PhD

Program in Molecular and Translational Medicine

DIMET

(XXVI Cycle, year 2011-2012)

University of Milano-Bicocca

Department of Biotechnology & Biosciences

"Role of Sox6 and COUP-TFII transcription factors in the regulation of hemoglobin switching"

Sudharshan Elangovan **Matr : 745216**

PhD Coordinator: Prof. Andrea Biondi

Principal Investigator: Prof. Antonella Ronchi

The research presented in this thesis was performed at the department of Biotechnology and Biosciences, University of Milano-Bicocca in the laboratory of genetics headed by Prof. Antonella Ronchi and co-tutored by Prof. Sergio Ottolenghi.

எல்லா புகழும் இறைவனுக்கே!

(All glory goes to God)

“I do not know what I may appear to the world, but to myself I seem to have been only like a boy playing on the sea-shore, and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary, whilst the great ocean of truth lay all undiscovered before me.”

— Isaac Newton

TABLE OF CONTENTS

Chapter 1:

Chapter 2:

MODEL SYSTEM FOR *IN VIVO* AND *EX VIVO* STUDIES OF GLOBIN GENE EXPRESSION

Chapter 3:

ROLE OF SOX6 AND COUP-TFII IN THE REGULATION OF HEMOGLOBIN SWITCHING

Chapter 4:

SOX-6, COUP-TFII AND BCL11A-XL INTERPLAY IN THE REGULATION OF HEMOGLOBIN SWITCHING

Chapter 5:

FINAL DISCUSSION: MOLECULAR AND CLINICAL CONSIDERATIONS

Chapter 1

GENERAL INTRODUCTION

HEMATOPOIESIS

The blood is one of the most fascinating tissues not only by the virtue of its unique structure but also due to the variety of cells composing the tissue. Each cell type has distinct appearance and performs specific and essential biological functions.

The blood is composed of different cell types: erythrocytes, megakaryocytes, lymphocytes, granulocytes, dendritic cells, macrophages and mast cells; but among all of them, erythrocytes are the most abundant, occupying about 45% of the total blood volume. They are small and enucleated cells, with a biconcave discoid shape, completely filled with hemoglobin, the molecule responsible for their capability of transporting oxygen and carbon dioxide throughout the body. The other cell types are less abundant, occupying about 1% of the blood volume, but performing important biological functions. Platelets are rod-shaped cell fragments derived from the differentiation of large cells denominated megakaryocytes, and have an essential role in blood coagulation. Lymphocytes, divided into B- and T-lymphocytes, play a crucial role in the specific immune response against microorganisms acting as direct killers either activating other cells that eliminate infected cells (T cells), or producing soluble antibodies that promote the elimination of the infected microorganism (B cells). Granulocytes are divided in neutrophils, eosinophils and basophils, and they represent the innate immune system having a role in the inflammatory response and phagocytosis. Macrophages and mast cells are not found in the circulation but within tissues. Macrophages arise from monocytes

which are circulating in the blood before they migrate to the tissues were they terminally differentiate. Mast cells originate from an unidentified progenitor in the bone marrow that migrates through the blood into mucosal tissues where they differentiate into mature mast cells.

All these specialized cells originate from a group of cells named hematopoietic stem cell (HSC), relatively quiescent cell, which rarely divide to generate both stem cell and progenitors. HSCs are highly proliferative cell that is able to initiate the whole hematopoiesis process (Fig.1). Because mature blood cells are predominantly short lived, stem cells are required throughout life to replenish multilineage progenitors and the precursors committed to individual hematopoietic lineages.

In mouse, HSCs are extremely rare cells and reside in the bone marrow (Abkowitz *et al.,* 2000; Jacobsen SE *et al.,* 2014). HSCs have the ability to renew themselves and to differentiate into all different blood cell populations. The progenitor's differentiation process occurs through a series of commitment steps, each leading to further restriction to a certain hematopoietic lineage. Which are these steps, and the existence of which lineages restricted progenitors, is a controversial issue, and different models have been proposed (Pronk *et al.,* 2007; Akashi *et al.,* 2000; Graf T *et al.,* 2009): the most widely accepted model claims that the first commitment step separates lymphoid from myeloid potential and is supported by the identification of a common lymphoid progenitor (CLP) and a common myeloid precursor (CMP).

CLPs are restricted to the lymphoid lineage and can give rise exclusively to B-, T-cells and natural killer cells, while CMPs can give rise to granulocytes, erythrocytes, megakaryocytes and macrophages (Pronk *et al.*, 2007; Akashi *et al.*, 2000; Johnson *et al.*, 1977). These CMPs will undergo further lineage restriction when the granulocytes/monocyte potential (pre-GM) is separated from erythroid/magakaryocyte potential (pre-MegE). The existence of CLP is now controversial because of the identification of myeloid/T cell (MTP) and myeloid/B-cell (MBP) bipotent precursors. The existence of such cells suggests an alternative model in which myeloid potential is maintained in the early commitment stages of all hematopoietic lineages. According to this model MPLs can give rise to common myeloid lymphoid precursors (CMPLs). CMPLs further commit into either MBPs or MTPs. Later on, the myeloid potential is separated from the megakaryocytic/erythroid and the B- and T-cell potential (Kawamoto H. and Katsura Y., 2009).

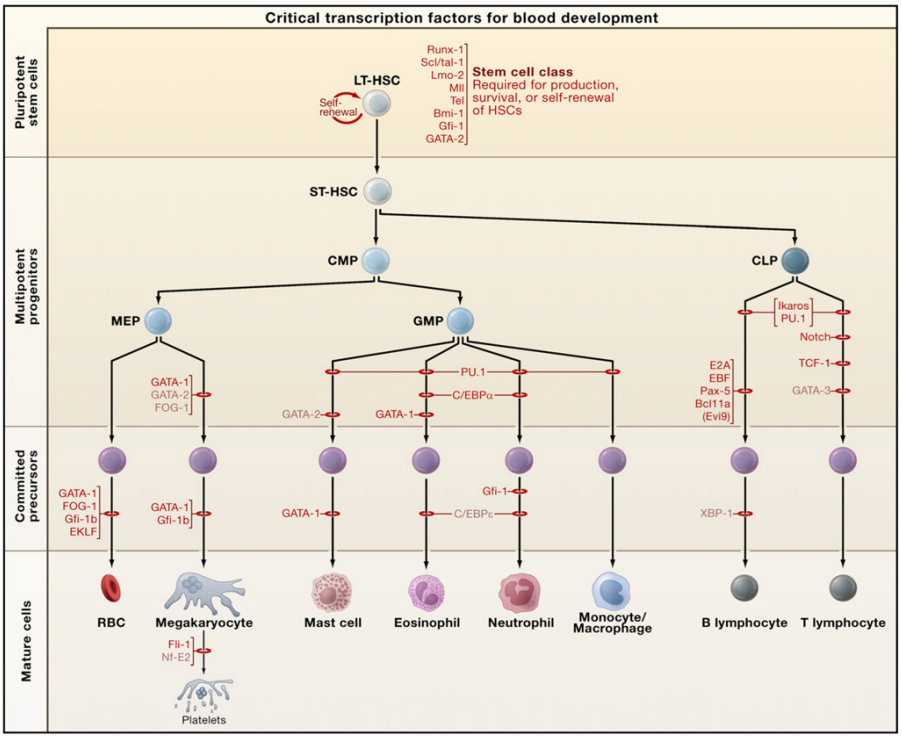


Figure 1.1 Schematic representations of Hematopoiesis with stages of hematopoietic differentiation and the principal transcription factors necessary for the commitment through each lineage. Figure adapted from Orkin SH and Zon LI, Cell. Feb 22, 2008.

During the last few years, however, the idea of this rigid hierarchy between the HSC and the different hematopoietic cells has been questioned. In fact, an increasing numbers of reports in particular "*in vitro*" conditions, suggest the occurrence, of "transdifferentiation" process between different hematopoietic precursors (Graf *et al.*, 2002) suggesting a higher level of plasticity.

Why a particular progenitor cell chooses its commitment is still a controversial issue, and is a matter of speculation. Two contrasting models have been put forward: the instructive model and the

permissive model. The instructive model states that specific exogenous signals induce the various serial commitment steps. These signals are ligands and cytokines that interact with specific receptors on cell surface and thus activate intracellular pathways leading to the expression of genes specific of a given lineage. This model reinforces the importance of the stroma and the niche in the cell differentiation. On the other hand, the permissive or selective model proposes that a cell-autonomous process, such as a stochastic change in the expression of some critical genes, drives commitment of blood cells to a distinct cell lineage. The role of the stroma, in this case, would be only to provide a selective environment for the growth and the survival of the committed progenitor cells (Enver T., Jacobsen S., 2009).

Hematopoiesis during development

In man and also in mouse, the site of hematopoiesis changes during embryonic development. In the early stage of mammalian development it takes place within the yolk sac. The first wave of hematopoiesis consists mainly in the production of large, nucleated, primitive erythrocytes that synthesize embryonic globins. Primitive hematopoiesis starts between embryonic day 7.0 (E7.0) and E7.5 in mice, or day 15-18 in humans, within the blood islands of the yolk sac. Blood islands are cellular clusters composed by a central macrophage surrounded by masses of cells that gradually differentiate into primitive erythroblasts. These primitive erythroblasts enter the vascular system of the embryo proper, where they continue to divide

for several days and finally enucleate into the blood stream (Palis J., 2008).

Definitive hematopoiesis is responsible of the production of all hematopoietic lineages present in the adult organism including the definitive erythrocytes, characterized by the absence of nucleus and by the expression of adult globins. The embryonic origin of definitive hematopoiesis and of the HSC is a controversial issue and different models have been proposed so far. It is now widely accepted that the yolk sac synthesizes a second transient wave of erythroid progenitors (that will give rise to "definitive erythrocytes"), that enter the bloodstream and seed the fetal liver. At the same time, hematopoietic stem cells emerging from the AGM (Aorta-Gonad-Mesonephros) region within the embryo, seed the liver and are the presumed source of long-term erythroid potential (Mikkola HKA.and Orkin SH., 2006).

In humans, the fetal liver becomes the site of definitive hematopoiesis around day 42, and in mouse around E10.5. Fetal definitive erythroid precursors mature in macrophage islands within the liver, enucleate, and enter the bloodstream as erythrocytes (McGrath K., Palis J., 2008). In mice, hematopoiesis also occurs in the spleen. Finally, around birth (22^{nd} week of human gestation), the HSCs migrate from the fetal liver to the bone marrow, which becomes the principal hematopoietic tissue throughout the whole adult life (Palis and Segel 1998). The process of hematopoiesis is generally conserved throughout vertebrate evolution.

ERYTHROPOIESIS

Human erythropoiesis is a complex multistep developmental process that begins at pluripotent stem cells (HSCs) and terminates in the production of erythrocytes or red blood cells (RBCs). Erythrocytes are the most abundant cells that are small, enucleated, biconcave discoid shaped cells and are extremely specialized to absolve their function of gaseous transport mediated by Hemoglobin. The average lifespan of erythrocytes is 60 days in mice and 120 in humans (Shemin D *et al.*, 1946; Robert SF., 2012). To guarantee a constant supply of oxygen to the body tissues, a continued generation of mature erythrocytes is necessary. The red blood cells volume is about the 45% of blood volume, and every day in our body 2 x 10^{11} red blood cells are produced, that means 2 million erythrocytes per second. Erythropoiesis involves several steps of programmed differentiation between the HSC and the mature erythrocyte (Fig. 1.2).

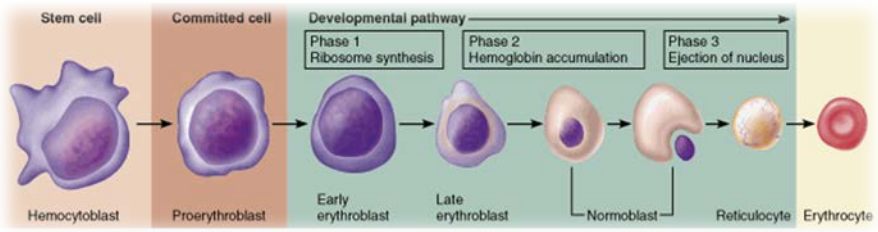


Figure 1.2 Overview of erythroid differentiation that begins with the proerythroblast, the first morphologically identifiable erythroid progenitor and terminates upon the hemoglobin protein accumulation and nucleus extrusion. Figure adapted from Benjamin Cummings 2001

In humans, the HSC differentiates into a Common Myeloid Progenitor (CMP). CMPs can give rise to MEPs, which are fully restricted to the megakaryocytic and erythroid lineages (Novershtern N *et al.*, 2011). In mouse, the HSCs give rise to Multipotent Progenitor Pool (MPP) that are heterogeneous and can be divided into colonogenic long-term self-renewing HSC (LT-HSC), transiently self-renewing HSC (short-term HSC), and non-self-renewing multipotent progenitors (MPP). Pre-MegE are intermediate progenitors that are derived from MPP and gives rise to Pre-CFU-E and BFU-E (Pronk CJH *et al.*, 2007). The Burst Forming Unit-Erythroid (BFU-E) is the most primitive erythroid-restricted progenitor, and we are able to identify it by functional assays (Wong *et al.*. 1986): in the presence of erythropoietin (Epo), interleukin 3 (IL 3), granulocyte-macrophage colony stimulating factor (GM-CSF), thrombopoietin (TPO) and stem cell factor (SCF). BFU-E can give rise to large colonies of some thousands of hemoglobinized erythroblasts after 5-7 (mouse cells) or 14-16 (human cells) days of culture.

The BFU-E further differentiates into the Colony Forming Unit-Erythroid (CFU-E), a more mature erythroid progenitor closely related to the proerythroblast, that, if plated in a semi-solid medium gives rise to small colonies, containing few hundreds hemoglobinized erythroblasts after 2-4 (mouse cells) or 5-8 (human cells) days of culture (Whyatt *et al.*, 2000; Wong *et al.*, 1986). The proerythroblast is a large cell (14-19 μm and 11-13 μ in diameter in human and mice respectively) with a large nucleus, visible nucleoli, surrounded by a basophilic cytoplasm. Basophilic erythroblasts are slightly smaller cells (12-17 μm and 9-10 μm in diameter, in human and mouse

respectively). A strong hemoglobinization occurs at the orthochromatic erythroblast stage. orthochromatic erythroblast cells are the smallest nucleated erythrocyte precursors (8-12 μm and 7-8 μm in diameter in human and mouse respectively) and their nuclei undergo pycnotic degeneration: the chromatin becomes very condensed and the nucleus becomes smaller. The nucleus is then extruded from the cell that becomes a reticulocyte.

Reticulocytes are a little bit larger than fully mature erythrocytes (7-8 μm in diameter in man and 3.5- 4.5 μm in diameter in mouse) have irregular shapes and still contain few cytoplasmic organelles, necessary to complete the globin protein synthesis (Allen *et al.*, 1982; Bondurantand Koury, 1998). It takes 48 to 72 hours from the proerythroblast to reticulocyte stage, and this process occurs in the erythroblastic island, a particular anatomic compartment able to provide a unique environment. Blood islands are a structure, consisting of a centrale macrophage surrounded by erythroid precursors at different stages of maturation. The more immature precursors are located to the centre, and as they mature, they move away from the body of the macrophage. Erythroid precursors keep in contact with the cytoplasmic extensions of the macrophage during all stages of maturation until occurrence of enucleation (Bessis *et al.*, 1983), and the macrophage has the final role to phagocytise the extruded nucleus after the terminal maturation of RBCs (Allen and Dexter, 1982; Allen and Testa, 1991).

THE HEMOGLOBIN SWITCHING

The hemoglobin protein:

The molecule of hemoglobin, the oxygen-carbon dioxide transporter of erythrocytes, is the best understood allosteric protein. It is a nearly spherical structure with a diameter of 55Aº consisting of four polypeptide chains, two α-like chains and two β-like chains, packed together in a tetrahedral array (fig 1.3b). Each polypeptide contains a heme group with a single oxygen binding site. The heme groups are located in cervices near the exterior of molecule providing the oxygen binding sites far apart from each other (Perutz *et al*.. 1978, 1984).

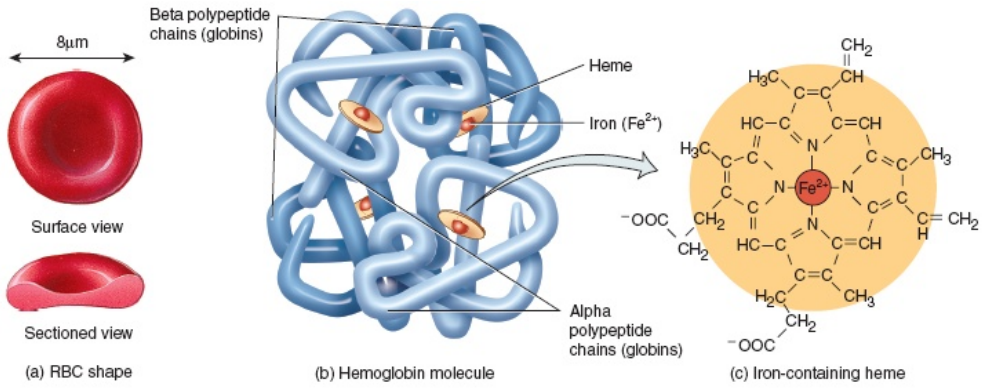


Figure 1.3 A). Shape and structure of a red blood cell (RBC). B). structure of hemoglobin molecule containing four polypeptide chains (blue) has one heme group (gold), which contains an iron ion (Fe^{2+}), shown in red. C). iron-containing heme composition. Figure adapted from the book John Wiley & Sons, 2011.

The capacity of hemoglobin to bind oxygen depends on the presence of the non-polypeptide unit of the heme group. It consists of an organic part (protoporphyrin) and an iron atom (fig 1.3c). Protoporphyrin is made up of four pyrrole rings linked together to

form tetrapyrrole ring. The iron atom binds to four nitrogen's in the center of the ring and can additionally form two extra bonds on either side of the heme plane. When bound to oxygen, the iron atom of the corresponding form of hemoglobin, called "ferrihemoglobin", has the ferrous (+3) oxidative state (Stryer L *et al.,* 1995). Thus, oxygen binding at one heme group facilitates the binding of oxygen at the other heme groups on the same tetramer, and vice versa: the unloading of oxygen at one heme group facilitates the unloading of oxygen at the others. The co-operative binding of oxygen by hemoglobin enables it to deliver 1.83 times as much oxygen under typical physiological conditions as it would be independent binding sites (Stryer *et al.*, 1995).

The globins genes:

The human α-globin cluster is located in chromosome 16 and is composed of ζ, α1, and α2–globin genes that are expressed sequentially and in a developmental specific way. In the mouse, the conserved α-globin locus resides on chromosome 11.

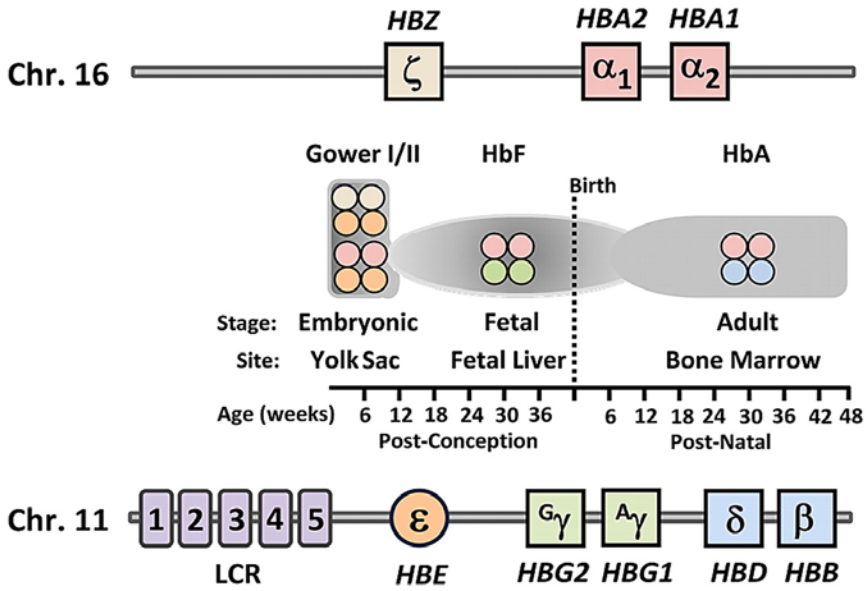


Figure 1.4: Genomic structural organization of the human α-globin and β-globin loci and temporal expression of the various hemoglobin types (Wilber et.al Blood 2011)

The β-globin locus genes in humans lie on chromosome 11 and are composed of 5'- ε-Gγ-Aγ-δ-β- 3' genes placed in the same order as their developmental expression (Fig 1.4). In the mouse, the β-globin locus genes are four and also positioned in the same order as they are expressed during mouse ontogeny; 5'-εy-βh1-βmaj-βmin-3'. Upstream to the β-globin locus lies a cluster of erythroid specific cis-regulatory elements named on the basis of their Dnase-I hypersensitivity, known as locus control region (LCR), ensures high-level expression of the linked globin genes (Grosveld F *et al.*. 1987).

Two globins genes switches occur during human development: the embryonic to fetal hemoglobin switch, which coincides with the transition from embryonic (yolk sac) to definitive (fetal liver)

hematopoiesis and the fetal to adult switch, which occurs at the perinatal period (Fig1.5, left panel). The switches from ε to γ, from γ to β and from ζ to α gene expression are controlled predominantly at the transcriptional level. Unlike humans, most species (for example the mouse) have only one switch, from embryonic to definitive globin expression (Fig 1.5 right panel), occurring early in development. The expression of an exclusive globin gene during the fetal period is a rather recent event, which took place 35 to 55 million years ago during primate evolution (Stamatoyannopoulos G., 2005). In mouse, the β -like globin genes cluster (ε_Y, βh1, β major, β minor) is located on chromosome 7, and the α-like globin genes (x, α_1, α_2) are on chromosome 11. The embryonic ε_Y and βh1 genes are expressed during embryonic life until they are substituted by adult β major and β minor genes between E11.5 – E13.5, when the mouse globins switching occurs (Fantoni *et al.*, 1969) (Fig.1.5). Although the mouse does not possess specific fetal globin genes, mice transgenic for the human globin locus linked to a LCR element express it "appropriately" during development: they start to transcribe fetal γ-globin gene around E10.5 and silence it around E16 (Strouboulis *et al.*, 1992). This suggests that human and mouse hemoglobin switching undergo similar molecular mechanisms and for this reason the mouse is a good model for studying the human globin switching.

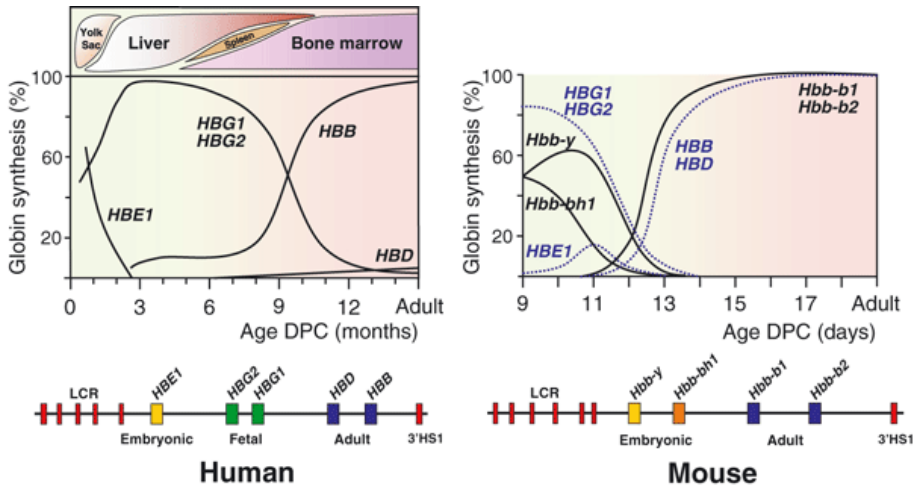


Figure 1.5: Developmental switching of the β-like globins genes expression in human (left) and mouse (right). The different developmental sites of erythropoiesis are shown above the graph. In the graph of mouse locus, the content of both endogenous mouse (black straight lines) and exogenous human β-like globins (blue dashed lines) in transgenic human β-globin locus mice are shown. Figure adapted from Sankaran VG *et al.*. 2010, Noordermeer & de Laat, 2008.

TRANSCRIPTIONAL CONTROL OF THE GLOBINS SWITCHING

The Locus Control Region (LCR):

The complex expression pattern of globins genes is controlled by regulatory cis-acting regions on the DNA, which reside thousands of nucleotides far from the globin genes in particular Locus Control Region (LCR). In fact, its absence leads to a complete loss of the

expression of the genes in the cluster in some β-thalassemic patients and to an inactive chromatin configuration of the locus (Forrester *et al.*, 1990). The LCR is located in man 22 kb upstream to the ε-globin gene, and is characterized by five DNAse I hypersensitive sites (HS), regions containing each a core of 200-300 nucleotides accessible to transcription factors and chromatin remodelling factors (Grosveld *et al.*, 1987). There are three other important distal HSs: one is located 20 kb downstream the human β -globin gene, and two are located 110 kb upstream the β-globin cluster (Bulger *et al.*, 2000).

Many models have been proposed to explain how the LCR acts in regulating globin genes expression. One model is that LCR forms loops interacting with distal HSs and creates an "active chromatin hub" (ACH) which, together with promoters of the globin genes, is able to assemble a "transcription factory" active during development (Fig. 1.6) (de Laat and Grosveld, 2003; Palstra *et al.*, 2003). Using the "chromosome conformation capture" (3C) technique it has been demonstrated that the transition from a "chromatin hub", to an ACH requires stage specific transcription factors: for example EKLF is necessary for the formation on an active chromatin hub on the β -globin gene promoter (Reviewed in Noordermeer D. Wouter de Laat W., 2008).

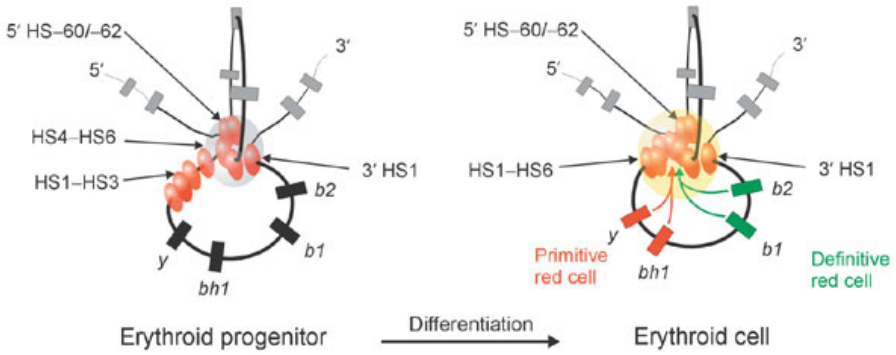


Fig 1.6: Two-dimensional representation of the mouse chromatin hub and its activation to ACH through erythroid differentiation. The core of the ACH is erythroid-specific and developmentally stable; when a developmental switch occurs, the different globin genes can enter the ACH and can thus be activated, as indicated by the arrows. Figure adapted from (Palstra *et al.* 2008).

In the murine primitive erythroid cells the ACH is mainly formed with embryonic globin genes promoters, whether in definitive erythroid cells it is formed with adult globin genes promoters. The globin switching is thus due to a chromatin remodelling that leads to the transcriptional activation in a mutually exclusive manner of globins genes that need to be expressed at different specific stages during the development (Palstra *et al.*, 2003).

TRANSCRIPTION FACTORS

In the last two decades, several transcription factors were demonstrated to bind different cis- regulatory region within the LCR

locus, and to every single globin gene promoter. Canonical transcriptional mechanisms involve the binding of sequence-specific *trans*-acting factors (transcription factors, TFs) to particular DNA sequences, the *cis*-elements, followed by recruitment of co-regulators via direct protein–protein interactions. Co-regulators typically form large multiprotein complexes and mediate either activation or repression, in a context-dependent manner. Among the variety of transcriptional factors and co-regulators of erythroid differentiation, there are some of particular importance.

- **GATA1** is a member of a family of six transcription factors, GATA-1 to GATA6, able to bind to the DNA consensus sequence (A/T)GATA(A/G) by two zinc-finger motifs. A common sequence motif [(A/T)GATA(A/G)] (GATA motif) exists within transcriptional regulatory regions of most, if not all, erythroid cell-specific genes. Within the hematopoietic system, GATA-1 is expressed in erythroid, megakaryocytic, eosinophil, and mast cell lineages. Targeted disruption of *Gata1* in mice provided evidence for its essential function in stimulating erythropoiesis (Pevny *et al.*, 1991, 1995; Weiss *et al.*, 1994). GATA-1 null mouse embryos die due to severe anemia between E10.5 and 11.5; GATA-1 knock-down embryos (GATA1.05) which express only about 5% of GATA-1 levels relatively to normal, show an arrest of the primitive erythropoiesis and die between E11.5 and E12.5 (Kim SI. Bresnick EH., 2007). At the cellular levels, high levels of GATA-1 arrest cellular proliferation Rylski *et al.*, 2003; Munugalavadla *et al.*, 2005); GATA-1 null erythroid cells fail to mature beyond the proerythroblast stage.

GATA-1-mediated survival of erythroid precursors is obtained by the expression of the Bcl-XL antiapoptotic protein (Weiss and Orkin, 1995; Gregory *et al.*, 1999) and by the activation of the transcription of Epo receptor (Chiba *et al.*, 1993), whose signalling is known to be an important progenitors survival pathway (Lacombe *et al.*, 1999).

- **KLF1**, Erythroid Kruppel-like transcription factor 1 binds to the CACCC motif on DNA in analogous to other KLFs that have diverse roles during cellular differentiation and development (Bieker JJ, 2001). EKLF occupies HS1–HS3 of the β-globin LCR and the β*major* promoter (Im *et al.*, 2005). The targeted disruption of EKLF in mice demonstrates that EKLF is crucial for erythropoiesis and for adult -like globin gene transcription (Nuez *et al.*, 1995; Perkins *et al.*, 1995). In fact, EKLF was initially considered to regulate exclusively the definitive erythropoiesis (Nuez *et al.*, 1995), but a recent work has demonstrated that EKLF also is able to promote primitive erythropoiesis (Hodge *et al.*, 2006). Transcriptional profiling studies revealed that EKLF regulates diverse genes, including those encoding heme biosynthesis enzymes and cytoskeletal proteins (Drissen *et al.*, 2005; Hodge *et al.*, 2006; Nilson *et al.*, 2006). Furthermore, EKLF is an important determinant of hemoglobin switching (Donze *et al.*, 1995; Perkins *et al.*, 1996; Wijgerde *et al.*, 1995; Gillemans *et al.*, 1998). The results of EKLF ablation in mouse and EKLF occupancy at the endogenous β-globin locus provide strong evidences that EKLF is an essential regulator of adult β–like globin gene regulation.

-p45/ NF-E2, consists of a hematopoietic-specific subunit, p45/NF-E2 (Ney *et al.,* 1993; Andrews *et al.,* 1993a), and of a member of the small Maf protein family (Andrews *et al.,* 1993b). NF-E2 is expressed in erythroid and megakaryocytic cells. Targeted disruption of *p45/NF-E2* in mice does not appear to significantly perturb erythropoiesis or β-globin transcription. However, the CB3 murine erythroleukemia cells, which lack p45/NF-E2 expression due to retroviral insertion within its locus, do not express β*major*, and p45/NF-E2 reintroduction in these cells reactivates β*major* expression (Kotkow and Orkin, 1995; Kiekhaefer *et al.,* 2004). Accordingly, NF-E2 expression in fact rescues Pol II occupancy at the β-globin promoter in CB3 cells, resulting in a considerable transcriptional activation (Johnson *et al.,* 2001). Endogenous p45/NF-E2 occupies the LCR, and the β*major* promoter in erythroid cells (Sawado *et al.,* 2001). Taken together these data indicate that NF-E2 is an important regulator of hemoglobin synthesis. The lack of a large defect in β-like globin expression in p45/NF-E2-null mice might be due to the existence of some NF-E2-related factors, able to compensate for the defect. On the contrary, the ablation of p45/NF-E2 in mice showed that p45/NF-E2 is crucial for megakaryopoiesis (Shivdasani *et al.,* 1995).

-Sox6 is a member of the Sry-related HMG box transcription factors. It is expressed in several tissues, including cartilage, testis, neuronal and erythropoietic tissues (Connor *et al.,* 1995, Takamatsu *et al.,* 1995, Lefebvre V *et al.,* 1998). Sox transcription factors bind to the minor groove of DNA and cause a drastic bend to about 75º leading to local conformational changes (Ferrari *et al.,* 1992). These properties

suggest that Sox proteins can act as an “architectural proteins” possibly by promoting the assembly of biologically active multiprotein complexes. The Sox6 DNA binding domain recognizes a very degenerate (A/T)(A/T)CAA(A/T)G consensus, making it very difficult to define the *in vivo* Sox6 targets on DNA. Sox6 lacks any activation / repression domain but it can act as an activator or a repressor, depending on its interacting protein and promoter context (Lefebvre *et al.,* 1998, Murakami *et al.,* 2001). Sox6 null mutant mice (p^{100H}) showed delayed growth, myopathy, arterioventricular heart block and die within 2 weeks following birth (Hagiwara N *et al.,* 2000).

Sox6 enhances definitive erythropoiesis in mice by stimulating erythroid cell survival, proliferation and terminal maturation (Dumitriu *et al.,* 2006). Recently it has been shown that Sox6 overexpression enhances erythroid differentiation in human erythroid progenitors and in K562 cells (Cantu C *et al.,* 2011). Sox6 represses εy-globin and to a lesser extent, βh1-globin expression in definitive erythropoiesis of adult mice (Cohen-Barak *et al.,* 2007). In particular, Sox6 directly silences εy-globin expression by binding to the double Sox6 binding site on the εy promoter (Yi Z *et al.,* 2006). Finally, it has been shown that Sox6 cooperates with BCL11A and act collaboratively in silencing the γ-globin gene in adult human erythroid cells (Xu J *et al.,* 2010).

-BCL11A (B-cell lymphoma/leukemia 11A) also known as Evi9 / CTIP1, is a Zinc-finger transcription factor was initially cloned as a myeloid or B cell proto-oncogene in mice and humans (Fell *et al.,*

1986, Li *et al.*, 1999, Suzuki *et al.*, 2002). Alternative pre-mRNA splicing of human BCL11A gene leads to a minimum of 4 transcripts predicted to yield protein isoforms designated as eXtra-Long (XL; 5.9 kb/125 kD), Long (L; 3.8 kb/100 kD), Short (S, 2.4 kb/35 kD) and eXtra-Short (XS, 1.5 kb/25 kD) (Liu H *et al.*, 2006). BCL11A-XL RNA is expressed at high levels in normal as well as malignant lymphoid tissues, including germinal centre (GC) B cells, B-CLL and follicular lymphoma (Satterwhite E *et al.*, 2001; Su Al *et al.*, 2002). BCL11A is required for normal lymphoid development (Liu *et al.*, 2003). BCL11A mutant embryos lack B cells and have alterations in several types of T cells.

Genome wide association studies (GWAS) originally led to the identification of a new HbF associated locus on chromosome 2, located within the gene BCL11A (Menzel *et al.*, 2007, Lettere *et al.*, 2008, Thein *et al.*, 2007, Uda *et al.*, 2008). Since this first finding, BCL11A has been shown to function as a regulator of HbF expression in humans (Sankaran VG *et al.*, 2008). Knockdown of BCL11A expression in adult human erythroid precursors leads to a robust induction of HbF (Sankaran VG *et al.*, 2008), consistent with a role of BCL11A as a repressor of HBG1 and HBG2. Knockdown of BCL11A in transgenic mice carrying the human β-globin locus prevents the proper silencing of the endogenous mouse β like embryonic genes in adult fetal liver erythroid cells (Sankaran *et al.*, 2009). Additionally, BCL11A interacts with the NuRD chromatin remodelling complexes, as well as erythroid transcription factors GATA1 and FOG1 in erythroid progenitors (Sankaran VG *et al.*, 2008) and recently it has been shown that BCL11A interacts physically with Sox6 in silencing

γ-globin transcription in adult human erythroid cells (Xu J *et al.*, 2010).

- COUP-TFII (Chicken Ovalbumin Upstream promoter-Transcription factor II) also called as NR2F2 / ARP-1 are Orphan nuclear receptors because ligands for COUP-TF have not been identified. COUP-TFs form homodimers and bind to a wide range of direct repeat AGGTCA motifs on DNA with a variety of spacing and orientation (Cooney AJ *et al.,* 1992, Kleiwer *et al.,* 1992). COUP-TFII has been implicated in many vital processes, such as organogenesis, angiogenesis, and metabolic homeostasis as well as in variety of developmental programs (Tang K *et al.,* 2006). Within the hematopoietic system, COUP-TFII is expressed in embryonic / fetal erythroid cell lines, mouse yolk sac and fetal liver (Filipe A *et al.,* 1999). COUP-TFII was proposed as a candidate repressor of ε- and γ- globin genes where it binds to the double CCAAT box regions on the ε- and γ-globin promoters and interferes with the binding of NF-Y (Ronchi A *et al.,* 1995, Liberati C *et al.,* 1998). COUP- TFII binds to the promoter of ε-globin by two sequences of the DR1 (Direct Repeat 1) type. The most of 3' site overlaps with the CCAAT box, recognized by NFY (Filipe A *et al.,* 1999). When COUP- TFII binds to the 5' site, it cooperates with NFY in the formation of a stable complex on DNA, which stimulates the transcription of the gene. Instead, when Coup -TFII binds to the site located at 3' end, competes with NFY and, directly or indirectly, causes a decrease in gene activity (Liberati C *et al.,* 2001). COUP-TFII might contribute to γ-globin regulation and to HPFH phenotype. Perturbation of Coup-TFII binding is a common effect of HPFH

mutations mapped in this region (Liberati C *et al.,* 2001). Gel shift experiments showed that COUP-TFII binds to the two CCAAT box (proximal and distal) on the γ-globin promoter sequence. It is hypothesized that NFY plays a role of strong activator and Coup -TF II to a weak activator or repressor: the occupancy of the site by one or the other factor, and the consequent activation or repression of the gene would change during development depending on their relative concentrations on the target sequence. Alternatively it can be assumed that this different behaviour could also be due to the availability of a ligand or a limiting subunit that interact with Coup -TFII , some of co-activator/co-repressor or variations in post-translational modifications of the regulatory factors involved (Liberati C *et al.,* 2001).

In human adult progenitors, expression of COUP-TFII was suppressed by SCF stimulation that in turn leads to the increase in γ-globin expression (Aerbajinai *et al.,* 2009). Of interest, BCL11A was originally isolated as CTIP1 (COUP-TFII interacting partner 1) (Avram *et al.,* 2000). Recently, it has been shown that NF-Y a transcriptional activator that stably binds at the CAAT boxes recruits and stabilizes the binding of BCL11A with COUP-TFII on the proximal γ-globin promoter region forming a “repressor hub” in silencing gamma globin expression (Zhu X *et al.,* 2012).

THE MEDICAL RELEVANCE:

HEMOGLOBINOPATHIES

During erythroid differentiation and maturation, the production of α-globin and β- globin chains in the right ratio is critical to form stable HbA. Imbalance or alterations of these components due to mutations or deletions in globin genes and /or their regulators can be deleterious for RBCs and their precursors and may lead to Hemoglobinopathies (Stamatoyannopoulos G *et al.*, 1992, Olivieri and Weatherall *et al.*, 1998). Hemoglobinopathies are the most common inherited monogenic diseases, and affect million people in the world; among them, in particular, β-thalassemia and Sickle cell anemia. These diseases are caused by mutations involving the adult β -globin gene, responsible for the synthesis of over 97% of the adult-type haemoglobin HBA (α2 β 2). Patients carrying heterozygous mutations on one β gene develop a mild anemia; mutations on both β genes cause a very severe anemia and require intensive clinical treatments (Stamatoyannopoulos and Nienhuis, 1994; Weatherall and Clegg, 1981; Forget, 1998; Thein, 1998). Hemoglobin disorders constitute a significant health problem in more than 160 countries worldwide. These countries are responsible for 89% of worldwide births. Over 330,000 affected infants are born annually (83% sickle cell, 17% thalassemia) (CDC 2012). The only possible clinical treatment is blood transfusion, which causes another series of symptoms, for example a toxic accumulation of Fe2+ within liver and heart: patients can overcome the surplus of Fe2+ only with daily administration of

specific iron-chelating drugs. Nowadays, researchers are trying to find different approaches for new therapeutical trials heading to heal thalassemia. Two different strategies are now considered. The first one is gene therapy, whose aim is to introduce a normal β-globin gene within hematopoietic stem cells of the patient by using viral vectors that can drive the expression of a correct variant of the gene in a tissue specific manner. At the moment, vectors derived from HIV backbone seems to be the most efficient way to transduce the β -globin gene into hematopoietic progenitors, but they still cannot be used for human gene therapy for safety reasons, although some progress toward safe and effective gene therapy have been recently reported (Lebensburger J., Persons DA., 2008). The second one is the maintenance of fetal γ-globin gene expression during the whole adult life. In fact, fetuses carrying genetic mutations that lead to β -thalassemia or sickle cell anemia, are not affected by the disease, since during the embryonic-fetal stages of development, other genes (ε and γ globin) are present with α-globin the largest majority of haemoglobin molecules α2ε2 (Gower2) and α2γ2 (HbF).

In a group of genetic conditions called hereditary persistence of fetal hemoglobin (HPFH), the expression of the γ-globin gene may persist at high levels in adult erythroid cells. Molecular studies of the HPFH syndromes have identified several important regulatory elements for the normal pattern of γ-globin gene expression. HPFH is usually due to deletions of different sizes involving the β-globin gene cluster or to point mutations in the γ-globin gene promoters. Moreover, Numerous clinical observations show how coinheritance of mutations causing high expression of γ- globin after birth (Hereditary

Persistence of Fetal Hemoglobin, HPFH), significantly ameliorates the clinical condition of patients affected by thalassemia or Sickle cell anemia (Stamatoyannopoulos and Nienhuis, 1994; Weatherall and Clegg, 1981; Forget, 1998; Thein, 1998). There are some Mendelian forms of HPFH that are rare and do not explain the common form of heterocellular HPFH, characterized by relatively low levels of fetal hemoglobin (2-4%), which represents the majority of normal HbF variation, and is clearly inherited as a quantitative genetic trait (Forget BG., 1998; Thein SL.*et al.*, 2009).

Many efforts have been made to increase γ-globin expression in β-thalassemic and HbS patients using drugs as butyrates, 5-azaC or their derivatives. Butyrates are non-specific inhibitors of histone deacetylase that lead to the maintenance of the histone acetylation on promoters, with the consequent transcriptional activation of the γ-globin gene. 5-azaC and its derivates inhibits the DNA methylation; since the γ- globin promoter is hypermethylated and indeed silenced during adult life, 5-azaC leads to an hypomethylation of this promoter and a consequent upregulation of γ-globin gene expression (Atweh *et al.*, 2003). The modest results obtained with these drugs encouraged the study of new therapeutical approaches with a particular attention to transcription factors and molecules that bind γ-globin promoter and that can modulate its transcriptional activity.

In the past our group identified some HPFH mutations within the γ- globin promoter that alter the binding of different nuclear factors (NFE3, GATA-1, Sp1, NF-Y) involved in the regulation of γ-globin gene expression (Nicolis *et al.*, 1989; Ottolenghi *et al.*, 1989; Ronchi *et al.*, 1995; Ronchi *et al.*, 1996; Ronchi *et al.*, 1989).

Recently genome-wide association studies in man identified three loci containing a series of five single-nucleotide polymorphisms (SNP), able to explain the 20% of variation in HbF expression (Menzel S., 2007). The SNP with the largest effect was localized in the locus of the transcription factor BCL11A, whose protein product was then demonstrated to be able to repress the γ- globin transcription by direct binding on its regulatory regions, in a stage-specific manner (Sankaran VG., 2008, Sankaran VG, 2009).

In conclusion, the identification of proteins involved in the regulation of globin genes through proteomics and functional genomics approaches, represents the rational therapeutical basis for possible future treatments of thalassemia. Strategies including the use of small molecule inhibitors or gene therapy knockdown approaches with shRNAs aimed at targeting TF's or other regulators of hemoglobin switching are promising avenues for the field.

SCOPE OF MY THESIS

The aim of my thesis is to elucidate the molecular mechanisms regulating the fetal (γ) to adult (β) globins switching. Basic research about erythropoiesis and globins chains transcriptional regulation may have a strong impact on the medical treatment β-thalassemias and sickle cell anemia (SCA). Many efforts have been done to study which genes are involved in the physiologic down the regulation of human fetal globin gene, in order to provide molecular approaches directed to

the reactivation of γ-globin in thalassemic and SCA patients, in which the adult β- globin production is impaired.

This potential approach was first suggested by the clinical observation on a group of genetic conditions called hereditary persistence of fetal haemoglobin (HPFH), in which the level of fetal γ-globin is maintained at high levels in the blood of adult individuals. The co-inheritance of HPFH in patients affected by β-thalassemia greatly ameliorates the clinical condition.

In my thesis, I would be focusing my attention on the transcription factors Sox6 and COUP-TFII and on their functional interaction (and interactors) in regulating the γ- to β- globin ratio.

- In **chapter 2**, I will describe the model systems employed in this project for *in vivo* and *ex vivo* studies of globin gene expression and switching. In particular, I will describe a new variant of the human K562 cell line that expresses both β- and γ- globin expression. Till date there are no other cell lines expressing good amount of β- globin that can be used to study by biochemical approaches the factors involved in the transcriptional control of the globin switching. To confirm the molecular mechanisms suggested by V-K562 studies and to focus on the switching time in a more physiological context, I have used mouse embryonic fetal liver cells from day E11.5-E13.5 (where the switching takes place). Finally, to validate the results, I used human adult peripheral blood cultured from healthy donor individuals (in collaboration with Prof. Paolo Moi & Dr. Giusi Marini, Cagliari, ITALY).

- In **chapter 3**, I will describe the molecular functions of the transcription factors Sox6 and COUP-TFII in modulating the γ- to β- globin ratio upon their lentiviral-mediated overexpression in V-K562 cells. I will show that both Sox6 and COUP-TFII bind on the human γ-globin promoter, *in vivo* by ChIP experiments. Finally, I will report the identification of a Sox6 binding site within the γ-globin promoter whose mutation increases γ-globin expression in Luciferase reporter assays in V-K562 cells. Taken together, these data suggest that Sox6 and COUP-TFII functionally interact with each other in controlling the γ to β -globins ratio.

- In **chapter 4**, I will present the results of Sox6 overexpression in V-K562 cells stably transfected with a pcDNA3_BCL11A-XL overexpression vector. I will also show by ChIP experiments that BCL11A-XL can bind to the ε- and γ-globin promoters only in the presence of Sox6. This suggests that BCL11A-XL can exert its repressive function only in the presence of Sox6. Finally, I will show the Sox6 overexpression correlates with changes in BCL11A isoforms suggesting that Sox6 could act as a splicing factor in splicing the BCL11A isoforms in both P19 cells and V-K562 cells

- Finally, in **chapter 5**, all the results presented in this thesis are briefly discussed as well as the future perspectives of this research in molecular and translational medicine.

REFERENCES

Abkowitz, J. L., D. Golinelli, D. E. Harrison and P. Guttorp (2000). "In vivo kinetics of murine hemopoietic stem cells." Blood **96**(10): 3399-3405.

Aerbajinai, W., Zhu, J., Kumkhaek, C., Chin, K. and Rodgers, G.P. (2009) SCF induces gamma-globin gene expression by regulating downstream transcription factor COUP-TFII. Blood, **114,** 187-94.

Akashi, K., D. Traver, T. Miyamoto, and I. L. Weissman. 2000. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. Nature **404**:193-197.

Allen, T. D. and T. M. Dexter. 1982. Ultrastructural aspects of erythropoietic differentiation in long-term bone marrow culture. Differentiation **21**:86-94.

Allen, T. D. and N. G. Testa. 1991. Cellular interactions in erythroblastic islands in long-term bone marrow cultures, as studied by time-lapse video. Blood Cells **17** :29-38.

Alvarez-Silva, M., P. Belo-Diabangouaya, J. Salaun, and F. Dieterlen-Lievre. 2003. Mouse placenta is a major hematopoietic organ. Development **130**:5437-5444.

Andrews, N. C., K. J. Kotkow, P. A. Ney, H. Erdjument-Bromage, P. Tempst, and S. H. Orkin. 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.

Atweh, G. F., J. DeSimone, Y. Sauntharajah, H. Fathallah, R. S. Weinberg, R. L. Nagel, M. E. Fabry, and R. J. Adams. 2003. Hemoglobinopathies.Hematology Am.Soc.Hematol.Educ.Program.14-39.

Avram, D., Fields, A., Pretty On Top, K., Nevrivy, D.J., Ishmael, J.E. and Leid, M. (2000) Isolation of a novel family of C(2)H(2) zinc

finger proteins implicated in transcriptional repression mediated by chicken ovalbumin upstream promoter transcription factor (COUP-TF) orphan nuclear receptors. J Biol Chem, **275,** 10315-22.

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

Bulger, M., M. A. Bender, J. H. van Doorninck, B. Wertman, C. M. Farrell, G. Felsenfeld, M. Groudine, and R. Hardison. 2000. Comparative structural and functional analysis of the olfactory receptor genes flanking the human and mouse beta-globin gene clusters. Proc.Natl.Acad.Sci.U.S.A **97**:14560-14565.

Cantu, C., Ierardi, R., Alborelli, I., Fugazza, C., Cassinelli, L., Piconese, S., Bose, F., Ottolenghi, S., Ferrari, G. and Ronchi, A. (2011) Sox6 enhances erythroid differentiation in human erythroid progenitors. Blood, **117,** 3669-79.

Cohen-Barak, O., Z. Yi, N. Hagiwara, K. Monzen, I. Komuro, and M. H. Brilliant. 2003. Sox6 regulation of cardiac myocyte development. Nucleic Acids Res. **31**:5941-5948.

Connor, F., Wright, E., Denny, P., Koopman, P. and Ashworth, A. (1995) The Sry-related HMG box-containing gene Sox6 is expressed in the adult testis and developing nervous system of the mouse. Nucleic Acids Res, **23,** 3365-72.

Cooney, A.J., Tsai, S.Y., O'Malley, B.W. and Tsai, M.J. (1992) Chicken ovalbumin upstream promoter transcription factor (COUP-TF) dimers bind to different GGTCA response elements, allowing COUP-TF to repress hormonal induction of the vitamin D3, thyroid hormone, and retinoic acid receptors. Mol Cell Biol, **12,** 4153-63.

Chiba, T., Y. Nagata, A. Kishi, K. Sakamaki, A. Miyajima, M. Yamamoto, J. D. Engel, and K. Todokoro. 1993. Induction of erythroid-specific gene expression in lymphoid cells. Proc.Natl.Acad.Sci.U.S.A **90**:11593-11597.

de Laat, W. and F. Grosveld. 2003. Spatial organization of gene expression: the active chromatin hub 130. Chromosome.Res. **11**:447-459.

de Laat, W. and F. Grosveld. 2003. Spatial organization of gene expression: the active chromatin hub. Chromosome.Res. **11**:447-459.

Donze, D., T. M. Townes, and J. J. Bieker. 1995. Role of erythroid Kruppel-like factor in human gamma- to beta-globin gene switching. J.Biol.Chem. **270**:1955-1959.

Drissen, R., von Lindern, M., Kolbus, A., Driegen, S., Steinlein, P., Beug, H., Grosveld, F. and Philipsen, S. (2005) The erythroid phenotype of EKLF-null mice: defects in hemoglobin metabolism and membrane stability. Mol Cell Biol, **25,** 5205-14.

Dumitriu, B., M. R. Patrick, J. P. Petschek, S. Cherukuri, U. Klingmuller, P. L. Fox, and V. Lefebvre. 2006. Sox6 cell-autonomously stimulates erythroid cell survival, proliferation, and terminal maturation and is thereby an important enhancer of definitive erythropoiesis during mouse development. Blood **108**:1198-1207.

Enver, T. and S. E. Jacobsen. 2009. Developmental biology: Instructions writ in blood. Nature **461**:183-184.

Fantoni, A., C. A. De la, D. Chui, R. A. Rifkind, and P. A. Marks. 1969. Control mechanisms of the conversion from synthesis of embryonic to adult hemoglobin. Ann.N.Y Acad.Sci. **165**:194-204.

Fell, H.P., Smith, R.G. and Tucker, P.W. (1986) Molecular analysis of the t(2;14) translocation of childhood chronic lymphocytic leukemia. Science, **232,** 491-4.

Fermi, G., Perutz, M.F., Shaanan, B. and Fourme, R. (1984) The crystal structure of human deoxyhaemoglobin at 1.74 A resolution. J Mol Biol, **175,** 159-74.

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,** 687-97.

Forget, B. G. 1998. Molecular basis of hereditary persistence of fetal hemoglobin. Ann.N.Y Acad.Sci. **850**:38-44.

Forrester, W. C., E. Epner, M. C. Driscoll, T. Enver, M. Brice, T. Papayannopoulou, and M. Groudine. 1990. A deletion of the human beta-globin locus activation region causes a major alteration in chromatin structure and replication across the entire beta-globin locus. Genes Dev. **4**:1637-1649.

Franco, R.S. (2012) Measurement of red cell lifespan and aging. *Transfus Med Hemother*, **39,** 302-7.

Gekas, C., E.Rhodes K., K. A. Mikkola H. 2008. Isolation and analysis of hematopoietic stem cells from the placenta. J Vis Exp. 10:3791/742.

Gillemans, N., R. Tewari, F. Lindeboom, R. Rottier, T. de Wit, M. Wijgerde, F. Grosveld, and S. Philipsen. 1998. Altered DNA-binding specificity mutants of EKLF and Sp1 show that EKLF is an activator of the beta-globin locus control region in vivo. Genes Dev. **12**:2863-2873.

Graf, T. and Enver, T. (2009) Forcing cells to change lineages. *Nature*, **462,** 587-94.

Graf, T. 2002. Differentiation plasticity of hematopoietic cells. Blood **99**:3089-3101.

Gregory, T., C. Yu, A. Ma, S. H. Orkin, G. A. Blobel, and M. J. Weiss. 1999. GATA-1 and erythropoietin cooperate to promote erythroid cell survival by regulating bcl-xL expression. Blood **94**:87-96.

Grosveld, F., G. B. van Assendelft, D. R. Greaves, and G. Kollias. 1987. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. Cell **51**:975-985.

Hagiwara, N., S. E. Klewer, R. A. Samson, D. T. Erickson, M. F. Lyon, and M. H. Brilliant. 2000. Sox6 is a candidate gene for p100H myopathy, heart block, and sudden neonatal death. Proc.Natl.Acad.Sci.U.S.A **97**:4180-4185.

Hock, H., M. J. Hamblen, H. M. Rooke, J. W. Schindler, S. Saleque, Y. Fujiwara, and S. H. Orkin. 2004. Gfi-1 restricts

proliferation and preserves functional integrity of haematopoietic stem cells. Nature **431**:1002-1007.

Hodge, D., E. Coghill, J. Keys, T. Maguire, B. Hartmann, A. McDowall, M. Weiss, S. Grimmond, and A. Perkins. 2006. A global role for EKLF in definitive and primitive erythropoiesis. Blood **107**:3359-3370.

Ingley, E., P. A. Tilbrook, and S. P. Klinken. 2004. New insights into the regulation of erythroid cells. IUBMB.Life **56**:177-184.

Johnson, G. R. and D. Metcalf. 1977. Pure and mixed erythroid colony formation in vitro stimulated by spleen conditioned medium with no detectable erythropoietin. Proc.Natl.Acad.Sci.U.S.A **74**:3879-3882.

Johnson, K. D., H. M. Christensen, B. Zhao, and E. H. Bresnick. 2001. Distinct mechanisms control RNA polymerase II recruitment to a tissue-specific locus control region and a downstream promoter. Mol.Cell **8**:465-471.

Katsura, Y. 2002. Redefinition of lymphoid progenitors. Nat.Rev.Immunol. **2**:127-132.

Kawamoto, H. and Katsura Y.. 2009. A new paradigm for hematopoietic cell lineages: revision of the classical concept of the myeloid–lymphoid dichotomy.Trends in Immunology. **30**:193-200

Kiekhaefer, C. M., M. E. Boyer, K. D. Johnson, and E. H. Bresnick. 2004. A WW domain-binding motif within the activation domain of the hematopoietic transcription factor NF-E2 is essential for establishment of a tissue-specific histone modification pattern. J.Biol.Chem. **279**:7456-7461.

Kim, S. I. and E. H. Bresnick. 2007. Transcriptional control of erythropoiesis: emerging mechanisms and principles. Oncogene **26**:6777-6794.

Kotkow, K. J. and S. H. Orkin. 1995. Dependence of globin gene expression in mouse erythroleukemia cells on the NF-E2 heterodimer. Mol.Cell Biol. **15**:4640-4647.

Lacombe, C. and P. Mayeux. 1999. The molecular biology of erythropoietin. Nephrol.Dial.Transplant. **14 Suppl 2**:22-28.

Lebensburger, J. and D. A. Persons. 2008. Progress toward safe and effective gene therapy for beta-thalassemia and sickle cell disease. Curr.Opin.Drug Discov.Devel. **11**:225-232.

Lefebvre, V., P. Li, and B. de Crombrugghe. 1998. A new long form of Sox5 (L-Sox5), Sox6 and Sox9 are coexpressed in chondrogenesis and cooperatively activate the type II collagen gene. EMBO J. **17**:5718-5733.

Lettre, G., Sankaran, V.G., Bezerra, M.A., Araujo, A.S., Uda, M., Sanna, S., Cao, A., Schlessinger, D., Costa, F.F., Hirschhorn, J.N. and Orkin, S.H. (2008) DNA polymorphisms at the BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. Proc Natl Acad Sci USA, **105,** 11869-74.

Liberati, C., Sgarra, R., Manfioletti, G. and Mantovani, R. (1998) DNA binding of NF-Y: the effect of HMGI proteins depends upon the CCAAT box. FEBS Lett, **433,** 174-8.

Liberati, C., Cera, M.R., Secco, P., Santoro, C., Mantovani, R., Ottolenghi, S. and Ronchi, A. (2001) Cooperation and competition between the binding of COUP-TFII and NF-Y on human epsilon- and gamma-globin gene promoters. J Biol Chem, **276,** 41700-9.

Liu, H., Ippolito, G.C., Wall, J.K., Niu, T., Probst, L., Lee, B.S., Pulford, K., Banham, A.H., Stockwin, L., Shaffer, A.L., Staudt, L.M., Das, C., Dyer, M.J. and Tucker, P.W. (2006) Functional studies of BCL11A: characterization of the conserved BCL11A-XL splice variant and its interaction with BCL6 in nuclear paraspeckles of germinal center B cells. Mol Cancer, **5,** 18.

McGrath, K. E., T. P. Bushnell, and J. Palis. 2008. Multispectral imaging of hematopoietic cells: where flow meets morphology. J.Immunol.Methods **336**:91-97.

Menzel, S., Garner, C., Gut, I., Matsuda, F., Yamaguchi, M., Heath, S., Foglio, M., Zelenika, D., Boland, A., Rooks, H., Best, S., Spector, T.D., Farrall, M., Lathrop, M. and Thein, S.L. (2007) A

QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. Nat Genet, **39,** 1197-9.

Mikkola H. and S. H. Orkin. The journey of developing hemopoietic stem cells. 2006. Development. **133**:3733-3744

Munugalavadla, V., L. C. Dore, B. L. Tan, L. Hong, M. Vishnu, M. J. Weiss, and R. Kapur. 2005. Repression of c-kit and its downstream substrates by GATA-1 inhibits cell proliferation during erythroid maturation. Mol.Cell Biol. **25**:6747-6759.

Murakami, A., S. Ishida, J. Thurlow, J. M. Revest, and C. Dickson. 2001. SOX6 binds CtBP2 to repress transcription from the Fgf-3 promoter. Nucleic Acids Res. **29**:3347-3355.

Ney, P.A., Andrews, N.C., Jane, S.M., Safer, B., Purucker, M.E., Weremowicz, S., Morton, C.C., Goff, S.C., Orkin, S.H. and Nienhuis, A.W. (1993) Purification of the human NF-E2 complex: cDNA cloning of the hematopoietic cell-specific subunit and evidence for an associated partner. Mol Cell Biol, **13,** 5604-12.

Nicolis, S., C. Bertini, A. Ronchi, S. Crotta, L. Lanfranco, E. Moroni, B. Giglioni, and S. Ottolenghi. 1991. An erythroid specific enhancer upstream to the gene encoding the cell-type specific transcription factor GATA-1. Nucleic Acids Res. **19** :5285-5291.

Nilson, D. G., D. E. Sabatino, D. M. Bodine, and P. G. Gallagher. 2006. Major erythrocyte membrane protein genes in EKLF-deficient mice. Exp.Hematol. **34**:705-712.

Noordermeer, D. and W. de Laat . 2008. Joining the loops: beta-globin gene regulation. IUBMB.Life **60**:824-833.

Novershtern, N., Subramanian, A., Lawton, L.N., Mak, R.H., Haining, W.N., McConkey, M.E., Habib, N., Yosef, N., Chang, C.Y., Shay, T., Frampton, G.M., Drake, A.C., Leskov, I., Nilsson, B., Preffer, F., Dombkowski, D., Evans, J.W., Liefeld, T., Smutko, J.S., Chen, J., Friedman, N., Young, R.A., Golub, T.R., Regev, A. and Ebert, B.L. (2011) Densely interconnected transcriptional circuits control cell states in human hematopoiesis. Cell, **144,** 296-309.

Nuez, B., D. Michalovich, A. Bygrave, R. Ploemacher, and F. Grosveld. 1995. Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. Nature **375**:316-318.

Olivieri, N.F. and Weatherall, D.J. (1998) The therapeutic reactivation of fetal haemoglobin. Hum Mol Genet, **7,** 1655-8.

Ottolenghi, S., R. Mantovani, S. Nicolis, A. Ronchi, N. Malgaretti, B. Giglioni, and J. Gilman. 1989. Altered binding to the gamma-globin promoter of two erythroid specific nuclear proteins in different HPFH syndromes. Prog.Clin.Biol.Res. **316A**:229-236.

Palis, J. and G. B. Segel. 1998. Developmental biology of erythropoiesis. Blood Rev. **12**:106-114.

Palis, J. 2008. Ontogeny of erythropoiesis. Curr.Opin.Hematol. **15**:155-161.

Palstra, R. J., B. Tolhuis, E. Splinter, R. Nijmeijer, F. Grosveld, and W. de Laat. 2003. The beta-globin nuclear compartment in development and erythroid differentiation. Nat.Genet. **35**:190-194.

Park, I. K., D. Qian, M. Kiel, M. W. Becker, M. Pihalja, I. L. Weissman, S. J. Morrison, and M. F. Clarke. 2003. Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. Nature **423**:302-305.

Perkins, A. C., A. H. Sharpe, and S. H. Orkin. 1995. Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. Nature **375**:318-322.

Perutz, M.F. (1978) Hemoglobin structure and respiratory transport. Sci Am, **239,** 92-125.

Pevny, L., M. C. Simon, E. Robertson, W. H. Klein, S. F. Tsai, V. D'Agati, S. H. Orkin, and F. Costantini. 1991. Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. Nature **349**:257-260.

Prohaska, S. S., D. C. Scherer, I. L. Weissman, and M. Kondo. 2002. Developmental plasticity of lymphoid progenitors. Semin.Immunol. **14**:377-384.

Pronk, C.J.H., Rossi, D.J., Månsson, R., Attema, J.L., Norddahl, G.L., Chan, C.K.F., Sigvardsson, M., Weissman, I.L. and Bryder, D. (2007) Elucidation of the phenotypic, functional, and molecular topography of a myeloerythroid progenitor cell hierarchy. Cell Stem Cell, **1,** 428-42.

Ronchi, A., Nicolis S., Santoro C., and Ottolenghi S.. 1989. Increased Sp1 binding mediates erythroid-specific overexpression of a mutated (HPFH) gamma-globulin promoter. Nucleic Acids Res. **17**:10231-10241.

Ronchi, A.E., Bottardi, S., Mazzucchelli, C., Ottolenghi, S. and Santoro, C. (1995) Differential binding of the NFE3 and CP1/NFY transcription factors to the human gamma- and epsilon-globin CCAAT boxes. J Biol Chem, **270,** 21934-41.

Ronchi, A., M. Berry, S. Raguz, A. Imam, N. Yannoutsos, S. Ottolenghi, F. Grosveld, and N. Dillon. 1996. Role of the duplicated CCAAT box region in gamma-globin gene regulation and hereditary persistence of fetal haemoglobin. EMBO J. **15**:143-149.

Rylski, M., J. J. Welch, Y. Y. Chen, D. L. Letting, J. A. Diehl, L. A. Chodosh, G. A. Blobel, and M. J. Weiss. 2003. GATA-1-mediated proliferation arrest during erythroid maturation. Mol.Cell Biol. **23**:5031-5042.

Sankaran, V. G., T. F. Menne, J. Xu, T. E. Akie, G. Lettre, B. Van Handel, H. K. Mikkola, J. N. Hirschhorn, A. B. Cantor, and S. H. Orkin. 2008. Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. Science **322**:1839-1842.

Sankaran, V. G., J. Xu, T. Ragoczy, G. C. Ippolito, C. R. Walkley, S. D. Maika, Y. Fujiwara, M. Ito, M. Groudine, M. A. Bender, P. W. Tucker, and S. H. Orkin. 2009. Developmental and species-divergent globin switching are driven by BCL11A. Nature **460**:1093-1097.

Satterwhite, E. (2001) The BCL11 gene family: involvement of BCL11A in lymphoid malignancies. Blood, **98,** 3413-3420.

Sawado, T., K. Igarashi, and M. Groudine. 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.

Shemin, D. and RittenbergD. (1946) The life span of the human red blood cell. *J Biol Chem*, **166,** 627-36.

Shivdasani, R. A., E. L. Mayer, and S. H. Orkin. 1995. Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. Nature **373**:432-434.

Stamatoyannopoulos, G., C. A. Blau, B. Nakamoto, B. Josephson, Q. Li, E. Liakopoulou, B. Pace, T. Papayannopoulou, S. W. Brusilow, and G. Dover. 1994. Fetal hemoglobin induction by acetate, a product of butyrate catabolism. Blood **84**:3198-3204.

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

Strouboulis, J., N. Dillon, and F. Grosveld. 1992. Developmental regulation of a complete 70-kb human beta-globin locus in transgenic mice. Genes Dev. **6**:1857-1864.

Stryer, L. 1995. Biochemistry. NewYork, WH Freeman and Company.

Su, A.I., Cooke, M.P., Ching, K.A., Hakak, Y., Walker, J.R., Wiltshire, T., Orth, A.P., Vega, R.G., Sapinoso, L.M., Moqrich, A., Patapoutian, A., Hampton, G.M., Schultz, P.G. and Hogenesch, J.B. (2002) Large-scale analysis of the human and mouse transcriptomes. Proc Natl Acad Sci USA, **99,** 4465-70.

Takamatsu, N., Kanda, H., Tsuchiya, I., Yamada, S., Ito, M., Kabeno, S., Shiba, T. and Yamashita, S. (1995) A gene that is related to SRY and is expressed in the testes encodes a leucine zipper-containing protein. Mol Cell Biol, **15,** 3759-66.

Thein, S. L. and J. E. Craig. 1998. Genetics of Hb F/F cell variance in adults and heterocellular hereditary persistence of fetal hemoglobin. Hemoglobin **22**:401-414.

Thein, S.L., Menzel S., Lathrop M., Garner C. 2009. Control of fetal hemoglobin: new insights emerging from genomics and clinical implications.Hum Mol Genet. 15;18(R2):R216-23.

Uda, M., Galanello, R., Sanna, S., Lettre, G., Sankaran, V.G., Chen, W., Usala, G., Busonero, F., Maschio, A., Albai, G., Piras, , S., Asunis, I., Deiana, M., Nagaraja, R., Perseu, L., Satta, S., Cipollina, M.D., Sollaino, C., Moi, P., Hirschhorn, J.N., Orkin, S.H., Abecasis, G.R., Schlessinger, D. and Cao, A. (2008) Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of beta-thalassemia. Proc Natl Acad Sci USA, **105,** 1620-5.

Weatherall, D. J., L. Pressley, W. G. Wood, D. R. Higgs, and J. B. Clegg. 1981. Molecular basis for mild forms of homozygous beta-thalassaemia. Lancet **1**:527-529.

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

Weiss, M. J. and S. H. Orkin. 1995. GATA transcription factors: key regulators of hematopoiesis. Exp.Hematol. **23**:99-107.

Whyatt, D., F. Lindeboom, A. Karis, R. Ferreira, E. Milot, R. Hendriks, M. de Bruijn, A. Langeveld, J. Gribnau, F. Grosveld, and S. Philipsen. 2000. An intrinsic but cell-nonautonomous defect in GATA-1-overexpressing mouse erythroid cells. Nature **406**:519-524.

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

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

Yi, Z., Cohen-Barak, O., Hagiwara, N., Kingsley, P.D., Fuchs, D.A., Erickson, D.T., Epner, E.M., Palis, J. and Brilliant, M.H.

(2006) Sox6 directly silences epsilon globin expression in definitive erythropoiesis. PLoS Genet, **2,** e14.

Xu, J., Sankaran, V.G., Ni, M., Menne, T.F., Puram, R.V., Kim, W. and Orkin, S.H. (2010) Transcriptional silencing of {gamma}-globin by BCL11A involves long-range interactions and cooperation with SOX6. Genes Dev, **24,** 783-98.

Zhu, X., Wang, Y., Pi, W., Liu, H., Wickrema, A. and Tuan, D. (2012) NF-Y recruits both transcription activator and repressor to modulate tissue- and developmental stage-specific expression of human gamma-globin gene. PLoS One, **7,** e47175.

Chapter 2

MODEL SYSTEMS FOR *in vivo* AND *ex vivo* STUDIES OF GLOBINS GENE EXPRESSION

INTRODUCTION

An understanding of the regulation of the human erythropoiesis is required to rationally approach the problem of hemoglobin switching and of the reactivation of fetal hemoglobin in adult hematopoietic cells. Cultured permanent cell lines maintain their neoplastic phenotype and behave like relatively homogenous cell populations. Each permanent cell line is used as a model system depending on its potential and its wide use in the study of *in vitro* erythropoiesis and hemoglobin switching. Here, I will describe a variant K562 cell line particularly suitable for the study of hemoglobin switching. The switching from γ- to β-globin during development involves both genes in a competitive model of alternating transcription (Wijgerde *et al.*, Gribnau *et al.,* 1996). A great number of studies have highlighted the complexity of this process and have pointed out the need of a flexible and a reliable model that facilitates mass usage and high throughput analyses. However, the use of cell lines has serious limitations and need to be accompanied by the use of more "physiological" cell systems such as erythroid cultures from human peripheral blood samples and *ex vivo* embryonic mouse fetal liver cells (Strouboulis J. *et al.*, 1992) a better model for the study of "hemoglobin switching".

Human Variant K562 (V-K562) cell line:

K562 cells are a human immortalized cell line derived from a female patient with chronic myelogenous leukemia (CML) (Lozzio CB and Lozzio BB., 1977). It is a widely used model for cell biology, biochemistry, and erythropoiesis. K562 are in suspension cells and bear the Philadelphia chromosome (9:22 bcr-Abl+ translocation). K562 cells behave like poorly differentiated early erythroid progenitors that can be induced to differentiate and to accumulate hemoglobin upon induction with hemin and other inducing agents (Villeval JL *et al.,* 1983, Tsiftsoglou A.S *et al.,* 2003). Differentiated K562 cells synthesize embryonic and fetal but not adult hemoglobin, despite the fact that these cells also bear the intact β-globin gene cluster. The variant K562 cell line that I used here is a kind gift by Dr. Giuliana Ferrari from Hospital San Rafaelle, Milan, ITALY. With the purpose of using this cell line for future experiments, I characterized the cells by comparing the traditional K562 cells with this new variant K562 cells (henceforth called **V-K562**) by gene expression profiling and signature DNA finger printing techniques.

Mouse Embryonic Fetal Liver:

Mouse models have been extremely instrumental for our understanding of hematopoiesis. Transgenic mice have been used extensively for the study of human hemoglobin switching. Although "fetal specific" genes are absent in mouse, (where fetal derived erythrocytes start synthesizing HbA) a transgenic human β-like globin

locus can be appropriately regulated in mouse (Strouboulis J. *et al.*, 1992), thus suggesting that the murine model may represent a useful tool for studying global changes of gene expression during the hemoglobin switch. Mouse fetal liver at embryonic day E11.5-E13.5 serves as a good model to study the molecular mechanisms involved in globins switching.

Primary cultures of human erythroid cells:

Primary culture of erythroid cells isolated from peripheral blood can be readily established from most normal individuals and patients. Their culture *ex vivo* represents more closely than cell lines the *in vivo* situations. Two-Phase liquid culture system (Phase I & II) was used (Fibach E. *et al.*, 1989) for transduction experiments from the normal healthy individual volunteers in Cagliari, Italy in the laboratory headed by our collaborator Prof. Paolo Moi & Dr. Giusi Marini. Peripheral blood cells are used in my experiments because it represents a more physiological system, their availability, the homogeneity of the peripheral blood erythroid progenitors (BFUe) differentiation and of the possibility to transduce them efficiently.

RESULTS

Characterization of V-K562 cell line:

The V-K562 cells morphologically appear to be the same as the original K562 cells by Hematoxylin / Eosin staining on the cytospin preparations (Fig.2.1).

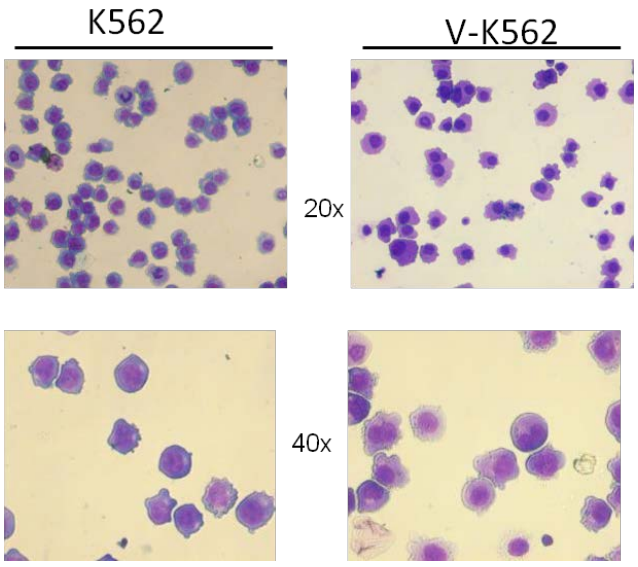


Figure 2.1: Phenotypic appearance of K562 and V-K562 cells upon staining with Haematoxylin and Eosin (H&E). The cellular morphology was determined by microscopy at the magnification 20x (upper panel) and 40x (lower panel) respectively.

qRT-PCR on globins genes expression revealed that V-K562 cells express reasonable levels of β-globin, corresponding to approx. 20% of β- like globins, whereas the original K562 cells do not express any β-globin (fig 2.2c). From Fig.2.2 a-b summarizes the results from semi

quantitative and quantitative RT-PCR for the globins expression. To better characterize this variant cell line, we first tested it for the presence of the typical bcr/abl p210 translocation in which fusion gene products gives a 210Kd protein (p210) (Price CM. *et al.*, 1988). To our surprise, as shown in figure 2.2c, V-K562 cells lack the typical p210 rearrangement. To further characterize these cells, we sent them to Nerviano Medical Sciences (NMS, Italy) to perform a DNA fingerprinting analysis to confirm the identity of V-K562 cell line. DNA fingerprinting takes advantage of hypervariable regions within DNA and was developed based on the short tandem repeat (STR) profiling of those hypervariable DNA regions (Butler JM. *et al.*, 2006). The V-K562 cell line show a 90% match with the parental K562 line suggesting that the V-K562 cells are genetically very close to the original K562 cells although they do express β-globins and lack the typical p210 rearrangement. The results of the STR analysis are summarized in Table.1 in appendix.

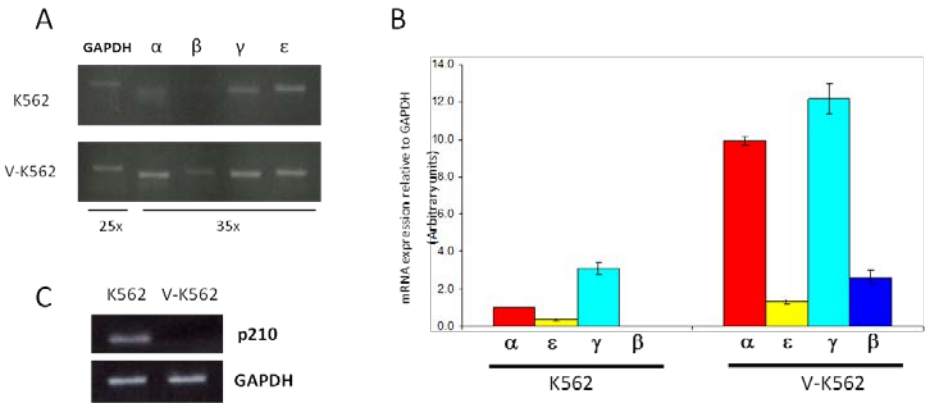


Figure 2.2: A-B). Semi-Quantitative and qRT- PCR on globin genes in K562 and V-K562 cells relative to glyceraldehyde-3-phosphate dehydrogenase (GAPDH). Blue histogram bar represents the level of β-globin in V-K562 cells (n=3). C). The

amplification of the fusion region typical of theBcr/Abl p210 translocation is absent in V-K562 cells.

MATERIALS & METHODS

Cell Cultures:

K562 and V-K562 cells were cultured in RPMI 1640 medium (Lonza) supplemented with 10% heat inactivated fetal bovine serum (Lonza), L-glutamine and antibiotics (Penicillin-Streptomycin 100U/100ug/ml) in a humidified 5% CO_2 atmosphere at 37°C. The morphology of K562 cells was observed after hematoxylin and eosin staining.

DNA Fingerprinting:

DNA fingerprints were obtained by using the AmpF STR Identifier PCR Amplification Kit (Applied Biosystems) according to the manufacturer's protocol. The kit amplifies the amelogenin gender-determining marker and 15 tetranucleotide repeat loci (listed in Table 1) in a single PCR amplification using 33 primers (the extra one is a degenerate primer targeting a mutation at the D8S1179 locus). Each of the STRs used in this study is composed of tetranucleotide repeat sequence. Allele calls were made from peak plots by comparing peaks to known standard fragment sizes using GeneMapper 4.0 (Applied Biosystems).

RNA Isolation and RT-PCR:

Total RNA from 10^6 cells (K562 and V-K562) were extracted with TRI Reagent (Applied Biosystem AM9738), treated with RQ1 DNase (Promega) for 30 min at 37°C and retrotranscribed (High Capacity cDNA Reverse Transcription Kit, Applied Biosystem cat n°4368814). Negative control reactions to detect the presence of genomic DNA (without Reverse Transcriptase) gave no signal. Real time analysis was performed using ABI Prism 7500, (PE Applied Biosystems). Primers were designed to amplify 100 to 150bp amplicons on the basis of sequences from the Ensembl database (http://www.ensembl.org/). Samples from three independent experiments were analyzed in triplicate. Specific PCR product accumulation was monitored by SYBR Green dye fluorescence in 25-µl reaction volume. Dissociation curves confirmed the homogeneity of PCR products. All primers used are listed in the table below:

Primers	Sequence
GAPDH Fw	ACGGATTTGGTCGTATTGGG
GAPDH Rev	TGATTTTGGAGGGATCTCGC
BCR+-ABL Fw	CGTCCACTCAGCCACAT
BCR+-ABL Rev	TCCAACGAGCGGCTTCAC
alpha globin Fw	GAGGCCCTGGAGAGGATGTTCC
alpha globin Rev	ACAGCGCGTTGGGCATGTCGTC
beta globin Fw	TACATTTGCTTCTGACACAAC
beta globin Rev	ACAGATCCCCAAAGGAC

gamma globin Fw	CTTCAAGCTCCTGGGAAATGT
gamma globin Rev	GCAGAATAAAGCCTACCTTGAAAG
epsilon globin Fw	GCCTGTGGAGCAAGATGAAT
epsilon globin Rev	GCGGGCTTGAGGTTGT

REFERENCES

Butler, J.M. (2006) Genetics and genomics of core short tandem repeat loci used in human identity testing. *J Forensic Sci*, **51,** 253-65.

Fibach, E., Manor, D., Oppenheim, A. and Rachmilewitz, E.A. (1989) Proliferation and maturation of human erythroid progenitors in liquid culture. *Blood*, **73,** 100-3.

Lozzio, B. B. and C. B. Lozzio . 1977. Properties of the K562 cell line derived from a patient with chronic myeloid leukemia. Int.J.Cancer **19**:136.

Price, C.M., Rassool, F., Shivji, M.K., Gow, J., Tew, C.J., Haworth, C., Goldman, J.M. and Wiedemann, L.M. (1988) Rearrangement of the breakpoint cluster region and expression of P210 BCR-ABL in a "masked" Philadelphia chromosome-positive acute myeloid leukemia. *Blood*, **72,** 1829-32.

Strouboulis, J., Dillon, N. and Grosveld, F. (1992) Developmental regulation of a complete 70-kb human beta-globin locus in transgenic mice. *Genes Dev*, **6,** 1857-64.

Tsiftsoglou, A.S., Pappas, I.S. and Vizirianakis, I.S. (2003) Mechanisms involved in the induced differentiation of leukemia cells. *Pharmacol Ther*, **100,** 257-90.

Villeval, J.L., Pelicci, P.G., Tabilio, A., Titeux, M., Henri, A., Houesche, F., Thomopoulos, P., Vainchenker, W., Garbaz, M., Rochant, H., Breton-Gorius, J., Edwards, P.A. and Testa, U. (1983) Erythroid properties of K562 cells. Effect of hemin, butyrate and TPA induction. *Exp Cell Res*, **146,** 428-35.

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,** 2894-902.

Chapter-3

ROLE OF SOX6 AND COUP-TFII IN THE REGULATION OF "HEMOGLOBIN SWITCHING"

(Manuscript in preparation)

INTRODUCTION

Several transcription factors are essential for erythroid commitment to occur and for differential globins gene expression during development: their absence is associated with a wide spectrum of phenotypes ranging from mild anemia to death due to a complete failure of erythropoiesis. Among them, Sox6 has been recently shown to stimulate erythroid cell survival, proliferation and terminal maturation of definitive murine erythropoiesis (Dumitriu B. *et al.* 2006, 2010). Sox6 has been proposed to directly act on globins genes by silencing the embryonic εy-gene in definitive erythroid cells by binding to its promoter (Yi Z. *et al.*, 2006). Accordingly with this proposed role, Sox6 null embryonic liver stem cells engrafted into lethally irradiated wild type adult mice show high levels of εy expression in bone marrow, spleen and blood (Cohen Barak O. *et al.*, 2007).

In our laboratory, we have demonstrated that Sox6 induces significantly erythroid differentiation in both K562 cells and in human primary erythroid cultures from cord blood purified $CD34^+$ cells (Cantú C. *et al.*, 2011). Sox6 overexpression in K562 forces their terminal maturation despite their erythroleukemic origin: K562 cells overexpressing Sox6 grow at a very low rate when compared with control cells, and die within about 6-9 days after transduction. Upon Sox6 overexpression, several erythroid specific transcripts are greatly increased, i.e. mRNAs for enzymes controlling the heme-biosynthetic pathway and for all globins (but of note, γ-globin expression is increased less than other globins), suggesting that Sox6 acts as a

general positive regulator of erythroid maturation and of erythroid genes expression. Accordingly to the phenotypic changes observed in K562 cells, Sox6 overexpression in primary cells is accompanied by an accelerated kinetics of maturation, with an increased number of cells that reach the final enucleation step (Cantú C. *et al.*, 2011). We know from literature that, the nuclear receptor COUP-TFII was shown to bind to the ε- and γ-globin gene promoters and was identified as the nuclear factor NF-E3. On the γ-globin promoter, COUP-TFII binds to a site overlapping the distal CCAAT box, competing with NF-Y, a strong activator of transcription recognizing the CCAAT box itself. The loss of COUP-TFII binding, observed as a consequence of some HPFH mutations mapping in this region, was suggested to contribute to the HPFH phenotype (Liberati C. *et al.*, 2001).

Among the number of multiprotein complexes binding to the γ-promoter region, I focused my study on Sox6 and COUP-TFII because both of them have been implicated in the repression of ε- and or γ-globin expression, with the aim of elucidating their role in the repression of embryonic genes and thus in the globins switching. In this chapter, I will describe the results of Sox6 and COUP-TFII overexpression in V-K562 cells and changes in the globin levels addressable to their enforced expression.I will also present a fluorescence based novel high content screening (HCS) suitable for multi parameter cell population studies. This approach allows, in addition to specific globin expression parameters, the collection of morphology and positional data at the single cell level. This technique was set in collaboration with Dr. Fabio Gasparri and Dr. Marta Durlak from Nerviano Medical Sciences (NMS), ITALY.

EXPERIMENTAL PROCEDURES

Sox6 and COUP-TFII overexpressing vector:

The Sox6 murine cDNA was kindly provided by Prof. Michiko Hamada-Kanazawa, Japan. The Sox6 cDNA was cloned in frame with a 3' FLAG epitope to generate a Sox6-FLAG cassette in which Sox6 cDNA lacks the 49 C-ter amino acids still retaining its biological proprieties (Hamada- Kanazawa *et al.*, 2004). The Sox6-FLAG cassette (EcoRI – KpnI blunted sites) was then cloned immediately upstream to the IRES - Emerald GFP cassette (blunted BamHI site) of the pHR SIN BX IR/EMW (derived from pHR SIN CSGW, (Demaison C *et al.*, 2002) lentiviral vector, a kind gift from Dr.Tariq Enver, Oxford. In this vector, the expression of the exogenous cDNA is driven by the SFFV promoter is highly active in hematopoietic cells. The two packaging plasmids psPAX2 and pMD-VSVG were used to produce Lentiviral pseudo-particles in HEK 293T cells (www.lentiweb.com). Alternatively, the Sox6 cDNA was also cloned in the same backbone containing IRES-ΔNGFR cassette to produce $Sox6_{\Delta NGFR}$. COUP-TFII cDNA was kindly provided by Prof. Michele Studer, France. The Full length COUP-TFII was initially cloned in (SacI- XhoI blunted sites) and was then cloned upstream to the IRES - Emerald GFP cassette (blunted BamHI site) the same way as Sox6 above ($COUP\text{-}TFII_{GFP}$).

The resulting cloned lentiviral vectors were checked for the expression exogenous mRNAs level and proteins in V-K562 cells, in parallel with their corresponding empty vector.

Lentivirus Harvesting Method:

Exponentially growing HEK293T cells were transfected with jetPEI™ reagent (Polyplus-Transfection) with the three vector lentiviral system. 72 hours after transfection, the supernatant containing the recombinant viruses were collected, filtered (0.45μm), centrifuged at 20,000g for 8 hours at 4ºC. The viral pellet were resuspended in PBS and stored in aliquots at -80ºC. Lentiviruses were titrated on HEK 293T cells by measuring the percentage of GFP positive cells or by staining them with Anti-NGFR (CD271) antibody conjugated PE/APC antibodies (for COUP-TFII$^{\Delta NGFR}$), by Flow Cytometry (FACS) analysis.

Cell Cultures:

V-K562 cells were cultured in RPMI 1640 medium (Lonza) supplemented with 10% heat inactivated fetal bovine serum (Lonza), L-glutamine (Euroclone) and antibiotics Penicillin-Streptomycin 100U/100ug/ml (Euroclone) in a humidified 5% CO_2 atmosphere at 37°C. Transduction was performed overnight, by adding the vector stock at multiplicity of infection (MOI) 30 for individual overexpression of Transcription factors.

RNA Isolation and RT-PCR:

Total RNA from 10^6 cells (V-K562) or mouse fetal liver E13.5 dpc erythroid cells were extracted with TRI Reagent (Applied Biosystems AM9738), treated with RQ1 DNase (Promega) for 30 min at 37°C and retrotranscribed (High Capacity cDNA Reverse Transcription Kit, Applied Biosystems cat n°4368814). Negative control reactions

(without Reverse Transcriptase) gave no signal. Real time analysis was performed using ABI Prism 7500, (PE Applied Biosystems). Primers were designed to amplify 100 to 150bp amplicons on the basis of sequences from the Ensembl / UCSC database. Samples from three independent experiments were analyzed in triplicate. Specific PCR product accumulation was monitored by SYBR Green dye fluorescence in 25-µl reaction volume. Dissociation curves confirmed the homogeneity of PCR products. All primers used are listed in the table below:

Primers	Sequence (5'-3')
GAPDH Fw	ACGGATTTGGTCGTATTGGG
GAPDH Rev	TGATTTTGGAGGGATCTCGC
ALAS-E Fw	CAACATCTCAGGCACCAGTA
ALAS-E Rev	CTCCACTGTTACGGATACCT
Fech Fw	ATCCAGCAGCTGGAGGGTCT
Fech Rev	TGAATCTTGGGGGTTCGGCG
CD117 (cKit) Fw	GATTTTGGTCTAGCCAGAGAC
CD117 (cKit) Rev	AAAATCCCATAGGACCAGAC
CD71 (TfR1) Fw	AAAATCCGGTGTAGGCACAG
CD71 (TfR1) Rev	CCTTTAAATGCAGGGACGAA
CD235a (GpA) Fw	CAGCTCATGATCTCAGGATG
CD235a (GpA) Rev	CACCTCAGTGGTACTTAATGC
CD44 Fw	GATCCACCCCAATTCCATCTGTGC
CD44 Rev	AACCGCGAGAATCAAAGCCAAGGCC
alpha globin Fw	GAGGCCCTGGAGAGGATGTTCC
alpha globin Rev	ACAGCGCGTTGGGCATGTCGTC

beta globin Fw	TACATTTGCTTCTGACACAAC
beta globin Rev	ACAGATCCCCAAAGGAC
gamma globin Fw	CTTCAAGCTCCTGGGAAATGT
gamma globin Rev	GCAGAATAAAGCCTACCTTGAAAG
epsilon globin Fw	GCCTGTGGAGCAAGATGAAT
epsilon globin Rev	GCGGGCTTGAGGTTGT
Sox6 Fw	GAGGCAGTTCTTTACTGTGG
Sox6 (endog) Rev	CCGCCATCTGTCTTCATAC
Sox6 (Flag) Rev	CTTATCGTCGTCATCCTTGTA
Coup-TFII (endog)Fw	TTGACTCAGCCGAGTACAGC
Coup-TFII (endog)Rev	AAAGCTTTCCGAATCTCGTC
Coup-TFII (exog) Fw	TCCAAGAGCAAGTGGAGAAG
Coup-TFII(exog) Rev	CTTCCAAAGCACACTGGGAC

Luciferase reporter Plasmids:

The human wildtype minimal γ-promoter and the mutant γ-promoter (generated by site directed mutagenesis) from nt -300 to +30 were cloned in pGL-2 vector upstream of a luciferase reporter. The plasmids were generated by two-step PCR. In the first PCR, mutations were introduced using an oligonucleotide containing the mutations and the other WT sequence annealing into the end of the sequence and carrying an external HindIII site for subsequent cloning. Resulting fragments were gel purified, annealed and then amplified again by PCR using the external oligonucleotides and Taq-polymerase for subsequent cloning into pGEM-T. Once cloned into p-GEM-T, the vector was HindIII digested and ligated into the HindIII restriction site

into the pGL-2 reporter vector (Promega). The amplified DNA regions were sequenced to avoid undesired mutations.

High Multiplex Fluorescent imaging:

For Immunoflourescence staining, V-K562 cells transduced with Empty Vector (EV), COUP-TFII and Sox6 after 72hrs, were collected and were fixed for 30 min by adding 100µl formaldehyde 7.4% solution to 100µl medium to prevent cell disruption and minimize cell detachment, then cells were washed with phosphate-buffered saline (PBS) and permeabilized with PBS solution containing 0.1% Triton X-100 and 1% bovine serum albumin (BSA; staining solution) for 30 min. Single and double Immunoflourescence were performed for all the samples. Briefly, conjugated primary antibodies (gamma and beta globins) (Santa Cruz) were added together at the recommended dilution in staining buffer containing 1µg/mL DAPI as well as Hoechst and incubated for 1h at 37 °C. Primary antibodies solution was removed, cells were washed twice with with PBS, and 100µl PBS was left in each well. All liquid handling was performed using a Biomek 2000 robotic station (Beckman Coulter, Fullerton, CA).

Image Acquisition and High Content Data Analysis:

The ArrayScan VTI High Content Screening reader (Thermo Fisher Scientific, Pittsburgh, PA) was used to perform single cell multiplexed immunofluorescence analysis. The platform is based on an inverted epifluorescence microscope (Zeiss, Thornwood, NY) equipped with a motorized stage, which automatically scans fields in 96 well

microplates. For each field, three fluorescence signals (blue, green, red) were collected with a 20x objective by exposing samples for fixed times and stored in three channels Ch1 (nuclear target recognition), Ch2 (cytoplasm green), and Ch3 (cytoplasm red) respectively. At least 10 fields were scanned per well and automatically analyzed within a single measurement by means of the Compartmental Analysis V3 Bio Application software (Thermo Fisher Scientific). This instrument allows quantitative analysis of multiple phenotypic parameters at the single cell level, including nuclear area and shape, DNA content (DAPI intensity), total and average fluorescence intensity in the nuclear and cytoplasmic compartments, and count and analysis of nuclear and cytoplasmic spots. Two-dimensional dot plots and fluorescence distribution histograms were generated by using Excel software (Microsoft, Redmond, WA). The entire fluorescence range (1‑4095) of ArrayScan data was subdivided into 200 bins. Visualizations and multidimensional analysis of cell subpopulations were performed using SpotFire software (Tibco, Palo Alto, CA).

Chromatin Immunoprecipitation (ChIP) assay:

Cells (10^6 V-K562 or $1x10^7$ primary cells) for each Immunoprecipitation reaction were fixed with 1% formaldehyde for 10 minutes at room temperature, and chromatin was sonicated (Branson™ Digic sonifier) to a size of about 100-500 bp in size. Immunoprecipitation was performed after overnight incubation with anti-Brn3 antibody (SC-6028), anti-FLAG antibody (Sigma F-7425)

or anti-Sox6 (abcam ab30455) and anti-COUP-TFII (abcam H7147) antibodies and subsequent incubation with protein-A agarose (Upstate biotechnology Millipore). Immunoprecipitated DNA was then analysed by amplifying an equivalent of 10^4 cells DNA with the following oligonucleotides:

Human Primers

Primers	Sequence (5'-3')
γ- globin promoter Fw	AAACGGTCCCTGGCTAAACT
γ- globin promoter Rev	GCTGAAGGGTGCTTCCTTTT
ε- globin promoter Fw	GAGCCTCAGGATCCAGCACAC
ε- globin promoter Rev	GATGCCAGGCCTGAGAGCTTGC
β- globin promoter Fw	GACAGGTACGGCTGTCATCA
β- globin promoter Rev	TAGATGGCTCTGCCCTGACT
Sox6 promoter Fw	TTTGAAAGAATACAGCCTCTG
Sox6 promoter Rev	ATGCATTAAGGTGGTTTGGTA
GAPDHexon Fw	CGGAGTCAACGGATTTGGTCGTAT
GAPDHexon Rev	AGCCTTCTCCATGGTGGTGAAGAC
COUP-TFII prom Fw	ATGGCAACGTGCGCTAAG
COUP-TFII prom Rev	GGCGCGGGAGGGAATGCG
LCR-HS-5 Fw	GACCTATATCTGGCAGGAC
LCR-HS-5 Rev	GTGATGTCTTACTAACTAGC
LCR-HS-4 Fw	TGGCATCTAGCGCAATGACTT
LCR-HS-4 Rev	GGGCAAGCCATCTCATAGCTG
LCR-HS-3 Fw	GGCAAGTGCCTTGACTCCTA
LCR-HS-3 Rev	TCTTCTGGAACTTGCCTGCT

LCR-HS-2 Fw	CCATAGTCCAAGCATGAGCA
LCR-HS-2 Rev	CTGGGGACCCAGATAGGAGT
LCR-HS-1 Fw	CCCCTA AGCTCCCAGAAAAC
LCR-HS-1 Rev	GGCTAAGGCATCTGTGAAGG
GpA Promoter Fw	TCATGAGCTGGTTCCTGAAG
GpA Promoter Rev	CCACTTTCATAGCCCCAAGA
NFE2 Promoter Fw	GAAAACCACGGCCACATATC
NFE2 Promoter Rev	GGGAAGCTGGTTGCATAACT

Mouse Primers

Primers	Sequence (5'-3')
εy- globin promoter Fw	CTCCACCCATGAGGACCACG
εy- globin promoter Rev	ACTGCTGCTAGAAGTGGTGG
Sox6 promoter Fw	GTCGTTTGCTTTTATTGTTTGGC
Sox6 promoter Rev	CACCACTCATCAGCATCTTCAC
Sox6 exon Fw	TTCCAAGCAAGCCACCTCTCC
Sox6 exon Rev	GTTGTGCATTATGGGGTGCAG
COUP-TFII prom Fw	ATGGCAACGTGCGCTAAG
COUP-TFII prom Rev	GGCGCGGGAGGGAATGCG
COUP-TFII 5'prom Fw	GATCCCGGCACCTTGCCACC
COUP-TFII 5'prom Rev	GGCTCTCGCAACTTGGCG
GpA Promoter Fw	TCATGAGCTGGTTCCTGAAG
GpA Promoter Rev	CCACTTTCATAGCCCCAAGA

Immunoprecipitation was repeated 2 times on independent Chromatin preparations.

Western Blot:

Total and nuclear extracts from V-K562 cells were prepared according to standard protocols (Schreiber E. *et al.,* 1989) and proteins were subjected to sodium dodecyl sulphate-polyacrylamide gel electrophoresis separation and blotting. The Sox6 protein was detected by anti-FLAG antibody (Sigma F-7425) or anti-Sox6 (abcam ab30455) and anti-COUP-TFII (abcam H7147) antibodies. Protein loading control was checked with an anti-U2AF antibody (Sigma-Aldrich). Antibody binding was detected by use of appropriate horseradish peroxidise-conjugated immunoglobulin G and revealed by enhanced chemiluminescence (LiteAblot; Euroclone).

Transfection Experiments:

A total of 1.5 x 10^5 exponentially growing V-K62 cells were transfected in 0.5ml of Opti-MEM medium (Invitrogen) with 2µl of Lipofectamine 2000 (Invitrogen), 100ng of reporter plasmids and increasing amounts (0.2 to 1µg) of the Sox6 expression plasmid (pCMV-Sox6Tag4b) per well. The pCMV-Tag4b empty vector was added to each transfection to equalize the total amount of DNA transfected in each reaction. After 48 hrs, total cellular extracts were prepared and luciferase activity was measured according to the Promega luciferase reporter system protocol. Non transfected cells were used as the background control for luciferase activity. Transfections were repeated in triplicates with 2 independent plasmid preparations.

Electrophoretic Mobility Shift Assay (EMSA):

^{32}P-labeled DNA probes were incubated with 3-6μg of nuclear protein extracts for 20 mins at 15ºC in a buffer that contained 5% glycerol, 50mM KCl, 20mM Tris, pH7.8, 0.5mM EDTA, 5mM $MgCl_2$, 1mM dithiothreitol, 100ng/μl poly (dG-dC), 50 ng/μl bovine serum albumin in a 15 μl final volume. The reaction mixture was loaded onto a 5% polyacrylamide gel (29:1 acrylamide-bisacrylamide ratio) and run at 4ºC at 150v for 3 hours. Gels were fixed in Acetic acid and methanol 10% solution, air dried and developed. Nuclear extracts were prepared according to standard proctocols (Schreiber E. *et al.*, 1989).

Flow cytometry analysis and FACS:

Transduced V-K562 cells were washed, fixed in 1% Paraformaldehyde and stained with PE-anti NGFR conjugated antibody (Biolegend) for 1 hr at 4ºC and analyzed by flow cytometer (Becton-Dickinson FACS Calibur). Freshly extracted mouse fetal liver cells from E13.5 cells were disaggregated to single cells and washed in phosphate-buffered saline (PBS), and incubated with conjugated antibodies: allophycocyanin (APC) conjugated anti-mouse CD117 (c-Kit), Phycoerythrin (PE) anti-mouse CD71 and FITC anti-mouse Ter119 all from Becton-Dickinson. The sorting was done by a MoFlo (DAKO-Cytomation) Cell sorter. The purity of the sorted cell populations were > 95%.

RESULTS

Sox6 and COUP-TFII overexpression in V-K562 cells

Since V-K562 cells express β-globin, they represent a good model to study the effect of transcription factors overexpression on the switching from γ- to β- globin expression. V-K562 cells were transduced with the following vectors: Sox-6 Flag (Fig 3.1a top), COUP-TFII (Fig 3.1a bottom) and the corresponding empty vector (EV-GFP). The efficiency of transduction was similar in all conditions with more than 95% GFP positivity (Fig.3.1c).

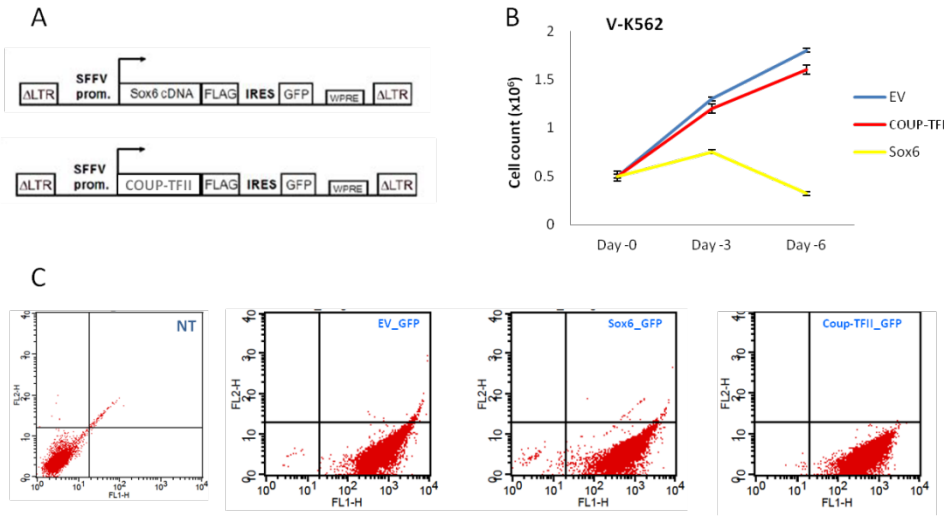


Figure 3.1: A). Schematic representation of the Sox6 and COUP-TFII overexpressing lentiviral vector used. LTR indicates Long Terminal Repeats; SFFV, Spleen Focus Forming Virus; WPRE, Woodchuck-Hepatitis-virus-Posttranscriptional-Regulatory Element. B). 1x 10^6 exponentially growing K562

cells were transduced at day 0 with Sox6, COUP-TFII or the corresponding EV. Cells were washed and re plated in fresh medium 24h after transduction. Sox6-transduced cells stop growing 3 days after transduction and the culture die within day6. Error bars refer to 2 independent experiments. C). FACS analysis showing the GFP positivity (FL1 channel) for untransduced, EV, Sox6 and COUP-TFII transduced cells, 72h after transduction.

Like in original K562 cells, in Sox6 overexpressing V-K562 cells, differentiation is associated with reduced proliferation, leading the culture to exhaustion within 6 days after transduction. According to the phenotypic changes observed in K562 cells, Sox6 overexpression is accompanied by an accelerated kinetics of maturation and by the increased number of cells that achieve enucleation. V-K562-Sox6 transduced cells show a red pellet upon centrifugation, indicating an accumulation of haemoglobin (Fig.3.2, panel A). In contrast, COUP-TFII overexpression does not induce any phenotypic changes or differentiation upon transduction. To better characterize Sox6 and Coup-TFII transduced cells, I performed a quantitative RT-PCR analysis 72 h after transduction, comparing the expression of a series of genes known to vary during erythroid maturation between EV-K562 cells and Sox6 / Coup-TFII transduced cells. The result of this analysis is summarized in Fig 3.2b. As expected on the basis of the observed increased hemoglobinization, genes encoding the heme biosynthesis pathways show a strong increase only upon Sox6 overexpression, whereas less significant changes in erythroid markers are observed upon COUP-TFII overexpression despite an increase in GpA (Fig 3.2).

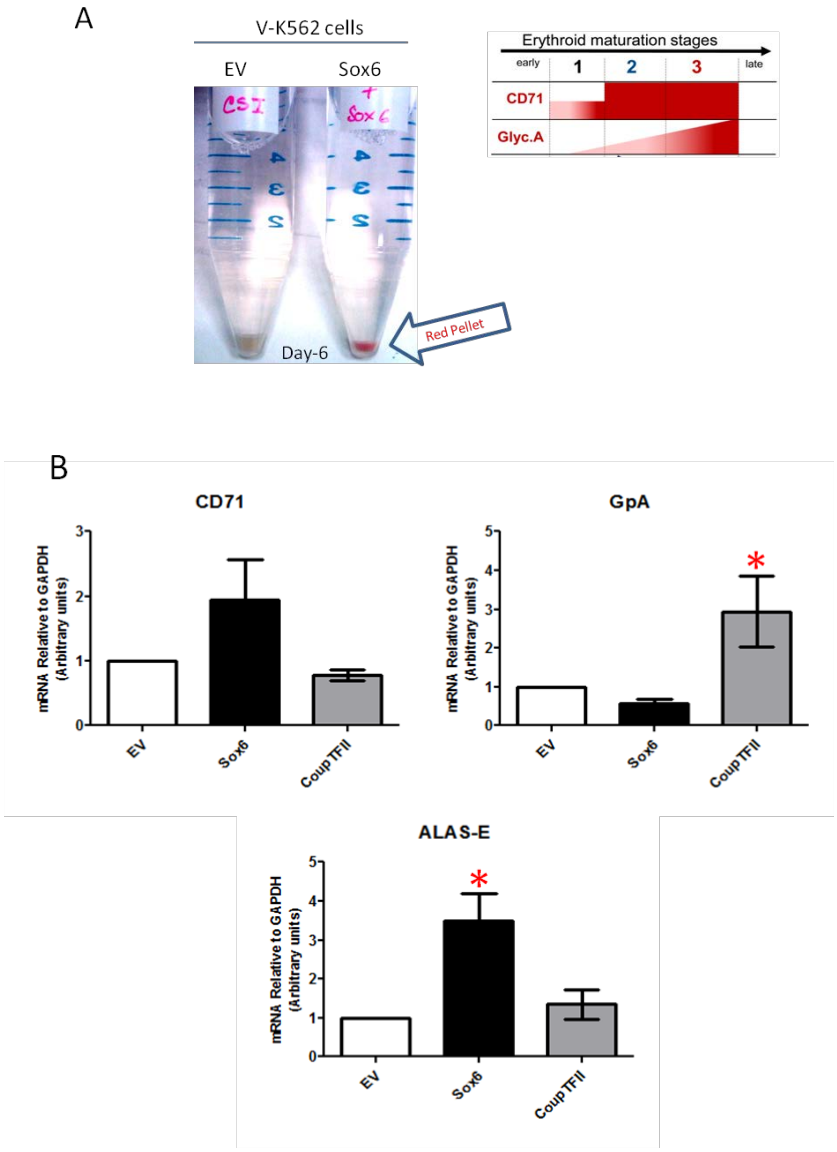


Figure 3.2: A). Phenotypic changes of V-K562 cells upon Sox6 overexpression. Sox6 transduced cells (right) show a red pellet after 6 days of transduction compared with the empty vector (EV) transduced cells (left). **B).** Quantitative RT-PCR analysis of erythroid markers in cells overexpressing Sox6 and COUP-TFII with respect to the empty vector. Histograms shows the relative levels of expression (mean ± SEM of at least 4 independent experiments) normalized with GAPDH considered as 1. Asterisk indicates significant P-value < 0.05.

Since both Sox6 and COUP-TFII have been proposed as candidate repressors of embryonic globins, I carefully analyzed by quantitative RT-PCR, the relative changes of globins genes transcription upon Sox6 and COUP-TFII overexpression considering GAPDH as an internal housekeeping standard. As shown in Figure.3.3, I see an increase in the expression level of all globins (α, ε, β, γ) upon Sox6 overexpression due to the general differentiation effect induced by Sox6 (Cantú C *et al.*, 2011). When setting the total amount of γ+β =1, the β/(γ+β) ratio increases from 0.2 in EV-V-K562 to 0.7 in Sox-6-V-K562, reflecting the larger increase (8.5x) of β-globin expression, with respect to the smaller (3.2x) increase in γ-globin expression (Fig 3.3).

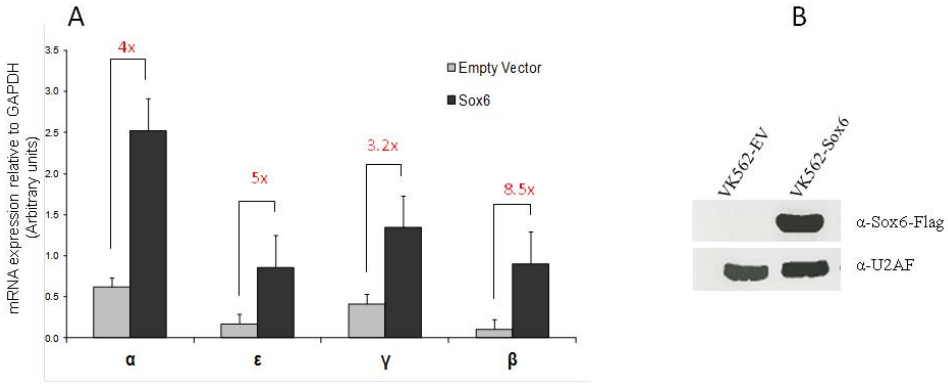


Figure 3.3: A). qRT-PCR analysis of globins on cDNAs from EV-VK562 and Sox6-VK562 cells. Sox6 increases the level of all globins, specifically β-globin. Histograms shows the relative level of expression to GAPDH (n=4, P<0.05) **B).** Western Blot analysis confirming the lentiviral-mediated Sox6 protein expression by using an anti-Flag antibody. Anti U2AF antibody is used as a loading control.

On the other hand, overexpression of COUP-TFII in V-K562 cells, induce an increase in relative levels of ε- and γ-globin levels compared to their levels in EV-K562 (Fig 3.4a). When setting γ+β =1, the β/(γ+β) ratio decreased from 0.2 (EV-VK562) to 0.08 (COUP-TFII), whereas there is little or no change in the expression of β-globin expression and the levels of α-globins are reduced by approximately 40% with respect to the EV-VK562 expression. This suggests that overexpression of COUP-TFII specifically increases γ-globin to substantial levels, as opposed to its originally proposed function of acting as a repressor of γ-globin. The observed specific increase in γ-globin expression suggests that this effect is not due to general erythroid differentiation induced upon COUP-TFII overexpression but it rather represents a specific effect on the γ-globin gene regulation.

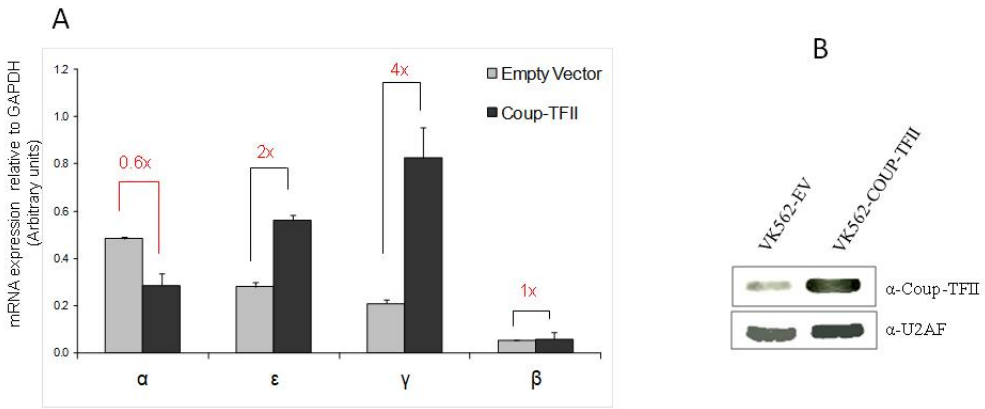


Figure 3.4: qRT-PCR analysis of globins from cDNAs of EV-VK562 and COUP-TFII-VK562 cells. **A).** COUP-TFII overexpression increases the level of ε- and more specifically of γ-globins expression. Histogram shows the level of expression relative to GAPDH (n=4, P<0.05) **B).** Western Blot analysis confirming the

lentiviral mediated COUP-TFII protein expression by using an anti-COUP-TFII antibody. Anti U2AF is used as a loading control.

Overall, these data show that whereas Sox6 overexpression induces a marked increase in β-globin transcription, COUP-TFII overexpression increases γ-globin. This opposite role now arose the question whether Sox6 and COUP-TFII could contribute to the γ- to β-globin switching. To test this hypothesis, I decided to co-transduce both Sox6 and COUP-TFII in V-K562 cells and to analyze the results of their simultaneous overexpression.

Co-transduction of V-K562 cells with Sox6 and COUP-TFII

To independently assess the level of transduction of Sox6 and COUP-TFII in co-transduction experiments, I cloned the Sox6 cDNA upstream to the IRES- ΔNGFR cassette instead of the IRES-Emerald GFP cassette (described more in detail in Materials & methods section). ΔNGFR can be detected through staining of the transduced cells by using a Phycoerythrin (PE) anti ΔNGFR antibody (visualized in flow cytometry analysis in the red FL2 channel). The corresponding Empty Vector (EV) containing the ΔNGFR cassette was used as a control.

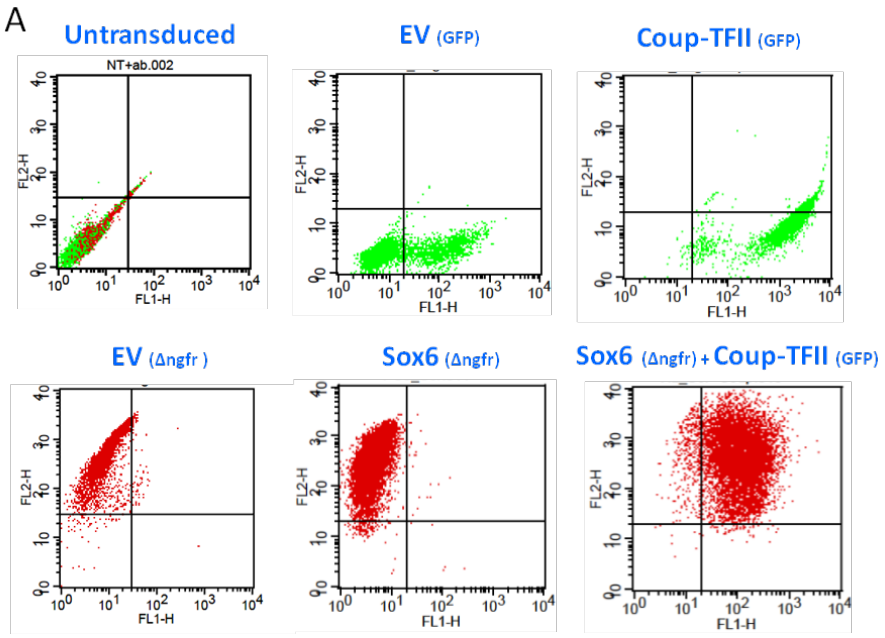
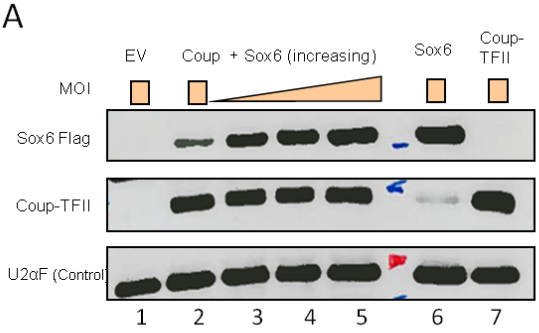


Figure 3.5: FACS plots showing the GFP positivity (FL1 green channel) for EV_{GFP} and $COUP\text{-}TFII_{GFP}$ transduced cells (single transduction) and PE positivity (FL2 red channel) for $EV_{\Delta NGFR}$ and $Sox6_{\Delta NGFR}$ transduced cells (single transduction) and for simultaneous $Sox6_{\Delta NGFR}$ and $COUP\text{-}TFII_{GFP}$ expression in double transduction.

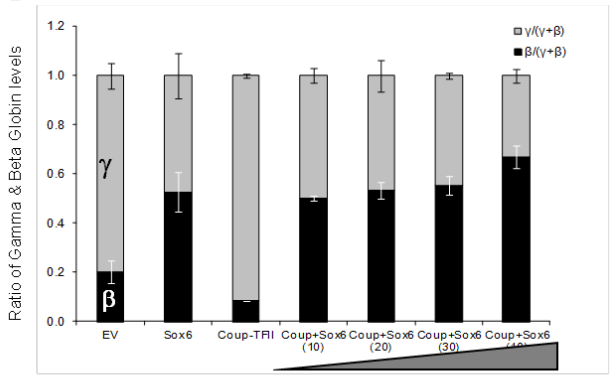
V-K562 cells were transduced with equal Multiplicity Of Infection (MOI) of both $Sox6_{\Delta NGFR}$ and $COUP\text{-}TFII_{GFP}$ lentiviruses and also at different relative concentration, in order to better understand how they control globins expression. To this end, I fixed the level of one transcription factor and increased the concentration of the other transcription factor by increasing the MOI of the corresponding expression vector. In a first series of experiments, the amount of COUP-TFII was fixed and Sox6 was increased correspondingly with the respective individual controls and the Empty vectors (fig 3.6). The efficiency of transduction was monitored after 72hrs by flow

cytometry and the expression of exogenous Sox6 and COUP-TFII were tested by Western blot (fig.3.6a). Flow cytometry analysis revealed more than 90% positivity for both single (GFP / ΔNGFR) and double positive ($Sox6_{\Delta NGFR}$ + $COUP\text{-}TFII_{GFP}$) transduced cells, along with the individual EV controls (fig 3.5).

A



B



C

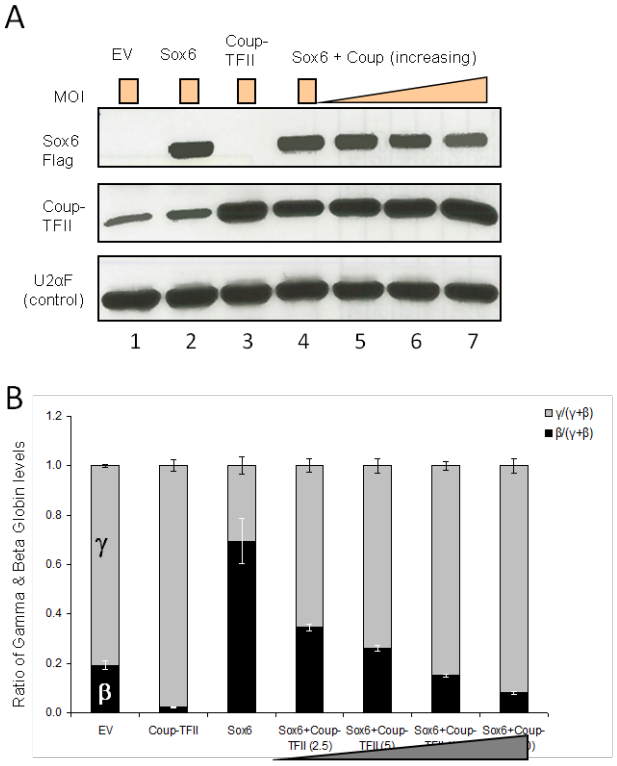
Globins	β/(γ+β)
EV	0.20
Sox6	0.53
Coup-TFII	0.08
Coup+S6(10)	0.50
Coup+S6(20)	0.53
Coup+S6(30)	0.55
Coup+S6(40)	0.67

Figure 3.6: Co-transduction of V-K562 cells with Sox6 and COUP-TFII overexpression **A).** Western blot analysis of Sox6 exogenous (anti-flag antibody) and COUP-TFII exogenous (anti-COUP-TFII antibody). U2AF levels are used as a loading control. Lanes 2-5 indicates the co-expression of COUP-TFII and Sox6 together with fixed COUP-TFII levels and increased Sox6. **B).** qRT-PCR to show the increase in the levels of β to γ ratio upon the Sox6 increase and fixed COUP-TFII. Black bars in the histogram represent the ratio of β over (γ+β) and the grey bars indicate the ratio of γ over (γ+β), being γ+β set equal to 1, normalized against GAPDH. **C).** Ratio of β over (γ+β) values is summarized in the table (n=2). The increasing MOIs are indicated within brackets in panel C.

qRT-PCR analysis on the double infected cells show that, upon fixed COUP-TFII and increasing Sox6 concentration, there is a concomitant increase in the level of β- globin expression (fig 3.6b). The β/(γ+β) ratio increased from 0.2 (EV-K562) to 0.53 (Sox6), with γ+β set equal to 1. In the presence of COUP-TFII + Sox6, the ratio of β/(γ+β) increases from 0.50 to a maximum of 0.67 in the presence of the highest Sox6 concentration (fig3.6c).

In a specular second set of experiments, the amount of Sox6 was fixed and COUP-TFII was increased correspondingly with the respective individual controls and the Empty vector. The efficiency of transduction was again monitored after 72hrs by flow cytometry and the expression of exogenous Sox6 and COUP-TFII were tested by Western blot (fig.3.7a). Flow cytometry analysis revealed more than 90% positivity for both single (GFP / ΔNGFR) and double positive ($Sox6_{\Delta NGFR}$ + $COUP\text{-}TFII_{GFP}$) cells along with the individual EV controls. qRT-PCR analysis on the double infected cells show that upon fixed Sox6 concentration and increasing COUP-TFII

concentration, there was a progressive dose-dependent increase in the proportion of γ-globin expression (fig 3.7b).



C

Globins	β/(γ+β)
EV	0.20
Sox6	0.70
Coup-TFII	0.02
S6+Coup(2.5)	0.34
S6+Coup(5)	0.26
S6+Coup(10)	0.15
S6+Coup(20)	0.08

Figure 3.7: Co-transduction of V-K562 cells with Sox6 and COUP-TFII overexpression **A).** Western blot analysis of Sox6 exogenous (anti-flag antibody) and COUP-TFII exogenous (anti-COUP-TFII antibody). U2AF levels are used as a loading control. Lanes 4-7 indicates the co-expression of COUP-TFII and Sox6

together with fixed Sox6 levels and increased COUP-TFII. **B).** qRT-PCR to show the decrease in the levels of β to γ ratio upon the COUP-TFII increase and fixed Sox6. Black bars in the histogram represent the ratio of β over (γ+β) and the grey bars indicate the ratio of γ over (γ+β) set to 1 normalized against GAPDH. **C).** Ratio of β over (γ+β) values is summarized in the table (n=2). The increasing MOIs are indicated within brackets in panel C.

The β/(γ+β) ratio decreased from 0.2 (EV-K562) to 0.02 (COUP-TFII) (with γ+β set to 1) and in the presence of COUP-TFII + Sox6 it decreases from 0.34 to 0.08 (fig3.7c), indicating that COUP-TFII promotes the increase in the proportion of γ-globin. The β/(γ+β) ratio decreased from 0.70 (Sox6 alone) to 0.08, when cells were co-transduced with the highest amount of COUP-TFII.

These data confirm that the relative levels of Sox6 and COUP-TFII are important modulators of the γ/β expression ratio, suggesting that changes in their level could contribute to tip the balance towards β- expression during the switching. According to this hypothesis, we should expect to see a variation in the relative expression of these two transcription factors around the switching time. We are currently addressing this question by analysing mouse fetal liver cells during the embryonic day E11.5-E13.5 when the switching takes place. Preliminary results confirm that COUP-TFII expression decreases from day E11.5-E13.5 while Sox6 expression increases gradually (data not shown, personal communication from Fugazza C).

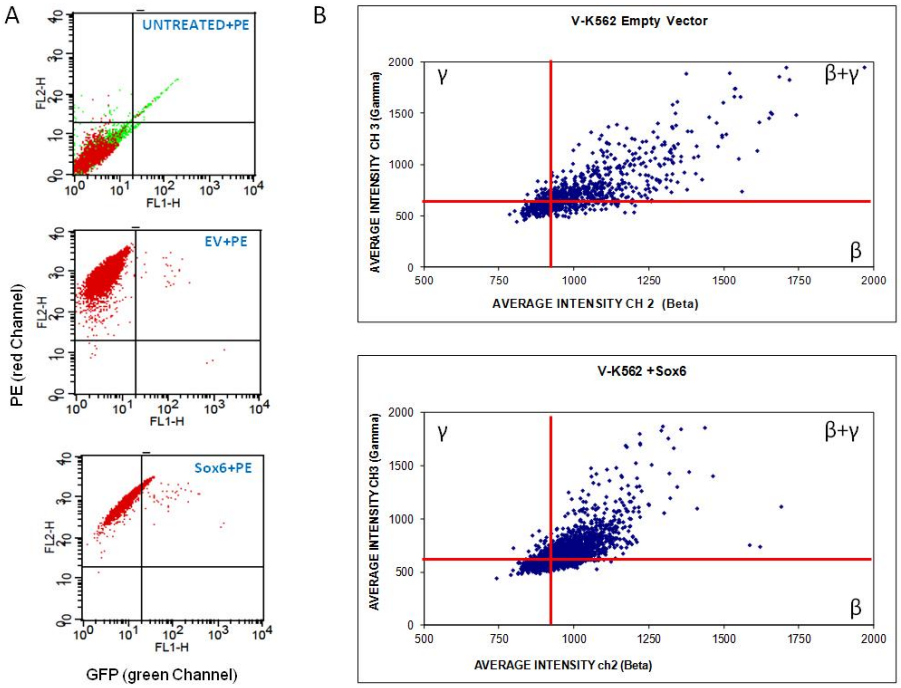
High Content Screening (HCS) assays to check for the globin expression in cells overexpressing Coup-TFII and Sox6

Fluorescence-based "high-content screening" (HCS) assays usually employ fluorescent dyes, antibodies, or fluorescent proteins to detect molecular markers or to report complex cellular processes. The greater the number of independent markers that can be evaluated simultaneously, the higher the probability of identifying phenotypically distinct subpopulations (Gasparri F. *et al.,* 2014).

In order to analyze the percentage of globins expressed at a single cell level, V-K562 cells were transduced with equal Multiplicity of Infection (MOI-30) of $EV_{\Delta NGFR}$, and $Sox6_{\Delta NGFR}$ lentiviruses. The purpose of using non -GFP vectors is to avoid cross talking and background signals in interfering with the conjugated fluorochrome antibodies during staining. The efficiency of transduction was above 95% of positive cells for all the vectors (Fig.3.8a). The transduced cells were fixed and labelled with fluorochrome conjugated β- globin (visualized in the green channel -Ch2) and γ-globin antibodies (visualized in the red channel -Ch3). High content analysis experiments were performed in 96-well microplates, by seeding cells at a density of 8000 / well in black clear-bottom plates.

Figure 3.8 panel-d shows the percentage of cells expressing β- and γ-globins. The total percentage of cells expressing β-globin in EV transduced cells is 89.3% (30.4% β-alone + 58.9% β+γ double expressing cells) and in Sox6 transduced cells is 84.2% (48.8% β-

alone and 35.4% β+γ double expressing cells) suggesting that upon Sox6 overexpression, double positive cells (β+γ) are completely switched to β-globin expression as indicated by the increase from 30.4% (EV) to 48.8% (Sox6) of cells single-positive for β-globin. These data are in line with what I observed in qRT-PCR analysis, further confirming that Sox6 overexpression induces a marked increase in β-globin transcription, as assayed by two independent methods.



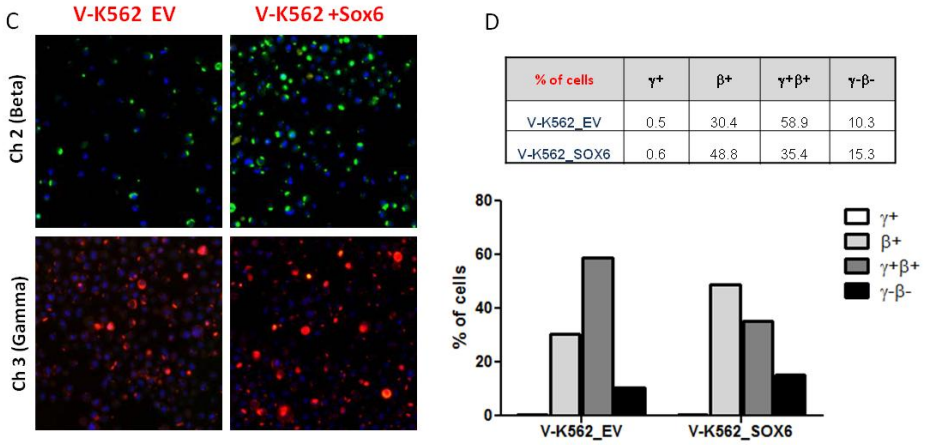


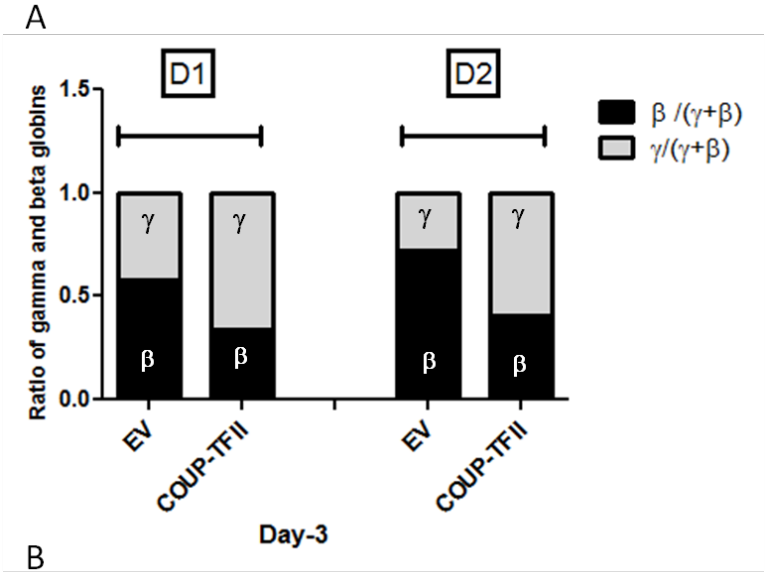
Figure 3.8: Transduction of V-K562 cells for High content Analysis (HCA) assay. **A).** FACS plot showing PE positivity (FL2- channel) for untreated, Empty vector ($EV_{\Delta NGFR}$), and $Sox6_{\Delta NGFR}$ transduced cells. More than 95% of cells are positive for PE (y-axis). **B).** Two-Dimensional dot plot images showing the percentage of γ- and β-globins. Nuclear average fluorescence intensity values (Average intensity Ch2) associated with β-globin and (Average Intensity Ch3) associated with γ-globin were acquired and reported. Two arbitrary thresholds (red line) were applied to determine the negative and positive stained cells. **C).** Representative ArrayScan images of transduced V-K562 cells (EV and Sox6) stained with DAPI (blue), anti- beta globin (green, Ch2), and anti-gamma globin (red, Ch3) antibodies. Magnification: 20μm **D).** Table shows the percentage of cells in each sector (γ^+, β^+ $\gamma^+\beta^+$ and $\gamma^-\beta^-$) calculated and reported within the graph (below) n=1.

COUP-TFII overexpression in cultured Human erythroid progenitor cells

Transduction of COUP-TFII in V-K562 cells showed promising results in increasing the levels of γ-globin expression. We then decided to test whether COUP-TFII can increase γ-globin expression in human erythroid cultures from $CD34^+$ cells purified from peripheral

blood (PB). These experiments are ongoing in a collaborative project with Prof. Paolo Moi and Dr. Giusi Marini in Cagliari, Italy. These adult erythroid cultures from PB purified from $CD34^+$ cells predominantly express adult hemoglobin (β-globin). BFUe cells from cultured Peripheral blood from two healthy donors (D1 and D2) were transduced with empty vector (EV-GFP) and COUP-TFII-GFP lentiviral vector on the day-1 of their differentiation phase with a multiplicity of infection (MOI) equal to 30. 72hrs post transduction, cells were collected for FACS analysis and for cDNA synthesis to check for the expression of globins. Flow cytometry analysis show 30% and 42% of GFP positivity for COUP-TFII transduced cells (D1 and D2) and 42% and 57% of GFP positivity for the corresponding empty vectors (EV-GFP) in D1 and D2. Preliminary data on globin expression levels analyzed by considering $\gamma+\beta = 1$ (fig3.9a), show that upon COUP-TFII infection, the $\beta/(\gamma+\beta)$ ratio decreased from 0.58 (EV-K562) to 0.34 (COUP-TFII) in (D1). In the second donor (D2), (42% GFP+ cells), the $\beta/(\gamma+\beta)$ ratio decreases from 0.72 (EV-K562) to 0.41 (COUP-TFII) indicating that even with a low percentage of COUP-TFII transduced cells, COUP-TFII promotes increased levels of γ-globin expression. These preliminary results encouraged us to set up an experiment, where transduced cells will be purified by immuno-magnetic selection in order to confirm this preliminary analysis. These experiments are currently ongoing.

A



B

β/(γ+β)	Donor 1	Donor 2
EV	0.58	0.72
COUP-TFII	0.34	0.41

Figure 3.9: A). qRT-PCR analysis on two healthy donors (D1 and D2) show the decrease in the levels of β to γ ratio upon COUP-TFII transduction. Black bars in the histogram represent the ratio of β over (γ+β) and the grey bars indicate the ratio of γ over (γ+β) set to 1 normalized against GAPDH. **C).** Ratio of β over (γ+β) values is summarized in the table (n=2).

Sox6 and COUP-TFII bind to the gamma globin promoter *in vitro* and *in vivo*

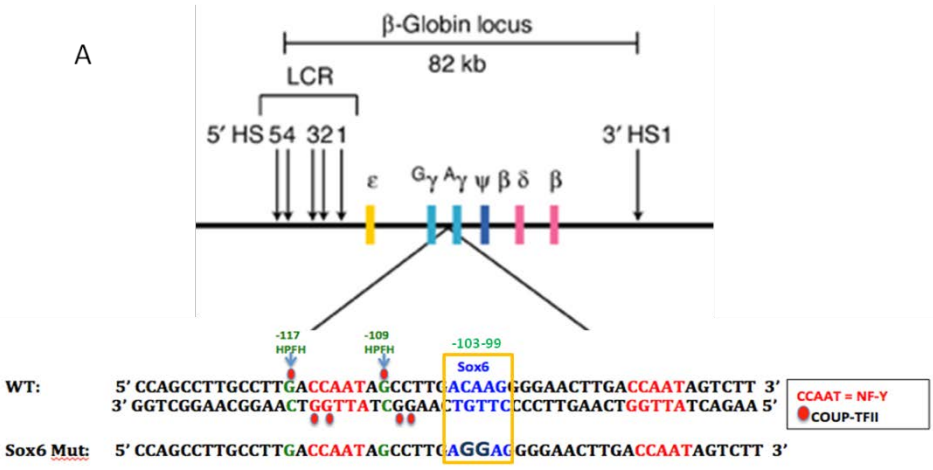
Sox6 has been shown to repress the εy- globin gene in mouse by direct binding on its promoter (Cohen-Barak *et al.*. 2007, Yi *et al.*. 2006) and it has also been demonstrated in this lab that Sox6 binds to human ε-gene and γ-globin gene promoters in ChIP experiments (Cantú C *et*

al., 2011a). However, the precise location at which the Sox6 consensus binds on the γ-globin promoter had to be elucidated. We used the sequence (A/T) (A/T)CAA(A/T)G Sox consensus sequence to search for putative Sox6 sites within the γ-globin promoter, taking advantage of the UCSC database (http://genome.ucsc.edu/). This approach revealed the presence of conserved Sox6 site at position -99 to -103 sites upstream to the transcription start site (TSS) of the γ-globin promoter (fig 3.10a). This region is crucial for γ-globin activity as it lies adjacent to NFY and COUP-TFII binding sites between the proximal and distal CCAAT box on the γ-globin promoter. Moreover, several HPFH are mapped in this region.

We set up an Electrophoretic Mobility Shift Assay (EMSA) experiment using as a probe the oligonucleotide encompassing the Wild type γ- globin double CCAAT box region and the corresponding oligo mutated in the putative Sox6 binding site to confirm that Sox6 is indeed able to bind to the identified consensus. As shown in fig. 3.10b the WT probe, when incubated with nuclear extracts from V-K562 overexpressing Sox6, gives a Sox6 band that is specifically supershifted by the anti-FLAG antibody (lane 2), whereas no detection of Sox6 is found on the mutant probe (lane 3 and 4), confirming that Sox6 binds to the identified consensus sequence *in vitro*.

To assess the regulatory relevance of the above identified consensus, we prepared a set of luciferase reporter constructs that were assayed by transfection assays in V-K562 cells. The 330bp region corresponding to the minimal γ-globin promoter sequence and containing the WT and mutated Sox6 binding site, was cloned upstream to the luciferase reporter gene in the pGL2 vectors to give

the γ-prom WT and γ-prom Mut (fig 3.10a) constructs. These constructs were transfected in V-K562 cells and assayed for luciferase activity 48hrs after transfection. The empty Vector (pGL2) alone served as a negative control (data not shown). As shown in fig 3.10c, the mutant form of γ- promoter shows an increased luciferase activity compared to the wild type (WT) at a different concentration (20, 50 and 100ng) of the reporter, suggesting that mutation of the Sox6 binding site abolishes the binding of Sox6, thereby increasing γ-globin promoter activity.



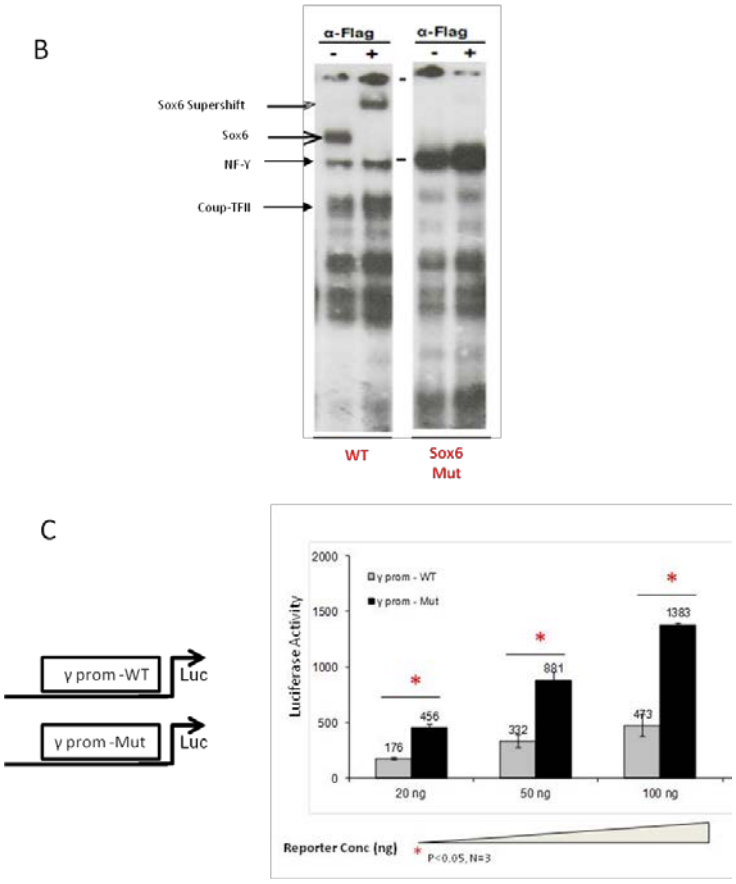


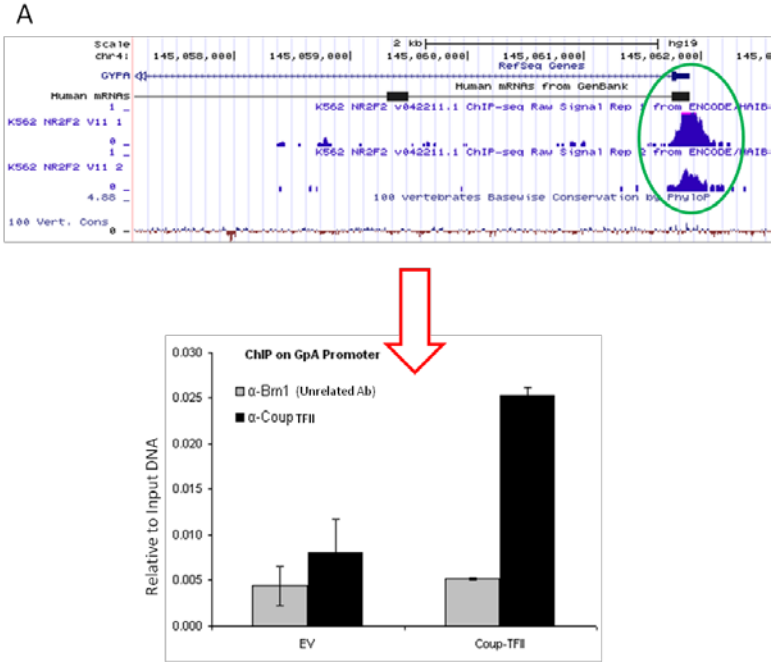
Figure 3.10: A). Schematic representation of the β-globin locus focussing on the -99-103 Sox6 binding site on the minimal γ-globin promoter and the mutant form of the γ-globin promoter (-102 and -101) denoted by big bold letters. B**).** EMSA experiment (data from Monclus IF) of Sox6 is supershifted by anti-Flag antibody on wildtype probe (lane 1 & 2). The probe mutated for Sox6 binding failed to give any band when tested in the same conditions. C**).** Transfection experiments in V-K562 cells. The constructs used are schematically represented in the left panel and the luciferase activities on the γ-globin promoter –WT & Mutant are shown on the right panel. The mutant γ-promoter always shows 2.5 fold (approx) higher activity when compared with the wild type construct. The luciferase activity is given in arbitrary units and the asterisk represents significant p-value <0.05 of 3 different experiments.

COUP-TFII and Sox6 share adjacent binding sites on the γ-globin promoter and might thus functionally interact with each other. Old data in literature (Liberati C *et al.* 1998, Filipe *et al.*1999) showed that COUP-TFII binds *in vitro* to the γ-globin CCAAT box promoter region, but no evidence was provided about its ability to bind *in vivo.*

In order to confirm the ability of COUP-TFII to bind to the γ-globin promoter *in vivo,* I setup a ChIP experiment in V-K562 cells. As a first step, I used a bioinformatic approach to fish out targets for COUP-TFII binding using the available ENCODE data on K562 cells. In order to search for the positive control target sequences, with the help of UCSC browser, I analyzed the two dataset available from ENCODE database on COUP-TFII in K562 cells. Basically, I took the list of regions showing enrichment for COUP-TFII (positive in both replicates) and filtered for promoter regions 1.5 kb upstream and downstream from transcription start site (TSS). Then the list of positive promoters showing the presence of COUP-TFII was annotated to their respective families of protein domains. Glycophorin A (GpA) promoter also showed a strong peak for COUP-TFII binding and was evolutionarily conserved within mammals. This site was used as a positive control in my ChIP experiment (fig 3.11b). With respect to COUP-TFII binding to the γ- globin promoter, ENCODE data show an enrichment of COUP-TFII binding in. To confirm this, I performed a Chromatin Immunoprecipitation experiment to ascertain whether COUP-TFII is binding on the γ- globin promoter on V-K562 cells. V-K562 cells overexpressing COUP-TFII show a positive enrichment for binding on the γ-globin promoter and GpA (as positive control) in comparison with the cells transduced with Empty vector (fig3.11a).

However, the same sequence is not immunoprecipitated by the same antibody when chromatin from V-K562 cells transduced with the empty vector is used, suggesting that the low physiological levels of COUP-TFII in V-K562 cells are not sufficient for an efficient binding. As a negative control, the unrelated anti-brn1 antibody fails to immunoprecipitate the γ-globin promoter region in both COUP-TFII V-K562 and empty vector- V-K562 chromatin samples.

A



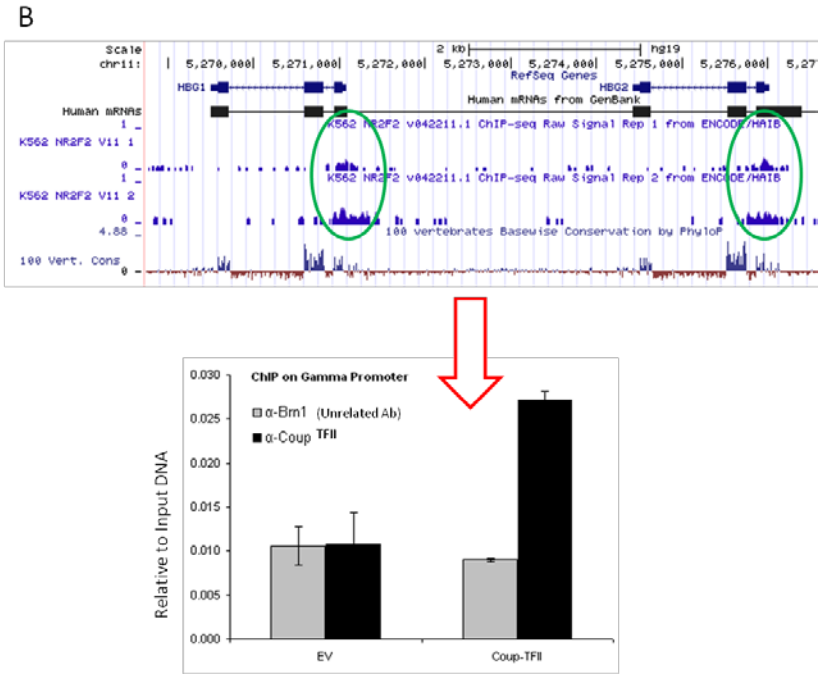


Figure 3.11: Chromatin Immunoprecipitation of COUP-TFII (NR2F2) on V-K562 cells. **A).** ChIP analysis on the GpA promoter region used as a positive control for COUP-TFII binding. Upper panel: Enrichment of COUP-TFII on the GpA promoter from the ENCODE data on K562 cells. Lower panel: ChIP of COUP-TFII on GpA promoter in V-K562 cells. **B).** ChIP analysis confirms the ability of COUP-TFII to bind on the γ-globin promoter. Upper panel: Enrichment of COUP-TFII on γ-globin promoter analyzed by ChIP-Seq ENCODE data on K562 cells. The minimal γ-globin promoter regions (Aγ and Gγ) are highlighted with circles. Lower panel: ChIP of COUP-TFII to the γ-promoter in V-K562 cells. α-brn1 serves as the unrelated antibody for negative control. Results were normalized relative to the Input DNA (from 10^5cells) levels (n=2).

Together, these data confirm that Sox6 and COUP-TFII is indeed able to bind both *in vitro* and *in vivo* to the γ-globin promoter in V-K562 cells.

COUP-TFII binds to DNaseI HS-3 and HS-4 within the LCR region on V-K562 cells

The locus control region (LCR) in human β-like globin genes consists of five DNase I hypersensitive sites (HSs 1-5) that play a critical role in the regulation of developmental stage-specific β-like genes expression. In particular, HS-2 functions as a powerful general erythroid-specific enhancer, whereas HS-3 is related to the regulation of ε-globin and γ-globin gene expression during embryonic period, and HS-4 play a major role in the expression of β-globin gene (Fraser P *et al.,* 1993). To obtain insight into whether COUP-TFII is also recruited in the chromatin regions of LCR, I checked for the enriched locations on the LCR through the ChIP-Seq ENCODE data. The results indicate strong association of COUP-TFII with the upstream HS-3 and HS-4 within the LCR region of the β-globin cluster in V-K562 cells (Fig 3.12a). ChIP in V-K562 cells overexpressing COUP-TFII confirmed the binding of COUP-TFII to the HS-3 and HS-4 sites (fig 3.11b) as compared with EV-K562, cells suggesting that this protein might function in a chromatin-associated multiprotein complex within the β-globin locus. At low physiological levels, COUP-TFII in V-K562 cells is not sufficient for an efficient binding. The presence of COUP-TFII on HS-3 and HS-4 site indicates that it might play a role in hemoglobin switching and is consistent with a potential role of HS-3 in mediating γ-globin transcription (Navas *et al.*, 1998; Fang *et al.,* 2005).

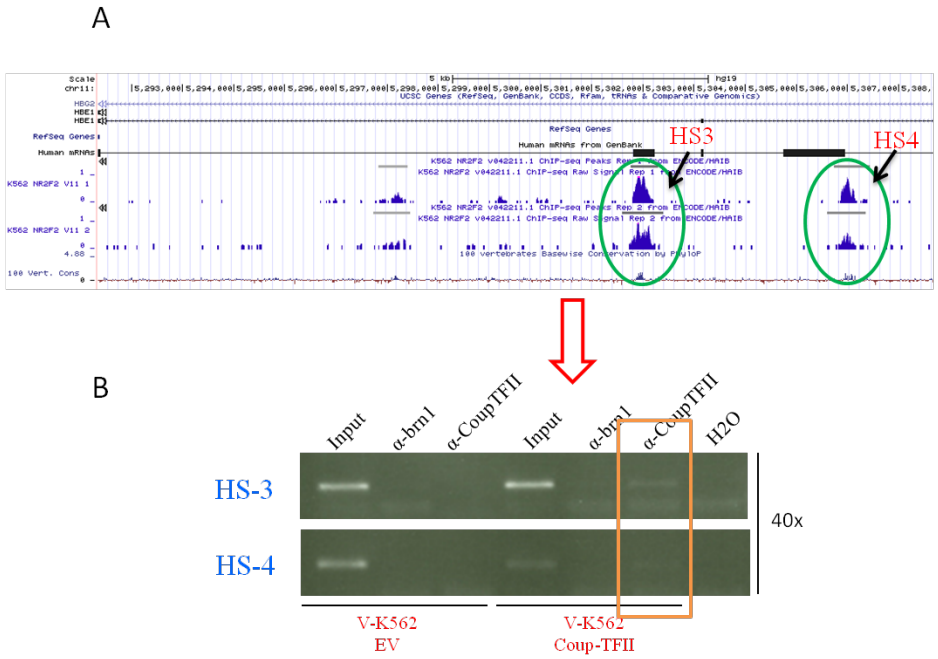


Figure 3.12: Chromatin Immunoprecipitation of COUP-TFII (NR2F2) in V-K562 cells. **A).** ChIP analysis confirms the ability of COUP-TFII to bind to HS-3 and HS-4 regions within the LCR, as seen from ENCODE data **B).** V-K562 cells transduced with COUP-TFII show enrichment in HS-3 and HS-4 relative to the Input DNA levels (10^5cells) and no enrichment seen in α-brn1, used as negative control. PCR was repeated twice to confirm the binding of COUP-TFII within the HS's regions of the LCR.

Mouse fetal liver embryonic day E13.5 shows the presence of Sox6 but not COUP-TFII binding *in vivo* on the embryonic εy-globin promoter

In mouse, the switch from embryonic to adult globins occurs between E11.5 and E13.5. Sox6 and COUP-TFII are co-expressed in mouse

during these time interval. To check for the binding of Sox6 and COUP-TFII on the εy- globin promoter I performed a ChIP on mouse fetal liver E13.5 cells. It has been shown that εy-globin promoter is a direct target for Sox6 whereas for COUP-TFII it has not been demonstrated before. NFE2 and GpA promoters, derived from the ChIP seq ENCODE data were again used as a positive control for COUP-TFII binding. Figure 3.13 shows that only Sox6 but not COUP-TFII binds to the εy- globin promoter at embryonic day E13.5, where the switch from embryonic to adult has already taken place (fig 3.13).

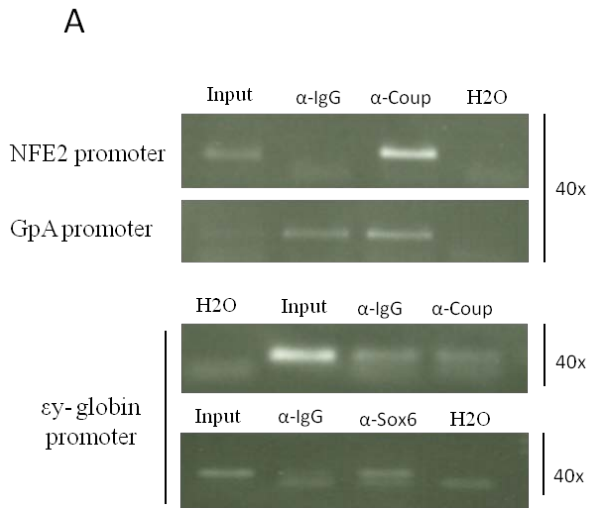


Figure 3.13: A). Chromatin Immunoprecipitation on mouse fetal liver cells at E13.5. Upper panel shows the enrichment of NFE2 and GpA promoters used as a positive control for COUP-TFII binding. Lower panel shows an enrichment for Sox6 binding on the εy-promoter, but not for COUP-TFII.

Sox6 overexpression activates COUP-TFII transcription

While performing overexpression experiments described in this chapter, we noticed that in Western blot experiments (Fig 3.6, lane 6) Sox6 overexpression in V-K562 cells is always accompanied by a mild activation of COUP-TFII expression at both the transcription and at the protein levels (fig 3.14) suggesting that Sox6 activates COUP-TFII upon overexpression in V-K562 cells. This observation raises the possibility that Sox6 activates COUP-TFII despite their opposite roles on γ-globin regulation, possibly suggesting a negative feedback regulatory loop. Therefore, I did a Chromatin Immunoprecipitation experiment on V-K562 cells to check whether Sox6 can directly bind on the COUP-TFII promoter regions and thus activate COUP-TFII expression. Through bioinformatic approaches, I mapped a potential double Sox6 binding site on the COUP-TFII promoter approx. 250 bp upstream to the transcription start site. However, this site does not show any enrichment in Sox6 binding, even upon overexpression in V-K562 cells excluding the involvement of the mapped Sox6 consensus in this effect.

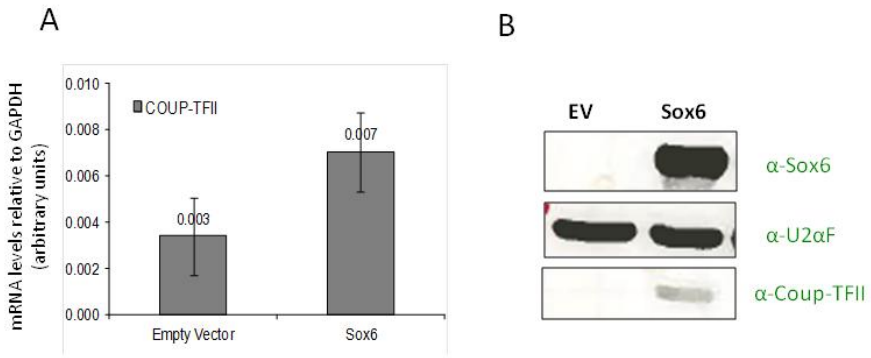


Figure 3.14: Sox6 activates COUP-TFII transcription when overexpressed in V-K562 cells. **A).** qRT-PCR shows an increase in the endogenous levels of COUP-TFII expression normalized against GAPDH. **B).** Western blot analysis of Sox6 overexpression in V-K562 transduced cells shows a mild increase in the levels of COUP-TFII as detected by α-COUP-TFII antibody. U2AF levels are used to check for protein loading.

DISCUSSION

In this chapter, I have described the outcome of Sox6 and COUP-TFII overexpression on erythroid differentiation and in the modulation of the ratio between γ- and β-globin expressions. The use of the variant K562 erythroleukemic cell line V-K562 expressing a substantial amount of β- globin helped us to decipher more in detail globins gene regulation upon overexpression of two different transcription factors: Sox6 and COUPT-TFII. Sox6 overexpression in V-K562 cells enhances and accelerates the terminal maturation of this culture, as clearly shown by differential cell counting and by sharp increase in the expression levels of erythroid markers CD71 and Glycophorin A (fig 3.2). On the contrary, overexpression of COUP-TFII in these cells does not induce any profound effects on the expression levels of the erythroid markers, excluding its direct involvement in erythroid differentiation of V-K562 cells, despite a marginal increase in GpA (Fig 3.2).

Transduction of V-K562 cells with Sox6 and COUP-TFII transcription factors influences γ- to β-globin ratio

With respect to the levels of globins expression, overexpression of Sox6 induces a general increase in all globins with a specific increase in the β-globin expression when compared to γ-globin increase (8.5x vs 3.2x, fig 3.3a). On the other hand, upon COUP-TFII overexpression, there is a specific increase only in fetal globins (ε and

γ) expression, whereas β-globin expression is unchanged (fig 3.4a). This result is apparently in contrast with previous reports proposing COUP-TFII as a repressor of γ-globin expression on the basis of *in vitro* experiments. This discrepancy is probably related to the different experimental settings used. In fact, co-transduction in V-K562 cells of Sox6 and COUP-TFII overexpressing vectors at various concentrations by keeping one factor constant and increasing the other and vice-versa, (fig 3.6 and 3.7) show that Sox6 and COUP-TFII have opposite effect on the β /(γ+β) ratio. This suggests that their differential expression could able to influence the switch between β and γ- globin expression. Precisely, the increase in the level of Sox6 would favour the increase in β-globin expression whereas increase in the levels of COUP-TFII would favour the increase in γ-globin expression. The profound changes in the levels of globins were phenotypically confirmed by immunostaining with fluorescent conjugated anti-globins antibodies visualized by a high end fluorescent imaging techniques at the single cell level (fig 3.8).

Forced COUP-TFII expression in erythroid cultures from human peripheral blood cells enhances fetal hemoglobin synthesis

In this chapter, I have also shown preliminary results (fig 3.9) showing that forced expression of COUP-TFII in erythroid cultures from $CD34^+$ cells purified from human peripheral blood correlates with a higher expression of γ-globin. Interestingly, with only about

30% (D1) and 42% (D2) of COUP-TFII transduced cells (as assayed by FACS analysis), I observed an increase in γ- globin expression from 0.58 (EV-K562) to 0.34 (COUP-TFII) in Donor1 (D1) and from 0.72 (EV-K562) to 0.41 (COUP-TFII) in Donor2 (D2), suggesting that a better transduction efficiency would result in an even more significant activation of γ-globin expression.

Finally, to address the underlying mechanism of the γ-globin gene regulation upon Sox6 and COUP-TFII overexpression, we next checked whether Sox6 and COUP-TFII are bound *in vivo* to the γ-globin gene promoter region to their adjacent sites on the γ-globin promoter. Through EMSA, Luciferase assay, and Chromatin Immmunoprecipitation experiments (fig 3.10), I demonstrated that Sox6 and COUP-TFII bind on the γ-globin promoter and that Sox6 acts as a γ- globin repressor, as previously suggested by data in literature. ENCODE ChIP seq data for COUP-TFII on K562 cells gave us some precise binding locations of COUP-TFII within the LCR and the γ-globin promoter. In particular, although COUP-TFII was already known to bind strongly to the distal and proximal CCAAT box *in vitro* (Liberati C *et al.* 1998), I have now also confirmed the binding through ChIP experiments in V-K562 cells (fig 3.11b) and in mouse E13.5 fetal liver cells (fig 3.13), suggesting that both Sox6 and COUP-TFII binds to the γ-globin promoter *in vivo*. Moreover, we mapped for the first time the presence of a Sox6 binding consensus within the tandem CCAAT box region of the γ- globin promoter. This is of particular relevance because this site lies in a region crucial for γ-globin gene regulation.

COUP-TFII binds to the HS-3 and HS-4 region within the LCR of the β-globin cluster, that mediate γ-globin transcription

In figure 3.12, I show that COUP-TFII binds to the DNase I HS-3 and HS-4 region within the LCR suggesting that COUP-TFII might interact with other protein components of the switch complex to regulate the γ-globin expression. It has been known that HS-3 and HS-4 are important for ε- and γ- globin regulation (Fraser P *et al.*, 1993). The presence of BCL11A-XL, SOX6, and GATA1 at the HS-3 site within the LCR is consistent with a potential role of HS-3 in mediating γ-globin transcription (Navas *et al.*. 1998; Fang *et al.*. 2005), likely through rearrangement of the configuration of the β-globin cluster by long-range interactions.

However, (fig 3.14) in V-K562 cells, Sox6 overexpression partially activates COUP-TFII expression, which might be due to the general erythroid differentiation effects elicited by Sox6 or could also be interpreted as a negative feedback regulatory loop that maintains a balance between the expression of γ and β- globin genes thus ensuring a "gradual" switching. A detailed study has to be carried out in order to have a clear understanding at the molecular level and to rule out the possibility of this effect being an artificial consequence of the high level of Sox6 expression achieved.

Sox6 and COUP-TFII molecular mechanism of action

The regulation of expression of the proximal γ-globin promoter complex in fetal liver and adult erythroid cells has been under tremendous investigation for decades. In particular, since the majority of HPFH mutations maps in the double CCAAT box region, this element has been the subject of extensive investigation. The identification of a Sox6 binding site exactly between the two CCAAT boxes, prompted us to further characterize the transcription factors binding here. The CCAAT boxes are the most conserved motifs found in the globin promoters. Sox6 binds on the γ-globin promoter between the two CCAAT boxes, whereas COUP-TFII binds to the region of proximal and distal CCAAT box overlapping the binding of NF-Y. SOX6 contains an evolutionarily conserved HMG domain that is essential for its DNA-binding activity (Grosschedl *et al.*. 1994). HMG domain factors are capable of binding to the minor groove of the DNA and to induce drastic DNA bending upon binding. One hypothesis is that, in the presence of high amount of Sox6, there is no cross talk between the two CCAAT box regions displacing the strong interaction of NF-Y which is a transcriptional activator, resulting in the repression of the γ-globin gene. In the absence of Sox6 or at equilibrium, the two CCAAT box region located within 22bp distance (two turns of DNA) would serve as a high affinity site for the binding of NF-Y and therefore for the activation of γ-globin expression. Alternatively, In the presence of high COUP-TFII, COUP-TFII at the distal CCAAT box would displace the binding of Sox6 (strong repressor) and stabilizes the binding of NF-Y complex, resulting in the activation of γ-globin. It must also be considered that Sox6 can

cooperate to recruit BCL11A-XL (Xu J *et al.,* 2010), in turn able to recruit chromatin repressors (see chapter 4).

COUP-TFII is a orphan nuclear receptor with no known ligand. COUP-TFII has been shown to act as transcriptional activator in a variety of cell types where it has a crucial role during development (Pereira FA. *et al.,* 2011). COUP-TFII cooperates with NFY when bound to the 5' site (Liberati C *et al.,* 2001) to form a more stable complex thus providing greater transcriptional activity. However, at the 3' COUP-TFII site overlapping to the CCAAT box, NFY and COUP-TFII might compete for binding as it has been proved to occur on the ε CCAAT box region. Thus, even if endowed with positive transcriptional activity, COUP-TFII bound to the 3' site might fail to increase (or might even repress) the activity of the promoter by preventing the binding and, thus, the activity of the strong activator NFY. Our functional data point to a more complex role, where the relative balance between the levels of COUP-TFII and Sox6 can act as the tip of balance switching the equilibrium from γ- (high COUP-TFII) to β- (high Sox6) globin expression (see also final discussion and models in chapter-5).

REFERENCES

Cantú, C., Grande, V., Alborelli, I., Cassinelli, L., Cantú, I., Colzani, M.T., Ierardi, R., Ronzoni, L., Cappellini, M.D., Ferrari, G., Ottolenghi, S. and Ronchi, A. (2011) A highly conserved SOX6 double binding site mediates SOX6 gene downregulation in erythroid cells. Nucleic Acids Res, **39,** 486-501.

Cantú, C., Ierardi, R., Alborelli, I., Fugazza, C., Cassinelli, L., Piconese, S., Bose, F., Ottolenghi, S., Ferrari, G. and Ronchi, A. (2011) Sox6 enhances erythroid differentiation in human erythroid progenitors. Blood, **117,** 3669-79.

Cohen-Barak, O., Erickson, D.T., Badowski, M.S., Fuchs, D.A., Klassen, C.L., Harris, D.T. and Brilliant, M.H. (2007) Stem cell transplantation demonstrates that Sox6 represses epsilon y globin expression in definitive erythropoiesis of adult mice. *Exp Hematol*, **35,** 358-67.

Demaison, C., Parsley, K., Brouns, G., Scherr, M., Battmer, K., Kinnon, C., Grez, M. and Thrasher, A.J. (2002) High-level transduction and gene expression in hematopoietic repopulating cells using a human immunodeficiency [correction of imunodeficiency] virus type 1-based lentiviral vector containing an internal spleen focus forming virus promoter. Hum Gene Ther, **13,** 803-13.

Dumitriu, B., M. R. Patrick, J. P. Petschek, S. Cherukuri, U. Klingmuller, P. L. Fox, and V. Lefebvre. 2006. Sox6 cell-autonomously stimulates erythroid cell survival, proliferation, and terminal maturation and is thereby an important enhancer of definitive erythropoiesis during mouse development. Blood **108**:1198-1207.

Fang, X., Sun, J., Xiang, P., Yu, M., Navas, P.A., Peterson, K.R., Stamatoyannopoulos, G. and Li, Q. (2005) Synergistic and additive properties of the beta-globin locus control region (LCR) revealed by 5'HS3 deletion mutations: implication for LCR chromatin architecture. Mol Cell Biol, **25,** 7033-41.

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,** 687-97.

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,** 106-13.

Grosschedl, R., Giese, K. and Pagel, J. (1994) HMG domain proteins: architectural elements in the assembly of nucleoprotein structures. Trends Genet, **10,** 94-100.

Hamada-Kanazawa, M., Ishikawa, K., Nomoto, K., Uozumi, T., Kawai, Y., Narahara, M. and Miyake, M. (2004) Sox6 overexpression causes cellular aggregation and the neuronal differentiation of P19 embryonic carcinoma cells in the absence of retinoic acid. FEBS Lett, **560,** 192-8.

Liberati, C., Cera, M.R., Secco, P., Santoro, C., Mantovani, R., Ottolenghi, S. and Ronchi, A. (2001) Cooperation and competition between the binding of COUP-TFII and NF-Y on human epsilon- and gamma-globin gene promoters. J Biol Chem, **276,** 41700-9.

Liberati, C., Cera, M.R., Secco, P., Santoro, C., Mantovani, R., Ottolenghi, S. and Ronchi, A. (2001) Cooperation and competition between the binding of COUP-TFII and NF-Y on human epsilon- and gamma-globin gene promoters. J Biol Chem, **276,** 41700-9.

Liberati, C., Sgarra, R., Manfioletti, G. and Mantovani, R. (1998) DNA binding of NF-Y: the effect of HMGI proteins depends upon the CCAAT box. FEBS Lett, **433,** 174-8.

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,** 4188-96.

Pereira, F.A., Tsai, M.J. and Tsai, S.Y. (2000) COUP-TF orphan nuclear receptors in development and differentiation. Cell Mol Life Sci, **57,** 1388-98.

Sankaran, V. G., T. F. Menne, J. Xu, T. E. Akie, G. Lettre, B. Van Handel, H. K. Mikkola, J. N. Hirschhorn, A. B. Cantor, and S. H. Orkin. 2008. Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. Science **322**:1839-1842.

Schreiber, E., Matthias, P., Muller, M.M. and Schaffner, W. (1989) Rapid detection of octamer binding proteins with 'mini-extracts', prepared from a small number of cells. Nucleic Acids Res, **17,** 6419.

Xu, J., Sankaran, V.G., Ni, M., Menne, T.F., Puram, R.V., Kim, W. and Orkin, S.H. (2010) Transcriptional silencing of {gamma}-globin by BCL11A involves long-range interactions and cooperation with SOX6. Genes Dev, **24,** 783-98.

Yi, Z., Cohen-Barak, O., Hagiwara, N., Kingsley, P.D., Fuchs, D.A., Erickson, D.T., Epner, E.M., Palis, J. and Brilliant, M.H. (2006) Sox6 directly silences epsilon globin expression in definitive erythropoiesis. PLoS Genet, **2,** e14.

Chapter-4

SOX6, COUP-TFII AND BCL11A-XL INTERPLAY IN THE REGULATION OF HEMOGLOBIN SWITCHING

INTRODUCTION

BCL11A is a highly conserved zinc finger transcription factor and is expressed in hematopoietic progenitors, but it is progressively down regulated during myeloid differentiation (Saiki Y *et al.,* 2000). Genome wide association studies (GWAS) led to the identification of a new HbF associated locus on chromosome 2, located within the gene BCL11A (Menzel *et al.,* 2007, Lettre *et al.,* 2008, Thein *et al.,* 2007, Uda *et al.,* 2008, Baur DE *et al.,* 2013). Alternative pre-mRNA splicing of human BCL11A (fig 4.1) leads to a minimum of 4 transcripts predicted to yield protein isoforms designated as eXtra-Long (XL; 5.9 kb/125 kD), Long (L; 3.8 kb/100 kD), Short (S, 2.4 kb/35 kD) and eXtra-Short (XS, 1.5 kb/25 kD). Exons 1 and 2 are common to all isoforms, whereas XL, L and S each utilize at least a portion of exon 4, leading to a variable number of C2H2 zinc fingers appended to the invariant C2HC zinc finger at the N-terminus (Liu H *et al.,* 2006). BCL11A-XL RNA is expressed at high levels in normal as well as malignant lymphoid tissues, including germinal centre (GC) B cells, B-CLL and follicular lymphoma (Satterwhite E *et al.,* 2001; Su Al *et al.,* 2002). The mouse similar "Evi9" locus is extraordinarily conserved at nucleotide (94%) and amino acids (95%) levels, placing BCL11A/Evi9 within a small subset of "ultra-conserved genes" (Bejerano G *et al.,* 2004). BCL11A has been shown to function as a regulator of HbF expression in humans (Sankaran VG *et al.,* 2008) and recently it has been proposed that BCL11A interacts physically with Sox6 in silencing γ-globin transcription in adult human erythroid cells (Xu J *et al.,* 2010).

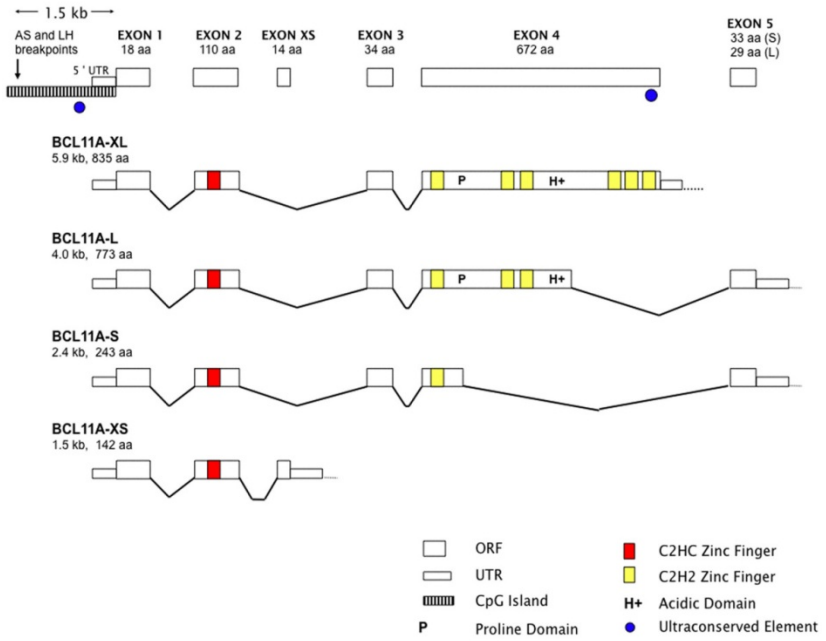


Figure 4.1: Schematic representation of Human BCL11A locus and predicted isoforms. Alternative splicing within the BCL11A locus (Entrez Gene ID53335) leads to the creation of four major protein isoforms: eXtra Long (XL; accession AJ404611), Long (L; accessionAJ404612), Short (S; accession AJ404613), and eXtra Short (XS; accession AY692278). Figure adapted from Liu *et al.*. Molecular Cancer 2006.

In my previous chapter, I have demonstrated that Sox6 and COUP-TFII functionally interact with each other and play a role in the regulation of γ-globin expression on V-K562 cells and possibly in preliminary cultures from peripheral blood in human erythroid cells. In mouse fetal liver, both transcription factors bind *in vivo* (ChIP experiments from Chapter 3) within the γ-globin promoter. The repression of γ-promoter mediated by Sox6 could involve the further recruitment of BCL11A-XL, recently identified as a stage specific

repressor of γ-globin (Sankaran VG *et al.,* 2008). In fact, Xu J and co-authors showed that BCL11A-XL interact with Sox6 that would thus recruit BCL11A-XL to the chromatin. Interestingly, BCL11A-XL was originally identified as COUP-TFII Interacting Protein 1 (CTIP1) because of its ability to physically interact with COUP-TFII (Avram D *et al.,* 2002). Despite all these data point to a functional interaction between Sox6, COUP-TFII and BCL11A-XL, the regulatory interplay between these transcription factors is still to be identified. High Resolution ChIP-Chip analysis on BCL11A reveals that XL isoform is the predominant form present in human erythroid progenitors and occupy discrete regions within the β-globin cluster in-vivo (Xu J *et al.*,2010).

However, of note, BCL11A-XL does not detectably bind on the promoter of γ-globin genes but it occupies HS-3 regions of the LCR, suggesting that BCL11A may perform its functions through long distance interactions or by being recruited to the DNA by other transcription factors (Xu J *et al.,* 2010). Paradoxically, a different study shows that BCL11A-XL binds to 5'-GGCCGG- 3' motif in nucleotide −56 to −51 on the γ-globin proximal promoter when transfected into K562 cells (Chen Z *et al.,* 2009). The difference in the above findings may be due in part due to sensitivity differences between the techniques used (i.e., ChIP-Chip experiments versus in vitro assays).

Bioinformatic searches show that the human β-globin locus contains three 5'-GGCCGG-3' canonical GC box binding sites for BCL11A, located near the HS-3 site in the LCR, the proximal

promoters Gγ- and Aγ- globin genes and on the promoters of ε-globin genes (GenBank U01317). By in vitro assays it is still not clear whether BCL11A-XL directly binds to the GC box region. BCL11A-XL binding at the γ-globin promoter site has been reported to detectably bind at lower levels (Zhu X *et al.*, 2012).

Finally, Kenji Ohe *et al.*, in 2002 have shown that Sox6 co-localizes with the splicing factors in the nuclear complex and acts directly in the process of pre-mRNA splicing, suggesting that Sox6 could act as an unconventional splicing factor. Our hypothesis was that Sox6 could play a role in the splicing of BCL11A isoforms such that the XL isoform is predominantly spliced as a major full length protein in adult erythroid cells thereby regulating the γ-globin gene expression, whereas the role and function of the other isoforms have not been studied.

The abundance of the different isoforms changes in the course of erythroid development/differentiation could thus likely have a role in the mechanism of globin switching. To elucidate this point, I first tested the relative abundance of the BCL11A isoforms by qRT-PCR in V-K562 cells and then I tested in a preliminary series of experiments whether Sox6 overexpression correlates with changes in BCL11A splicing isoforms.

EXPERIMENTAL PROCEDURES

BCL11A isoforms overexpression vectors:

The BCL11A gene isoforms (XS, S, L and XL) were kindly provided by Dr. Philip.W.Tucker, USA and were cloned as expression plasmids pcDNA3/3xFlag/BCL11A–XL, –L, and –S.

Lentiviral Harvesting Method:

Exponentially growing HEK293T cells were transfected with jetPEI^{TM} reagent (Polyplus-Transfection) with the three vector lentiviral system. 72 hours after transfection, the supernatant containing the recombinant viruses were collected, filtered (0.45μm), centrifuged at 20,000g for 8 hours at 4ºC. The viral pellet were resuspended in PBS and stored in Aliquots at -80ºC. Lentiviruses were titrated on HEK 293T cells by measuring the percentage of GFP positive cells by Flow Cytometry analysis.

Cell Cultures:

V-K562 and P19 cells were cultured in RPMI 1640 medium (Lonza) supplemented with 10% heat inactivated fetal bovine serum (Lonza), L-glutamine (Euroclone) and antibiotics Penicillin-Streptomycin 100U/100ug/ml (Euroclone) in a humidified 5% CO_2 atmosphere at 37°C. Transduction was performed overnight, by adding the vector stock at multiplicity of infection (MOI) 30 for individual overexpression of Transcription factors.

RNA Isolation and RT-PCR:

Total RNA from 10^6 cells (V-K562) were extracted with TRI Reagent (Applied Biosystems AM9738), treated with RQ1 DNase (Promega) for 30 min at 37°C and retrotranscribed (High Capacity cDNA Reverse Transcription Kit, Applied Biosystems cat n°4368814). Negative control reactions (without Reverse Transcriptase) gave no signal. Real time analysis was performed using ABI Prism 7500, (PE Applied Biosystems). Primers were designed for the different specific BCL11A isoforms to amplify 100 to 150bp amplicons on the basis of sequences from the NCBI / UCSC database. Samples from three independent experiments were analyzed in triplicate. Specific PCR product accumulation was monitored by SYBR Green dye fluorescence in 25-µl reaction volume. Dissociation curves confirmed the homogeneity of PCR products. All primers used are listed in the table below:

Primers	Sequence (5'-3')
GAPDH Fw	ACGGATTTGGTCGTATTGGG
GAPDH Rev	TGATTTTGGAGGGATCTCGC
BCL11A-XL Fw	ATGCGAGCTGTGCAACTATG
BCL11A-XL Rev	GTAAACTCCTTCCCCACCT
BCL11A-L Fw	CAGCTCAAAAGAGGGCAGAC
BCL11A-L Rev	GAGCTTCCATCCGAAAACTG
BCL11A-S Fw	AGAATATGCCCCGCAGGG
BCL11A-S Rev	TCAGAACTTAAGGGCTCTC

BCL11A-XS Fw	CCCTTCACCAATCGAGATGA
BCL11A-XS Rev	CACAAATACATCCTCCAGTTCAG
alpha globin Fw	GAGGCCCTGGAGAGGATGTTCC
alpha globin Rev	ACAGCGCGTTGGGCATGTCGTC
beta globin Fw	TACATTTGCTTCTGACACAAC
beta globin Rev	ACAGATCCCCAAAGGAC
gamma globin Fw	CTTCAAGCTCCTGGGAAATGT
gamma globin Rev	GCAGAATAAAGCCTACCTTGAAAG
epsilon globin Fw	GCCTGTGGAGCAAGATGAAT
epsilon globin Rev	GCGGGCTTGAGGTTGT

Chromatin Immunoprecipitation (ChIP) assay:

V-K562 Cells ($1x10^6$) for each Immunoprecipitation reaction were fixed with 1% formaldehyde for 10 minutes at room temperature, and chromatin was sonicated (Branson^{TM} Digic sonifier) to a size of about 100-500 bp in size. Immunoprecipitation was performed after overnight incubation with Rabbit anti-IgG antibody (Santa Cruz SC-2027) and anti-BCL11A-XL (Novus Biologicals NP600-261), and subsequent incubation with protein-A agarose (Upstate biotechnology Millipore). Immunoprecipitated DNA was then analysed by amplifying an equivalent of 10^4 cells DNA with the following oligonucleotides:

Primers	Sequence (5'-3')
γ- globin promoter Fw	AAACGGTCCCTGGCTAAACT
γ- globin promoter Rev	GCTGAAGGGTGCTTCCTTTT
ε- globin promoter Fw	GAGCCTCAGGATCCAGCACAC
ε- globin promoter Rev	GATGCCAGGCCTGAGAGCTTGC
LCR-HS-3 Fw	GGCAAGTGCCTTGACTCCTA
LCR-HS-3 Rev	TCTTCTGGAACTTGCCTGCT
GAPDHexon Fw	CGGAGTCAACGGATTTGGTCGTAT
GAPDHexon Rev	AGCCTTCTCCATGGTGGTGAAGAC

Western Blot:

Total and nuclear extracts from V-K562 cells were prepared according to standard protocols (Schreiber E *et al.*, 1989) and proteins were subjected to sodium dodecyl sulphate- polyacrylamide gel electrophoresis separation and blotting. The Sox6 protein was detected by anti-FLAG antibody (Sigma F-7425) or anti-Sox6 (abcam ab30455) antibody and anti-BCL11A-XL (Novus Biologicals NP600-261) antibody. Protein loading control was checked with an anti-U2AF antibody (Sigma-Aldrich). Antibody binding was detected by use of appropriate horseradish peroxidise-conjugated immunoglobulin G and revealed by enhanced chemiluminescence (LiteAblot; Euroclone).

RESULTS

BCL11A isoforms and their expression in V-K562 cells

BCL11A isoforms were checked for their individual expression in V-K562 cells. K562 cells in general express very low levels of BCL11A. As seen in Figure 4.3, the shorter BCL11A-S isoform is predominantly expressed when compared to the other larger (-L and -XL) isoforms.

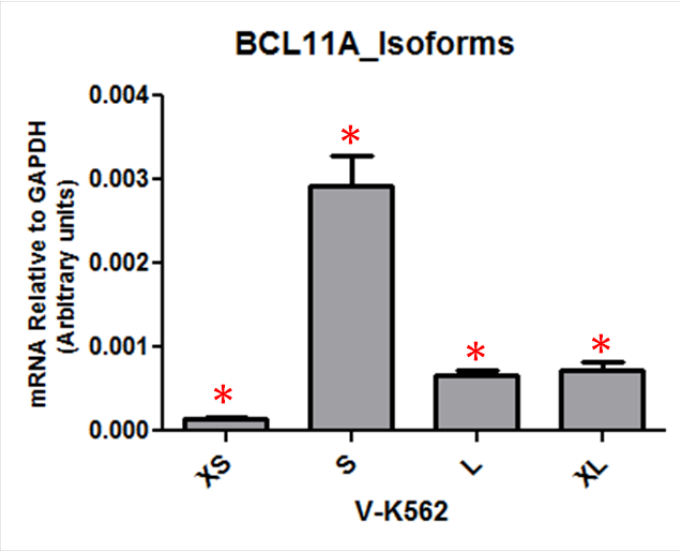


Figure 4.3: Expression of BCL11A isoforms in V-K562 cells. qRT-PCR analysis on V-K562 cells shows that S isoform is the predominant BCL11A isoform. Histograms shows the relative levels of expression normalized with respect to GAPDH. Asterisk indicates significant P-value < 0.05 +/- SEM (n=3).

BCL11A- XL overexpression in V-K562 cells

BCL11A has been shown to function as a transcriptional repressor. However, the biological consequences of its trans repression effects remain unknown. Since BCL11A-XL isoform has been predominantly expressed in adult erythroid cells, we decided to focus our attention on this isoform and its effects in globin gene regulation. It has been demonstrated that overexpression of BCL11A-XL can induce necrosis and apoptotic cell death mediated through mitochondrial pathway (Liu H *et al.,* 2006). To overexpress BCL11A-XL but to minimize the cell death upon BCL11A-XL overexpression, we decided to stably transfect V-K562 cells with pcDNA3 empty vector and pcDNA3-BCL11A–XL vector, under G418-Neomycin resistance selection. The stably transfected BCL11A-XL cells showed 3.5 times increase in the levels of BCL11A-XL overexpression (Fig 4.4), at both mRNA and protein level.

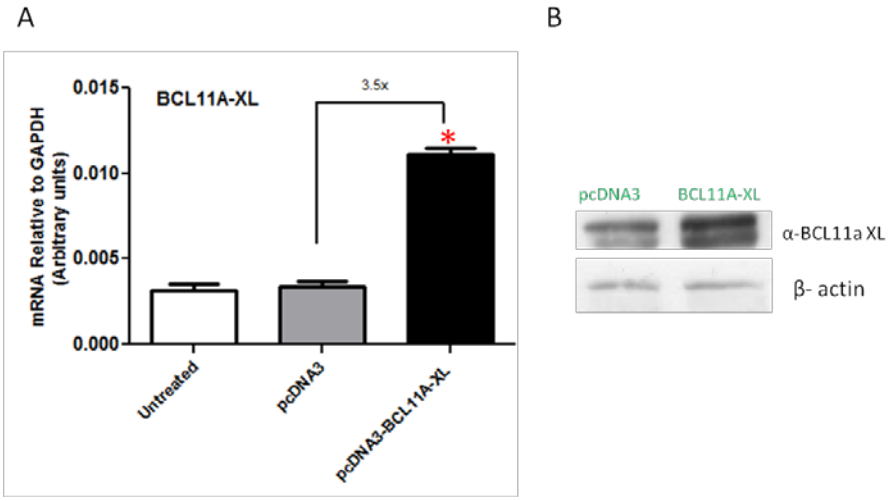


Figure 4.4: Overexpression of BCL11A-XL in V-K562 cells **A).** qRT-PCR analysis on cells overexpressing BCL11A-XL with respect to the empty vector pCDNA3 and untransfected cells. Histograms shows the relative levels of expression normalized with respect to GAPDH. Asterisk indicates significant P-value < 0.001 +/- SEM (n=3). **B).** Western Blot analysis showing BCL11A-XL protein expression by using an anti-BCL11A-XL isoform specific antibody. The BCL11A-XL band was always seen as two discrete bands. Anti β-actin was used as a loading control.

At this modest level of overexpression of BCL11A-XL, cells do not die and continue to grow well. This allowed a careful analysis of the levels of globins upon BCL11A-XL overexpression. Unexpectedly, stably transfected BCL11A-XL overexpressing cells show an increase in all globins (fig 4.5), except for ε-globin expression. However, there is no decrease in the γ-globin expression as expected on the basis of its reported function as a repressor. This suggests that, in V-K562 cells, the modest level of BCL11A-XL overexpression compatible with V-K562 survival does not *per se* have the ability to repress γ-globin expression. Instead, it seems that in these conditions BCL11A-XL promotes a modest increase in α-, β- and γ- globins genes.

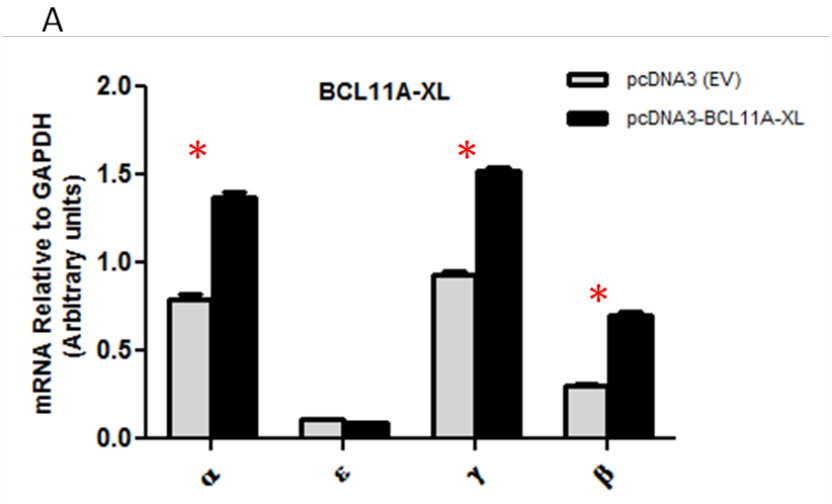


Figure 4.5: A). qRT-PCR analysis on globins with cells overexpressing BCL11A-XL with respect to the empty vector pCDNA3. Histograms show the relative levels of expression normalized with respect to GAPDH. Asterisk indicates significant P-value < 0.001 +/- SEM (n=3).

Lentiviral mediated Sox6 overexpression in V-K562 cells stably transfected with BCL11A-XL

It has been shown that BCL11A-XL cooperates with Sox6 and together represses γ-globin expression (Xu J *et al.*, 2010). However, the mechanism by which Sox6 recruits BCL11A-XL is still to be explored. In order to see whether this interaction either mediates a general effect on globins genes regulation or a specific effect in γ-globin expression, I overexpressed Sox6 by lentiviral mediated transduction experiments in V-K562 cells stably overexpressing BCL11A-XL. In chapter 3, I have described that overexpression of Sox6 leads to a general increase in all globins, as a consequence of a general induction in erythroid dufferentation, despite being a repressor of γ-globin. On the contrary, the overexpression of Sox6 in stably transfected V-K562-BCL11A-XL cells induces a reduction in the level of γ-globin expression (fig 4.6, last graph), with respect to the EV-transduced V-K562-BCL11A-XL (black histogram bar). This is in agreement with the paper by Xu J *et al.*, 2010, suggesting that BCL11A-XL and Sox6 cooperate in repressing γ-globin expression.

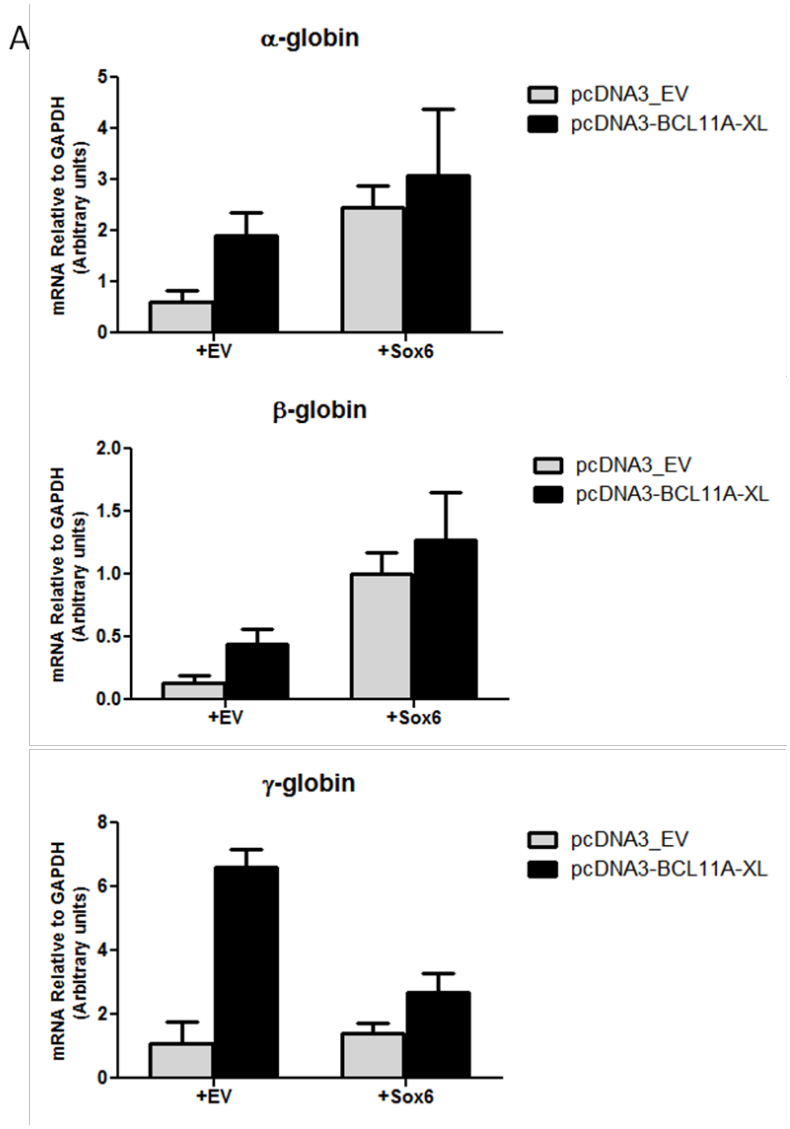
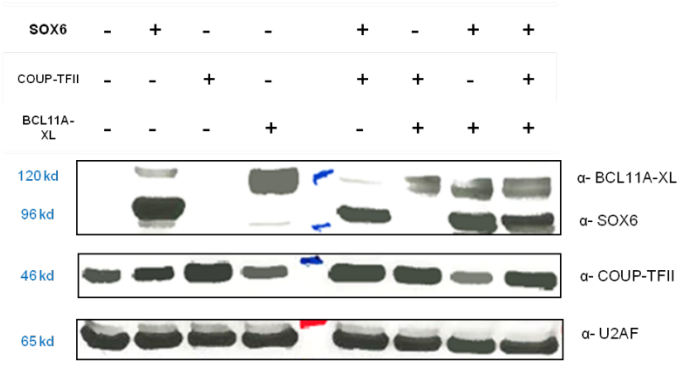


Figure 4.6: Overexpression of Sox6 in V-K562 cells stably transfected with BCL11A-XL. **A).** Level of globins in Empty vector and Sox6 transduced cells show a slight decrease in the level of γ-globin expression upon Sox6 overexpression. Histograms shows the relative levels of expression normalized with respect to GAPDH. P-value < 0.05 +/- SEM (n=4).

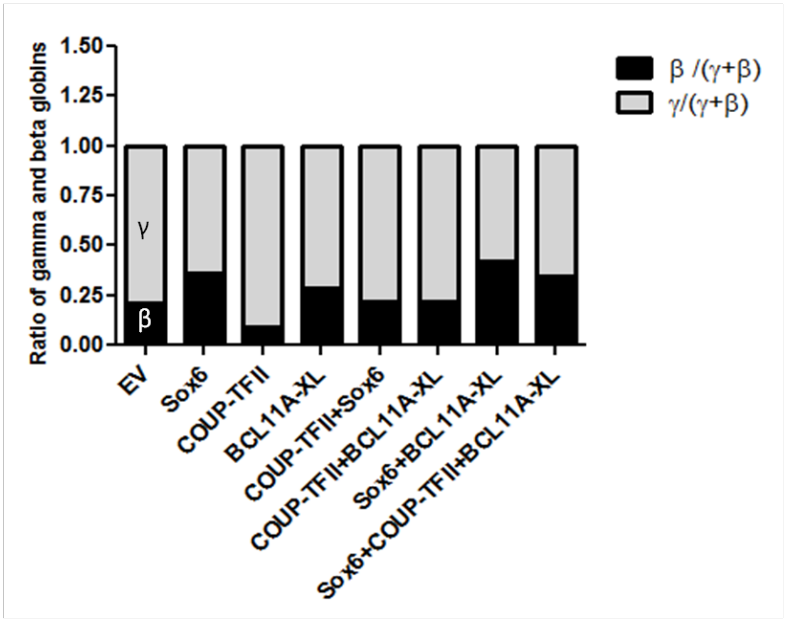
Co-transduction of Sox6, COUP-TFII and BCL11A-XL in V-K562 cells

In Chapter 3, I have shown that co-transduction of COUP-TFII and Sox6 alters the levels of γ- and β-globin expression, suggesting that COUP-TFII and Sox6 could play a role in switching. Since BCL11A-XL is also important for hemoglobin switching, I cloned the BCL11A-XL gene upstream to the GFP in the same way as I cloned for COUP-TFII and Sox6 respectively. Since we know that overexpression of BCL11A-XL induces apoptosis in cells after three days of transduction, I planned to overexpress BCL11A-XL in a short-term for three days and quickly analyse their effects. As a preliminary experiment, I co-transduced BCL11A-XL with Sox6, BCL11A-XL with COUP-TFII and BCL11A-XL with Sox6 and COUP-TFII all together in equimolar concentrations of the virus with a MOI-30. The corresponding empty vector (EV) and the individual vector transductions were also carried out as controls. The efficiency of transduction was monitored after 72hrs by flow cytometry and the expression of exogenous Sox6, COUP-TFII and BCL11A-XL were checked by western blot (Fig.4.7a).

A



B



C

Vector	β/(γ+β)
EV	**0.21**
Sox6	**0.36**
Coup-TFII	**0.09**
BCL11A-XL	**0.29**
Coup-TFII+Sox6	0.22
Coup+BCL11A-XL	0.22
Sox6+BCL11A-XL	0.42
Sox6+Coup+BCL11A-XL	0.35

Figure 4.7: Co-transduction of BCL11A-XL, Sox6 and COUP-TFII in various combination in V-K562 cells. **A).** Western blot confirming the exogenous overexpression of transcription factors: the first 4 lanes (1-4) representing the individual overexpression of each factors and lanes 4-8 show different combination of factors. Anti-U2AF is used as a loading control. **B).** qRT-PCR shows the the ratio of beta and gamma globins. Black bars in the histogram represent the ratio of β over (γ+β) and the grey bars indicate the ratio of γ over (γ+β) where the ratio of (γ+β) was set equal to 1 normalized against GAPDH. **C).** Ratio of β over (γ+β) values is summarized in the table (n=1).

So far a single experiment has been performed and a more careful experimentation has to be carried out in order to obtain a sound data. However, this preliminary experiment suggests that in the presence of BCL11A and COUP-TFII there is no change in the level of globins when compared with the empty vector (EV) whereas in the presence of BCL11A, COUP-TFII and Sox6 there is an increase in β-globin expression confirming that BCL11A-XL indeed acts as a repressor of γ-globin but it requires Sox6 to perform its repressor function. If these data will be confirmed, they would suggest that the balance between these three different transcription factors, more than the single factors themselves, controls γ-globin transcription.

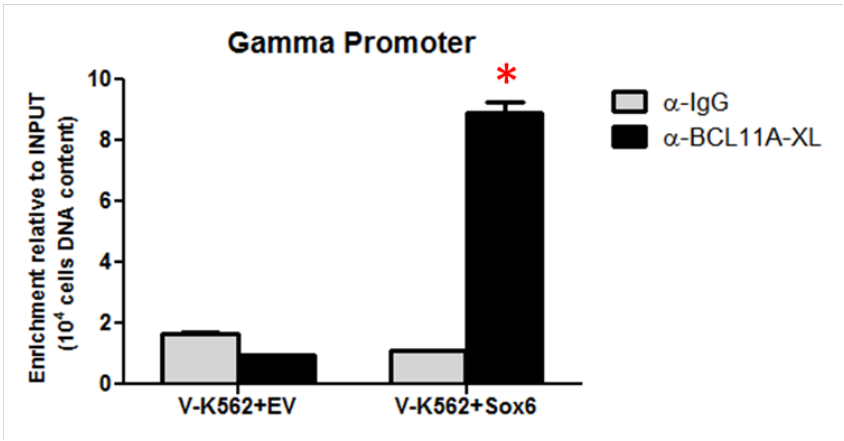
Chromatin Immunoprecipitation on V-K562 cells overexpressing Sox6

It is unclear whether BCL11A-XL directly represses the γ-globin promoter by binding to the target sequences or whether it requires Sox6 to be recruited to DNA. In fact, it has been proposed that BCL11A-XL directly binds to the canonical 5'-GGCCGG-3' (-56) region at the γ-globin promoter site at low detectable levels (Zhu X *et*

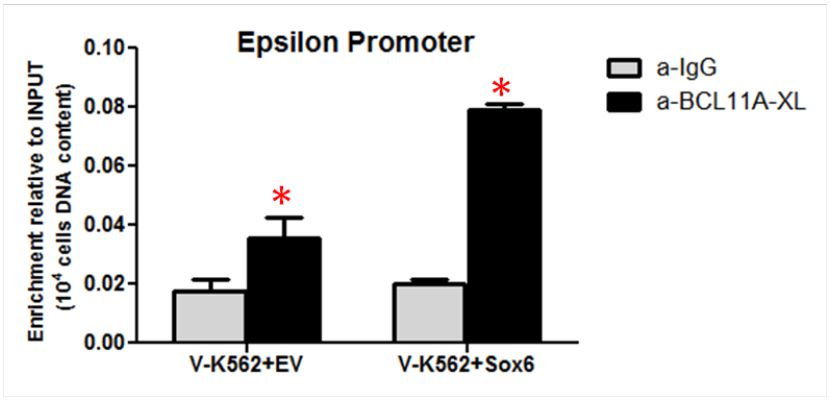
al., 2012). However, this data has not been confirmed. To address this issue, I set up a series of ChIP experiment to test whether BCL11A-XL is able to bind to the γ-globin promoter by itself or whether it requires the presence of Sox6. Cells transduced either with EV or with Sox6 overexpressing vector were immunoprecipated with anti-BCL11A-XL antibody; rabbit anti-IgG was used as a negative control. ChIP-qPCRs show that enrichment of BCL11A-XL occupancy on the γ-globin and ε-globin promoters occurs only in the cells overexpressing Sox6 (Fig 4.8a-b) suggesting that BCL11A-XL requires Sox6 to bind to the γ-globin and ε-globin promoter regions.

BCL11A-XL has also been shown to bind on the LCR of the β-globin cluster, specifically with DNaseI HS-3 region (Fig.4.8c). I also found an increased enrichment of BCL11A-XL upon Sox6 overexpression. This result strongly argues that BCL11A-XL is able to bind the ε-globin and γ-globin promoters only in the presence of Sox6, whereas it is able to directly bind to HS-3 even in the absence of high levels of Sox6.

A


Gamma Promoter
Enrichment relative to INPUT
(10^4 cells DNA content)
10
8
6
4
2
0
V-K562+EV
V-K562+Sox6
α-IgG
α-BCL11A-XL
*

B


Epsilon Promoter
Enrichment relative to INPUT
(10^4 cells DNA content)
0.10
0.08
0.06
0.04
0.02
0.00
V-K562+EV
V-K562+Sox6
a-IgG
a-BCL11A-XL
*

C

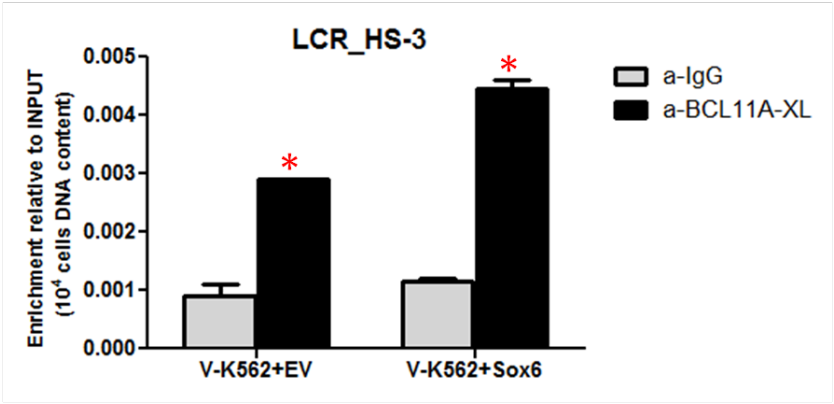

LCR_HS-3
Enrichment relative to INPUT
(10^4 cells DNA content)
0.005
0.004
0.003
0.002
0.001
0.000
V-K562+EV
V-K562+Sox6
a-IgG
a-BCL11A-XL
*

Figure 4.8: Chromatin Immunoprecipitation in V-K562 cells. **A-B).** BCL11A-XL enrichment on the human gamma and epsilon promoters is seen only in cells overexpressing Sox6. **C).** Since BCL11A-XL is known to bind on the DnaseI HS-3 region of the LCR, this was used as a positive control for BCL11A-XL binding. Grey histogram bar represents α-IgG used as negative control and black bars represents pull down using α-BCL11aXL antibody. Results were normalized relative to Input DNA from 10^4 cells. Asterisk indicates significant P-value <0.001+/- SEM (n=3).

BCL11A-XL expression on mouse embryonic fetal liver cells

To assess the levels of BCL11A-XL expression in mouse fetal liver cells during the switching time, I checked for its expression at day E11.5 to E13.5, where the switching occurs. Fig 4.9 shows that BCL11A-XL is expressed in all stages of fetal liver development although at very low levels (see y-axis in comparison with endogenous GAPDH). In line with this, very recently BCL11A-XL was reported to be absent at the mRNA and protein levels in human fetal liver cells (Zhu X *et al.*, 2012). These evidences suggest that, although BCL11A-XL is present during the time of switching, it is unlikely be directly responsible for the relative changes in embryonic to adult expression of globin genes, taking place at these stages.

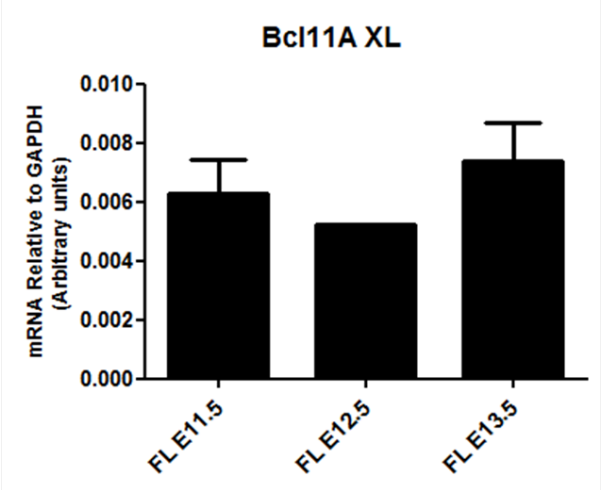


Figure 4.9: qRT-PCR analysis of BCL11A-XL on mouse fetal liver embryonic day E11.5-E13.5 (switching time) showing similar levels of expression at all time points. Significance cannot be calculated as the experiment was performed only once.

BCL11A isoforms and alternative splicing in P19 cells and V-K562 cells

As Sox6 has been reported to have some effects on splicing (Ohe *et al.*. 2009), I wanted to check whether Sox6 overexpression could modulate BCL11A splicing. I performed initial experiments in mouse P19 teratocarcinoma cells, as these cells express high levels of BCL11A and its isoforms (Fig 4.10a). In order to check whether Sox6 could play a role in splicing the BCL11A isoforms, I then overexpressed Sox6 in P19 cells. As seen from Fig 4.2b, Sox6 overexpression induces a major shift in the balance of the BCL11A -XL and -L (larger) and S (smaller) RNA isoforms, with an almost complete suppression of the L and XL isoform, relative to the S

isoform. Although these experiments were done in a neural cell line, preliminary experiments suggest that Sox6 induction correlates with changes in the BCL11A isoforms, suggesting that Sox6 could possibly act as a splicing factor in splicing the BCL11A isoforms.

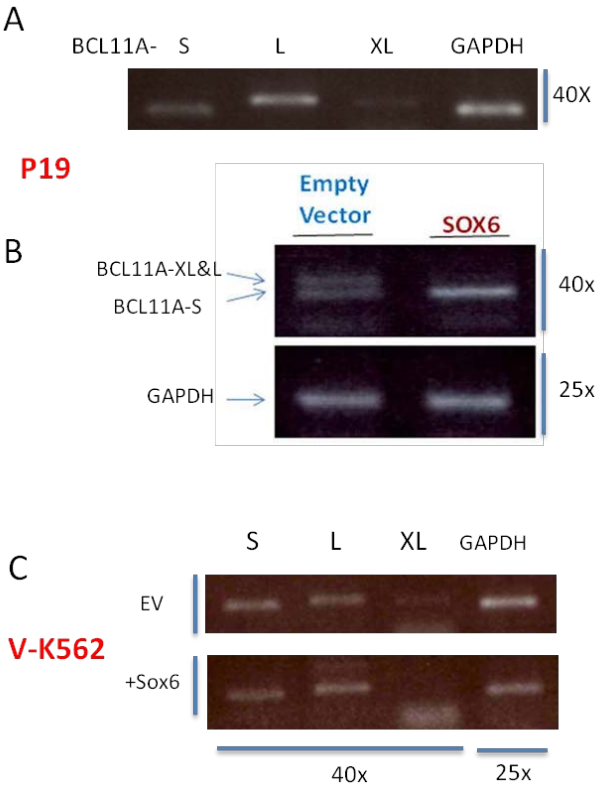


Figure 4.10: Semi-quantitative PCR analysis. **A).** BCL11A and its isoforms in P19 cells. **B).** In the presence of Sox6, BCL11A-L isoform is completely absent or spliced to produce only the S isoform suggesting that Sox6 plays a role in splicing BCL11A. (n=3, representative experiment 1 out of 3 shown).

In order to better understand the role of Sox6 in splicing, I next checked whether Sox6 exerts the same function in V-K562 cells that correlates with the changes in BCL11A splicing isoforms. BCL11A

isoforms expression was checked in V-K562 cells transduced with EV and Sox6 lentiviruses. From Figure 4.10c, I show that Sox6 overexpression reduces the expression of BCL11A-XL isoform and a slight increase in L isoform, as observed qualitatively with respect to empty vector transduced cells. This suggests that Sox6 could directly or indirectly play a role in the splicing of BCL11A isoforms. Although preliminary, these results along with the P19 data are important because the larger isoforms (L and XL) are known to be involved in the γ-globin expression. More careful studies on the role of Sox6 in splicing will help us to understand if BCL11A-XL is predominantly spliced as a major full length protein in adult erythroid cells thereby regulating the γ-globin gene expression.

DISCUSSION

In this chapter, I show that V-K562 cells, expresses BCL11A at very low levels and among its isoforms, only the shorter BCL11A–S isoform is predominantly expressed when compared to other larger isoforms (Fig 4.3). However, BCL11A-XL isoform is predominantly expressed in adult erythroid cells.

Recently, it has been shown that Sox6 interacts with BCL11A-XL in repressing (Xu J *et al.,* 2010) γ-globin expression. However, the biological consequences of its repression effects remain unknown. In order to functionally characterize the role of Sox6 and BCL11A in the regulation of γ-globin expression, I stably overexpressed BCL11A-XL in V-K562 cells. It has been known that overexpression of BCL11A in K562 cells causes apoptotic cell death (Liu H *et al.,* 2006). Keeping this is mind, I moderately expressed BCL11A-XL (3.5 times) in V-K562 cells at levels enough for the cell survival (fig 4.4). I show that moderate overexpression of BCL11A-XL in V-K562 cells increases the expression of all globins except ε-globin expression (Fig 4.5). This suggests that BCL11A-XL *per se* does not have the ability to repress γ-globin. Instead, in these conditions, it acts as an activator. However, transduction of BCL11A-XL with Sox6 induces a reduction in the γ-globin (Fig 4.6), in agreement with the paper by Xu J *et al.,* 2010, suggesting that BCL11A-XL and Sox6 co-operate in repressing the γ-globin expression. I next checked whether COUP-TFII would play a role along with BCL11A-XL and Sox6 in modulating γ to β ratio, as COUP-TFII was known to physically interact with BCL11A-XL (Avram D *et al.,* 2002). Co-transduction results of BCL11A-XL with

COUP-TFII show no repression or activation of γ-globin (Fig 4.7c). When BCL11A-XL is co-transduced with Sox6 and COUP-TFII, there is a increase in the level of β/(γ+β) from 0.21 in EV to 0.35, at levels lower than the effect upon BCL11A-XL and Sox6 alone (from 0.21 in EV to 0.42. This suggests that BCL11A-XL synergises with Sox6 to exert its repressor function.

Further, to assess whether BCL11A-XL directly binds to the γ-globin promoter, ChIP experiments on V-K562 cells were performed. These experiments show that BCL11A-XL binds to the γ-globin and ε-globin promoters only in the presence of Sox6 overexpression. Notably, V-K562 cells do express Sox6 at a very low level and this could be the reason why we do not see BCL11A-XL binding on both the γ-globin and ε-globin promoters. Of interest, it has been shown that BCL11A-XL can bind to the strong LCR HS-3 site in ChIP-Chip and ChIP-qPCR experiments (Zhu X *et al.*, 2012). Here, I show that COUP-TFII also binds strongly to the LCR HS-3 region (Chapter 3, Fig3.12), suggesting that COUP-TFII and BCL11A-XL could functionally interact on chromatin.

I also show that the mouse P19 teratocarcinoma cell line expresses good amount of BCL11A and its isoforms. Preliminary transduction experiments show that BCL11A-L isoform is completely shut down upon Sox6 overexpression suggesting that Sox6 could mediate the splicing of BCL11A isoforms (Fig 4.10).

Moreover, mouse embryonic fetal liver cells at day E11.5-E13.5 (when the switching takes place), show very low levels of BCL11A-XL expression at all three days, during the switching time. This suggests that BCL11A-XL is unlikely to be directly responsible for the relative changes in embryonic to adult expression of globins genes around the switching time.

The schematic model (Fig 4.11) shows the interaction between Sox6, BCL11A-XL and COUP-TFII as described in literature suggesting that they can either act synergistically at the same time or independently, in a stage specific manner, in regulating the hemoglobin switching.

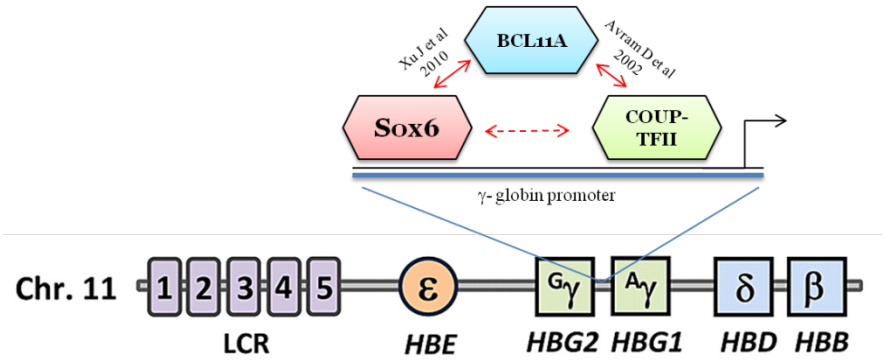


Figure 4.11: Schematic representation of the three transcription factors interacting with each other on the γ-globin promoter. The interaction (dashed lines) between Sox6 and COUP-TFII has been studied in this work.

REFERENCES

Avram, D., Fields, A., Senawong, T., Topark-Ngarm, A. and Leid, M. (2002) COUP-TF (chicken ovalbumin upstream promoter transcription factor)-interacting protein 1 (CTIP1) is a sequence-specific DNA binding protein. Biochem J, **368,** 555-63.

Bauer, D.E., Kamran, S.C., Lessard, S., Xu, J., Fujiwara, Y., Lin, C., Shao, Z., Canver, M.C., Smith, E.C., Pinello, L., Sabo, P.J., Vierstra, J., Voit, R.A., Yuan, G.-C., Porteus, M.H., Stamatoyannopoulos, J.A., Lettre, G. and Orkin, S.H. (2013) An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. Science, **342,** 253-7.

Chen, Z., Luo, H.Y., Steinberg, M.H. and Chui, D.H. (2009) BCL11A represses HBG transcription in K562 cells. Blood Cells Mol Dis, **42,** 144-9.

Lettre, G., Sankaran, V.G., Bezerra, M.A., Araujo, A.S., Uda, M., Sanna, S., Cao, A., Schlessinger, D., Costa, F.F., Hirschhorn, J.N. and Orkin, S.H. (2008) DNA polymorphisms at the BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. Proc Natl Acad Sci USA, **105,** 11869-74.

Liu, H., Ippolito, G.C., Wall, J.K., Niu, T., Probst, L., Lee, B.S., Pulford, K., Banham, A.H., Stockwin, L., Shaffer, A.L., Staudt, L.M., Das, C., Dyer, M.J. and Tucker, P.W. (2006) Functional studies of BCL11A: characterization of the conserved BCL11A-XL splice variant and its interaction with BCL6 in nuclear paraspeckles of germinal center B cells. Mol Cancer, **5,** 18.

Menzel, S., Garner, C., Gut, I., Matsuda, F., Yamaguchi, M., Heath, S., Foglio, M., Zelenika, D., Boland, A., Rooks, H., Best, S., Spector, T.D., Farrall, M., Lathrop, M. and Thein, S.L. (2007) A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. Nat Genet, **39,** 1197-9.

Ohe, K., Tamai, K.T., Parvinen, M. and Sassone-Corsi, P. (2009) DAX-1 and SOX6 molecular interplay results in an antagonistic effect in pre-mRNA splicing. Dev Dyn, **238,** 1595-604.

Saiki, Y., Yamazaki, Y., Yoshida, M., Katoh, O. and Nakamura, T. (2000) Human EVI9, a homologue of the mouse myeloid leukemia gene, is expressed in the hematopoietic progenitors and down-regulated during myeloid differentiation of HL60 cells. *Genomics,* **70,** 387-91.

Sankaran, V. G., J. Xu, T. Ragoczy, G. C. Ippolito, C. R. Walkley, S. D. Maika, Y. Fujiwara, M. Ito, M. Groudine, M. A. Bender, P. W. Tucker, and S. H. Orkin. 2009. Developmental and species-divergent globin switching are driven by BCL11A. Nature **460**:1093-1097.

Sankaran, V. G., T. F. Menne, J. Xu, T. E. Akie, G. Lettre, B. Van Handel, H. K. Mikkola, J. N. Hirschhorn, A. B. Cantor, and S. H. Orkin. 2008. Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. Science **322**:1839-1842.

Satterwhite, E. (2001) The BCL11 gene family: involvement of BCL11A in lymphoid malignancies. Blood, **98,** 3413-3420.

Su, A.I., Cooke, M.P., Ching, K.A., Hakak, Y., Walker, J.R., Wiltshire, T., Orth, A.P., Vega, R.G., Sapinoso, L.M., Moqrich, A., Patapoutian, A., Hampton, G.M., Schultz, P.G. and Hogenesch, J.B. (2002) Large-scale analysis of the human and mouse transcriptomes. Proc Natl Acad Sci USA, **99,** 4465-70.

Thein, S.L., Menzel, S., Peng, X., Best, S., Jiang, J., Close, J., Silver, N., Gerovasilli, A., Ping, C., Yamaguchi, M., Wahlberg, K., Ulug, P., Spector, T.D., Garner, C., Matsuda, F., Farrall, M. and Lathrop, M. (2007) Intergenic variants of HBS1L-MYB are responsible for a major quantitative trait locus on chromosome 6q23 influencing fetal hemoglobin levels in adults. Proc Natl Acad Sci USA, **104,** 11346-51.

Uda, M., Galanello, R., Sanna, S., Lettre, G., Sankaran, V.G., Chen, W., Usala, G., Busonero, F., Maschio, A., Albai, G., Piras, M.G., Sestu, N., Lai, S., Dei, M., Mulas, A., Crisponi, L., Naitza, S., Asunis, I., Deiana, M., Nagaraja, R., Perseu, L., Satta, S.,

Cipollina, M.D., Sollaino, C., Moi, P., Hirschhorn, J.N., Orkin, S.H., Abecasis, G.R., Schlessinger, D. and Cao, A. (2008) Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of beta-thalassemia. Proc Natl Acad Sci USA, **105,** 1620-5.

Xu, J., Sankaran, V.G., Ni, M., Menne, T.F., Puram, R.V., Kim, W. and Orkin, S.H. (2010) Transcriptional silencing of {gamma}-globin by BCL11A involves long-range interactions and cooperation with SOX6. Genes Dev, **24,** 783-98.

Zhu, X., Wang, Y., Pi, W., Liu, H., Wickrema, A. and Tuan, D. (2012) NF-Y recruits both transcription activator and repressor to modulate tissue- and developmental stage-specific expression of human gamma-globin gene. PLoS One, **7,** e47175.

Chapter 5

FINAL DISCUSSION:

MOLECULAR AND CLINICAL

CONSIDERATIONS

SUMMARY

The aim of my project was to elucidate the molecular mechanisms by which the two transcription factors Sox6 and COUP-TFII contribute to the regulation of the fetal to adult hemoglobin switching.

To define this interplay functionally, I had used a novel cellular model system, a variant K562 cell line (V-K562) that expresses all globins, including β- globin. This cell line represents an attractive tool to study the switching from γ- to β- globin expression because almost all human cell lines available so far, show an "embryonic" like pattern of expression (with no β- globin), thus hampering biochemical analysis of the γ to β switching. However, being a cell line, V-K562 represents a non "physiological" model for the switching. For this reason, the results obtained in this cellular model were further validated in mouse embryonic fetal liver cells at day E11.5 to E13.5, period in which the globins switching occurs and in human erythroid cultures from peripheral blood.

In the first set of experiments in V-K562, the overexpression of Sox6 predominantly increases the level of β-globin (Fig3.3a; 8.5x increase with respect to cells transduced with empty vector) whereas COUP-TFII predominantly increases the level of γ-globin expression (Fig3.4a; 4x increase with respect to empty vector). Co-transduction of both Sox6 and COUP-TFII alters the balance between the γ- and the β-globin expression, suggesting that the relative levels of Sox6 and COUP-TFII could be involved in the hemoglobin switching.

Both Sox6 and COUP-TFII bind to the human γ-globin promoter in both *in vitro* and *in vivo* experiments, suggesting an interaction between the two transcription factors on the γ-globin promoter (Fig 3.10 and 3.11b). Finally, preliminary experiments in human erythroid cultures from peripheral blood indicate that COUP-TFII overexpression increases γ-globin levels (Fig3.9b). The β/(γ+β) ratio decreased from 0.58 (EV-K562) to 0.34 (COUP-TFII) in (D1) and from 0.72 (EV-K562) to 0.41 (COUP-TFII). These results points to COUP-TFII as a prospective target whose modulation can increase γ-globin expression.

Sox6 and COUP-TFII in the hemoglobin switching

Sox6 was reported as a repressor of the εy-globin gene in mouse, where Sox6 directly binds to the εy-globin promoter (Yi Z. *et al.*, 2006). Very recently it has been shown that Sox6 cooperates with BCL11A-XL in repressing the γ-globin expression in adult human erythroid progenitors (Xu J *et al*., 2010). COUP-TFII, on the other hand, was identified as a modulator of ε- and γ- globin genes, where it binds to the CCAAT box regions. Once bound to these regions, it can either compete or cooperate with the binding of NF-Y (Ronchi A *et al.*, 1995, Liberati C *et al.*, 1998). COUP- TFII binds to the promoter of ε-globin by two sequences of the type DR1 (Direct Repeat 1), of which most of the 3' overlaps the CCAAT box, recognized by NFY (Filipe A *et al.*, 1999) and was shown that COUP- TF II competes with NFY and, directly or indirectly, causes a decrease in gene activity in transfection experiments. Of interest, loss of Coup-TFII

binding is a common effect of HPFH mutations mapped in the region around the distal CCAAT box of the γ promoter (Liberati C *et al.*, 2001).

Taking advantage of the V-K562 model, carrying the native β-globin locus, I showed that COUP-TFII overexpression specifically leads to an increase in the levels of ε- and γ-globin gene expression (Fig 3.4), suggesting a more complex regulation. In these experiments, Sox6 behaves as general activator of globins but specifically activates β-globin expression, whereas, COUP-TFII specifically activates only ε- and γ-globin expression. Co-expression of Sox6 and COUP-TFII at different concentrations (Chapter-3, figs 3.6 and 3.7) results in the modulation of the γ to β ratio, suggesting that these two transcription factors could play a role in the switching mechanism.

Taken together, these data suggest that Sox6 could either act as an activator of β-globin or repressor of γ-globin while COUP-TFII acts an activator of γ-globin expression.

Sox6 and COUP-TFII binds to the γ-globin promoter

Although it has been shown that Sox6 binds to the γ-globin promoter directly (Cantú C *et al.*, 2011), the exact location to which Sox6 binds was not known. To clarify this issue, I performed a ChIP analysis on the V-K562 cells overexpressing Sox6 to confirm the ability of Sox6 to bind on the γ-globin promoter *in vivo*. I then mapped the putative Sox6 consensus binding sequence at position -103-99 upstream to the γ-globin promoter. Interestingly, this region is known to be crucial for

γ-globin promoter regulation. *In vitro* studies by EMSA and transfection assay confirmed the ability of Sox6 to bind to this site (fig 3.9). COUP-TFII has specific binding sites on the γ-globin promoter where it overlaps with NF-Y on the double CCAAT box region. An additional half site is present adjacent to the Sox6 binding region (Fig 5.1a). ChIP analysis confirmed the *in vivo* binding of COUP-TFII on the γ-promoter region (chapter 3, fig 3.10) as well as to the DNaseI HS-3 and HS-4 subunit of the β-globin LCR region, suggesting that COUP-TFII could play a role in the switching mechanism (Chapter 3, fig 3.11). With respect to εy-globin promoter, ChIP on mouse fetal embryonic day E13.5 shows that Sox6 but not COUP-TFII binds to the εy-globin promoter (Fig 3.13a), pointing to a specific role of COUP-TFII in activating γ-globin expression.

Taken together, these results suggest that both Sox6 and COUP-TFII bind to the γ-globin promoter region and perform their activation / repression function by binding to adjacent sites and possibly by directly interacting with each other or through other factors.

Increased γ-globin expression in erythroid cultures from $CD34^+$ purified peripheral blood cells

Since we found that COUP-TFII acts as an activator of γ-globin upon overexpression in V-K562 cells, and with the hypothesis that higher dosage of COUP-TFII could contribute to reactivate the silenced γ-globin expression in the adult, we transduced CD34+ purified human erythroid peripheral blood cells derived from two different healthy

donors with COUP-TFII lentivirus. Three days after transduction, the γ- to β-globin ratio increased in cells overexpressing COUP-TFII, when compared with the empty vector (EV) transduced cells despite only 40-50% of the cells being transduced with COUP-TFII. Although very preliminary, these results suggest that forced expression of COUP-TFII could significantly increase γ-globin expression (Chapter 3, Fig 3.9).

Proposed molecular mechanism of the Sox6 and COUP-TFII competition at the molecular level

In this study, we propose a molecular mechanism of Sox6 and COUP-TFII interaction and of the possibility, that they could facilitate the switching from the fetal to the adult stage in human erythropoiesis. In the fetal stage of erythropoiesis, COUP-TFII is more expressed and binds to the γ-globin promoter, thereby stabilizing the activator NF-Y Complex. The presence of COUP-TFII and NF-Y would form a strong activator complex sustaining γ-globin transcription. During the switching, the increased expression of Sox6 would lead it to bind to its consensus between the two CCAAT box regions thus weakening the binding of COUP-TFII. The Sox6 binding might also weaken or disrupt the interaction between the two NF-Y complexes, thereby hampering their activator function. Once bound, Sox6 could recruit other transcription factors like BCL11A-XL and could silence the γ-globin by forming a strong repressor hub (fig 5.1).

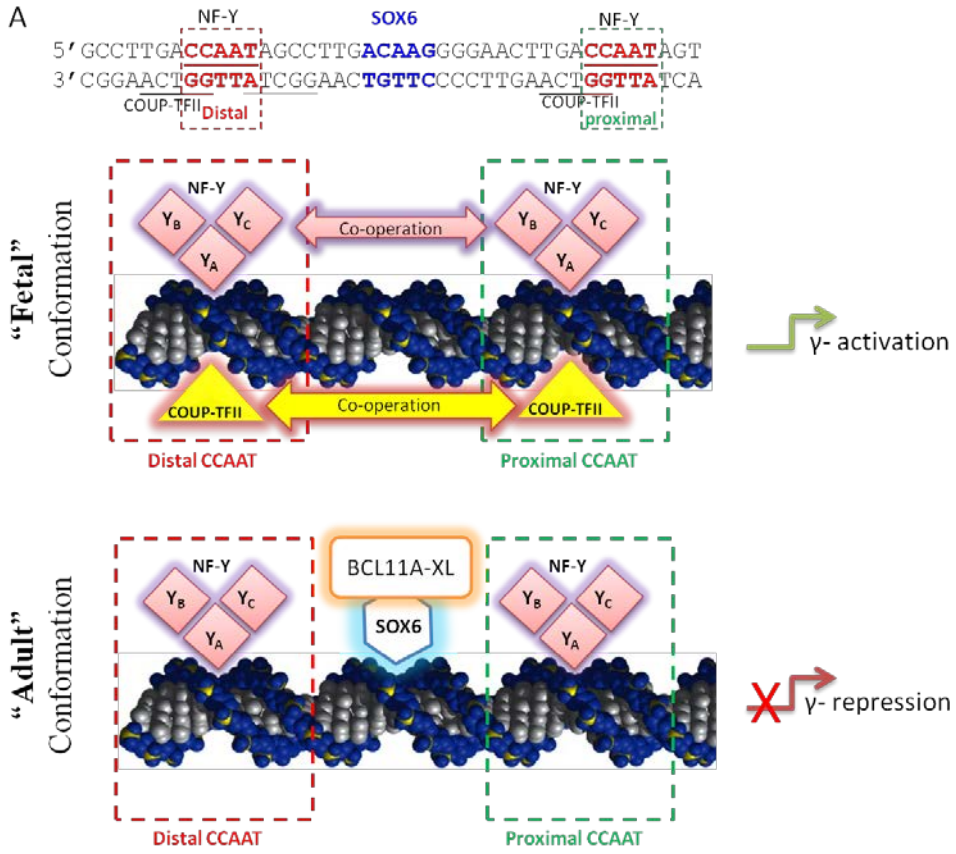


Figure 5.1a: Sequence of the minimal γ-globin promoter from nucleotides -122 to -82 containing both the proximal (green dashed box) and distal CCAAT box region (red dashed box). Sox6 binding site (-106-100, in blue colour) and COUP-TFII binding sites are underlined in black, overlapping with the NF-Y sequence (in red colour). Below the sequence is a graphical model representing the two factors binding during fetal and adult stage. In the fetal stage Sox6 is very low and COUP-TFII interaction with NF-Y promotes γ-globin activation, whereas during the switching, Sox6 accumulates and binds to the proposed Sox6 binding site, thus disrupting the interaction between NF-Y and COUP-TFII. At this stage, other co-repressors, such as BCL11A-XL could be recruited to reinforce the repression.

In the scenario of forced COUP-TFII expression, the chance is that COUP-TFII may bind to the two CCAAT boxes as well as to the

weak half site adjacent to the distal binding site, thus resulting in an active complex for γ expression. In the situation, where both Sox6 and COUP-TFII are present at equimolar concentrations, we expect a milder repression of γ-globin as the binding of Sox6 disturbs the conformational structures of NF-Y binding despite the presence of COUP-TFII thereby, destabilising the NF-Y and COUP-TFII interaction (Fig 5.2). In addition, Sox6 expression could also favour the expression of β-globin by possibly binding to the β-globin promoter.

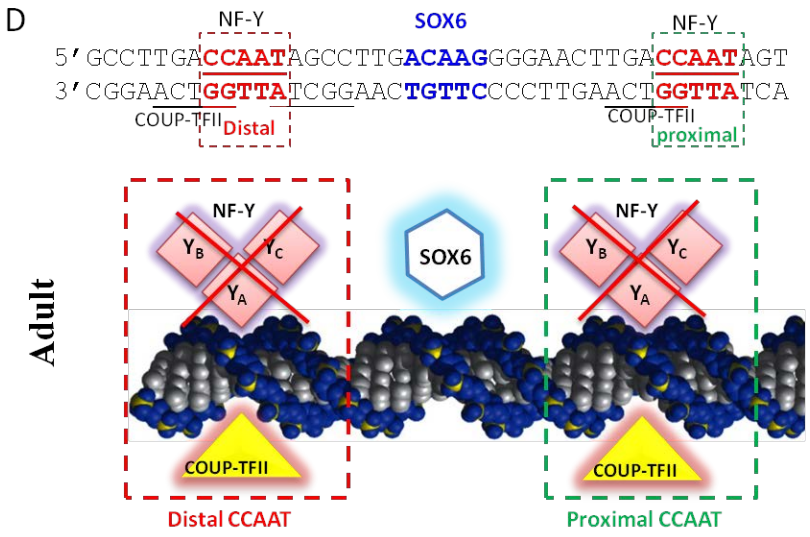


Figure 5.2: Model describing the binding of both Sox6 and COUP-TFII at the same time resulting in a repression of γ-globin expression as the interaction between NF-Y and COUP-TFII is impaired.

Consistent with this hypothesis is the occurrence of several mutations (mostly point mutations) in alleles associated with persistence of fetal hemoglobin in adults (HPFH). Several such alterations, HPFH -117(G>A), -114(C>T), and -110(A>C) and Δ13

involve the distal CCAAT box region. From literature it has been known that -110 (A>C) and -117 (G>A) mutation abolishes the binding of COUP-TFII, but not of NFY still allowing the cooperation between the two CCAAT box resulting in an active state of γ-globin transcription. -114 (G>A) mutation abolishes only NFY but not COUP-TFII where COUP-TFII can bind to the two CCAAT boxes as well as to the weak half site adjacent to the distal binding site and could stabilize the NFY binding to the proximal CCAAT box *per se* a weak NFY binder, resulting again in active complex for gamma globin transcription.

A large number of multiprotein complexes are bound to the γ-promoter region. With the aim of elucidating their role functionally, in the future, we planned to perform gel shift experiments to better understand the binding of one factors with one another at different stages in human erythropoiesis.

THE CLINICAL RELEVANCE

In the last four decades, the major goal of haematologists working on β-thalassemia and sickle-cell anemia has been to discover targets to reactivate fetal haemoglobin. In fact, clinical evidences have so far demonstrated that the maintenance of γ-globin expression through adult life (even modest), is sufficient to ameliorate the clinical condition caused by β-thalassemia and sickle-cell anemia (Forget B., 1998). Till date, a safe and potent HbF-inducing therapy remains to be discovered. Clinical trials with drugs such as 5'-azacytidine, Sodium butyrate and hydroxyurea show a robust HbF inductive response, however the use of these drugs are limited due to their side effects, mutagenicity and was effective only in a small subset of patients (Ley TJ *et al.,* 1982; Platt OS *et al.,* 2008).

The best opportunity for a cure is hematopoietic stem cell (HSC) transplantation, but this treatment is only available to persons who have human leukocyte antigen-matched donors (Gaziev J *et al.*, 2005) making it less successful and the need for more gene therapy methods. Gene therapy is one of the possible approaches for the cure of β-thalassemia, following β-globin gene transfer into hematopoietic stem cells (HSCs) (May C *et al.*, 2000, 2002; Rivella S *et al.,* 2002). However this approach exhibits several critical issues in the control of transgene expression, differentiation and erythroid specificity etc. Despite these, clinical trials based on gene therapy on β-thalassemic patients is currently ongoing in a restricted number of patients. Moreover, the high cost for gene therapy could not be made available

to all patients making it difficult and raises the need for safer alternative methods.

DNA-based approaches to increase γ-globin production have been optimized, including treatment of target cells with lentiviral vectors carrying γ-globin genes (Gambari R, 2012). It was showed that a lentiviral vector encoding a short-hairpin RNA targeting the γ-globin gene repressor BCL11A was able to increase HbF levels from 33% to 45% in β-thalassemic erythroid cells, without compromising erythroid differentiation (Wilber A *et al.*. 2011). Similarly, the use of vectors carrying the β-globin gene together with an HbF inducer (Eg. Mithramycin) leads to forced *de novo* accumulation of β-globin and increased production of γ-globin (Zuccato C *et al.*, 2012) expression. On the basis of these findings, sequences driving the production of shRNAs targeting mRNA encoding a repressor of human γ- globin gene transcription would be of interest for therapeutic purposes.

Direct genome editing is another possible therapeutic modality (Cheng LT *et al.*, 2012). Sequence-specific nucleases such as zinc-finger nucleases or transcription activator-like effector nucleases (TALENs) are tools that can introduce targeted double-stranded breaks in the genome to produce mutations, such as frame shifts or deletions, or to stimulate homologous recombination. For therapeutic purposes, genome editing would require exquisite specificity to prevent off-target mutagenic events.

Hemoglobin switching is indeed a very complex developmental phenomenon involving many players with diverse roles. Some, such as COUP-TFII, Sox6 etc, although having

significant roles in organogenesis and development, also have a very specific role in the switching processes. One among them, COUP-TFII as I have showed in this thesis is an attractive tool to modulate HbF synthesis. Increasing the levels of COUP-TFII expression in an erythroid specific way may be considered to increase fetal hemoglobin in thalassemic patients. Even after identification of a suitable target, the modality to address the target remains a formidable challenge. For multiprotein complexes, critical interfaces between partner protein subunits may be targeted, for example by stabilized peptides (Moellering RE *et al.*, 2009).

Alternatively, identification and characterization of other remodelling complexes and their interactions with transcription factors (Sox6, BCL11A, and GATA-1 etc.) should provide additional details in the γ-globin gene regulation and in the hemoglobin switching. Although targets that directly influence HbF silencing have been more or less identified, translating this knowledge to clinical therapies faces a lot of hurdles.

REFERENCES

Cantú, C., Grande, V., Alborelli, I., Cassinelli, L., Cantú, I., Colzani, M.T., Ierardi, R., Ronzoni, L., Cappellini, M.D., Ferrari, G., Ottolenghi, S. and Ronchi, A. (2011) A highly conserved SOX6 double binding site mediates SOX6 gene downregulation in erythroid cells. Nucleic Acids Res, **39,** 486-501.

Cantú, C., Ierardi, R., Alborelli, I., Fugazza, C., Cassinelli, L., Piconese, S., Bose, F., Ottolenghi, S., Ferrari, G. and Ronchi, A. (2011) Sox6 enhances erythroid differentiation in human erythroid progenitors. Blood, **117,** 3669-79.

Cheng, L.T., Sun, L.T. and Tada, T. (2012) Genome editing in induced pluripotent stem cells. Genes Cells, **17,** 431-8.

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,** 687-97.

Forget, B.G. (1998) Molecular basis of hereditary persistence of fetal hemoglobin. Ann N Y Acad Sci, **850,** 38-44.

Gambari, R. (2012) Alternative options for DNA-based experimental therapy of beta-thalassemia. Expert Opin Biol Ther, **12,** 443-62.

Gaziev, J. and Lucarelli, G. (2005) Stem cell transplantation and gene therapy for hemoglobinopathies. Curr Hematol Rep, **4,** 126-31.

Gaziev, J. and Lucarelli, G. (2005) Stem cell transplantation for thalassaemia. Reprod Biomed Online, **10,** 111-5.

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,** 1469-75.

Liberati, C., Cera, M.R., Secco, P., Santoro, C., Mantovani, R., Ottolenghi, S. and Ronchi, A. (2001) Cooperation and competition between the binding of COUP-TFII and NF-Y on human epsilon- and gamma-globin gene promoters. J Biol Chem, **276,** 41700-9.

Liberati, C., Sgarra, R., Manfioletti, G. and Mantovani, R. (1998) DNA binding of NF-Y: the effect of HMGI proteins depends upon the CCAAT box. FEBS Lett, **433,** 174-8.

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,** 82-6.

May, C., Rivella, S., Chadburn, A. and Sadelain, M. (2002) Successful treatment of murine beta-thalassemia intermedia by transfer of the human beta-globin gene. Blood, **99,** 1902-8.

Moellering, R.E., Cornejo, M., Davis, T.N., Del Bianco, C., Aster, J.C., Blacklow, S.C., Kung, A.L., Gilliland, D.G., Verdine, G.L. and Bradner, J.E. (2009) Direct inhibition of the NOTCH transcription factor complex. Nature, **462,** 182-8.

Platt, O.S. (2008) Hydroxyurea for the treatment of sickle cell anemia. N Engl J Med, **358,** 1362-9.

Ronchi, A., M. Berry, S. Raguz, A. Imam, N. Yannoutsos, S. Ottolenghi, F. Grosveld, and N. Dillon. 1996. Role of the duplicated CCAAT box region in gamma-globin gene regulation and hereditary persistence of fetal haemoglobin. EMBO J. **15**:143-149.

Ronchi, A.E., Bottardi, S., Mazzucchelli, C., Ottolenghi, S. and Santoro, C. (1995) Differential binding of the NFE3 and CP1/NFY transcription factors to the human gamma- and epsilon-globin CCAAT boxes. J Biol Chem, **270,** 21934-41.

Rivella, S. and Sadelain, M. (2002) Therapeutic globin gene delivery using lentiviral vectors. Curr Opin Mol Ther, **4,** 505-14.

Wilber, A., Hargrove, P.W., Kim, Y.S., Riberdy, J.M., Sankaran, V.G., Papanikolaou, E., Georgomanoli, M., Anagnou, N.P., Orkin, S.H., Nienhuis, A.W. and Persons, D.A. (2011) Therapeutic levels of fetal hemoglobin in erythroid progeny of beta-thalassemic CD34+ cells after lentiviral vector-mediated gene transfer. Blood, **117,** 2817-26.

Xu, J., Sankaran, V.G., Ni, M., Menne, T.F., Puram, R.V., Kim, W. and Orkin, S.H. (2010) Transcriptional silencing of {gamma}-globin by BCL11A involves long-range interactions and cooperation with SOX6. Genes Dev, **24,** 783-98.

Yi, Z., Cohen-Barak, O., Hagiwara, N., Kingsley, P.D., Fuchs, D.A., Erickson, D.T., Epner, E.M., Palis, J. and Brilliant, M.H. (2006) Sox6 directly silences epsilon globin expression in definitive erythropoiesis. PLoS Genet, **2,** e14.

Zuccato, C., Breda, L., Salvatori, F., Breveglieri, G., Gardenghi, S., Bianchi, N., Brognara, E., Lampronti, I., Borgatti, M., Rivella, S. and Gambari, R. (2012) A combined approach for beta-thalassemia based on gene therapy-mediated adult hemoglobin (HbA) production and fetal hemoglobin (HbF) induction. Ann Hematol, **91,** 1201-13.

APPENDIX

STR profiling of V-K562 performed at NMS, ITALY

CELL_LINE	K-562	K-562	V-K562
BARCODE	Reference	10716823	10798423
FILE_GENEMAPPER		DNA_13042010	DNA_10052012
AMEL	X X	X X	X X
CSF1PO	9 10	9 10	8 10
D13S317	8 8	8 8	8 8
D16S539	11 12	11 12	11 12
D18S51	15 16	15 16	16 16
D19S433	14 14.2	14 14.2	14 14.2
D21S11	-	29 30 31	30 31
D2S1338	17 17	17 17	17 17
D3S1358	16 16	16 16	16 16
D5S818	11 12	11 12	11 12
D7S820	9 11	9 11	9 11
D8S1179	12 12	12 12	12 12
FGA	21 24	21 24	21 21
TH01	9.3 9.3	9.3 9.3	9.3 9.3
TPOX	8 9	8 9	8 9
VWA	16 16	16 16	16 16
EV % Matching		100%	89.20%

Table 1.0: STR profiling of V-K562 cell line (Column 3) show almost 90% match with the original K562 cells (Column 2) suggesting that the V-K562 cells are genetically very close to the original K562 cells.

ABBREVIATIONS

AGM	-	Aorta-Gonad Mesonephros
APC	-	AlloPhycoCyanin
BCL11A	-	B-cell Lymphoma/Leukemia 11A
BFU-E	-	Blast Forming Unit – Erythroid
CFU-E	-	Colony Forming Unit – Erythroid
CFU-S	-	Colony Forming Unit –Spleen
ChIP	-	Chromatin Immmunoprecipitation
CLP	-	Common Lymphoid Progenitor
CML	-	Chronic Myelogenous Leukemia
CMP	-	Common Myeloid Progenitor
COUP-TFII	-	Chicken Ovalbumin Upstream Promoter-Transcription Factor II
ENCODE	-	Encyclopaedia of DNA Elements
EPO	-	Erythropoietin
FITC	-	Fluorescein Isothiocyanate
GFP	-	Green Fluorescence Protein
GM-CSF	-	Granulocyte Macrophage –Colony stimulating Factor
GWAS	-	Genome Wide Association Studies
HCS	-	High Content Screening

HMG	-	High Mobility Group
HPFH	-	Hereditary Persistence of Fetal Hemoglobin
HSC	-	Hematopoietic Stem Cell
IL-3	-	Interleukin-3
IRES	-	Internal Ribosome Entry Site
LCR	-	Locus Control Region
MBP	-	Myeloid B-cell Precursor
MEP	-	Megakaryocyte Erythroid Progenitor
MLP	-	Myeloid Lymphoid progenitor
MOI	-	Multiplicity Of Infection
MTP	-	Myeloid T-cell Precursor
NGFR	-	Nerve Growth Factor Receptor
PB	-	Peripheral Blood
PE	-	PhycoErythrin
RBC	-	Red Blood Cell
SFFV	-	Spleen Focus Forming Virus
TALEN	-	Transcription Activator Like Effector Nuclease
WPRE	-	Woodchuck-hepatitis-virus Posttranscriptional Regulatory Element

ACKNOWLEDGEMENT

ACKNOWLEDGEMENT

Three and half years back, when I first landed in ITALY, I was a lonely man, ten thousand miles away from home. It has been a long road, but here I am, at the end of my PhD, having gained a wealth of knowledge and experiences not only in science but also in personal life. I am grateful to many people who have helped in innumerable situations to adapt, grow up and made me a better person in life.

First and Foremost, I would like to express my sincere gratitude to my mentor, Prof. Antonella Ronchi for the continuous support during my PhD study and research. I admire her for her immense knowledge, hard-work, enthusiasm, motivation, and tremendous patience to handle a person like me. Her guidance helped me during my research and writing of this thesis. Anto, I'm highly indebted and very thankful for your support inside and outside of the lab. I consider myself very lucky and could not have hoped for a better advisor and mentor in my life. I thank you from the bottom of my heart.

Besides my advisor, I would also like to thank my co-tutor Prof. Sergio Ottolenghi for his guidance, valuable advice, constructive criticism and his extensive discussions of my work. I still remember and admire the amount of efforts he had put in even to sketch out my holiday plans.

I take this opportunity to sincerely acknowledge my program coordinator Prof. Andrea Biondi and the DIMET board for selecting me and providing financial assistance in the form of graduate fellowship. I also thank Prof. Silvia Barabino for her timely help and support.

Some of the results described in this thesis would not have been obtained without collaboration with few laboratories. I owe my gratitude to Dr. Fabio Gaspari and Marta Durlak from NMS, Italy and Prof. Paolo Moi and Dr. Maria Giuseppe Marini from UNICA, Cagliari, Italy.

The road to my Ph.D started in Lab 3021 and it is great fun to work here. I was a part of a small family of my fellow lab mates: Gloria (my bench neighbour), Isaura, Giovanni, Francesca, Domenico and Cristina. I have to thank you for all the special moments we have shared together. You were always supportive and available when I needed you. I extend my warm thanks to my ex-members of the lab Ilaria, Maria, Ileana from early days of my research tenure for training me in cell cultures and other lab techniques.

I also express my sincere thanks to my predecessor Dr. Claudio Cantú for having spent his vacations, teaching me ChIP experiments and supporting me in this work. I want to thank my friends Sriganesh Jammula and Shruti Sinha from IEO for helping me with bioinformatics analysis.

I would like to thank all the members of 3018, 3040 and 3048 lab for your assistance and support and for the memorable experiences we had together. General thanks to all the other people who are not mentioned here, everybody was very helpful and each one very special.

Last but not the least, I would like to convey my regards to my family for their inspiration and support throughout my career. This journey would not have been possible without their support and blessings. I owe everything to them.