armstrong, L., Lako, M., Dean, W. and Stojkovic, M. (2006) Epigenetic modification is central to genome reprogramming in somatic cell nuclear transfer. *Stem Cells*, **24**, 805-814.
308. Morgan, H.D., Santos, F., Green, K., Dean, W. and Reik, W. (2005) Epigenetic reprogramming in mammals. *Hum Mol Genet*, **14 Spec No 1**, R47-58.
309. Ragozy, 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.
310. 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.
311. Ryan, T.M., Behringer, R.R., Martin, N.C., Townes, T.M., Palmiter, R.D. and Brinster, R.L. (1989) A single erythroid-specific DNase I super-hypersensitive site activates high levels of human beta-globin gene expression in transgenic mice. *Genes Dev*, **3**, 314-323.
312. Bresnick, E.H. and Tze, L. (1997) Synergism between hypersensitive sites confers long-range gene activation by the beta-globin locus control region. *Proc Natl Acad Sci U S A*, **94**, 4566-4571.
313. Jackson, J.D., Petrykowska, H., Philipsen, S., Miller, W. and Hardison, R. (1996) Role of DNA sequences outside the cores of DNase hypersensitive sites (HSs) in functions of the

- beta-globin locus control region. Domain opening and synergism between HS2 and HS3. *J Biol Chem*, **271**, 11871-11878.
314. Efstratiadis, A., Posakony, J.W., Maniatis, T., Lawn, R.M., O'Connell, C., Spritz, R.A., DeRiel, J.K., Forget, B.G., Weissman, S.M., Slightom, J.L. *et al.* (1980) The structure and evolution of the human beta-globin gene family. *Cell*, **21**, 653-668.
315. Blank, V., Kim, M.J. and Andrews, N.C. (1997) Human MafG is a functional partner for p45 NF-E2 in activating globin gene expression. *Blood*, **89**, 3925-3935.
316. 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.
317. Liang, S.Y., Moghimi, B., Crusselle-Davis, V.J., Lin, I.J., Rosenberg, M.H., Li, X., Strouboulis, J., Huang, S. and Bungert, J. (2009) Defective erythropoiesis in transgenic mice expressing dominant-negative upstream stimulatory factor. *Mol Cell Biol*, **29**, 5900-5910.
318. 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.
319. Sims, R.J., 3rd, Belotserkovskaya, R. and Reinberg, D. (2004) Elongation by RNA polymerase II: the short and long of it. *Genes Dev*, **18**, 2437-2468.
320. Lin, I.J., Zhou, Z., Crusselle-Davis, V.J., Moghimi, B., Gandhi, K., Anantharaman, A., Pantic, D., Huang, S., Jayandharan, G., Zhong, L. *et al.* (2009) Calpeptin increases the activity of upstream stimulatory factor and induces high level globin gene expression in erythroid cells. *J Biol Chem*, **284**, 20130-20135.
321. Miles, J., Mitchell, J.A., Chakalova, L., Goyenechea, B., Osborne, C.S., O'Neill, L., Tanimoto, K., Engel, J.D. and Fraser, P. (2007) Intergenic transcription, cell-cycle and the developmentally regulated epigenetic profile of the human beta-globin locus. *PLoS One*, **2**, e630.
322. 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.
323. Bottardi, S., Ross, J., Pierre-Charles, N., Blank, V. and Milot, E. (2006) Lineage-specific activators affect beta-globin locus chromatin in multipotent hematopoietic progenitors. *Embo J*, **25**, 3586-3595.
324. Levings, P.P. and Bungert, J. (2002) The human beta-globin locus control region. *Eur J Biochem*, **269**, 1589-1599.

325. Nagai, T., Igarashi, K., Akasaka, J., Furuyama, K., Fujita, H., Hayashi, N., Yamamoto, M. and Sassa, S. (1998) Regulation of NF-E2 activity in erythroleukemia cell differentiation. *J Biol Chem*, **273**, 5358-5365.
326. Song, S.H., Hou, C. and Dean, A. (2007) A positive role for NLI/Ldb1 in long-range beta-globin locus control region function. *Mol Cell*, **28**, 810-822.
327. Kim, S.I. and Bresnick, E.H. (2007) Transcriptional control of erythropoiesis: emerging mechanisms and principles. *Oncogene*, **26**, 6777-6794.
328. Bieker, J.J. (2001) Kruppel-like factors: three fingers in many pies. *J Biol Chem*, **276**, 34355-34358.
329. Burdon, T., Smith, A. and Savatier, P. (2002) Signalling, cell cycle and pluripotency in embryonic stem cells. *Trends Cell Biol*, **12**, 432-438.
330. Koledova, Z., Raskova Kafkova, L., Calabkova, L., Krystof, V., Dolezel, P. and Divoky, V. (2009) Cdk2 inhibition prolongs G1 phase progression in mouse embryonic stem cells. *Stem Cells Dev*.
331. Jenuwein, T. and Allis, C.D. (2001) Translating the histone code. *Science*, **293**, 1074-1080.
332. Bulger, M. (2005) Hyperacetylated chromatin domains: lessons from heterochromatin. *J Biol Chem*, **280**, 21689-21692.
333. Wen, Y.D., Cress, W.D., Roy, A.L. and Seto, E. (2003) Histone deacetylase 3 binds to and regulates the multifunctional transcription factor TFII-I. *J Biol Chem*, **278**, 1841-1847.
334. Cook, P.R. (1999) The organization of replication and transcription. *Science*, **284**, 1790-1795.

BIOGRAPHICAL SKETCH

Zhuo Zhou was born on April 13, 1979 in the beautiful city of Harbin in Heilongjiang Province, China. She grew up in the neighborhood of Harbin Medical University, where her parents were working, and had decided to work on something related to medical science since she was little. She graduated from Middle School Attached to Harbin Normal University in 1997 and then attended Peking University, China with a grade in the top 0.01% on the National Matriculation Examination. In 2001, Zhuo was awarded a B.S. degree in Physiology and Biophysics from the College of Life Science at Peking University. She then worked as research assistant in Key Laboratory of Cell Biology and Genetics in College of Life Sciences, Peking University until 2004. In 2004, Zhuo attended the IDP as a Ph.D. candidate in College of Medicine, University of Florida, mentored by Dr. Jörg Bungert. In May 2010, she graduated with her Ph.D. degree from Department of Biochemistry and Molecular Biology, College of Medicine, University of Florida. She is also expecting to receive her M.S. degree in Statistics in August 2010. Zhuo intends to acquire a post-doctoral fellowship investigating the role of chromatin structure on gene regulation. In the long term, she would like to pursue work in academia.