γ-Globin Reactivation

Farzin Pourfarzad

The research presented in this thesis was performed at the Department of Cell Biology within the Erasmus University Medical Center, Rotterdam, The Netherlands. The Department is member of Medisch Genetisch Centrum Zuid-West Nederland (MGC).

The research has been supported by NIH grant "Molecular basis of γ-globin gene activation", HL 073455, NGI, NOW, Landsteiner foundation for Blood Transfusion Research (NL), the Center for Biomedical Genetics (NGI, NL) and the FP7 EuTRACC consortium of the European Commission.

The studies described in this thesis were partly supported by a grant from the BSIK Innovation programme; Stem Cells in Development and Disease (SCDD, BSIK).

Financial support by SCDD for the publication of this thesis is gratefully acknowledged.

Financial support by the Erasmus MC, SCDD and J.E. Jurriaanse Stichting for the publication of this thesis is gratefully acknowledged.

Cover Design: Mona Hassanlou Shams
Inside Art by Nazaire: Chapter 1 "Ten Redefined", Chapter 3 and 6 "Hope" Chapter 4 "Need not Suffer Alone" from http://www.sicklecellart.com/
Layout: Sima Kheradmand Kia and Parham Solaimani

Printed by Wöhrmann Print Service, Zutphen

γ-Globin Reactivation

γ-Globine Reactivering

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de rector magnificus

Prof. dr. H. G. Schmidt
en volgens besluit van het College voor Promoties

De openbare verdediging zal plaatsvinden op
Woensdag 28 april 2010 om 13.30

door

Farzin Pourfarzad

geboren te Teheran, Iran

Promotie commissie

Promotor:	Prof.dr. F.G. Grosveld
Overige leden:	Prof.dr. J.N.J. Philipsen Prof.dr. D. Meijer Dr. M. von Lindern

توانا بود هر که دانا بود

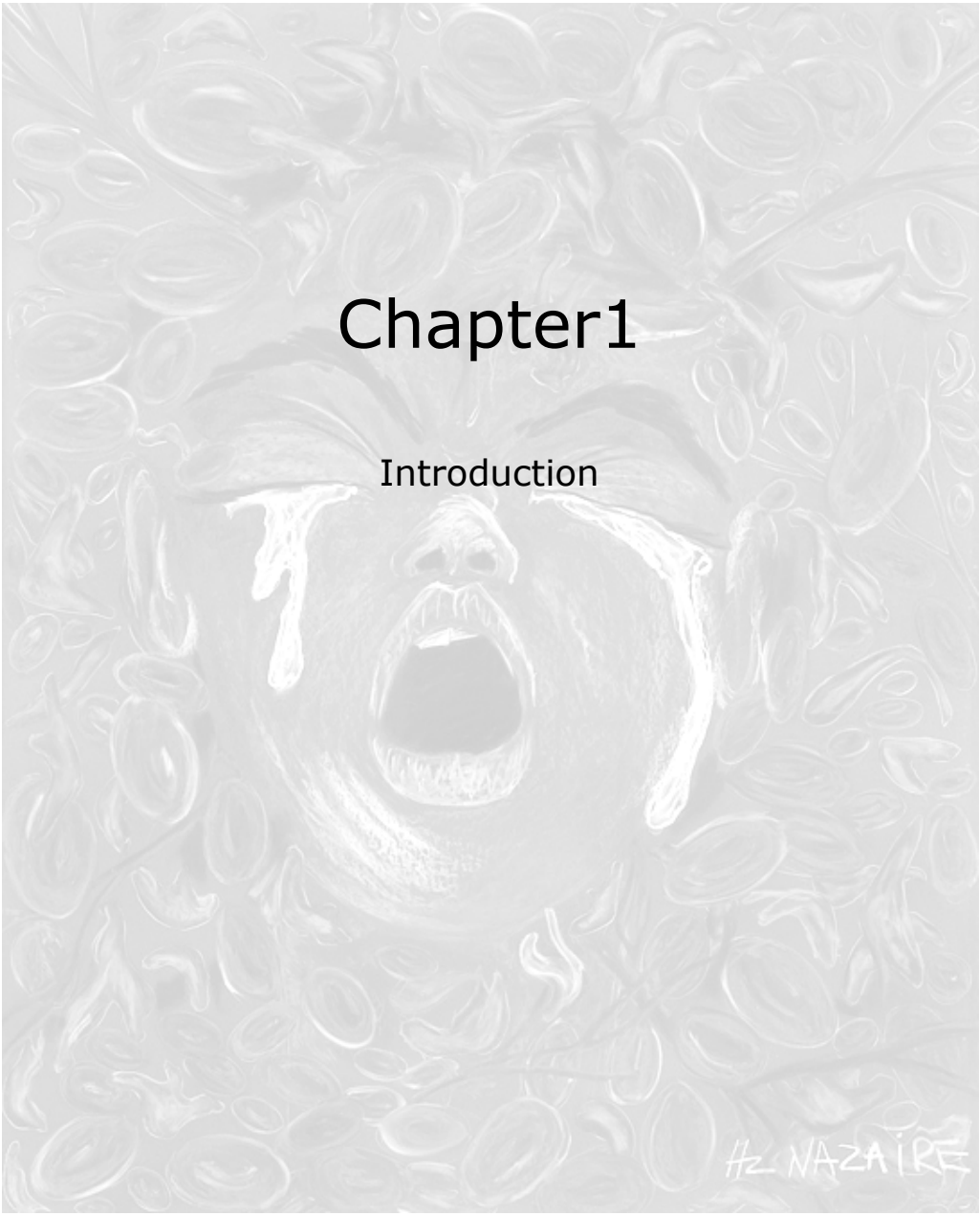
ز دانش دل پیر برنا بود

Capable is he who is wise - Happiness from wisdom will arise

(Hakīm Abu'l-Qāsim Firdawsī Tūsī; Persian poet, 935–1020)

To my parents and
my grandmother

Contents



Chapter1

Introduction

1. Blood

Blood is the tissue mainly responsible for gas exchange (oxygen and carbon dioxide), food (nutrients and minerals), waste traffic and the immune response. It is composed of red blood cells (RBC) or erythrocytes, leukocytes and platelets suspended in plasma. Leukocytes are mainly involved in immune response and platelets are responsible for coagulation in case of tissue damage. Plasma is a solution of proteins, hormones, nutrient, minerals and other necessary substances for the cells of the organism. Plasma also transports waste and carbon dioxide away from the same cells. Erythrocytes are the most abundant cells in the blood and they transport oxygen. Most of the volume of a red blood cell is filled with a protein complex called hemoglobin.

2. Hemoglobin

Hemoglobin (abbreviated as Hb or Hgb) is a heme containing globular protein complex (globin). Hemoglobin was discovered by Hünefeld in 1840 (Hünefeld, 1840). It transports oxygen from the lungs to the tissues. In 1962 Max Perutz and John Kendrew shared the Nobel prize for the solving of the molecular structure of hemoglobin by X-ray crystallography (Perutz, 1960; Perutz et al., 1960).

The Hemoglobin molecule contains four subunits, each containing one heme group embedded in globin polypeptide chain (Fig. 1 left). The heme group contains one iron atom

that is able to bind one oxygen molecule. Hemoglobin is continuously produced from the proerythroblast to the reticulocyte. In mammals, but not in other species such as birds, reticulocytes enucleate. Although residual ribosomal bound mRNA allows further hemoglobin synthesis until the reticulocytes enter the vasculature.

In adult humans, the most common hemoglobin type is called adult hemoglobin, HbA, which is composed of 4 subunit proteins, two α and two β subunits ($\alpha 2\beta 2$) made of 141 and 146 amino acid residues, respectively. These polypeptide chains are non-covalently bound to each other (Fig. 1 right). In the fetus, the hemoglobin molecule is made from 2 α chains and 2 γ chains and called fetal hemoglobin (HbF, $\alpha 2\gamma 2$). After birth, the γ chains are gradually replaced by β chains as the infant grows (Stamatoyannopoulos and Grosveld, 2001). HbF binds oxygen with a higher affinity than adult hemoglobin and as a result, fetal blood in the placenta is able to take up oxygen from maternal blood.

Hemoglobin is saturated with oxygen molecules (oxyhemoglobin) during respiration in the alveoli of the lungs. The oxygen then travels in the blood to be dropped off at cells for use in aerobic glycolysis. At this stage hemoglobin becomes deoxygenated (deoxyhemoglobin) and travels back to the lungs for a new cycle of oxygenation.

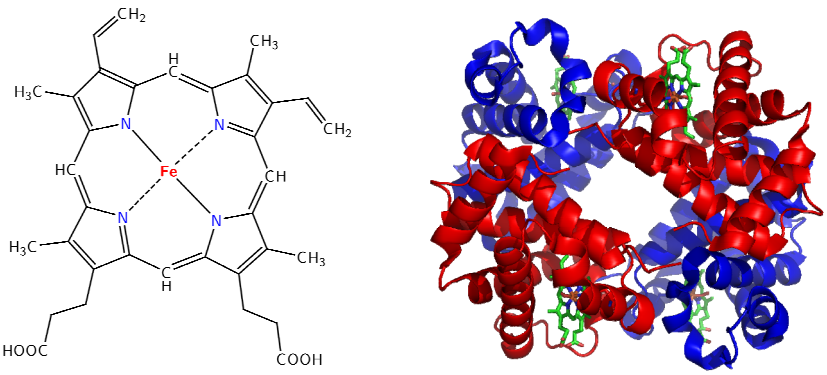


Figure 1. Molecular structure of Heme group (left). 3D structure of Hemoglobin Molecule (right).

2.1. Hemoglobin variants

During development different variants of hemoglobin are produced including pathogenic variants of hemoglobin that are resulted from mutations in globin genes. The best known of these is hemoglobin S (HbS) which causes sickle cell disease (SCD). The most important hemoglobin variants are listed below.

Embryonic globins:

- Gower 1 ($\zeta 2\varepsilon 2$)
- Gower 2 ($\alpha 2\varepsilon 2$)

Fetal globins:

- Hemoglobin F ($\alpha 2\gamma 2$)

Adult globins:

- Hemoglobin A ($\alpha 2\beta 2$) - The most common (over 95%)
- Hemoglobin A2 ($\alpha 2\delta 2$) - 1.5-3.5% of total hemoglobin

Pathologic Variants of Hemoglobin:

- Hemoglobin H ($\beta 4$)- β chain tetramer found in α thalassemia.
- Hemoglobin Barts ($\gamma 4$)- γ chain tetramer found in α thalassemia.
- Hemoglobin S ($\alpha 2\beta^S 2$)- mutation in the β-globin gene found in sickle cell disease.
- Hemoglobin C ($\alpha 2\beta^C 2$) and Hemoglobin E ($\alpha 2\beta^E 2$) – due to mutations in the β- globin gene and cause chronic hemolytic anemia (Bunn, 2001).

3. Hemoglobinopathies

Hemoglobinopathies are genetic disorders that result from mutations in one of the globin genes and hence result in production of abnormal globin polypeptides. Hemoglobinopathies are inherited single-gene disorders. The most common hemoglobinopathies are thalassemias and sickle-cell disease (SCD). Approximately 7% (420 million) of the world's population are carrier of different types of hemoglobinopathies. Hemoglobinopathies are common in ethnic populations from Africa, the Mediterranean basin and Southeast Asia (Weatherall, 1996). Hemoglobinopathies can be the result of either a decreased level of the hemoglobin (anemia) or a decreased efficiency of hemoglobin molecule to bind and transfer oxygen molecule (Newman Dorland, 2007).

3.1. Thalassemia and sickle cell disease

Sickle cell disease (SCD) and β-thalassemia are the most frequent human β-hemoglobinopathies and together with the α-thalassemias, they are the most common human genetic single gene disorders. The high frequency of the SCD and β-thalassemia are due to their selective advantage in malaria infection (Aidoo et al., 2002; Nagel and Roth, 1989). In contrast the homozygous condition shortens the lifespan of the patients. Thalassemias and sickle cell disease occur among individuals of Indian, Mediterranean area, African, Middle East, Southeast Asian or Hispanic origin (Fig. 2) (Mabaera et al., 2008).

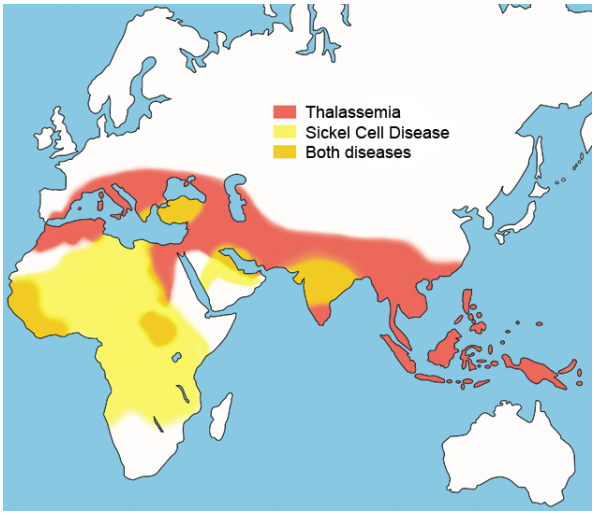


Figure 2. Geographical distribution of Thalassemia and sickle cell disease.

β-thalassemia is a genetic disease due to mutations in β-globin locus causing abnormal, decreased or absence of β-globin production chain of hemoglobin A (HbA). β-thalassemia major occurs in patients homozygous for mutations affecting β-globin synthesis. Infants with β-thalassemia are usually healthy and may survive up to 2-3 years. After this age they need blood transfusions combined with iron chelation therapy to prevent iron overload from lysed red blood cells. To date more than 200 β-thalassemia mutations in β-globin gene have been reported (Table 1)(Patrinos et al., 2004b). Inefficient β-globin production results in free α-globin accumulation and precipitation in erythrocytes. This causes oxidative damage to the membrane of the red blood cells and erythroid precursors and results in apoptosis (Sheridan et al., 1976). These effects cause severe anemia and thereby splenomegaly, skeletal abnormalities and growth retardation.

Affected individuals usually die in their third decade of life. Most of the patients with β-thalassemia die because of cardiac failure from iron overload due to long-term transfusion therapy. Despite advances in iron chelation, median survival is 49 years for patients with optimal chelation regimens (Delea et al., 2007). The cost of treatment is very high, estimated at $30,000 per year. Individuals who are heterozygous for thalassemia may have mild hypochromic anemia.

Sickle Cell Disease (SCD) results from an A-to-T mutation in codon 6 of the β-globin gene. Because of this mutation, a valine residue replaces the normal glutamate of the β-globin protein. This mutated variant of β-globin is called β^S globin and replaces wild type β-globin in HbA molecule. The resulting hemoglobin molecule ($\alpha 2\beta^S 2$) is known as HbS or sickle

hemoglobin (Ingram, 1956; Pauling, 1949). Sickle cell disease is defined as the homozygous condition for the HbS mutation.

Incorporation of β^S globin into the hemoglobin tetramer causes hemoglobin polymerization under deoxygenated condition. Heterozygous carriers of the sickle hemoglobin (HbS) gene, in the absence of another β-globin gene mutation are said to have the sickle cell trait. Polymer formation is less likely in carriers, because each cell contains about 30–40% HbS. The HbS polymer dissociates when hemoglobin is oxygenated and reforms when the molecule is deoxygenated. Cycles of polymerization and de-polymerization cause irreversible damage to the sickle erythrocyte membrane (Lux, 1976), which causes adhesion of sickle cells to the endothelium of the microvasculature (Hebbel, 1991; Hebbel, 1980; Hoover, 1979). This starts a cascade of events that affects erythrocytes, neutrophils, platelets, endovascular cells, and finally the coagulation system resulting in vascular occlusion and the characteristic SCD complications (Frenette and Atweh, 2007).

	Mutations	Anemia	HbA	HbA2	HbF	Reference
Normal	None	None	Normal	Normal	1-2%	
β+ Thalassemia	more than 200 mutations including point mutations, frame shifts, splice variant, deletion, insertions	Severe	Decreased	Normal	5-10 %	(Patrinos et al., 2004b)
β° Thalassemia		Severe	Absent	Normal	5-10 %	(Patrinos et al., 2004b)
(δβ)° Thalassemia	deletion of δ and β – globin genes	Mild	Absent	Absent	10-30 %	(Ottolenghi, 1982)
Deletion HPFH*	Large deletions begin 3' to the Aγ gene and extend to 3' of the structural β-globin gene; the deletion includes intergenic γ-δ sequences and δ and β globin genes	None	Absent	Absent	85-100%	(Katsantoni et al., 2003; Martin et al., 1989; Stamatoyannopoulos and Grosveld, 2001)
Non-deletion HPFH	Aγ -114 C to T, Aγ -117 G to A, Aγ -158 C to T , Aγ -175 T to C, Aγ -195 C to G, Aγ -196 C to T, Aγ -198 T to C, Aγ -202 C to T, Gγ -114 C to T, Gγ -175 T to C, Gγ -200 +C, Gγ -202 C to G	None	Absent	Absent	85-100%	(Berry et al., 1992; Collins, 1985; Gumucio et al., 1991)
Corfu	7.2-kb deletion extends from the γ-δ region upstream of δ gene to the 5' end of the structural δ-globin gene	Mild	Decreased	Absent	85-100%	(Chakalova et al., 2005)

Table 1: Hemoglobin defects in β-thalassemia and related disorders

SCD patients are subjected to many complications, including anemia, acute chest syndrome, painful crises, strokes, splenic and renal dysfunction, bone and joint symptoms, retinopathy, pulmonary hypertension, cholecystitis, hepatic dysfunction, life-threatening infections and skin ulcerations.

The survival for SCD patients in the United States was into their 40s (Platt et al., 1994), while in Sub-Saharan Africa, only half of affected children survive more than 5 years (World Health Organization 2006).

3.2. Hereditary Persistence Fetal Hemoglobin (HPFH)

HPFH is a condition in which a significant amount of fetal hemoglobin is produced during adulthood. HPFH condition is subdivided into deletion and non-deletion HPFH (Table 1). In deletion HPFH, different large regions downstream of the γ-globin genes are deleted whereas in non-deletion HPFH, point mutations in the γ-globin promoter are the cause of elevated levels of HbF. The most widely accepted explanation for this elevation in deletion HPFH is the juxtaposition of downstream enhancer regulatory sequences next to the γ-globin genes (Arcasoy et al., 1997; Feingold and Forget, 1989; Katsantoni et al., 2003).

Co-inheritance of HPFH with β-thalassemia or sickle cell disease can ameliorate the severity of these diseases (Fessas and Stamatoyannopoulos, 1964; Prchal, 1981; Stamatoyannopoulos et al., 1975). High γ-globin expression inhibits α-globin precipitation or HbS polymerization in β-thalassemia and sickle cell disease, respectively (Platt et al., 1994; Poillon, 1993). Because of the ameliorating effect of HbF in HPFH individuals, the genetic mutations resulting to HPFH phenotype is of great interest for the mechanism of γ-globin regulation.

4. β-thalassemia and sickle cell disease treatment

Treatment of SCD and thalassemia major includes chronic blood transfusion therapy, iron chelation and splenectomy. Transfusion of red cells can be lifesaving and useful for treating some complications. This approach produces iron overload and has the risk of viral infection (e.g. HIV, HBV i.e. infected blood products) (Steinberg and Nagel, 2009).

Hematopoietic stem cell transplantation has been successfully used for SCD and thalassemia patients (Bhatia and Walters, 2007; Shenoy, 2007). Despite promising results, stem cell transplantation is only limited to a small proportion of hemoglobinopathy patients, due to age, care and cost considerations. The approach is more likely to be successful in the patients less than 17 years of age (Lucarelli and Gaziev, 2008), and matched sibling donors are available to only 15% of the patients (Bhatia and Walters, 2007).

Correction of the defected β-globin gene or introducing a γ-globin gene is another approach to cure β-thalassemia and SCD. Gene therapy has "cured" sickle cell disease and thalassemia in mice engineered to have these disorders (May et al., 2002; Pawliuk, 2001). This approach has been difficult largely because of inefficient viral transduction of human

hematopoietic stem cells and the required high-level gene expression (Nienhuis, 2008; Sadelain, 2006). Recently a human trial gene transfer protocol for sickle cell disease and β-thalassemia using lentiviral vectors has been introduced successfully by the group of Philipe Leboulch (Cooley's Anemia Symposium 2010, New Nork) and others are about to start organized by the group of Michel Sadelain at Sloan Kettering (Cooley's Anemia Symposium 2010, New Nork).

Gene correction by gene-specific targeting has also been carried out in a humanized sickle cell disease mouse model using reprogrammed autologous induced pluripotent stem (ips) cell derived from fibroblast transduced with retroviral vectors expressing combinations of four transcription factors Oct4, Sox2, Klf4, and c-Myc (Hanna et al., 2007) and offers hope for future application in humans (Carey et al., 2009).

All of the approaches mentioned above require extremely sophisticated medical care that is limited to a number of highly specialized centres and extensive financial resources that are not available to the vast majority of SCD and thalassemia patients.

4.1. γ-globin pharmacological reactivation

Infants with sickle cell disease were reported to have a very mild phenotype when accompanied by high HbF levels (Watson, 1948). Newborns from diabetic mothers had higher fetal haemoglobin levels than normal (Bard and Prosmanne, 1985) and the switch from fetal to adult globins was delayed in mothers treated with insulin (Perrine et al., 1985). It was found that higher concentrations of butyric acid analogues in the mothers augmented γ-globin and inhibited β-globin expression (Perrine et al., 1987). Despite close to 80% sickle hemoglobin (HbS), compound sickle cell trait and HPFH individuals were reported clinically normal with more than 20% HbF (Conley et al., 1963; Edington, 1955). This is because of the ability of γ-globin chains to inhibit HbS polymerization in SCD patients (Goldberg et al., 1977) or to lessen the degree of α-globin chain imbalance and its subsequent precipitation in β-thalassemia erythrocytes (Nathan and Gunn, 1966). Such observations led to the use of pharmacologic compounds to increase HbF production in patients (Benz 1982).

Although any HbF increase is clinically beneficial and reduces mortality risk, at least 20% HbF in all erythrocytes is needed to reduce most clinical symptoms. Unfortunately, this is not within reach of the known pharmacological agents (Papayannopoulou, 1976; Platt et al., 1994; Poillon, 1993).

Many therapeutic compounds have since been analysed for their ability to increase HbF by different means (Haley et al., 2003; Mabaera et al., 2008). Their mechanisms of action are not clear and may overlap. There are basically three different groups of compounds that potentially effectuate γ-globin induction in β-thalassemia and SCD patients which have been studied extensively. Limited clinical trials have been done, but only one drug, hydroxyurea, is routinely used in clinical practice.

5-azacytidine (5-Aza) a nucleoside analogue, was the first HbF inducing agent used initially in baboons and later in patients. (DeSimone, 1982; Dover, 1983; Ley 1983; Ley, 1982). Because of its toxicity an analogue of 5-azacytidine, 5-aza-2'-deoxycytidine (decitibine) with less cytotoxic effects can be used for treatment. Decitibine can induce HbF synthesis close to 7% of total hemoglobin (Koshy et al., 2000).

5-Aza is a DNA methyltransferase (DNMT) inhibitor and two hypotheses have been proposed to explain its activity. The first one is based on its general DNA methylation inhibition, including the γ-globin genes, resulting in their induction (Lavelle et al., 2006b; Saunthararajah et al., 2004). The second hypothesis is based on 5-Aza's cytotoxic effects on differentiating erythroid cells. In this model, the treated bone marrow responds with rapid erythroid proliferation and an increased proportion of HbF containing cells (Stamatoyannopoulos et al., 1985).

Short chain fatty acids like butyrate are histone deacetylase (HDAC) inhibitors and they can induce HbF expression by binding to transcriptionally active elements, so-called butyrate response elements, in the 5'-flanking region of the gene (Glauber et al., 1991; Ikuta et al., 1998; Pace et al., 1996; Weinberg et al., 2005). This effect leads to histones H3 and H4 hyperacetylation at the γ-globin genes promoters (Fathallah et al., 2007). Butyrate also increases translational efficiency of γ-globin mRNA (Weinberg et al., 2005). It has been demonstrated that sodium butyrate increased HbF in baboons and erythroid cells of β-thalassemia and SCD patients (Constantoulakis et al., 1989; Perrine et al., 1989; Perrine et al., 1987). The result of butyrate trials in sickle cell disease and β-thalassemia infused with 2- to 3-week intervals was inconsistent, probably due to its antiproliferative effects (Blau et al., 1993; Dover et al., 1994; Perrine et al., 1993; Sher et al., 1995). It restrains cell growth by inhibiting histone deacetylase, cyclin D1, D and E activity (Boosalis et al., 2001).

Hydroxyurea (HU) is a ribonucleotide reductase inhibitor and the sole drug for treating sickle cell disease approved by the US Food and Drug Administration (FDA)(Charache et al., 1995). It is known that HU is a DNA replication inhibitor, but how it induces γ-globin gene expression is poorly understood. Most patients respond to HU treatment and about 10–20% of patients appear to be non-responders. This difference in response may be genetically regulated, which will be discussed later in this chapter (Kumkhaek et al., 2008; Ma et al., 2007; Sedgewick et al., 2008).

HU treatment reduces the need for blood transfusion, hospitalization and acute chest syndrome in more than 40% of the cases (Charache et al., 1995) and can increase HbF production close to 20% (Ferster et al., 2001; Gulbis et al., 2005; Kinney et al., 1999; Wang et al., 2001; Zimmerman et al., 2004). Hydroxyurea increases the red cell volume and presumably alters cellular hydration, and also reduces the white cell count. All of these effects may play a role in reducing painful crises (Scott et al., 1996). Several mechanisms have been proposed for the induction of HbF by HU including rapid erythroid regeneration, increased

erythropoietin (EPO) production (Stamatoyannopoulos, 1990), apoptosis, nitric oxide (NO) production (Cokic et al., 2008), affecting the guanylate cyclase (sGC) (Cokic et al., 2006) and p38 MAPK pathway (Park et al., 2001) which in turn can modulate γ-globin expression.

Erythroid progenitors have higher HbF levels compared to terminally differentiated cells (Papayannopoulou, 1976, 1977). In stressed conditions more F-cell are produced (Blau et al., 1993) however, cytotoxic agents like HU and 5-Aza kill late erythroid progenitor cells and trigger stressed erythroid regeneration which leads to more F cell formation and higher HbF levels (Dover et al., 1986; Torrealba-de Ron et al., 1984).

Unfortunately, none of these γ-globin inducing agents is efficient and safe enough that would make them applicable to most hemoglobinopathy patients worldwide. Except HU, administration of these drugs are very inconvenient, for example butyrate treatment requires infusions over several days each month or many pills in the case of 5-Aza. In most cases their effect is pleiotropic, reversible and suppress erythropoiesis (Fathallah and Atweh, 2006). On the other hand, since they create mutations and nonspecific epigenetic modification, they cause changes in global gene expression patterns and hence carry an increased risk of cancer development or other diseases when used over extended periods of time.

5. Erythropoiesis and switching

During embryogenesis there are two waves of hematopoiesis, namely primitive and definitive hematopoiesis. The first population consists of large, nucleated red cells originated in the yolk sac. These primitive red cells are replaced by the second population of definitive red cells that are smaller and enucleated and produced in fetal liver and bone marrow. In humans, these two waves are associated with the production of three different types of hemoglobin: embryonic (ζ2ε2) during the first few weeks of life, fetal (α2γ2) during the rest of gestation until around birth and adult (α2β2) during the rest of life. These hemoglobin switches are parallel to the changes in the site of hematopoiesis, (Wood, 1977). Primitive erythropoiesis is first detectable in the yolk sac day 14 to 19 post-conception and persists in this organ until the 9th week of gestation. Yolk-sac-derived primitive erythrocytes undergo a partial hemoglobin switch at week 5. Before this time yolk sac erythroblasts synthesize Hb Gower I (ζ2ε2), but at weeks 6 to 8, they also synthesize large amounts of Hb Gower II (α2ε2) and HbF (α2γ2)(Gilmour, 1941; Peschle, 1985) (Fgure 3 left).

The definitive wave of erythropoiesis starts in the aorta-gonado-mesonephros of the embryo (Tavian et al., 1996) and is first detectable during week 6 in the fetal liver. Erythroblasts produced in fetal liver express α, γ, and small amounts of β-globin but the ε and ζ-globin genes are rapidly silenced while the α and γ-globin genes remain expressed at high level until around birth. Bone marrow erythropoiesis which is first detectable around week 11 becomes the major site of erythropoiesis and expression of the β-globin gene (and a low level δ-globin gene expression), which almost completely replaces γ-globin gene expression (Fig. 3 left).

6. The β-globin locus and its regulation

The human β-globin locus on chromosome 11 is one of the best characterized loci in the human genome. It was the first mammalian sequenced gene (Proudfoot, 1976); one of the first eukaryotic cDNA's cloned in a bacterial vector (Rougeon et al., 1975); the first mammalian gene discovered to have intervening sequences (introns) (Jeffreys and Flavell, 1977); the first transfected gene expressing RNA from its own *cis*-regulatory sequences (Mantei, 1979); the first locus shown to have a Locus Control Region (LCR) (Grosveld et al., 1987); the first gene shown to be transcribed intermittently (Milot et al., 1996); the first locus shown to be organised in loops around chromatin hubs (Tolhuis et al., 2002); the first locus shown to change loops (long range interactions) during development (Palstra et al., 2003) and the first locus used to show the formation of the chromatin hub dependence on specific factors (Drissen et al., 2004).

The human β-globin locus contains five functional β-like globin genes and an upstream regulatory element, the locus control region (LCR) which is composed of five DNase I hypersensitive sites (HS1-HS5; Fig. 4 top). These genes are arranged spatially in the order of their expression during ontogeny, 5'-ε-Gγ-Aγ-δ-β-3', in a tissue- and developmental- specific manner (Fig. 3 left and 4 top).

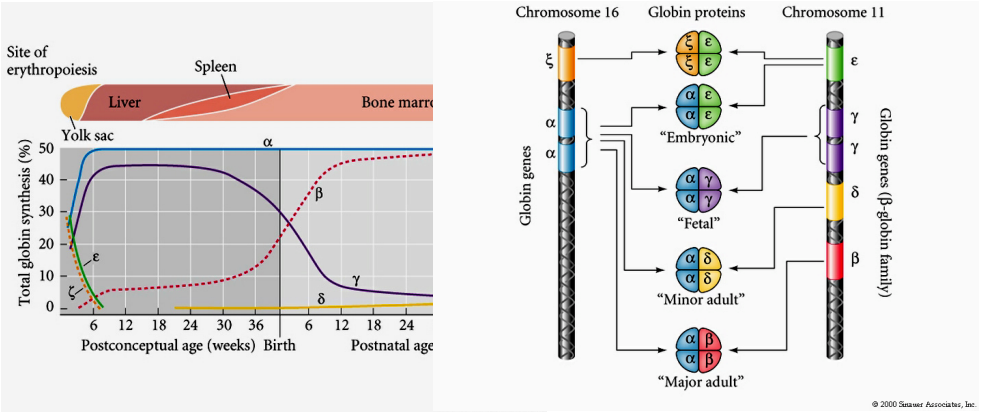


Figure 3. Site of hematopoiesis and globin gene expression switching during development (right). Different globin gene loci and Hemoglobin composition (left).

6.1. The locus control region (LCR)

LCR, originally termed a dominant control region, is located between 5-25 kb of the most upstream globin gene, ε-globin. The LCR is a strong enhancer and regulator of expression of the downstream genes (Fig. 4 top). It overcomes heterochromatin suppression

thus conferring integration position-independent expression to the human globin transgenes (Fraser and Grosveld, 1998). Another property of the LCR is its ability to provide tissue specific expression to linked genes (van Assendelft et al., 1989). The LCR is composed of 5 DNase I hypersensitive sites (5'HS1-5), which are short stretches of DNA. Every individual HS is DNA sequence with the length of approximately 300 bp (Lowrey et al., 1992; Philipsen, 1990; Pruzina et al., 1991; Stamatoyannopoulos et al., 1995; Talbort and Grosveld, 1991) where one or two nucleosomes have been displaced. The individual HSs contain clusters of different binding motifs for several erythroid specific transcription factors like GATA-1, EKLF, NF-E2 and other ubiquitous DNA-binding proteins (Fig. 4 bottom).

These HSs interact with each other (Ellis, 1996; Tolhuis et al., 2002), and with additional 5' HS-111 and 3'HS1 and specific active globin genes to create a chromatin hub (CH) by looping out the intervening sequences (see blow, Fig. 7) (Carter et al., 2002; de Laat and Grosveld, 2003; Fang et al., 2007; Patrinos et al., 2004a; Tolhuis et al., 2002). When the gene is active additional loops are formed to generate an Active Chromatin Hub (ACH). Spatial structure formation of ACH is dependent on many erythroid-specific and general transcription factors including GATA, FOG, EKLF, IKAROS, NF-E2 p18, CTCF and BRG1 (Drissen et al., 2004; Du et al., 2008; Keys, 2008; Kim et al., 2009; Splinter et al., 2006; Vakoc et al., 2005). During erythroid differentiation these interactions establish contacts between the LCR and active genes and coincide with high levels of β-globin gene expression (Palstra et al., 2003).

LCR 5'HS2 is a strong enhancer (Ney et al., 1990; Tuan et al., 1989) mediated by binding of NF-E2/AP-1 (Talbort and Grosveld, 1991). In addition, transactivation by 5'HS2 is mediated by the histone acetyltransferase (HAT) CREB binding protein (CBP) (Johnson et al., 2002). Johnson et al showed that NF-E2 recruits CBP as a critical step in transactivation, but additional components of 5'HS2 are required to achieve maximal enhancer activity.

LCR Associated Remodeling Complex (LARC) is another complex that is recruited to 5'HS2 NF-E2 recognition element. LARC consists of heterogeneous nuclear ribonucleoprotein C1/C2, nucleosome remodeling SWI/SNF, and nucleosome remodeling and deacetylating (NuRD)/MeCP1 as a single heterogeneous complex. It is not clear whether LARC and NF-E2 compete or cooperate for the binding to the same site. It has also been shown that LARC binds to the γ and β globin promoters (Mahajan et al., 2005). The LARC complex is interesting since it contains activator and repressor complexes at the same time. It suggests that LARC is recruited to the 5'HS2 LCR via its NuRD/MeCP1 complex and interacts with globin promoters via SWI/SNF complex to facilitate transcriptional machinery recruitment.

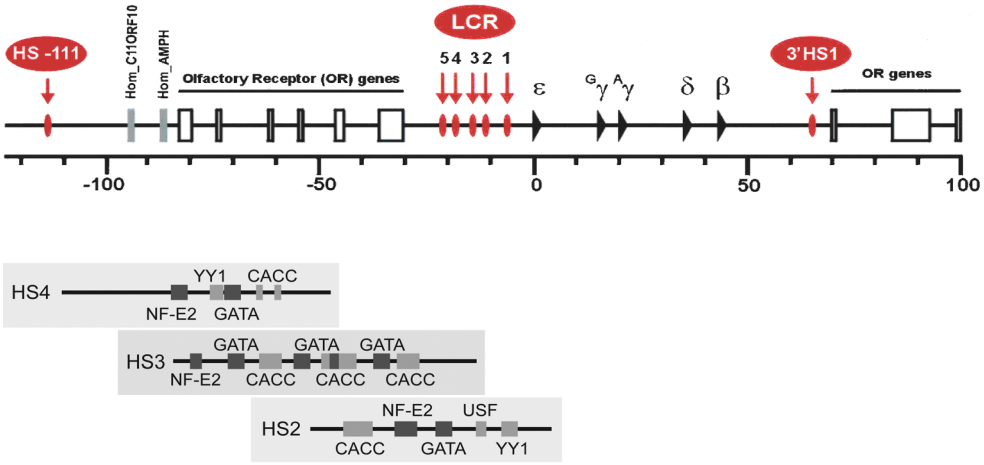


Figure 4. Schematic representation of the human β-globin locus (top) (Patrinos et al., 2004a). Transcription binding sites in individual Hypersensitive sites (bottom).

LCR 5'HS3 is the most powerful element to confer active chromatin structure (Ellis, 1996). Erythroid Krupple Like Factor (EKLF) is the active factor at HS3. It induces DNase I hypersensitivity and cooperates with the SWI/SNF family of proteins to modify chromatin structure (Gillemans et al., 1998).

LCR 5'HS4 deletion in transgenic mice carrying the human globin locus had no effect on globin gene expression during embryonic erythropoiesis but significantly decreased γ- and β-globin gene expression during definitive erythropoiesis. Thus, it seems to be required for globin gene expression only during definitive erythropoiesis. Interestingly it contains a CTCF binding site that is involved in looping before the genes are expressed (Tolhuis et al., 2002). Absence of the 5'HS4 core element may limit the ability of the LCR to provide an open chromatin domain (Navas et al., 2001).

LCR 5'HS5 and 3'HS1 are two DNase I hyper sensitive sites located on the far 5' and 3' ends of the β-globin locus**.** 5'HS5 of the LCR functions as a chromatin insulator *in vivo* and protects a position-sensitive Aγ-globin gene against position effects in transgenic mice (Li et al., 2002). Conserved CTCF sites were found at 5'HS5 and 3'HS1. All of these sites bound to CTCF *in vitro* and also function as enhancer blocker (Farrell et al., 2002; Hou et al., 2008). Human HS5 is the homologue of the Chicken β-globin 5'HS4 (cHS4) (Hardison et al., 1997; Li et al., 1999). cHS4 has both enhancer blocking and insulator (barrier) activity. cHS4 marks the border between the active chromatin and heterochromatic region (Hebbes et al., 1994; Litt et al., 2001; Prioleau et al., 1999)and alleviate the silencing effects by binding to USF1

(Huang et al., 2007). cHS4 also blocks the interaction between an enhancer and promoter via CTCF binding (Bell et al., 1999; Chung et al., 1993). 5'HS5 is believed to have a similar insulator activity on the same fragment just like chicken 5'HS4, because it has enhancer blocking and transgene protection activities (Farrell et al., 2002; Li and Stamatoyannopoulos, 1994; Li et al., 1999; Tanimoto et al., 2003). On the the other hand HS5 deletion had no effect on the linked globin genes transcription(Li et al., 2002; Reik et al., 1998; Wai et al., 2003). It is suggested that HS5 functions as border elements with developmental specificity (Wai et al., 2003). 3'HS1 and 5'HS5 are thought to function as insulators that prevent inappropriate interactions between β-globin regulatory elements and those of neighboring domains or subdomains, such as olfactory receptor genes. However actual evidence for this proposal has not been reported. It is known that CTCF is directly involved in chromatin architecture and mediates enhancer-promoter interactions in the globin locus (Splinter et al., 2006).

6.2. Globin Gene Promoters

All of the globin genes have a number of regulatory elements that are important for their tissue and developmental stage specific expression (Fig. 5). It has been shown that these elements are important for interaction between LCR and globin gene promoters (Palstra et al., 2003; Patrinos et al., 2004a; Tolhuis et al., 2002). The regulatory elements contain binding motifs for several erythroid specific and ubiquitously expressed transcriptional activators or suppressors. These factors interact with each other forming multimeric complexes that result in chromatin structure changes to allow the formation/activation of the transcription machinery at the 5'end of the genes.

In addition to erythroid specific binding motifs like GATA, EKLF and others, globin gene promoters contain a TATA box at -30, a CAAT box at around -70 with the exception of γ globin genes which has two CAAT boxes at -85 and -115, and a single CAC box at -110 in ε promoter and at -145 in the γ promoters. The β-globin gene has two CAC boxes at -90 and -110 (Fig. 5)(Stamatoyannopoulos and Grosveld, 2001).

6.2.1-The γ-globin gene promoters and their regulators

The regulation of these genes is of special interest because expression of the γ-globin genes in β-thalassemia and sickle cell disease ameliorates the effects of these diseases. Several natural mutations and transgenic mouse experiments have indicated that the first 300bp of the γ globin gene promoters play a primary role in the suppression of the genes in the adult. The mutations are categorized as deletion and non-deletion mutations (Table 1) and the deletion mutations are further subdivided into different types of thalassemia and deletion Hereditary Persistence Fetal Hemoglobin (HPFH). Non-deletion mutations consist of point mutations in the γ-globin promoter and cause non-deletion HPFH. Most of the non-deletion HPFH mutations

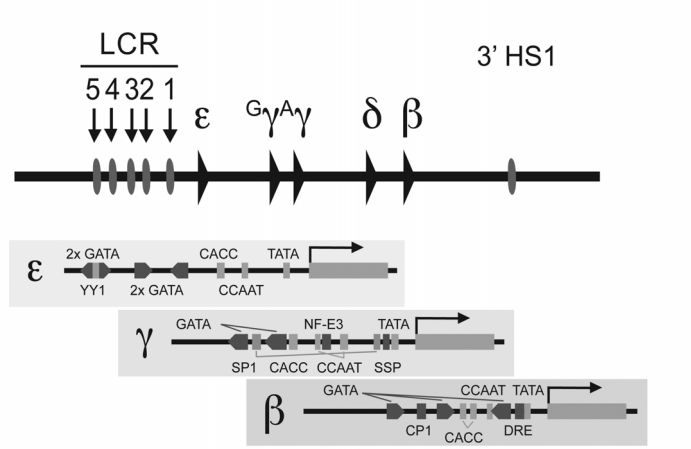


Figure 5. Regulatory elements in β-globin like genes.

occur in transcription factor binding sites. They create new binding motifs or destroy existing sites (Liu et al., 2005).

The two γ-globin promoters are identical in sequence. Several transcription factor-binding motifs exist within the γ-gene promoter. They contain a TATA box, a duplicated CAAT box and a single CACCC box (Fig. 6).

The sequence between the TATA box and CAAT is GC reach that influences Sp1 and Stage selector protein (SSP) to the γ-globin gene promoter. It has also been suggested that CpG dinucleotide methylation of this region modulate SSP and Sp1 binding. DNA methylation is mediated by methyl-CpG binding domain 2 protein (MBD2) (Rupon et al., 2006). It has been shown that CpG de-methylation increases Sp1 binding activity by 10 fold. Dinucleotide methylation in the adult is proposed to result in SSP replacement as an activator by Sp1 binding as a repressor. At the early stage SSP would give a competitive advantage to γ-globin genes over β globin (Jane et al., 1993b; Jane et al., 1992). SSP contains the erythroid-specific protein NF-E4, which has a short (14 kDa) and a long (22 kDa) isoform, and the ubiquitous transcription factor CP2. While the long isoform acts as an activator, the short isoform has been shown to repress γ-globin gene expression (Zhao et al., 2006). Forced p22NF-E4 expression in β-globin locus containing transgenic mouse delays the γ-globin switching (Zhou et al., 2004), however the switch is fully completed in the adult bone marrow suggesting that p22 NF-E4 is capable of influencing human globin gene expression but not able to override the intrinsic mechanisms of γ-globin gene silencing.

MBD2 is a subunit of methyl-CpG binding protein repressor complex 1 (MeCP1), it interacts with the NuRD complex and directs it to methylated DNA (Zhang et al., 1999). In addition, the level of γ-globin expression is higher in MBD2-/- mice in embryonic days 14.5- and 16.5 in fetal liver erythroblasts suggesting a role for MBD2 in embryonic/fetal erythroid

development. DNA methylation levels are modestly decreased in MBD2-/- mice. MBD2 does not bind to the γ-globin promoter region to maintain γ-globin silencing (Rupon et al., 2006). Changes in MBD2 that would result in a differential regulation of γ-globin expression could be an indirect effect via other protein factors that are NuRD interactors and supposedly have a role in γ-globin silencing such as GATA-1, BCL11A, IKAROS and others.

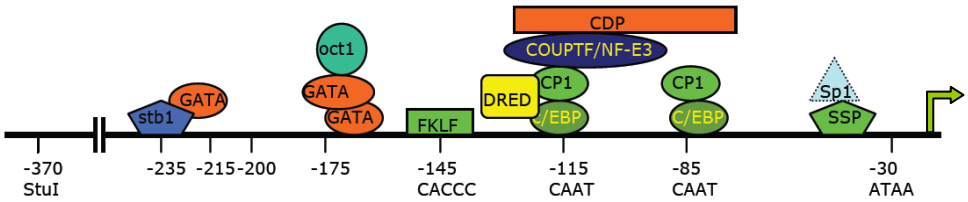


Figure 6. Schematic representation of human γ-globin gene promoter with (possibly) bound factors.

The duplicated CCAAT box is a unique feature of the γ-globin genes and they are mutation hot spots in HPFH (Table 1). A number of proteins may bind to the CCAAT motif, such as C/EBP (CCAAT/enhancer-binding protein), and may act as a positive regulators (Zafarana G and Grosveld F. unpublished data)(Wall et al., 1996). NF-Y/CBF (CCAAT-binding factor, also called CP1) also acts as an activator (Fucharoen et al., 1990), while CDP (CCAAT displacement protein) could bind as a repressor by competitive replacement of the activators (Skalnik et al., 1991). NF-Y/CBF requires a high degree of conservation of the CCAAT sequence and is considered to be the major protein recognizing the CCAAT box. NF-Y is one of the factors that specifically bind to the duplicated CCAAT boxes in vivo (Duan et al., 2001). Furthermore, NF-Y recruitment to the duplicated CCAAT box region may facilitate the recruitment of the basal transcription machinery. NF-Y binding to the CAAT box also appears to participate in the communication between the γ-gene promoter and the LCR complex (Duan et al., 2001).

CCAAT-box region of globin promoters also contain one or two direct repeats of a short motif analogous to DR-1 binding sites (TGACCAATAGCCT) for non-steroid nuclear hormone receptors. The NF-E3 complex, that is related to orphan nuclear receptors, can bind to these repeats and acts as a developmental repressor (Liberati et al., 2001). In addition, the structure and abundance of NF-E3/COUP-TF complexes vary during fetal liver development. These results suggest that NF-E3 may play a role in globin gene expression suggesting that nuclear hormone receptors are involved in the control of globin gene switching (Berry et al., 1992; Filipe et al., 1999; Gumucio et al., 1988; Ronchi et al., 1995).

Direct Repeat Erythroid Definitive (DRED) is an erythroid ε -globin gene repressor activity that binds with high affinity to DR1 sites in the human ε- and γ-globin promoters. It is a protein complex that contains the nuclear orphan receptors TR2 and TR4. The –117 Greek HPFH mutation shows reduced TR2/TR4 binding in vitro and its overexpression shows reduced

γ-globin expression in murine cells (Tanabe et al., 2002). However chromatin precipitation analysis of DRED shows that it is not bound to the suppressed γ globin gene promoter in vivo (K. Kolodziej, unpublished data). TR2 and TR4 also bind to DR element on the GATA1 hematopoietic enhancer (G1HE) and repress GATA1 transcription (Tanabe et al., 2007b). The slight γ-globin upregulation observed in TR2/TR4 knock out mice (Tanabe et al., 2007a) could also be due to an indirect effect on GATA1 expression rather than direct effect on the γ-globin promoter.

The **TATA and CCAAT** box of the γ-globin gene appear to have distinct roles in development. These promoter elements are essential for high-level γ gene expression in definitive erythropoiesis, they appear to play a minor role in embryonic erythroid cells. The CCAAT and TATA mutations in the human γ-globin genes promoter impair recruitment of TBP, TFIIB, and Pol II in adult splenic erythroblasts, but not in embryonic erythroid cells (Duan et al., 2002; Fang et al., 2004).

The **CACCC box** was shown to be bound by Sp1 and Sp3 in vitro (Marin et al., 1997). FKLF and FKLF-2 are also able to bind to the CACCC box but the role of any of these factors in γ-globin expression remains to be determined (Asano et al., 1999).

There are a number of other sites in the upstream region of γ-globin gene promoter that have regulatory roles in γ-globin gene expression. Most of these are known from the analysis of their effect in non deletion HPFH individuals. The most important ones are at -175 (Stoming et al., 1989; Surrey et al., 1988) which alters a GATA-1 binding site (Mantovani et al., 1988), -200, -198, -202 (Table 1).

The **PYR complex** binds to a 250-bp polypyrimidine (PYR) rich DNA sequence 1 kb upstream of the human δ-globin gene. The PYR DNA-binding site is included in the Corfu deletion (O'Neill, 1991), a 7.2 kb deletion in the region between γ and δ globin genes. Patients with the Corfu deletion express very high levels of HbF, they have very mild anemia and no transfusion requirements (Chakalova et al., 2005; Kulozik et al., 1988; Wainscoat et al., 1985). The transcription factor IKAROS which is predominantly expressed in adult hematopoietic cells, is the DNA-binding subunit of the PYR complex (O'Neill et al., 2000). Deletion of 511 bp of DNA including the PYR binding site upstream of the human δ-globin gene or IKAROS knock out mice delays human γ to β switching in human β-globin locus transgenic mice (Enver et al., 1990; Lopez et al., 2002; O'Neill et al., 1999). In addition to IKAROS the PYR complex contains the SWI/SNF complex that activates gene transcription and the HDAC-containing NURD repressor complex (O'Neill et al., 2000; O'Neill et al., 1999). However experiments in transgenic mice containing the LCR and a γ-globin gene lacking the downstream region of γ-globin genes demonstrated that γ-globin silencing is autonomous (Arcasoy et al., 1997; Calzolari et al., 1999; Dillon and Grosveld, 1991; Katsantoni et al., 2003). It should be noted that deletion of the PYR region can only cause a mild delay in switching (O'Neill et al., 1999) and deletion of a 2.5 kb region downstream of the γ-globin gene which includes the Corfu deletion does not influence γ-globin silencing, instead β-globin expression is decreased (Calzolari et al., 1999). Moreover, the IKAROS (Ikzf1) null mouse has

different hematopoietic defects in addition to a delay in γ-globin switching (Lopez et al., 2002). Whether the increased γ-globin expression in the transgenic mice is due to a direct or an indirect effect on hematopoiesis is not clear.

7. Genetic determinants of HbF production

Genome wide association studies in β thalassemia and sickle cell patients showed an association of HbF level with several quantitative trait loci (QTL) including BCL11A, the HBS1L-MYB region on 6q23, the *Xmn*I polymorphism in the proximal promoter of the Gγ-*globin* gene, chromosome 8 and the X-chromosome (Chang et al., 1995; Chang et al., 1997; Garner et al., 2002; Labie et al., 1985a; Labie et al., 1985b; Lettre et al., 2008; Solovieff et al., 2009; Thein et al., 1994; Uda et al., 2008; Wyszynski et al., 2004).

The most important of these is BCL11A. The potential role of BCL11A in haemoglobin switching was first suggested by genetic association studies in human (Lettre et al., 2008; Menzel et al., 2007; Uda et al., 2008). BCL11A shows stage specific expression in human and mouse. The protein has short and long isoforms in human and only a long isoform in mouse. The short isoform is expressed in fetal and embryonic hematopoietic tissues, the long isoform in definitive hematopoietic tissues. BCL11A down regulation increases fetal globin gene expression in human erythroid progenitor cells (Sankaran et al., 2008) . Bcl11a knock out mouse fails to restrict the mouse embryonic and fetal globins (εy and βh1) to the primitive hematopoietic tissues. Moreover, Bcl11a knock out mouse containing the complete human β-globin locus expresses high level of γ-globin in definitive hematopoietic cells. It is suggested that the differences in BCL11A expression patterns between man and mouse might be responsible for the divergence of globin switching between these two species during the course of evolution (Sankaran et al., 2008; Sankaran et al., 2009).

BCL11A binds to the LCR HS3 and γ-δ globin intragenic region but not to the promoters (Chen et al., 2009; Sankaran et al., 2008). Many mutations in the globin locus are reported that reverse γ-globin silencing in adults, regardless of the high expression of BCL11A repressive isoform in man and mouse. This suggests that, although BCL11A has a very important role in the switching process, it is not the only factor involved. BCL11A physically interacts with the components of the NuRD complex, GATA and FOG, but whether these interactions are important for the switching process remains to be investigated.

8. Epigenetic modifications of the β-globin locus

As observed with other genes, epigenetic modifications and DNA methylation also play role in the regulation of globin genes. In embryonic day 11.5 yolk sac blood, the LCR and the ε-γ domain are highly enriched for acetylated histone 3 (H3ac), and methylated histone 3 lysine 4 (H3K4me3). These active histone modification marks are less frequently found in δ-β region. The H3K4me3 and H3ac enrichment in ε-γ region compare to δ-β region correlates well with the high transcription levels of ε- and γ-globin at embryonic day 11.5. Adult erythroid

cells are very different compare to embryonic day 11.5 cells. The LCR and the δ-β domain region are highly enriched for H3K4me3, H3K4me2 and H3ac while these histone modifications are absent at the ε-γ domain. The presence of these histone modifications in the LCR and δ-β domains are consistent with active transcription of these regions in definitive erythropoiesis of adult cells (Fathallah et al., 2008; Lavelle et al., 2006a; Miles et al., 2007; Yin et al., 2007).

Methylation patterns of the -50 region of γ-globin promoter regions inversely correlate with γ-globin expression (Jane et al., 1993a; Mabaera et al., 2007; Sengupta et al., 1994). There are a number of compounds that are able to increase γ-globin gene expression by inhibition of DNA methyl transferases (5-Azacytidine) or inhibition of histone deacetylases (sodium butyrate) (DeSimone, 1982; Dover, 1983; Fathallah et al., 2007; Ley 1983; Ley, 1982). These data show that epigenetic regulation plays an active role in γ-globin gene expression. The role of MBD2 in DNA methylation is mentioned above (Rupon et al., 2006). Protein arginine methyltransferase (PRMT5) is involved in symmetric methylation of histone H4 arginine 3 (H4R3me2s) and is required for subsequent DNA methylation by the DNA methyltransferase (DNMT3A) (Zhao et al., 2009).

9. The γ-globin gene promoter: LCR interaction

It has been shown that only one of the globin genes is active at any particular moment in time because the LCR alternates between the fetal and adult genes for their activation during γ-globin gene switching (Wijgerde et al., 1995). This suggests that LCR regulatory elements can only contact a single gene at a given time and that (at least part of the) loops are unstable and continuously form and fall apart. The 3C studies on β-globin locus demonstrates that in primitive erythroid cells the chromatin hub (CH) interacts alternatively with the promoter of the active embryonic globin genes form a labile active chromatin hub (ACH). Later in development, during definitive erythropoiesis and after globin switching, the interactions switch to adult globin genes (Palstra et al., 2003). The presence of the LCR hypersensitive sites, but not the globin promoters, are important for ACH formation (Fig. 7)(Fang et al., 2007; Patrinos et al., 2004a).

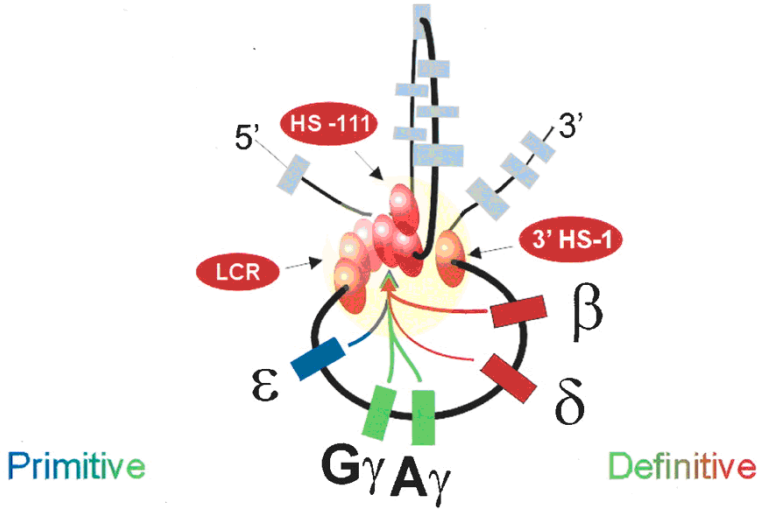


Figure 7. Two-dimensional representation of the 3-D interactions between the HSs and the globin genes in erythroid cells. The LCR, together with the most 5' and 3' HSs form the CH (indicated as a yellow sphere), with which the globin genes interact to form the ACH resulting in high-levels of transcription. The genes are depicted in different colors, indicative of their interaction with the LCR within the ACH during development (from Palstra et al., 2003; Patrinos et al., 2004a). Orange ellipse indicate HS sites, blue boxes olfactory receptor genes.

Our knowledge about the mechanism of γ-globin switching and the molecules involved has expanded in the recent years. Several factors have been described with the potential to release γ-globin suppression. The most important of these is BCL11A, although it is presently not clear with what factors it interacts to exert its function on the γ-globin gene promoter. Are there many factors as different complexes in physical contact with each other? Are they all necessary for γ-globin silencing? At what stage do they take part in the switching process? These questions can only be answered once factors that bind to the suppressed promoter have been identified. It would therefore be important to isolate the suppressed γ-globin promoter chromatin *in vivo*, identify the bound proteins by proteomic analysis followed by their functional validation.

References

(World Health Organization 2006). Sickle-cell anaemia, Report by the Secretariat, pp. 1-5.

Aidoo, M., Terlouw, D.J., Kolczak, M.S., McElroy, P.D., ter Kuile, F.O., Kariuki, S., Nahlen, B.L., Lal, A.A., and Udhayakumar, V. (2002). Protective effects of the sickle cell gene against malaria morbidity and mortality. The Lancet *359*, 1311-1312.

Arcasoy, M.O., Romana, M., Fabry, M.E., Skarpidi, E., Nagel, R.L., and Forget, B.G. (1997). High levels of human gamma-globin gene expression in adult mice carrying a transgene of deletion-type hereditary persistence of fetal hemoglobin. Mol Cell Biol *17*, 2076-2089.

Asano, H., Li, X.S., and Stamatoyannopoulos, G. (1999). FKLF, a Novel Kruppel-Like Factor That Activates Human Embryonic and Fetal beta -Like Globin Genes. Mol Cell Biol *19*, 3571-3579.

Bard, H., and Prosmanne, J. (1985). Relative rates of fetal hemoglobin and adult hemoglobin synthesis in cord blood of infants of insulin-dependent diabetic mothers. Pediatrics *75*, 1143-1147.

Bell, A.C., West, A.G., and Felsenfeld, G. (1999). The Protein CTCF Is Required for the Enhancer Blocking Activity of Vertebrate Insulators. Cell *98*, 387-396.

Benz , E.J.J.r. (1982). Clinical management of gene expression. N Engl J Med 307, 1515–1516 *307*, 1515-1516.

Berry, M., Grosveld, F., and Dillon, N. (1992). A single point mutation is the cause of the Greek form of hereditary persistence of fetal haemoglobin. Nature *358*, 499-502.

Bhatia, M., and Walters, M.C. (2007). Hematopoietic cell transplantation for thalassemia and sickle cell disease: past, present and future. Bone Marrow Transplant *41*, 109-117.

Blau, C.A., Constantoulakis, P., Shaw, C.M., and Stamatoyannopoulos, G. (1993). Fetal hemoglobin induction with butyric acid: efficacy and toxicity. Blood *81*, 529-537.

Boosalis, M.S., Bandyopadhyay, R., Bresnick, E.H., Pace, B.S., Van DeMark, K., Zhang, B., Faller, D.V., and Perrine, S.P. (2001). Short-chain fatty acid derivatives stimulate cell proliferation and induce STAT-5 activation. Blood *97*, 3259-3267.

Bunn, H.F. (2001). Human hemoglobins: Sickle hemoglobin and other mutants. In The molecular basis of blood diseases (3rd ed), G. Stamatoyannopoulos, Majeurs P., Perlmutter, P., Varmus H., ed. (Philadelphia, W.B. Saunders), pp. 227-262.

Calzolari, R., McMorrow, T., Yannoutsos, N., Langeveld, A., and Grosveld, F. (1999). Deletion of a region that is a candidate for the difference between the deletion forms of hereditary persistence of fetal hemoglobin and delta-beta-thalassemia affects beta- but not gamma-globin gene expression. EMBO J *18*, 949-958.

Carey, B.W., Markoulaki, S., Hanna, J., Saha, K., Gao, Q., Mitalipova, M., and Jaenisch, R. (2009). Reprogramming of murine and human somatic cells using a single polycistronic vector. Proc Natl Acad Sci U S A *106*, 157-162.

Carter, D., Chakalova, L., Osborne, C.S., Dai, Y.-F., and Fraser, P. (2002). Long-range chromatin regulatory interactions in vivo. Nature Genetics *32*, 623-626.

Chakalova, L., Osborne, C.S., Dai, Y.-F., Goyenechea, B., Metaxotou-Mavromati, A., Kattamis, A., Kattamis, C., and Fraser, P. (2005). The Corfu delta-beta- thalassemia deletion disrupts gamma-globin gene silencing and reveals post-transcriptional regulation of HbF expression. Blood *105*, 2154-2160.

Chang, Y.C., Smith, K.D., Moore, R.D., Serjeant, G.R., and Dover, G.J. (1995). An analysis of fetal hemoglobin variation in sickle cell disease: the relative contributions of the X-linked factor, beta-globin haplotypes, alpha-globin gene number, gender, and age. Blood *85*, 1111-1117.

Chang, Y.P., Maier-Redelsperger, M., Smith, K.D., Contu, L., Ducroco, R., de Montalembert, M., Belloy, M., Elion, J., Dover, G.J., and Girot, R. (1997). The relative importance of the X-linked FCP locus and beta-globin haplotypes in determining haemoglobin F levels: a study of SS patients homozygous for beta S haplotypes. Br J Haematol *96*, 806-814.

Charache, S., Terrin, M.L., Moore, R.D., Dover, G.J., Barton, F.B., Eckert, S.V., McMahon, R.P., Bonds, D.R., and The Investigators of the Multicenter Study of Hydroxyurea in Sickle Cell, A. (1995). Effect of Hydroxyurea on the Frequency of Painful Crises in Sickle Cell Anemia. N Engl J Med *332*, 1317-1322.

Chen, Z., Luo, H.-y., Steinberg, M.H., and Chui, D.H.K. (2009). BCL11A represses HBG transcription in K562 cells. Blood Cells, Molecules, and Diseases *42*, 144-149.

Chung, J.H., Whiteley, M., and Felsenfeld, G. (1993). A 5' element of the chicken [beta]-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila. Cell *74*, 505-514.

Cokic, V.P., Andric, S.A., Stojilkovic, S.S., Noguchi, C.T., and Schechter, A.N. (2008). Hydroxyurea nitrosylates and activates soluble guanylyl cyclase in human erythroid cells. Blood *111*, 1117-1123.

Cokic, V.P., Beleslin-Cokic, B.B., Tomic, M., Stojilkovic, S.S., Noguchi, C.T., and Schechter, A.N. (2006). Hydroxyurea induces the eNOS-cGMP pathway in endothelial cells. Blood *108*, 184-191.

Collins, F.S., Metherall, J.E., Yamakawa, M., Pan, J., Weissman, S.M., Forget, B.G. (1985). A point mutation in the A gamma-globin gene promoter in Greek hereditary persistence of fetal haemoglobin. Nature *313*, 325-326.

Conley, C.L., Weatherall, D.J., Richardson, S.N., Shepard, M.K., and Charache, S. (1963). Hereditary Persistence of Fetal Hemoglobin: A Study of 79 Affected Persons in 15 Negro Families in Baltimore. Blood *21*, 261-281.

Constantoulakis, P., Knitter, G., and Stamatoyannopoulos, G. (1989). On the induction of fetal hemoglobin by butyrates: in vivo and in vitro studies with sodium butyrate and comparison of combination treatments with 5-AzaC and AraC. Blood *74*, 1963-1971.

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

Delea, T.E., Edelsberg, E., Sofrygin, O., Thomas, S.K., Baladi, J.F., Phatak, P.D., and Coates, T.D. (2007). Consequences and costs of noncompliance with iron chelation therapy in patients with transfusion-dependent thalassemia: a literature review. Blood *47*, 1919-1929.

DeSimone, J., Heller, P., Hall, L., Zwiers, D. (1982). 5-Azacytidine stimulates fetal hemoglobin synthesis in anemic baboons. . Proc Natl Acad Sci U S A *79*, 4428-4431.

Dillon, N., and Grosveld, F. (1991). Human gamma-globin genes silenced independently of other genes in the beta-globin locus. Nature *350*, 252-254.

Dover, G.J., Brusilow, S., and Charache, S. (1994). Induction of fetal hemoglobin production in subjects with sickle cell anemia by oral sodium phenylbutyrate. Blood *84*, 339-343.

Dover, G.J., Charache, S.H., Boyer, S.H., Talbot, C.C. Jr, Smith, K.D. (1983). 5-Azacytidine increases fetal hemoglobin production in a patient with sickle cell disease. . Prog Clin Biol Res *134*, 475-488.

Dover, G.J., Humphries, R.K., Moore, J.G., Ley, T.J., Young, N.S., Charache, S., and Nienhuis, A.W. (1986). Hydroxyurea induction of hemoglobin F production in sickle cell disease: relationship between cytotoxicity and F cell production. Blood *67*, 735-738.

Drissen, R., Palstra, R.-J., Gillemans, N., Splinter, E., Grosveld, F., Philipsen, S., and de Laat, W. (2004). The active spatial organization of the Î²-globin locus requires the transcription factor EKLF. Genes Dev *18*, 2485-2490.

Du, M.-J., Lv, X., Hao, D.-L., Zhao, G.-W., Wu, X.-S., Wu, F., Liu, D.-P., and Liang, C.-C. (2008). MafK/NF-E2 p18 is required for [beta]-globin genes activation by mediating the proximity of LCR and active [beta]-globin genes in MEL cell line. The International Journal of Biochemistry & Cell Biology *40*, 1481-1493.

Duan, Z.-J., Fang, X., Rohde, A., Han, H., Stamatoyannopoulos, G., and Li, Q. (2002). Developmental specificity of recruitment of TBP to the TATA box of the human beta-globin gene. Proc Natl Acad Sci U S A *99*, 5509-5514.

Duan, Z., Stamatoyannopoulos, G., and Li, Q. (2001). Role of NF-Y in In Vivo Regulation of the gamma-Globin Gene. Mol Cell Biol *21*, 3083-3095.

Edington, G.M., Lehmann, H. (1955). Expression of the sickle-cell gene in Africa. Br Med J *1*, 1308-1311.

Ellis, J., Tan-Un, K.C., Harper, A., Michalovich, D., Yannoutsos, N., Philipsen, S., Grosveld, F. (1996). A dominant chromatin-opening activity in 5' hypersensitive site 3 of the human beta-globin locus control region. EMBO J *15*, 562-568.

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

Fang, X., Han, H., Stamatoyannopoulos, G., and Li, Q. (2004). Developmentally Specific Role of the CCAAT Box in Regulation of Human Î³-Globin Gene Expression. J Biol Chem *279*, 5444-5449.

Fang, X., Xiang, P., Yin, W., Stamatoyannopoulos, G., and Li, Q. (2007). Cooperativeness of the Higher Chromatin Structure of the [beta]-Globin Locus Revealed by the Deletion Mutations of DNase I Hypersensitive site 3 of the LCR. Journal of Molecular Biology *365*, 31-37.

Farrell, C.M., West, A.G., and Felsenfeld, G. (2002). Conserved CTCF Insulator Elements Flank the Mouse and Human beta-Globin Loci. Mol Cell Biol *22*, 3820-3831.

Fathallah, H., and Atweh, G.F. (2006). Induction of Fetal Hemoglobin in the Treatment of Sickle Cell Disease. Hematology Am Soc Hematol Educ Program *2006*, 58-62.

Fathallah, H., Portnoy, G., and Atweh, G.F. (2008). Epigenetic analysis of the human alpha- and beta-globin gene clusters. Blood Cells, Molecules, and Diseases *40*, 166-173.

Fathallah, H., Weinberg, R.S., Galperin, Y., Sutton, M., and Atweh, G.F. (2007). Role of epigenetic modifications in normal globin gene regulation and butyrate-mediated induction of fetal hemoglobin. Blood *110*, 3391-3397.

Feingold, E.A., and Forget, B.G. (1989). The breakpoint of a large deletion causing hereditary persistence of fetal hemoglobin occurs within an erythroid DNA domain remote from the beta-globin gene cluster. Blood *74*, 2178-2186.

Ferster, A., Tahriri, P., Vermylen, C., Sturbois, G., Corazza, F., Fondu, P., Devalck, C., Dresse, M.F., Feremans, W., Hunninck, K.*, et al.* (2001). Five years of experience with hydroxyurea in children and young adults with sickle cell disease. Blood *97*, 3628-3632.

Fessas, P., and Stamatoyannopoulos, G. (1964). Hereditary Persistence of Fetal Hemoglobin in Greece. A Study and a Comparison. Blood *24*, 223-240.

Filipe, A., Li, Q., Deveaux, S., Godin, I., Romeo, P.-H., Stamatoyannopoulos, G., and Mignotte, V. (1999). Regulation of embryonic/fetal globin genes by nuclear hormone receptors: a novel perspective on hemoglobin switching. EMBO J *18*, 687-697.

Fraser, P., and Grosveld, F. (1998). Locus control regions, chromatin activation and transcription. Current Opinion in Cell Biology *10*, 361-365.

Frenette, P.S., and Atweh, G.F. (2007). Sickle cell disease: old discoveries, new concepts, and future promise. The Journal of Clinical Investigation *117*, 850-858.

Fucharoen, S., Shimizu, K., and Fukumaki, Y. (1990). A novel C-T transition within the distal CCAAT motif of the G{gamma}-globin gene in the Japanese HPFH: implication of factor binding in elevated fetal globin expression. Nucleic Acids Res *18*, 5245-5253.

Garner, C.P., Tatu, T., Best, S., Creary, L., and Thein, S.L. (2002). Evidence of Genetic Interaction between the [beta]-Globin Complex and Chromosome 8q in the Expression of Fetal Hemoglobin. The American Journal of Human Genetics *70*, 793-799.

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

Gilmour, J.R. (1941). Normal haematopoiesis in intrauterine and neonatal life. J Pathol Bact *52*, 25.

Glauber, J.G., Wandersee, N.J., Little, J.A., and Ginder, G.D. (1991). 5'-flanking sequences mediate butyrate stimulation of embryonic globin gene expression in adult erythroid cells. Mol Cell Biol *11*, 4690-4697.

Goldberg, M.A., Husson, M.A., and Bunn, H.F. (1977). Participation of hemoglobins A and F in polymerization of sickle hemoglobin. J Biol Chem *252*, 3414-3421.

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

Gulbis, B., Haberman, D., Dufour, D., Christophe, C., Vermylen, C., Kagambega, F., Corazza, F., Devalck, C., Dresse, M.-F., Hunninck, K.*, et al.* (2005). Hydroxyurea for sickle cell disease in children and for prevention of cerebrovascular events: the Belgian experience. Blood *105*, 2685-2690.

Gumucio, D.L., Rood, K.L., Blanchard-McQuate, K.L., Gray, T.A., Saulino, A., and Collins, F.S. (1991). Interaction of Sp1 with the human gamma globin promoter: binding and transactivation of normal and mutant promoters. Blood *78*, 1853-1863.

Gumucio, D.L., Rood, K.L., Gray, T.A., Riordan, M.F., Sartor, C.I., and Collins, F.S. (1988). Nuclear proteins that bind the human gamma-globin gene promoter: alterations in binding produced by point mutations associated with hereditary persistence of fetal hemoglobin. Mol Cell Biol *8*, 5310-5322.

Haley, J.D., Smith, D.E., Schwedes, J., Brennan, R., Pearce, C., Moore, C., Wang, F., Petti, F., Grosveld, F., Jane, S.M.*, et al.* (2003). Identification and characterization of mechanistically distinct inducers of [gamma]-globin transcription. Biochemical Pharmacology *66*, 1755-1768.

Hanna, J., Wernig, M., Markoulaki, S., Sun, C.-W., Meissner, A., Cassady, J.P., Beard, C., Brambrink, T., Wu, L.-C., Townes, T.M.*, et al.* (2007). Treatment of Sickle Cell Anemia Mouse Model with iPS Cells Generated from Autologous Skin. Science *318*, 1920-1923.

Hardison, R., Slightom, J.L., Gumucio, D.L., Goodman, M., Stojanovic, N., and Miller, W. (1997). Locus control regions of mammalian beta-globin gene clusters: combining phylogenetic analyses and experimental results to gain functional insights. Gene *205*, 73-94.

Hebbel, R.P. (1991). Beyond hemoglobin polymerization: the red blood cell membrane and sickle disease pathophysiology. Blood *77*, 214-237.

Hebbel, R.P., Boogaerts, M.A., Eaton, J.W., Steinberg, M.H. (1980). Erythrocyte adherence to endothelium in sickle-cell anemia. A possible determinant of disease severity. N Engl J Med 1980 May 1;302(18):992-5 *302*, 992-995.

Hebbes, T.R., Clayton, A.L., Thorne, A.W., and Crane-Robinson, C. (1994). Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain. EMBO J *15*, 1823-1830.

Hoover, R., Rubin, R., Wise, G., Warren, R. (1979). Adhesion of normal and sickle erythrocytes to endothelial

monolayer cultures. Blood 54, 872–876. Blood *54*, 872-876.

Hou, C., Zhao, H., Tanimoto, K., and Dean, A. (2008). CTCF-dependent enhancer-blocking by alternative chromatin loop formation. Proc Natl Acad Sci U S A *105*, 20398-20403.

Huang, S., Li, X., Yusufzai, T.M., Qiu, Y., and Felsenfeld, G. (2007). USF1 Recruits Histone Modification Complexes and Is Critical for Maintenance of a Chromatin Barrier. Mol Cell Biol *27*, 7991-8002.

Hünefeld, F.L. (1840). Die Chemismus in der thierischen Organization. Leipzig.

Ikuta, T., Kan, Y.W., Swerdlow, P.S., Faller, D.V., and Perrine, S.P. (1998). Alterations in Protein-DNA Interactions in the gamma -Globin Gene Promoter in Response to Butyrate Therapy. Blood *92*, 2924-2933.

Ingram, V.M. (1956). A specific chemical difference between the globins of normal human and sickle-cell anaemia

haemoglobin. Nature *178*, 792–794.

Jane, S.M., Gumucio, D.L., Ney, P.A., Cunningham, J.M., and Nienhuis, A.W. (1993a). Methylation-enhanced binding of Sp1 to the stage selector element of the human gamma-globin gene promoter may regulate development specificity of expression. *13*, 3272-3281.

Jane, S.M., Gumucio, D.L., Ney, P.A., Cunningham, J.M., and Nienhuis, A.W. (1993b). Methylation-enhanced binding of Sp1 to the stage selector element of the human gamma-globin gene promoter may regulate development specificity of expression. Mol Cell Biol *13*, 3272-3281.

Jane, S.M., Ney, P.A., Vanin, E.F., Gumucio, D.L., and Nienhuis, A.W. (1992). Identification of a stage selector element in the human gamma-globin gene promoter that fosters preferential interaction with the 5' HS2 enhancer when in competition with the beta-promoter. EMBO J *11*, 2961-2969.

Jeffreys, A.J., and Flavell, R.A. (1977). The rabbit [beta]-globin gene contains a large insert in the coding sequence. Cell *12*, 1097-1108.

Johnson, K.D., Norton, J.E., and Bresnick, E.H. (2002). Requirements for utilization of CREB binding protein by hypersensitive site two of the beta-globin locus control region. Nucleic Acids Res *30*, 1522-1530.

Katsantoni, E.Z., Langeveld, A., Wai, A.W.K., Drabek, D., Grosveld, F., Anagnou, N.P., and Strouboulis, J. (2003). Persistent gamma-globin expression in adult transgenic mice is mediated by HPFH-2, HPFH-3, and HPFH-6 breakpoint sequences. Blood *102*, 3412-3419.

Keys, J.R., Tallack, M.R., Zhan, Y., Papathanasiou, P., Goodnow, C.C., Gaensler, K.M., Crossley, M., Dekker, J., Perkins, A.C. (2008). A mechanism for Ikaros regulation of human globin gene switching. Br J Haematol *141*, 398-406.

Kim, S.-I., Bultman, S.J., Kiefer, C.M., Dean, A., and Bresnick, E.H. (2009). BRG1 requirement for long-range interaction of a locus control region with a downstream promoter. Proc Natl Acad Sci U S A *106*, 2259-2264.

Kinney, T.R., Helms, R.W., O'Branski, E.E., Ohene-Frempong, K., Wang, W., Daeschner, C., Vichinsky, E., Redding-Lallinger, R., Gee, B., Platt, O.S*., et al.* (1999). Safety of Hydroxyurea in Children With Sickle Cell Anemia: Results of the HUG-KIDS Study, a Phase I/II Trial. Blood *94*, 1550-1554.

Koshy, M., Dorn, L., Bressler, L., Molokie, R., Lavelle, D., Talischy, N., Hoffman, R., van Overveld, W., and DeSimone, J. (2000). 2-deoxy 5-azacytidine and fetal hemoglobin induction in sickle cell anemia. Blood *96*, 2379-2384.

Kulozik, A.E., Yarwood, N., and Jones, R.W. (1988). The Corfu delta beta zero thalassemia: a small deletion acts at a distance to selectively abolish beta globin gene expression [published erratum appears in Blood 1988 May;71(5):1509]. Blood *71*, 457-462.

Kumkhaek, C., Taylorj, J.G., Zhu, J., Hoppe, C., Kato, G.J., and Rodgers, G.P. (2008). Fetal haemoglobin response to hydroxycarbamide treatment and sar1a promoter polymorphisms in sickle cell anaemia. Br J Haematol *141*, 254-259.

Labie, D., Dunda-Belkhodja, O., Rouabhi, F., Pagnier, J., Ragusa, A., and Nagel, R.L. (1985a). The -158 site 5' to the G gamma gene and G gamma expression. Blood *66*, 1463-1465.

Labie, D., Pagnier, J., Lapoumeroulie, C., Rouabhi, F., Dunda-Belkhodja O, Chardin P, Beldjord C, Wajcman H, Fabry ME, and RL., N. (1985b). Common haplotype dependency of high G gamma-globin gene expression and high Hb F levels in beta-thalassemia and sickle cell anemia patients. Proc Natl Acad Sci U S A *28*, 2111-2114.

Lavelle, D., Vaitkus, K., Hankewych, M., Singh, M., and DeSimone, J. (2006a). Developmental changes in DNA methylation and covalent histone modifications of chromatin associated with the epsilon-, gamma-, and beta-globin gene promoters in Papio anubis. Blood Cells, Molecules, and Diseases *36*, 269-278.

Lavelle, D., Vaitkus, K., Hankewych, M., Singh, M., and DeSimone, J. (2006b). Effect of 5-aza-2'-deoxycytidine (Dacogen) on covalent histone modifications of chromatin associated with the epsilon-, gamma-, and beta-globin promoters in Papio anubis. Experimental Hematology *34*, 339-347.

Lettre, G., Sankaran, V.G., Bezerra, M.A.C., Araújo, A.S., Uda, M., Sanna, S., Cao, A., Schlessinger, D., Costa, F.F., Hirschhorn, J.N., *et al.* (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 U S A *105*, 11869-11874.

Ley , T.J., Anagnou, N.P., Noguchi, C.T., Schechter, A.N., DeSimone, J., Heller, P., Nienhuis, A.W. (1983). DNA methylation and globin gene expression in patients treated with 5-azacytidine. . Prog Clin Biol Res *134*, 457-474.

Ley, T.J., DeSimone, J., Anagnou, N.P., Keller, G.H., Humphries, R.K., Turne,r P.H., Young, N.S., Keller ,P., Nienhuis, A.W. (1982). 5-azacytidine selectively increases gamma-globin synthesis in a patient with beta thalassemia. N Engl J Med *307*, 1469-1475.

Li, Q., and Stamatoyannopoulos, G. (1994). Hypersensitive site 5 of the human beta locus control region functions as a chromatin insulator. Blood *84*, 1399-1401.

Li, Q., Zhang, M., Duan, Z., and Stamatoyannopoulos, G. (1999). Structural Analysis and Mapping of DNase I Hypersensitivity of HS5 of the [beta]-Globin Locus Control Region. Genomics *61*, 183-193.

Li, Q., Zhang, M., Han, H., Rohde, A., and Stamatoyannopoulos, G. (2002). Evidence that DNase I hypersensitive site 5 of the human beta-globin locus control region functions as a chromosomal insulator in transgenic mice. Nucleic Acids Res *30*, 2484-2491.

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*, 41700-41709.

Litt, M.D., Simpson, M., Recillas-Targa, F., Prioleau, M.N., and Felsenfeld, G. (2001). Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci, pp. 2224 - 2235.

Liu, L., R,, Du, Z., W,, Zhao, H., L,, Liu, X., L,, Huang, X., D,, Shen, J., Ju, L., M,, Fang, F., D,, and Zhang, J., W, (2005). T to C Substitution at -175 -173 of the Îbeta-Globin Promoter Affects GATA-1 and Oct-1 Binding in Vitro Differently but Can Independently Reproduce the Hereditary Persistence of Fetal Hemoglobin Phenotype in Transgenic Mice. J Biol Chem *280*, 7452-7459.

Lopez, R.A., Schoetz, S., DeAngelis, K., O'Neill, D., and Bank, A. (2002). Multiple hematopoietic defects and delayed globin switching in Ikaros null mice. Proc Natl Acad Sci U S A *99*, 602-607.

Lowrey, C.H., Bodine, D.M., and Nienhuis, A.W. (1992). Mechanism of DNase I hypersensitive site formation within the human globin locus control region. Proc Natl Acad Sci U S A *89*, 1143-1147.

Lucarelli, G., and Gaziev, J. (2008). Advances in the allogeneic transplantation for thalassemia. Blood Reviews *22*, 53-63.

Lux, S.E., John, K.M., Karnovsky, M.J. (1976). Irreversible deformation of the spectrin-actin lattice in irreversibly sickled cells. J Clin Invest *58*, 955-963.

Ma, Q., Wyszynski, D.F., Farrell, J.J., Kutlar, A., Farrer, L.A., Baldwin, C.T., and Steinberg, M.H. (2007). Fetal hemoglobin in sickle cell anemia: genetic determinants of response to hydroxyurea. Pharmacogenomics J *7*, 386-394.

Mabaera, R., Richardson, C.A., Johnson, K., Hsu, M., Fiering, S., and Lowrey, C.H. (2007). Developmental- and differentiation-specific patterns of human {gamma}- and {beta}-globin promoter DNA methylation. Blood *110*, 1343-1352.

Mabaera, R., West, R.J., Conine, S.J., Macari, E.R., Boyd, C.D., Engman, C.A., and Lowrey, C.H. (2008). A cell stress signaling model of fetal hemoglobin induction: what doesn't kill red blood cells may make them stronger. Experimental Hematology *36*, 1057-1072.

Mahajan, M.C., Narlikar, G.J., Boyapaty, G., Kingston, R.E., and Weissman, S.M. (2005). Heterogeneous nuclear ribonucleoprotein C1/C2, MeCP1, and SWI/SNF form a chromatin remodeling complex at the beta-globin locus control region. Proc Natl Acad Sci U S A *102*, 15012-15017.

Mantei, N., Boll, W., Weissmann, C. (1979). Rabbit beta-globin mRNA production in mouse L cells transformed with cloned rabbit beta-globin chromosomal DNA. Nature *281*, 40-46.

Mantovani, R., Malgaretti, N., Nicolis, S., Ronchi, A., Giglioni, B., and Ottolenghi, S. (1988). The effects of HPFH mutations in the human {gamma}-globin promoter on binding of ubiquitous and erythroid specific nuclear factors. Nucleic Acids Res *16*, 7783-7797.

Marin, M., Karis, A., Visser, P., Grosveld, F., and Philipsen, S. (1997). Transcription Factor Sp1 Is Essential for Early Embryonic Development but Dispensable for Cell Growth and Differentiation. Cell *89*, 619-628.

Martin, D.I.K., Tsai, S.-F., and Orkin, S.H. (1989). Increased gamma-globin expression in a nondeletion HPFH mediated by an erythroid-specif ic DNA-binding factor. Nature *338*, 435-438.

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*, 1902-1908.

Menzel, S., Garner, C., Gut, I., Matsuda, F., Yamaguchi, M., Heath, S., Foglio, M., Zelenika, D., Boland, A., Rooks, H*., et al.* (2007). A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. Nat Genet *39*, 1197-1199.

Miles, J., Mitchell, J.A., Chakalova, L., Goyenechea, B., Osborne, C.S., O'Neill, L., Tanimoto, K., Engel, J.D., and Fraser, P. (2007). Intergenic Transcription, Cell-Cycle and the Developmentally Regulated Epigenetic Profile of the Human Beta-Globin Locus. PLoS ONE *2*, e630.

Milot, E., Strouboulis, J., Trimborn, T., Wijgerde, M., de Boer, E., Langeveld, A., Tan-Un, K., Vergeer, W., Yannoutsos, N., Grosveld, F*., et al.* (1996). Heterochromatin Effects on the Frequency and Duration of LCR-Mediated Gene Transcription. Cell *87*, 105-114.

Nagel, R.L., and Roth, E.F., Jr. (1989). Malaria and red cell genetic defects. Blood *74*, 1213-1221.

Nathan, D.G., and Gunn, R.B. (1966). Thalassemia: The consequences of unbalanced hemoglobin synthesis. The American Journal of Medicine *41*, 815-830.

Navas, P.A., Peterson, K.R., Li, Q., McArthur, M., and Stamatoyannopoulos, G. (2001). The 5'HS4 core element of the human beta-globin locus control region is required for high-level globin gene expression in definitive but not in primitive erythropoiesis. Journal of Molecular Biology *312*, 17-26.

Newman Dorland, W.A. (2007). Hemoglobin. In Dorland's Medical Dictionary (Elsevier publication).

Ney, P.A., Sorrentino, B.P., McDonagh, K.T., and Nienhuis, A.W. (1990). Tandem AP-1-binding sites within the human beta-globin dominant control region function as an inducible enhancer in erythroid cells. Genes Dev *4*, 993-1006.

Nienhuis, A.W. (2008). Development of gene therapy for blood disorders. Blood *111*, 4431-4444.

O'Neill, D., Bornschlegel, K., Flamm, M., Castle, M., Bank, A. (1991). A DNA-binding factor in adult hematopoietic cells interacts with a pyrimidine-rich domain upstream from the human delta-globin gene. Proc Natl Acad Sci U S A *88*, 8953-8957.

O'Neill, D.W., Schoetz, S.S., Lopez, R.A., Castle, M., Rabinowitz, L., Shor, E., Krawchuk, D., Goll, M.G., Renz, M., Seelig, H.-P*., et al.* (2000). An Ikaros-Containing Chromatin-Remodeling Complex in Adult-Type Erythroid Cells. Mol Cell Biol *20*, 7572-7582.

O'Neill , D., Yang, J., Erdjument-Bromage, H., Bornschlegel, K., Tempst, P., and Bank, A. (1999). Tissue-specific and developmental stage-specific DNA binding by a mammalian SWI/SNF complex associated with human fetal-to-adult globin gene switching. Proc Natl Acad Sci U S A *96*, 349-354.

Ottolenghi, S., Giglioni, B. (1982). The deletion in a type of delta 0-beta 0-thalassaemia begins in an inverted AluI repeat. Nature *300*, 770-771.

Pace, B.S., Li, Q., and Stamatoyannopoulos, G. (1996). In vivo search for butyrate responsive sequences using transgenic mice carrying A gamma gene promoter mutants. Blood *88*, 1079-1083.

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

Papayannopoulou, T.H., Brice, M., Stamatoyannopoulos G. (1976). Stimulation of fetal hemoglobin synthesis in bone marrow cultures from adult individuals. Proc Natl Acad Sci U S A *73*, 2033-2037.

Papayannopoulou, T.H., Brice, M., Stamatoyannopoulos G. (1977). Hemoglobin F synthesis in vitro: evidence for control at the level of primitive erythroid stem cells. Proc Natl Acad Sci U S A *74*, 2923-2927.

Park, J.-I., Choi, H.-S., Jeong, J.-S., Han, J.-Y., and Kim, I.-H. (2001). Involvement of p38 Kinase in Hydroxyurea-induced Differentiation of K562 Cells. Cell growth & Differentiation *12*, 481-486.

Patrinos, G.P., de Krom, M., de Boer, E., Langeveld, A., Imam, A.M.A., Strouboulis, J., de Laat, W., and Grosveld, F.G. (2004a). Multiple interactions between regulatory regions are required to stabilize an active chromatin hub. Genes Dev *18*, 1495-1509.

Patrinos, G.P., Giardine, B., Riemer, C., Miller, W., Chui, D.H.K., Anagnou, N.P., Wajcman, H., and Hardison, R.C. (2004b). Improvements in the HbVar database of human hemoglobin variants and thalassemia mutations for population and sequence variation studies. Nucleic Acids Res *32*, D537-541.

Pauling, L., Itano, H., Singer, S.J., and Wells, I.C. (1949). Sickle cell anemia a molecular disease. Science *110*, 543-548.

Pawliuk, R., Westerman ,K.A., Fabry ,M.E., Payen, E., Tighe, R., Bouhassira, E.E., Acharya, S.A., Ellis, J., London, I.M., Eaves, C.J., Humphries, R.K., Beuzard, Y., Nagel, R.L., Leboulch, P. (2001). Correction of sickle cell disease in transgenic mouse models by gene therapy. Science *294*, 2368-2371.

Perrine, S.P., Ginder, G.D., Faller, D.V., Dover, G.H., Ikuta, T., Witkowska, H.E., Cai, S.-p., Vichinsky, E.P., and Olivieri, N.F. (1993). A Short-Term Trial of Butyrate to Stimulate Fetal-Globin-Gene Expression in the beta-Globin Disorders. N Engl J Med *328*, 81-86.

Perrine, S.P., Greene, M.F., and Faller, D.V. (1985). Delay in the fetal globin switch in infants of diabetic mothers. N Engl J Med 1985 Feb 7;312(6):334-8 *312*, 334-338.

Perrine, S.P., Miller, B.A., Faller, D.V., Cohen, R.A., Vichinsky, E.P., Hurst, D., Lubin, B.H., and Papayannopoulou, T. (1989). Sodium butyrate enhances fetal globin gene expression in erythroid progenitors of patients with Hb SS and beta thalassemia. Blood *74*, 454-459.

Perrine, S.P., Miller, B.A., Greene, M.F., Cohen, R.A., Cook, N., Shackleton, C., and Faller, D.V. (1987). Butryic acid analogues augment gamma globin gene expression in neonatal erythroid progenitors. Biochemical and Biophysical Research Communications *148*, 694-700.

Perutz, M.F. (1960). Structure of hemoglobin. Brookhaven Symp Biol *13*, 165-183.

Perutz, M.F., Rossmann, M.G., Cullins, A.F., Muirhead, H., Will, G., and North, A.C. (1960). Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5-A. resolution, obtained by X-ray analysis. Nature *185*, 416-422.

Peschle, C., Mavilio, F., Care, A. Migliaccio, G., Migliaccio, A.R., Salvo, G., Samoggia, P., Petti, S., Guerriero, R., Marinucci, M., et al. (1985). Haemoglobin switching in human embryos: asynchrony of zeta to alpha and epsilon to gamma-globin switches in primitive and definite erythropoietic lineage. Nature *313*, 235-238.

Philipsen, S., Talbot, D., Fraser, P., Grosveld, F. (1990). The beta-globin dominant control region: hypersensitive site 2. EMBO J *9*, 2159-2167.

Platt, O.S., Brambilla, D.J., Rosse, W.F., Milner, P.F., Castro, O., Steinberg, M.H., and Klug, P.P. (1994). Mortality In Sickle Cell Disease -- Life Expectancy and Risk Factors for Early Death. N Engl J Med *330*, 1639-1644.

Poillon, W.N., Kim, B.C., Rodgers, G.P., Noguchi, C.T., Schechter, A.N. (1993). Sparing effect of hemoglobin F and hemoglobin A2 on the polymerization of hemoglobin S at physiologic ligand saturations. Proc Natl Acad Sci U S A *90*, 5039-5043.

Prchal, J., Stamatoyannopoulos, G. (1981). Two siblings with unusually mild homozygous beta-thalassemia: a didactic example of the effect of a nonallelic modifier gene of the expressivity of a monogenic disorder. Am J Med Genet *10*, 291-300.

Prioleau, M.-N., Nony, P., Simpson, M., and Felsenfeld, G. (1999). An insulator element and condensed chromatin region separate the chicken [beta]-globin locus from an independently regulated erythroid-specific folate receptor gene. EMBO J *18*, 4035-4048.

Proudfoot, N.J. (1976). Nucleotide sequence from the coding region of rabbit {beta}-globin messenger RNA. Nucl Acids Res *3*, 1811-1822.

Pruzina, S., Hanscombe, O., Whyatt, D., Grosveld, F., and Philipsen, S. (1991). Hypersensitive site 4 of the human {beta} globin locus control region. Nucleic Acids Res *19*, 1413-1419.

Reik, A., Telling, A., Zitnik, G., Cimbora, D., Epner, E., and Groudine, M. (1998). The Locus Control Region Is Necessary for Gene Expression in the Human beta -Globin Locus but Not the Maintenance of an Open Chromatin Structure in Erythroid Cells. Mol Cell Biol *18*, 5992-6000.

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 Î³- and Ïµ-Globin CCAAT Boxes. J Biol Chem *270*, 21934-21941.

Rougeon, F., Kourilsky, P., and Mach, B. (1975). Insertion of a rabbit {beta}-globin gene sequence into an E.coli plasmid. Nucl Acids Res *2*, 2365-2378.

Rupon, J.W., Wang, S.Z., Gaensler, K., Lloyd, J., and Ginder, G.D. (2006). Methyl binding domain protein 2 mediates gamma-globin gene silencing in adult human betaYAC transgenic mice. Proc Natl Acad Sci U S A *103*, 6617-6622.

Sadelain, M. (2006). Recent advances in globin gene transfer for the treatment of beta-thalassemia and sickle cell anemia. Current Opinion in Hematology *13*, 142-148.

Sankaran, V.G., Menne, T.F., Xu, J., Akie, T.E., Lettre, G., Van Handel, B., Mikkola, H.K.A., Hirschhorn, J.N., Cantor, A.B., and Orkin, S.H. (2008). Human Fetal Hemoglobin Expression Is Regulated by the Developmental Stage-Specific Repressor BCL11A. Science *322*, 1839-1842.

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

Sauntharajah, Y., Lavelle, D., and DeSimone, J. (2004). DNA hypo-methylating agents and sickle cell disease. Br J Haematol *126*, 629-636.

Scott, J.P., Hillery, C.A., Brown, E.R., Misiewicz, V., and Labotka, R.J. (1996). Hydroxyurea therapy in children severely affected with sickle cell disease. The Journal of Pediatrics *128*, 820-828.

Sedgewick, A.E., Timofeev, N., Sebastiani, P., So, J.C.C., Ma, E.S.K., Chan, L.C., Fucharoen, G., Fucharoen, S., Barbosa, C.G., Vardarajan, B.N*., et al.* (2008). BCL11A is a major HbF quantitative trait locus in three different populations with [beta]-hemoglobinopathies. Blood Cells, Molecules, and Diseases *41*, 255-258.

Sengupta, P.K., Lavelle, D., and DeSimone, J. (1994). Increased binding of Sp1 to the gamma-globin gene promoter upon site-specific cytosine methylation. Am J Hematol 1994 Jul;46(3):169-72 *46*, 169-172.

Shenoy, S. (2007). Has stem cell transplantation come of age in the treatment of sickle cell disease? Bone Marrow Transplant *40*, 813-821.

Sher, G.D., Ginder, G.D., Little, J., Yang, S., Dover, G.J., and Olivieri, N.F. (1995). Extended Therapy with Intravenous Arginine Butyrate in Patients with beta-Hemoglobinopathies. N Engl J Med *332*, 1606-1610.

Sheridan, B.L., Weatherall, D.J., Clegg, J.B., Pritchard, J., Wood, W.G., Callender, S.T., Durrant, I.J., McWhirter, W.R., Ali, M., Partridge, J.W. , and Thompson, E.N. (1976). The patterns of fetal haemoglobin production in leukaemia. Br J Haematol *32*, 487.

Skalnik, D.G., Strauss, E.C., and Orkin, S.H. (1991). CCAAT displacement protein as a repressor of the myelomonocytic-specific gp91-phox gene promoter. J Biol Chem *266*, 16736-16744.

Solovieff, N., Milton, J.N., Hartley, S.W., Sherva, R., Sebastiani, P., Dworkis, D.A., Klings, E.S., Farrer, L.A., Garrett, M.E., Ashley-Koch, A*., et al.* (2009). Fetal hemoglobin in sickle cell anemia: genome-wide association studies suggest a regulatory region in the 5' olfactory receptor gene cluster. Blood, blood-2009-2008-239517.

Splinter, E., Heath, H., Kooren, J., Palstra, R.-J., Klous, P., Grosveld, F., Galjart, N., and de Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the Î²-globin locus. Genes Dev *20*, 2349-2354.

Stamatoyannopoulos, G., and Grosveld, F. (2001). Hemoglobin switching. In The molecular basis of blood diseases (3rd ed), G. Stamatoyannopoulos, Majeurs P., Perlmutter, P., Varmus H., ed. (Philadelphia, W.B. Saunders), pp. 135-182.

Stamatoyannopoulos, G., Umemura, T., al-Khatti, A., Spadaccino, E., Abels, R.I., Fritsch, E.F., Papayannopoulou ,T. (1990). Modulation of HBF production by erythropoietin. Prog Clin Biol Res *316B*, 269-280.

Stamatoyannopoulos, G., Veith, R., Galanello, R., and Papayannopoulou, T. (1985). Hb F production in stressed erythropoiesis: observations and kinetic models. Ann N Y Acad Sci *445*, 188-197.

Stamatoyannopoulos, G., Wood, W.G., Papayannopoulou, T., and Nute, P.E. (1975). A new form of hereditary persistence of fetal hemoglobin in blacks and its association with sickle cell trait. Blood *46*, 683-692.

Stamatoyannopoulos, J.A., Goodwin, A., Joyce, T., and Lowrey, C.H. (1995). NF-E2 and GATA binding motifs are required for the formation of DNase I hypersensitive site 4 of the human beta-globin locus control region. EMBO J *14*, 106-116.

Steinberg, M.H., Forget, B.G., Higgs, D.R., and Nagel, R.L. (2009). . In Disorders of Hemoglobin: Genetics, Pathophysiology, and Clinical Management, M.H. Steinberg, Forget, B.G., Higgs, D.R., and D.J. Weatheral, eds. (Cambridge, Cambridge University Press), pp. 979-1027.

Stoming, T.A., Stoming, G.S., Lanclos, K.D., Fei, Y.J., Altay, C., Kutlar, F., and Huisman, T.H. (1989). An A gamma type of nondeletional hereditary persistence of fetal hemoglobin with a T----C mutation at position -175 to the cap site of the A gamma globin gene. Blood *73*, 329-333.

Surrey, S., Delgrosso, K., Malladi, P., and Schwartz, E. (1988). A single-base change at position -175 in the 5'-flanking region of the G gamma-globin gene from a black with G gamma-beta+ HPFH. Blood *71*, 807-810.

Talbort, D., and Grosveld, F. (1991). The 5'HS2 of the globin locus control region enhances transcription through the interaction of a multimeric complex binding at two functionally distinct NF-E2 binding sites. EMBO J *10*, 1391-1398.

Tanabe, O., Katsuoka, F., Campbell, A.D., Song, W., Yamamoto, M., Tanimoto, K., and Engel, J.D. (2002). An embryonic/fetal beta-type globin gene repressor contains a nuclear receptor TR2/TR4 heterodimer. EMBO J *21*, 3434-3442.

Tanabe, O., McPhee, D., Kobayashi, S., Shen, Y., Brandt, W., Jiang, X., Campbell, A.D., Chen, Y.-T., Chang, C.s., Yamamoto, M.*, et al.* (2007a). Embryonic and fetal beta-globin gene repression by the orphan nuclear receptors, TR2 and TR4. EMBO J *26*, 2295-2306.

Tanabe, O., Shen, Y., Liu, Q., Campbell, A.D., Kuroha, T., Yamamoto, M., and Engel, J.D. (2007b). The TR2 and TR4 orphan nuclear receptors repress Gata1 transcription. Genes Dev *21*, 2832-2844.

Tanimoto, K., Sugiura, A., Omori, A., Felsenfeld, G., Engel, J.D., and Fukamizu, A. (2003). Human beta-Globin Locus Control Region HS5 Contains CTCF- and Developmental Stage-Dependent Enhancer-Blocking Activity in Erythroid Cells. Mol Cell Biol *23*, 8946-8952.

Tavian, M., Coulombel, L., Luton, D., Clemente, H.S., Dieterlen-Lievre, F., and Peault, B. (1996). Aorta-associated CD34+ hematopoietic cells in the early human embryo. Blood *87*, 67-72.

Thein, S.L., Sampietro, M., Rohde, K., Rochette, J., Weatherall, D.J., Lathrop, G.M., and Demenais, F. (1994). Detection of a major gene for heterocellular hereditary persistence of fetal hemoglobin after accounting for genetic modifiers. Am J Hum Genet *54*, 214-228.

Tolhuis, B., Palstra, R.-J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and Interaction between Hypersensitive Sites in the Active [beta]-globin Locus. Molecular Cell *10*, 1453-1465.

Torrealba-de Ron, A.T., Papayannopoulou, T., Knapp, M.S., Fu, M.F., Knitter, G., and Stamatoyannopoulos, G. (1984). Perturbations in the erythroid marrow progenitor cell pools may play a role in the augmentation of HbF by 5-azacytidine. Blood *63*, 201-210.

Tuan, D.Y., Solomon, W.B., I.M., L., and Lee, D.P. (1989). An erythroid-specific, developmental-stage-independent enhancer far upstream of the human "beta-like globin" genes. Proc Natl Acad Sci U S A *86*, 2554-2558.

Uda, M., Galanello, R., Sanna, S., Lettre, G., Sankaran, V.G., Chen, W., Usala, G., Busonero, F., Maschio, A., Albai, G.*, et al.* (2008). Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of beta-thalassemia. Proc Natl Acad Sci U S A *105*, 1620-1625.

Vakoc, C.R., Letting, D.L., Gheldof, N., Sawado, T., Bender, M.A., Groudine, M., Weiss, M.J., Dekker, J., and Blobel, G.A. (2005). Proximity among Distant Regulatory Elements at the [beta]-Globin Locus Requires GATA-1 and FOG-1. Molecular Cell *17*, 453-462.

van Assendelft, G.B., Hanscombe, O., Grosveld, F., and Greaves, D.R. (1989). The [beta]-globin dominant control region activates homologous and heterologous promoters in a tissue-specific manner. Cell *56*, 969-977.

Wai, A.W.K., Gillemans, N., Raguz-Bolognesi, S., Pruzina, S., Zafarana, G., Meijer, D., Philipsen, S., and Grosveld, F. (2003). HS5 of the human beta-globin locus control region: a developmental stage-specific border in erythroid cells. EMBO J *22*, 4489-4500.

Wainscoat, J.S., Thein, S.L., Wood, W.G., Weatherall, D.J., Metaxotou-Mavromati, A., Tzotos, S., Kanavakis, E., and Kattamis, C. (1985). A novel deletion in the beta-globin gene complex. Ann N Y Acad Sci *445*, 20-27.

Wall, L., Destroismaisons, N., Delvoye, N., and Guy, L.-G. (1996). CAAT/Enhancer-binding Proteins Are Involved in Î²-Globin Gene Expression and Are Differentially Expressed in Murine Erythroleukemia and K562 Cells. J Biol Chem *271*, 16477-16484.

Wang, W.C., Wynn, L.W., Rogers, Z.R., Scott, J.P., Lane, P.A., and Ware, R.E. (2001). A two-year pilot trial of hydroxyurea in very young children with sickle-cell anemia. The Journal of Pediatrics *139*, 790-796.

Watson, J. (1948). The significance of the paucity of sickle cells in newborn Negro infants. Am J Med Sci *215*, 419-423.

Weatherall, D.J., Ledingham J.G.G., Warrell DA. (1996). In Oxford textbook of medicine 3rd ed, L. Weatherall DJ, and W.D. JGG, eds. (Oxford, Oxford University Press).

Weinberg, R.S., Ji, X., Sutton, M., Perrine, S., Galperin, Y., Li, Q., Liebhaber, S.A., Stamatoyannopoulos, G., and Atweh, G.F. (2005). Butyrate increases the efficiency of translation of gamma-globin mRNA. Blood *105*, 1807-1809.

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

Wood, W.G., Clegg, J.B., Weatherall, D.J. (1977). Developmental biology of human hemoglobins. Prog Hematol *10*, 43-90.

Wyszynski, D.F., Baldwin, C.T., Cleves, M.A., Amirault, Y., Nolan, V.G., Farrell, J.J., Bisbee, A., Kutlar, A., Farrer, L.A., and Steinberg, M.H. (2004). Polymorphisms near a chromosome 6q QTL area are associated with modulation of fetal hemoglobin levels in sickle cell anemia. Cell Mol Biol (Noisy-le-grand) *50*, 23-33.

Yin, W., Barkess, G., Fang, X., Xiang, P., Cao, H., Stamatoyannopoulos, G., and Li, Q. (2007). Histone acetylation at the human beta-globin locus changes with developmental age. Blood *110*, 4101-4107.

Zhang, Y., Ng, H.-H., Erdjument-Bromage, H., Tempst, P., Bird, A., and Reinberg, D. (1999). Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. Genes Dev *13*, 1924-1935.

Zhao, Q., Rank, G., Tan, Y.T., Li, H., Moritz, R.L., Simpson, R.J., Cerruti, L., Curtis, D.J., Patel, D.J., Allis, C.D*., et al.* (2009). PRMT5-mediated methylation of histone H4R3 recruits DNMT3A, coupling histone and DNA methylation in gene silencing. Nat Struct Mol Biol *16*, 304-311.

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

Zhou, W., Zhao, Q., Sutton, R., Cumming, H., Wang, X., Cerruti, L., Hall, M., Wu, R., Cunningham, J.M., and Jane, S.M. (2004). The Role of p22 NF-E4 in Human Globin Gene Switching. J Biol Chem *279*, 26227-26232.

Zimmerman, S.A., Schultz, W.H., Davis, J.S., Pickens, C.V., Mortier, N.A., Howard, T.A., and Ware, R.E. (2004). Sustained long-term hematologic efficacy of hydroxyurea at maximum tolerated dose in children with sickle cell disease. Blood *103*, 2039-2045.

Scope of this thesis

β-thalassemia and sickle cell disease are major human genetic health problem in many parts of the world. Available treatments are not satisfactory as none of them exhibit the optimal combination of safety, efficiency and convenience of use that would make them applicable to most hemoglobinopathy patients, especially to those who lack access to modern medical facilities. The observation that induction of γ-globin gene expression ameliorates the disease phenotype led to the proposal to induce γ-globin gene expression for the treatment of these diseases. Screening of a small molecule libraries (186000 compounds) has been done but new reagents with a higher HbF inducing effect and reduced cytotoxicity than those already known (butyrates, HU) were not found. Despite all the attempts made during the last 30 years, it is still not known how the γ-globin genes are switched off after birth, e.g. whether it is absence of activating factors or presence of suppressing factors or combination of both that leads to the down regulation of the γ-globin genes in the adult. There is only very limited information about the factors that are bound to the γ-globin promoters in vivo, certainly in its repressed state. It is likely that factors other than those described in the literature (at the start of this work) are essential for the suppression process. We therefore wanted to develop new strategies that specifically target fetal gene activation without cytotoxicity, widespread epigenetic alterations or difficult to manage side effects through the identification of the relevant transcription factors acting at the promoter.

To this end we designed a strategy to isolate and identify protein factors bound *in vivo* to the suppressed human γ-globin gene promoter and to study their effect on γ-globin regulation (chapter 3). This targeted, *in vivo* single gene promoter chromatin purification has been carried out for the first time and is a completely novel strategy. To optimize such a purification protocol, different parameters have been tested including the use of various purification tags, reagents, buffers, etcetera (presented in chapter2).

An alternative approach would be to identify target molecules within pathways that are involved in γ-globin gene expression. These pathways can be studied in the context of high and low HbF expressing β-thalassemia and sickle cell disease patients (chapter 4). There are many reports describing the induction of different HbF levels in response to hydroxyurea treatments in patients. We therefore wanted to understand the mechanism by which HU induces γ-globin in patients using expression analysis between so-called 'responder' and 'non-responder' patients, i.e. between those showing high γ-globin production versus low γ-globin production. The results of this study are presented in chapter4.

Finally in chapter 5 we present a novel set of genetic markers that is associated with increased γ-gobin expression in β-thalassemia patients followed by a discussion of future prospects in Chapter 6.

Chapter2

Optimal use of tandem biotin and V5 tags in ChIP assays

Katarzyna E Kolodziej, **Farzin Pourfarzad**, Ernie de Boer, Sanja Krpic, Frank Grosveld and John Strouboulis.

BMC Mol Biol. 2009 Feb 5;10:6.

BMC Molecular Biology

BioMed Central

Methodology article

Open Access

Optimal use of tandem biotin and V5 tags in ChIP assays

Katarzyna E Kolodziej[1,2], Farzin Pourfarzad[1], Ernie de Boer[1], Sanja Krpic[1], Frank Grosveld*[1] and John Strouboulis*[1,3]

Address: [1]Department of Cell Biology, Erasmus MC, Dr Molewaterplein 50, 3015GE Rotterdam, the Netherlands, [2]Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, the Netherlands and [3]Institute of Molecular Oncology, Biomedical Sciences Research Center Alexander Fleming, PO Box 74145, 16602 Varkiza, Greece

Email: Katarzyna E Kolodziej - k.kolodziej@nki.nl; Farzin Pourfarzad - f.pourfarzad@erasmusmc.nl; Ernie de Boer - ernie.deboer@erasmusmc.nl; Sanja Krpic - s.krpic@erasmusmc.nl; Frank Grosveld* - f.grosveld@erasmusmc.nl; John Strouboulis* - strouboulis@fleming.gr

* Corresponding authors

Published: 5 February 2009

BMC Molecular Biology 2009, **10**:6 doi:10.1186/1471-2199-10-6

Received: 20 September 2008
Accepted: 5 February 2009

This article is available from: http://www.biomedcentral.com/1471-2199/10/6

Abstract

Background: Chromatin immunoprecipitation (ChIP) assays coupled to genome arrays (Chip-on-chip) or massive parallel sequencing (ChIP-seq) lead to the genome wide identification of binding sites of chromatin associated proteins. However, the highly variable quality of antibodies and the availability of epitopes in crosslinked chromatin can compromise genomic ChIP outcomes. Epitope tags have often been used as more reliable alternatives. In addition, we have employed protein in vivo biotinylation tagging as a very high affinity alternative to antibodies. In this paper we describe the optimization of biotinylation tagging for ChIP and its coupling to a known epitope tag in providing a reliable and efficient alternative to antibodies.

Results: Using the biotin tagged erythroid transcription factor GATA-1 as example, we describe several optimization steps for the application of the high affinity biotin streptavidin system in ChIP. We find that the omission of SDS during sonication, the use of fish skin gelatin as blocking agent and choice of streptavidin beads can lead to significantly improved ChIP enrichments and lower background compared to antibodies. We also show that the V5 epitope tag performs equally well under the conditions worked out for streptavidin ChIP and that it may suffer less from the effects of formaldehyde crosslinking.

Conclusion: The combined use of the very high affinity biotin tag with the less sensitive to crosslinking V5 tag provides for a flexible ChIP platform with potential implications in ChIP sequencing outcomes.

Background

Affinity tags have been widely used for the study of protein interactions and the isolation of protein complexes. Such tags are also increasingly used in ChIP assays in detecting the in vivo binding of transcription factors and associated co-factors to their target genes in chromatin. In searching for the optimal affinity tag for ChIP applications, three criteria are important: (a) tags must have high binding affinity; (b) tags should be preferably small and not strongly charged so as to minimize possible interference with transcription factor function (c) tags should be fairly insensitive to formaldehyde fixation. The latter is

true for most tags that contain no or few lysine, arginine or histidine residues [1-3].

The biotin/(strept)avidin affinity system fulfils the above criteria due to its unique characteristics [4], which include: (a) the very tight and specific binding of biotin by avidin (or streptavidin) which, with a K_d of 10^{15} $L*mol^{-1}$, is one of the highest non covalent interactions known in nature, close to almost 10^3 – 10^6times greater than the interaction of epitopes with their specific antibodies. Once formed, the biotin-streptavidin complex is not disturbed by changes in pH, introduction of detergents or high salt concentration, thus remaining stable even under very stringent washing conditions; (b) biotin is a very small molecule and is not known to affect the biological activity of tagged proteins [5,6]; (c) there are few (mostly cytoplasmic) naturally biotinylated proteins in mammalian cells, as a result the non-specific background binding of nuclear extract is low [7].

We have previously used [7,8] a short (23aa) biotinylatable tag [9,10] for the purification of GATA-1 protein complexes from nuclear extracts of erythroid cells. GATA-1 is a DNA sequence-specific zinc finger transcription factor that is essential for the differentiation of erythroid, megakaryocytic, eosinophil and mast cell lineages [11,12]. N-terminally tagged GATA-1 was co-expressed with the E. coli BirA biotin ligase in mouse erythroleukemic (MEL) cells and subsequently purified from nuclear extracts together with interacting proteins by high affinity binding to streptavidin beads [7]. In this way, a number of known and novel GATA-1 protein partners were identified [8]. We also tested the utility of the biotin tag and streptavidin binding in ChIP assays and provided preliminary evidence that it can be successfully applied in place of antibodies in ChIPs of GATA-1 target genes [7,13]. Subsequent work in other labs has provided further supporting evidence for the application of biotinylation tagging in ChIP and Chip-on-chip assays [14-16]. Thus, despite the fact that biotin contains groups that are crosslinkable by formaldehyde, it can be successfully employed in ChIP assays

In this manuscript we present steps for improving the efficiency of biotinylation tagging in ChIP applications, using biotin-tagged GATA-1 in combination with known target genes [8] as an example. We first show that different streptavidin beads are not equally efficient in ChIP assays. We also show that effective blocking with fish skin gelatin and omission of SDS during chromatin sonication are important factors in reducing background signals, which is a major concern in ChIP using complex chromatin from mammalian cells. Furthermore, we explored the utility of double affinity tags in ChIP assays. Different tags may be used in tandem, separated by a protease cleavage site to allow for differential purification using either tag or for two sequential affinity purification steps using both tags to lower the background of non-specific proteins. At the same time, this approach can greatly enhance the ability to purify the complexes to homogeneity (by using several factors of the same complex differentially tagged and co-expressed) for other applications. To these ends, we combined the biotin tag with the V5 tag and show that the V5 tag antibody mediated ChIP is as efficient as the biotin streptavidin ChIP. These results have important implications when it comes to selecting an optimal strategy for genomic ChIP and proteomic analyses of transcription factor functions.

Results

We previously showed using biotin-tagged GATA-1 that streptavidin binding of crosslinked chromatin can substitute for antibodies in enriching for GATA-1 target genes in ChIP assays [7]. Due to the potential advantages offered by the very high affinity of streptavidin for binding to biotin and the importance of having multiple tags that can be used in the same cell, we wanted to extend these observations in developing an optimized protocol for the streptavidin binding of chromatin from cells expressing biotin-tagged GATA-1. In doing so, we used the EKLF and c-myb promoters as examples of GATA-1 gene targets that are upregulated or repressed, respectively, in erythroid cells. Figure 1A (EKLF) and Additional File 1A (myb) show the location of primers used for the EKLF and myb promoters and negative control sequences. Primer sequences are listed in Table 1.

Comparison of different types of streptavidin beads

Biotinylation of biological substrates is frequently used in a variety of different applications and hence many manufacturers offer a wide range of immobilized streptavidin matrices. We have previously used paramagnetic M280 streptavidin Dynabeads for the isolation of protein complexes [7,8,17-19]. We also tested the performance of M280 beads in chromatin immunoprecipitations and compared them to three other available matrices: streptavidin agarose, streptavidin mutein and NeutrAvidin. Streptavidin agarose was used to test whether an immobilization matrix different to that of paramagnetic particles would give better yields. NeutrAvidin is a streptavidin derivative without carbohydrate side chains which is predicted to reduce background binding. Streptavidin mutein is a mutated recombinant streptavidin which binds biotin with a lower affinity thus allowing elution of bound material under gentler conditions by using biotin. Chromatin from biotin-tagged GATA-1 cells and control cells expressing BirA ligase only, was bound to different types of beads under identical conditions (overnight binding and subsequent washes). Biotin-tagged GATA-1 was eluted from the beads by decrosslinking, except for the mutein beads

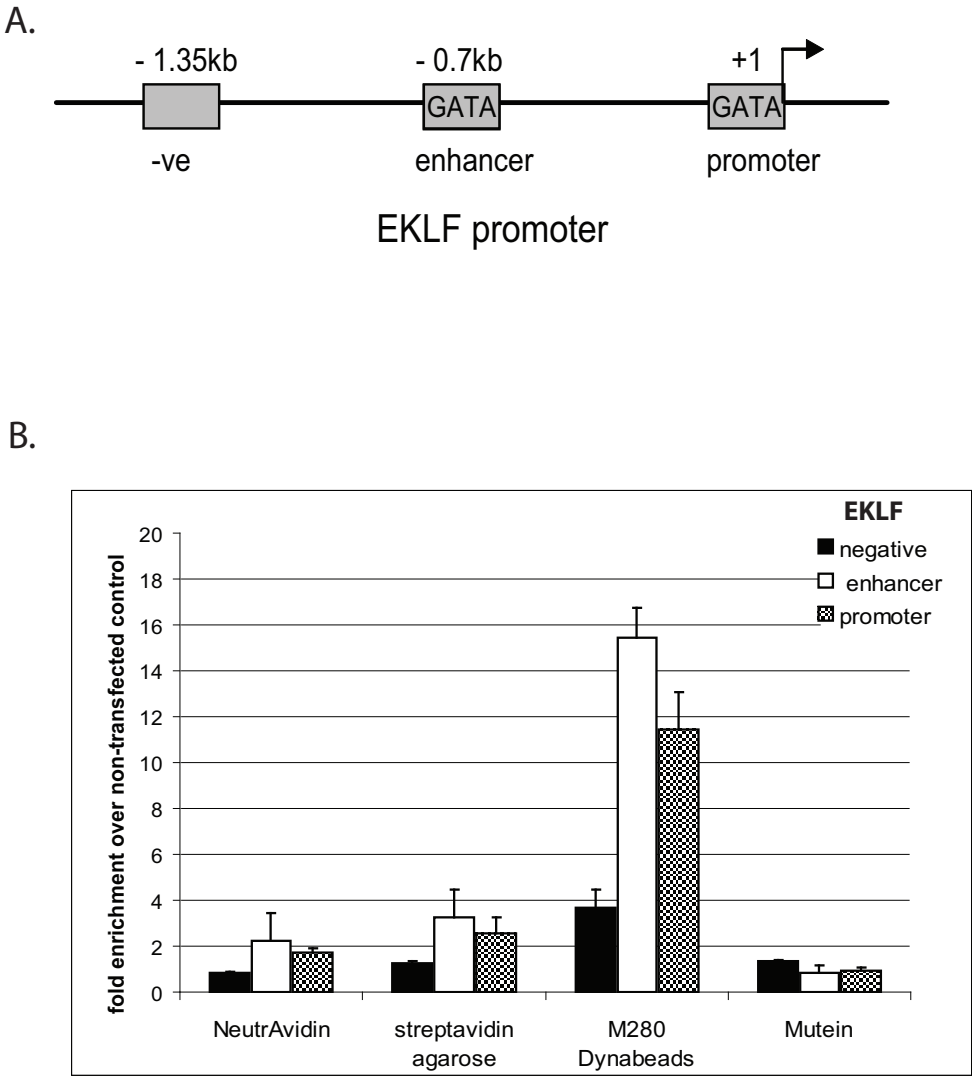


Figure 1
A) Location of the ChIP primers in the EKLF gene. "GATA" boxes indicate GATA-1 binding sites. B) Comparison of different derivatives of immobilized streptavidin: NeutrAvidin, streptavidin agarose, streptavidin mutein and M280 Dynabeads. Relative enrichment for EKLF sequences was calculated over negative control chromatin isolated from cells expressing BirA biotin ligase, but not tagged GATA-1.

Table 1:

Primer name	Sequence
Myb prom FOR	ACTGCAGGGGCGCCAGATTT
Myb prom REV	GGAGAAAGGGGAGGAGAAGGAGGTA
Neg myb FOR	GAAGTAGAGGCAGGATAATCAGGAA
Neg myb REV	AGGATGAACCAGGGCTAATGC
EKLF Upstream FOR	CTGGCCCCCCTACCTGAT
EKLF Upstream REV	GGCTCCCTTTCAGGCATTATC
EKLF Promoter FOR	TATCGCACACACCCCTCCTT
EKLF Promoter REV	CCCACATCTGATTGGCTGTCT
Neg EKLF FOR	TGCTCCCCACTATGATAATGGA
Neg EKLF REV	GCCACAACCAAAGAAGACATTTT
Necdin FOR	GGTCCTGCTCTGATCCGAAG
Necdin REV	GGGTCGCTCAGGTCCTTACTT
Amylase FOR	CTCCTTGTACGGGTTGGT
Amylase REV	AATGATGTGCACAGCTGAA

where we used biotin for elution. We find that the M280 streptavidin Dynabeads are the most efficient in capturing biotin-tagged GATA-1 bound to the EKLF (Figure 1B) and myb (Additional File 1B) promoters. Using M280 beads, we also find clear enrichment of GATA-1 binding to regulatory elements of the repressed GATA-2 locus (Additional File 2B). As a result, the M280 Dynabeads were used in all subsequent experiments.

Pre-clearing chromatin

Pre-clearing of chromatin is one of the methods used to decrease background binding in ChIP assays using antibodies. We tested this by preclearing chromatin with Protein G paramagnetic beads (Dynal) for 1 hour at 4°C. As shown in Figure 2, this resulted in lower background and improved enrichment of EKLF sequences bound by biotin tagged GATA-1. Similar results were also obtained with the c-myb promoter (Additional File 1C).

Blocking with fish skin gelatin

Among various blocking compounds (e.g. BSA, Chicken Egg Albumin etc.) fish skin gelatin (FGEL) has been shown to be very effective for blocking in Western and ELISA experiments [20,21]. In addition, FGEL is a very inexpensive reagent compared to other blocking reagents such as BSA. We therefore investigated whether the use of FGEL for blocking would improve the performance of M280 beads in a streptavidin ChIP. Figure 3 shows that addition of as little as 0.5% (final concentration) of FGEL (together with salmon sperm DNA) can significantly improve the yield of EKLF target sequences bound by GATA-1. Similar results were also observed with the c-myb promoter and GATA-2 locus sequences (Additional Files 1C and 2, respectively). Thus, blocking the beads for 1 hour with FGEL and salmon sperm DNA reduces the background compared to blocking with salmon sperm DNA alone. In addition to blocking the beads, we also added 1% FGEL to chromatin during binding to the beads and obtained similar results to those when FGEL was used for blocking the beads only (not shown). As a result, we have included 1% FGEL in blocking the beads in all subsequent experiments.

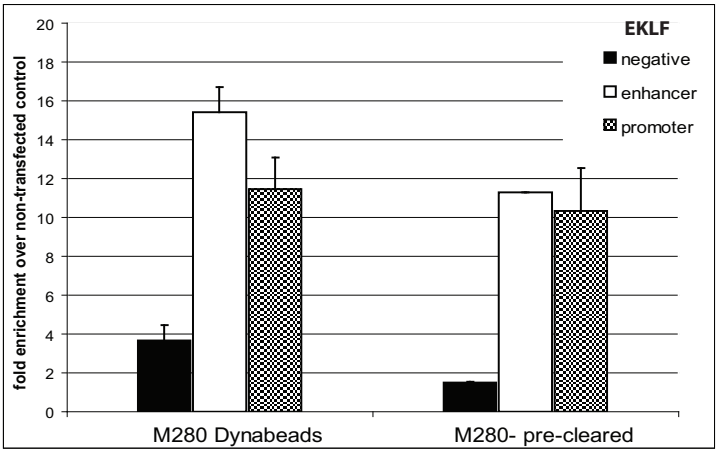
It has been shown previously that biotin tagging allows more stringent washes (containing up to 3% SDS) compared to other affinity tags [15]. For example, urea and thiourea are reagents widely used in proteomics to resuspend hydrophobic proteins. We therefore tested whether the background binding of hydrophobic proteins can be reduced by washing in urea/thiourea/SDS. We found that the additional wash did not significantly lower the background or increase the specific binding signals of the EKLF or c-myb promoters (data not shown) and this parameter was not investigated further.

The non-covalent binding of biotin to streptavidin is one of the strongest known in nature [22,23]. However this presents a drawback when eluting bound chromatin from the beads as is usually done in ChIP using antibodies. We were indeed unsuccessful in eluting the biotinylated protein from the streptavidin beads. The only way it could be removed was by boiling, which may result in some background due to the co-elution of non-specifically bound proteins. Alternatives would be the inclusion of protease (TEV or PreScission) cleavage sites [24-27], or the use of double tags (see below).

Sonication without SDS

Most ChIP protocols, including those used in our laboratory, are based on the Upstate (now Millipore) ChIP protocol which includes sonicating chromatin in a buffer containing 1% SDS. Addition of SDS introduces stringent conditions and helps prevent the aggregation of insoluble protein complexes. However, high SDS concentration may affect optimal binding of chromatin by the antibody or beads and, in some approaches, inclusion of SDS is not compatible with further experimental procedures, for example in chromatin fractionation by CsCl gradient centrifugation [28,29].

A.



B.

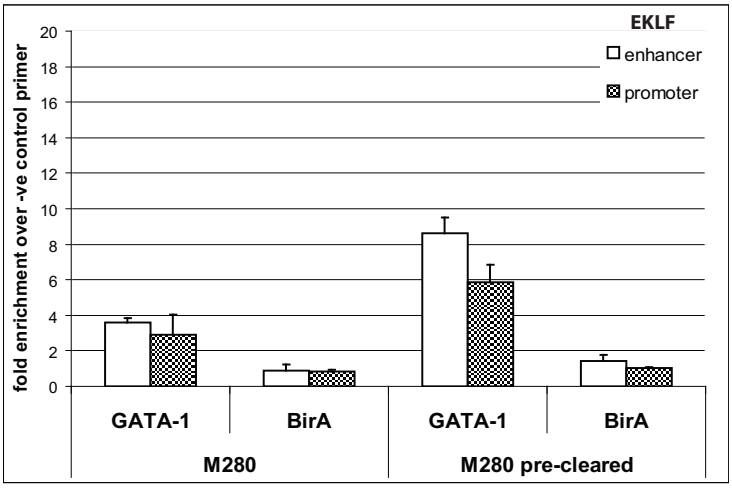


Figure 2
Preclearing chromatin with Protein G Dynabeads. A) Relative enrichment of EKLF sequences calculated over chromatin from control cells. The enrichment of the specific EKLF promoter elements appears lower after preclearing presumably due to some loss of chromatin in the additional preclearing step. B) Relative enrichment of biotin-tagged GATA-1 binding at EKLF promoter and enhancer calculated over the negative control sequence for biotin-tagged GATA-1 chromatin and chromatin from BirA expressing cells as negative control.

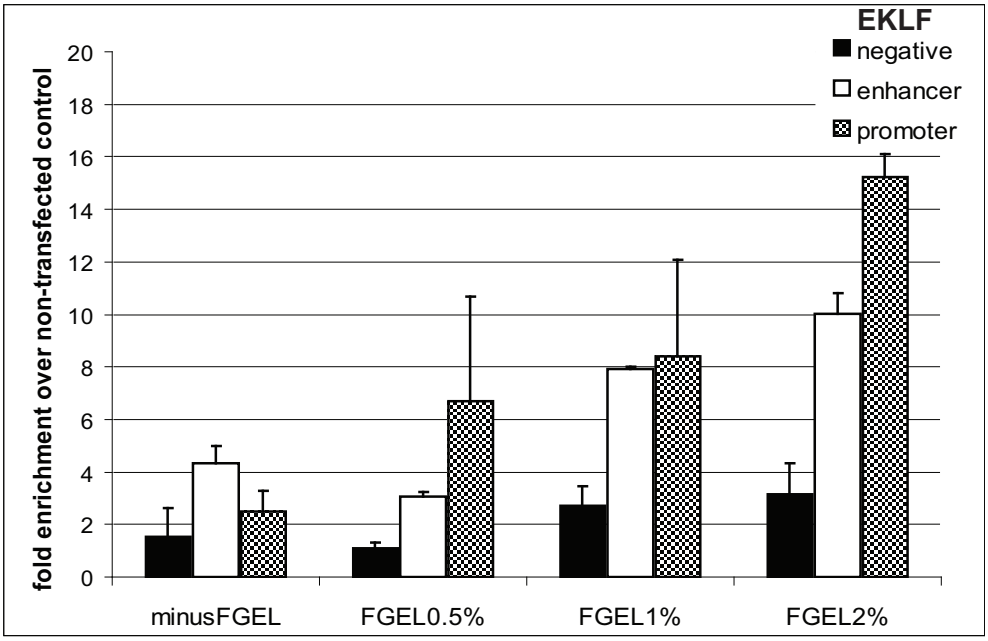


Figure 3
The effect of using different concentrations of FGEL in blocking the beads. Relative enrichment was calculated for EKLF sequences over chromatin from non transfected control cells. Low enrichment of the EKLF promoter and enhancer sequences is observed when blocking with salmon sperm DNA alone (minus FGEL control) due to the high background binding obtained with chromatin from the non-transfected control cells.

We therefore tested whether sonicating chromatin without SDS would improve the efficiency of a streptavidin ChIP. Omission of SDS did not affect the efficiency of DNA shearing. Sonicating chromatin without SDS resulted in higher enrichment of the EKLF (Figure 4) and c-myb (Additional File 1D) promoter sequences, albeit with a small increase in the background binding of the EKLF negative control sequence. Thus, omitting SDS from the sonication buffer improves the yield of a streptavidin pull-down significantly. We next tested whether the omission of SDS would also improve a ChIP in a regular antibody precipitation. For this purpose we tested the precipitation of a RAD21 gene target which, under the standard SDS conditions, can be enriched 4–5 fold over background. This represents a borderline enrichment for further analysis by ChIP-sequencing (SK unpublished). When a RAD21 ChIP to a site in the β globin locus in I/11 erythroid cells was carried out with or without SDS, the sample without SDS gave a considerable improvement in enrichment (Figure 5).

Comparison of biotin and V5 epitope tags to anti GATA-1 N6 and M20 antibodies

The experiments described above were carried out with an N-terminally biotin-tagged GATA-1 [7]. We also generated a second construct containing a tandem affinity tag created by fusing the short (14 aa) biotin tag [10] with the 14aa V5 tag to the C-terminus of GATA-1 (Figure 6A). V5 is a short peptide sequence derived from the C-terminus of the P and V proteins of Simian virus 5 [30,31]. This construct can be used in the two-step affinity purification of tagged protein complexes, thus reducing background binding. Alternatively, one of the tags can be used on its own in cases where the second tag is inefficient, for example due to reduced accessibility in crosslinked chromatin. These are important considerations in ChIP experiments particularly in applications involving ChIP-on-chip or

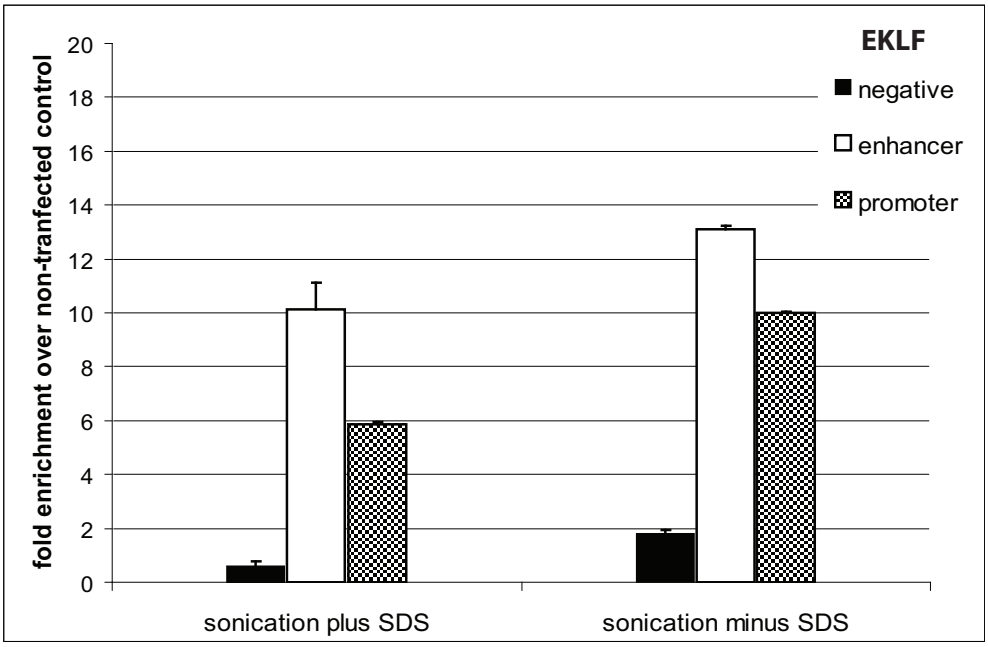


Figure 4
Comparison of sonication buffer with or without the addition of SDS. Relative enrichment of EKLF sequences was calculated over chromatin from control cells.

ChIP sequencing. This construct also allows comparison with the streptavidin-ChIP results obtained with the biotin tag fused to the N-terminus of GATA-1.

We obtained a MEL stable transfectant expressing GATA-1-V5-bio at approximately equal levels to the endogenous GATA-1 protein (Figure 6B). We compared the efficiency of ChIP by streptavidin binding or V5 antibody immunoprecipitation to two anti-GATA-1 antibodies: the N6 rat monoclonal antibody against the N-terminus of GATA-1 and the M20 goat polyclonal antibody against the C-terminus of GATA-1. We find that both the streptavidin pull-down with M280 Dynabeads and the ChIP with V5 antibodies are at least as good as or more efficient in enriching for EKLF sequences compared to the anti-GATA-1 N6 and M20 antibodies, when compared to IgG controls or chromatin from cells expressing BirA only (Figure 6C and 6D).

The V5-ChIP also works very efficiently, as it gives 12-fold higher enrichment in specific binding to the EKLF enhancer in comparison to the non-specific binding to non-related sequence (Figure 6C). In fact, the V5 tag appears to work at least as good as streptavidin binding in ChIP. When normalising to control cells expressing only BirA, the V5 ChIP gives actually a slightly better enrichment compared to the streptavidin ChIP (Figure 6C), albeit with a slightly higher background binding to the negative control sequence (Figure 6C and 6D). However, we cannot exclude that the elution from the anti-V5 agarose beads is more efficient than that from the M280 streptavidin beads.

Formaldehyde crosslinking affects the biotin-tag more than the V5 tag

Formaldehyde cross-linking as first introduced by Salomon et al in *Drosophila* [32], has been widely used to study the binding of proteins to DNA elements in intact cells. Formaldehyde crosslinks proteins primarily through lysine, glutamine, asparagine, arginine, tryptophan, tyrosine and histidine residues [1]. Biotinylation tagging takes place through the addition of a biotin moiety to a single lysine residue present in the peptide tag, thus rendering this lysine residue unavailable for crosslinking, However, the biotin molecule has two nitrogens in the ring structure

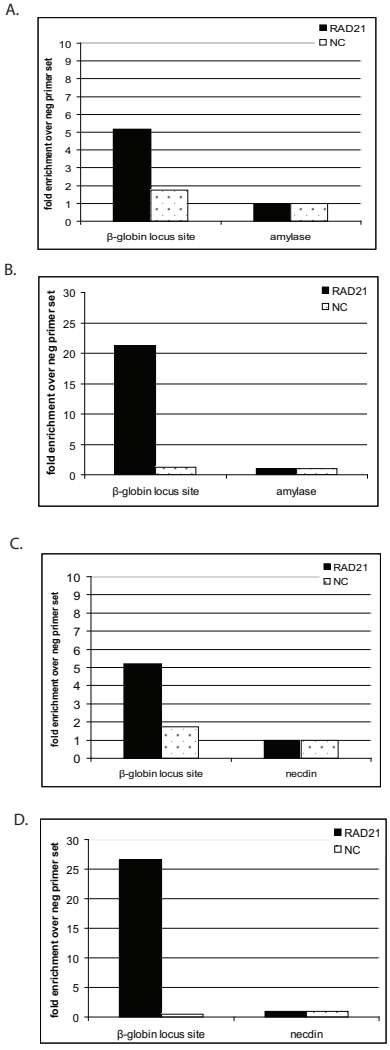


Figure 5
A) Chromatin from I/11 cells prepared with SDS in sonication buffer. Binding of RAD 21 and negative control (NC) to a β-globin site and the amylase gene (control for non-specific binding). B) Chromatin from I/11 cells prepared without SDS in sonication buffer. Binding of RAD 21 to specific site in β-globin locus is significantly improved in comparison to the result in panel A. C) Chromatin from I/11 cells prepared with SDS in sonication buffer. Binding of RAD 21 and negative control (NC) to a β-globin site and the necdin gene (control for non-specific binding). D) Chromatin from I/11 cells prepared without SDS in sonication buffer. Binding of RAD 21 to specific site in β-globin locus is significantly improved in comparison to result in panel C. Note that the binding of negative control (NC) to β-globin locus site is reduced.

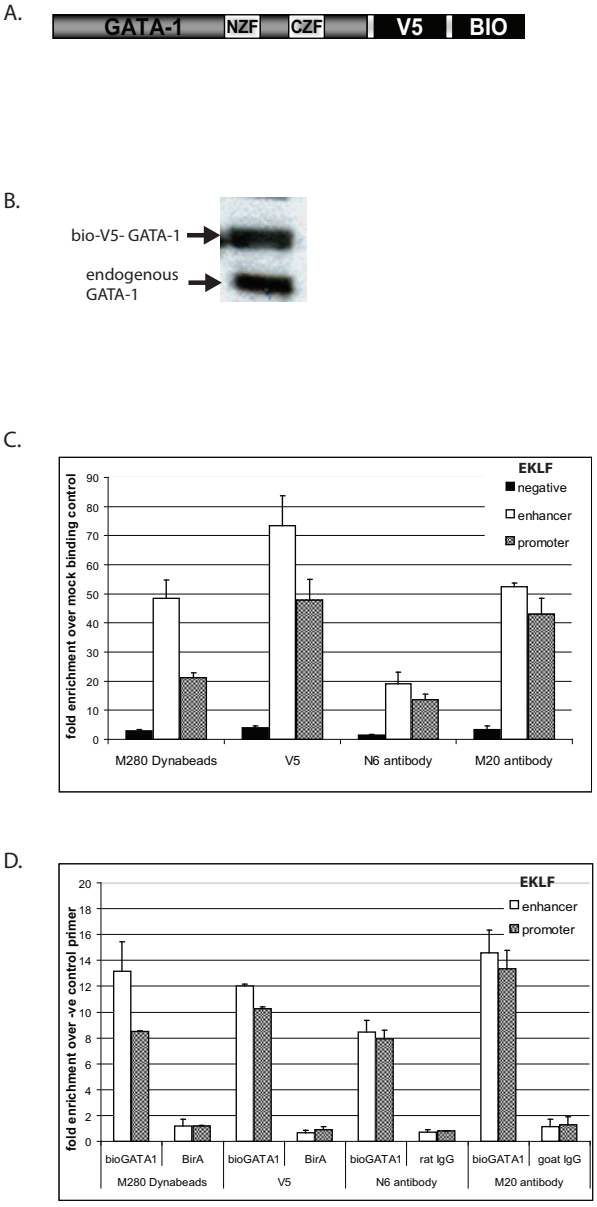


Figure 6 *(see legend on next page)*

Figure 6 *(see previous page)*
A) Schematic representation of the C-terminally biotin-V5-tagged GATA-1. NZF and CZF: N-terminal and C-terminal zinc fingers, respectively. V5 and biotin (BIO) tags are not drawn to scale. B) Western blot detected with anti-GATA-1 N6 antibody showing the relative amounts of biotin-V5-GATA-1 and endogenous GATA-1 in the cells used for ChIP. C) Comparison of V5, M280 and two different anti-GATA-1 antibodies, N6 and M20, tested for the binding of GATA-1 to the EKLF promoter. The enrichment is calculated over a BirA-only transfected control or over IgG negative control, respectively. V5 ChIP gives the highest yield in the EKLF enhancer and promoter elements, streptavidin ChIP with M280 Dynabeads gives comparable yield in the upstream enhancer element as M20 anti-GATA-1 antibody. The M20 antibody can enrich for more GATA-1 bound to the EKLF promoter than the M280 Dynabeads. The N6 antibody precipitates the least amount of GATA-1 bound to EKLF promoter elements. D) Comparison of V5, M280 and two different anti-GATA-1 antibodies N6 and M20, tested for the binding of GATA-1 to the EKLF promoter. Enrichment of the specific binding to EKLF promoter and enhancer was calculated over the negative primer set (-1.35 kb element in EKLF promoter). V5 agarose and M280 Dynabeads bring down comparable amounts of GATA-1 bound to EKLF enhancer and promoter sequences. The M20 antibody enriches the most for GATA-1 bound to the EKLF upstream enhancer, though this also included sequences in vivo bound by endogenous GATA-1 protein which can not be bound by M280 or anti-V5 beads. Rat and goat IgGs as well as BirA control show similarly low enrichments of specific primer sets.

that are crosslinkable (Figure 7A). With this in mind, we compared how the biotin tag or the V5 tag, a 14aa long tag containing lysine and asparagine residues (GKPIPNPLLGLDST), are affected by formaldehyde crosslinking. To this end, the extracts containing equal amounts of GATA-1-V5-bio (see input lanes 1 and 4 in figure 7B, top panel) were bound under identical conditions. Input, bound and unbound fractions were loaded on an SDS-PAGE gel and the western blots were detected with an anti GATA-1 (N6) antibody followed by detection with streptavidin-HRP. The results showed that the binding of crosslinked material to the M280 Dynabeads was good but less efficient compared to non-crosslinked protein extract, since there was more GATA-V5-bio found in unbound fraction (compare the amounts of GATA-1 in lane 2 and lane 5 of unbound fractions 7B, top panel). In addition, anti-GATA-1 antibody as well as streptavidin-HRP detection (Fig. 7B, both panels) showed that there was less GATA-1-V5-bio bound to the M280 Dynabeads in cross-linked extract pull-down than in non-crosslinked material (compare lanes 3 with 6 in Figure 7B, both panels). These results would be expected considering that some of the biotin tagged GATA-1 protein would be inaccessible for binding by M280-streptavidin beads in the crosslinked chromatin, as has also been observed with epitope access by antibodies in ChIP.

We next tested how formaldehyde crosslinking affects binding to anti-V5 agarose. The western blot data presented in Figure 7C suggest that crosslinked and non-crosslinked material can be efficiently bound by anti-V5 beads (Figure 7C top panel, compare lanes 2 with 3 and 5 with 6). The data suggest that the efficiency of binding of crosslinked chromatin using the V5 tag is higher than that of streptavidin ChIP (compare lanes 2, 3 of Figure 7C top panel with lanes 5, 6 of Figure 7B top panel). However, efficient recovery of chromatin from the V5 beads relies on efficient elution with the V5 peptide. In addition, in other cases we have not seen this difference (E. Soler, C. Andrieu, unpublished observations) suggesting that the structure of the target protein may also have an influence.

Discussion

We describe optimised conditions for the application of streptavidin-ChIP using crosslinked chromatin from cells expressing a biotin-tagged transcription factor. We show that pre-clearing chromatin and blocking streptavidin beads with fish skin gelatin reduces the background binding. In addition, avoiding SDS in the sonication buffer appears to increase enrichment for bio-GATA-1 binding to specific DNA sequences. We also compared streptavidin-ChIP to antibody ChIP using two different anti-GATA-1 antibodies and showed the former to be more efficient. The real difference in efficiencies between streptavidin-ChIP and antibody ChIP is even bigger because the anti-GATA-1 antibodies precipitate both the tagged and endogenous GATA-1, whereas streptavidin binds only tagged GATA-1. This means that the enrichments obtained with antibodies should be hypothetically approximately twice as high as with affinity tags pull-downs. The same applies for the efficiency of V5 ChIP, since, as mentioned above, the relative levels of bio-V5- tagged GATA-1 and endogenous GATA-1 in the extract used in this experiment are similar (Figure 6B). Thus, the results obtained with anti-GATA-1 antibodies are an overestimate when compared to the tag-based ChIPs described here.

A potential drawback of biotinylation tagging in streptavidin ChIP is related to previous reports of histones being naturally biotinylated [33]. Though it remains unknown what proportion of histones is biotinylated in vivo, the fact that previous studies have shown nuclear biotin to account for less than 1% of total cellular biotin [33,34] suggests that it is very little. In addition, we have per-

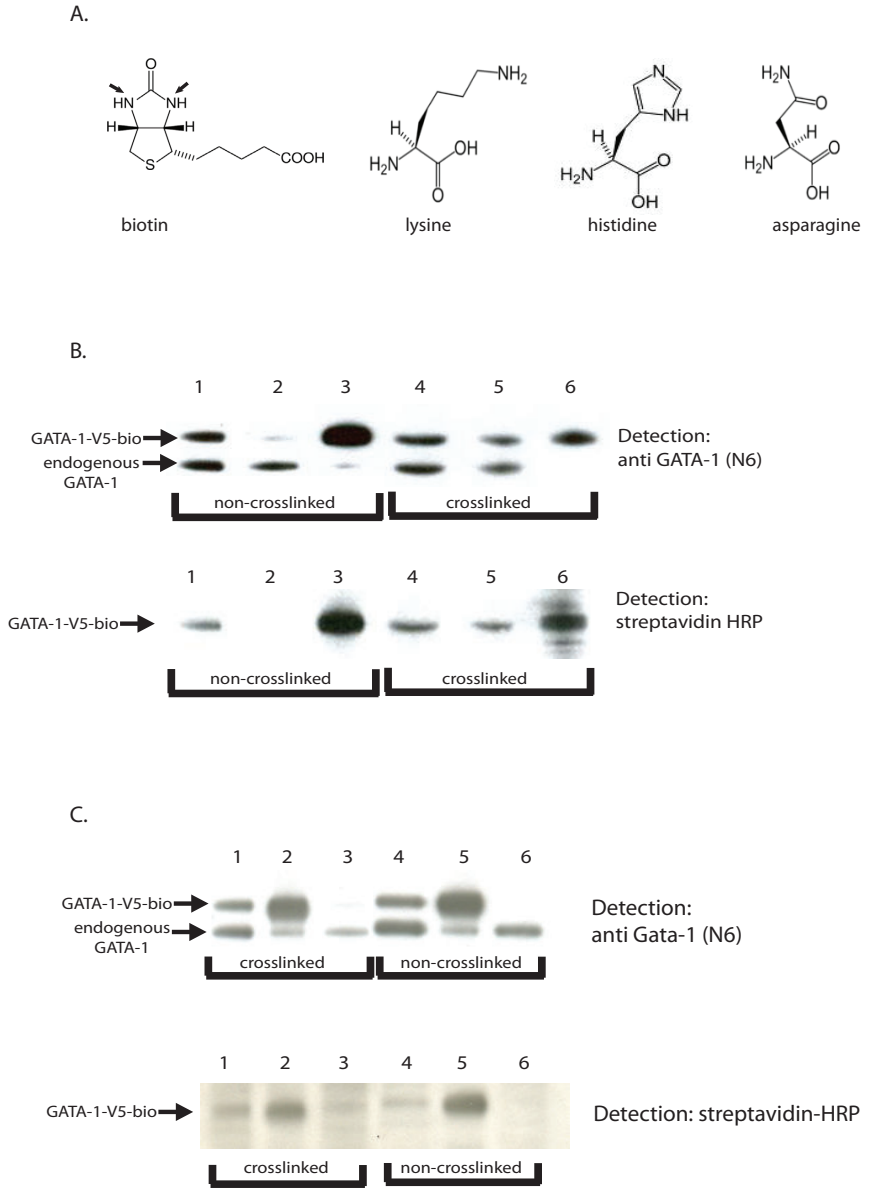


Figure 7 *(see legend on next page)*

Figure 7 *(see previous page)*
A) Structure of the formaldehyde sensitive amino acids and sensitive groups in biotin (arrows). B) Western blot analysis of two binding experiments where non-crosslinked nuclear extracts (lanes 1, 2, 3) and formaldehyde crosslinked chromatin (lanes 4, 5, 6) were tested. Input and supernatant fractions (lanes 1 and 2, 4 and 5 respectively) represent 1% of total material, while bound lanes (3 and 6) represent 25% of total material. The biotin-V5-tagged GATA-1 was detected with N6 anti-GATA-1 antibody (top panel). After stripping, the same membrane was incubated with streptavidin-HRP (lower panel). C) Western blot showing comparison of binding of two different extracts: crosslinked chromatin (lanes 1–3) and non-crosslinked nuclear extract (lanes 4–6) precipitated by anti-V5 agarose. The biotin-V5-tagged GATA-1 was detected with N6 anti-GATA-1 antibody (top panel), stripped and re-probed with streptavidin-HRP (lower panel). Input and supernatant fractions (lanes 1 and 3, 4 and 6 respectively) represent 1% of total material, while bound lanes (2 and 4) represent 50% of total material. Note that V5-tagged GATA-1 is completely depleted in both experiments as the supernatant lanes are empty.

formed extensive protein analyses by mass spectrometry and, whereas we find histones co-purifying with biotin tagged transcription factors, they do not represent any significant proportion of the background, nor did we notice an increase of histone peptides upon expression of BirA in cells [7]. Similarly, the purification of a histone variant protein complex by biotinylation tagging and streptavidin purification showed negligible histone binding [35]. Lastly, biotinylation tagging has been employed in ChIP-on-chip approaches of transcription factors and of the histone H3.3 variant, again with no evidence of background due to endogenous histone biotinylation [14,36]. Taken together, this evidence suggests that any background due to histone biotinylation is likely to be very low. In fact, one of the advantages of employing biotinylation tagging of different factors with streptavidin ChIP is that background will be the same in all cases, whereas different antibodies present with different backgrounds.

We also provide evidence that a tandem affinity tag composed of the biotin tag and the V5 epitope tag works very efficiently in ChIP. The biotin tag binding is fast and very tight, while the binding to V5 epitope is reversible (bound material can be eluted from the beads using V5 peptide), thus each tag can be used to advantage. On the basis of this evidence, we propose a scheme whereby optimal binding of chromatin, under the conditions described here, first by anti-V5 agarose, followed by elution using V5 peptide and re-binding by streptavidin beads, followed by elution/reversal of crosslinks provides a convenient and rapid purification method. Our preliminary ChIP-sequencing data (E. de Boer, unpublished) show that biotin and V5 tagging can both be very effectively used for transcription factor target sequence mapping and that the (very low) background seen in these experiments contains mostly non-specific DNA fragments that can be easily distinguished from specific target sites.

Conclusion

The optimal conditions for streptavidin ChIP described here and the use of biotin-V5 tandem affinity tagging of transcription factors offers an easy, rapid and effective way for comparative and functional studies of different transcription complexes.

Methods

Cells and constructs

MEL cells were cultured as previously described [35]. Constructs and stably transfected cell lines were described previously [7]. I/11 cells were cultured as previously described [37].

Chromatin crosslinking

Approximately 2×10^7 induced MEL cells were crosslinked with 1% formaldehyde for 10 minutes at room temperature and processed for sonication essentially as described in the Upstate (now Millipore) protocol http://www.millipore.com/techpublications/tech1/mcproto407. Chromatin was sonicated on ice with a Sanyo Soniprep 150 sonicator at amplitude 6 using 10 cycles of 15 sec "on" and 45 sec "off" to a DNA fragment size in the range of 300 to 800 nucleotides. The alternative "no SDS" sonication buffer is: 10 mM Tris, 1 mM EDTA and 0.5 mM EGTA. All buffers were supplemented with Complete Protease Inhibitor Cocktail (Roche). Aliquots of sonicated chromatin of 10×10^6 cells were stored at -80°C.

Streptavidin ChIP

Chromatin pull-downs with streptavidin beads were carried out overnight using 20 μl of streptavidin Dynabeads M280 (Invitrogen) or 20 μl of UltraLink Immobilized NeutrAvidin Protein (Pierce) per chromatin aliquot. For streptavidin agarose (Sigma) or Streptavidin mutein (Roche) chromatin pull-downs, 60 μl of agarose slurry or beads were used per aliquot. All the beads/slurry were blocked with 400 μg sonicated salmon sperm DNA for 1 h at 4°C. Pre-clearing of chromatin prior to binding to M280 streptavidin beads was done using 20 μl of Protein G Dynabeads (Invitrogen) preblocked with salmon sperm DNA. Chromatin incubation with beads was carried out in a total reaction volume of 1 mL supplemented with Complete Protease Inhibitor, at 4°C overnight on a rotating wheel. After binding, beads were washed with 1 mL of low salt, high salt, LiCl and TE (10 mM Tris-HCl pH 8.0,

1 mM ETDA) wash buffers, 3–5 minutes each. An additional urea wash (5 M urea/2 M thiourea/1% TritonX100) was carried out after LiCl buffer wash and before the TE washes (see Results). After the washes, bound chromatin was eluted by resuspending Neutravidin, Mutein, streptavidin agarose and M280 beads in 500 μl 0.1 M sodium carbonate, 1%SDS, 0.2 M NaCl elution buffer transferring to a fresh tube and decrosslinking with shaking at 65°C for at least 5 h. Thermal elution of chromatin from M280 Dynabeads was carried out by resuspending beads in 500 μl 95% formamide, 0.1 M sodium bicarbonate and boiling at 95°C for 10 min. M280 Dynabeads were subsequently separated from the buffer using a magnetic rack. Eluted chromatin was transferred to a fresh tube and decrosslinking was carried out as described above. Decrosslinked samples were deproteinized as described by the Upstate protocol. DNA was recovered by phenol/chloroform/isoamyl alcohol extraction and isopropanol precipitation using 20 μg glycogen as carrier.

Antibody ChIP

GATA-1 ChIP has been previously described [8]. Anti-GATA-1 antibodies used were N6 and M20 (Santa Cruz), anti-RAD21 antibodies were ab992-50 (Abcam). Anti GATA-1 antibody immunoprecipitates were eluted from the beads by incubation in elution buffer (0.1 M sodium carbonate, 1%SDS) twice for 15 minutes each at room temperature. The V5 ChIP was carried out with V5 beads (Sigma A7345). 60 μl of beads were spun and taken up in 1 ml PBS containing 200 μg sonicated salmon sperm DNA and 1.5% fish skin gelatin (Sigma G7765) mixed for 1 hour RT and washed with PBS. Diluted chromatin was taken up to 1 ml with Upstate ChIP dilution buffer and bound overnight to the blocked V5 beads at 4°C on a rotating wheel. Washing was carried out as described in the Upstate protocol http://www.millipore.com/techpublications/tech1/mcproto407. Elution was done in 40 μl HENG(10 mM HEPES-KOH pH 9.0, 1.5 mM $MgCl_2$, 0.25 mM EDTA, 20% glycerol), 250 mM KCl, 0.3% NP40 and 0.5 mg/ml V5 peptide repeated 3 times 20 min each. Eluates were pooled.

Blocking with Cold Sea Fish Skin Gelatin

A 45% Fish Skin gelatin (FGEL) stock solution (Sigma G7765) was used to block beads at 0.5%, 1% or 2% final concentrations together with sonicated salmon sperm DNA by incubation for 1 h at room temperature. Where indicated, FGEL was also added to chromatin and beads to a 1% final concentration during overnight binding.

Real time PCR

This was done in an Opticon I (MJ Research) thermal cycler using SYBR Green and Platinum Taq polymerase (Invitrogen) as described previously [8]. Primers (listed in Table 1) were designed using Primer Express 2.0 (PE Applied Biosystems). For each experiment at least two runs were done with each sample loaded in duplicate. PCR conditions: 95°C for 10 min, 40 cycles of 30 sec at 95°C, 60 sec at 60°C, 15 sec at 75°C. Enrichment for a specific DNA sequence was calculated using the comparative C_T method, as previously described [38]. The enrichment of bound DNA over input is calculated using the formula $2^{Ct(IP)-Ct(Ref)}$. Enrichment over the negative primer set or negative control chromatin from MEL cells expressing BirA only was subsequently calculated by dividing.

Nuclear extracts and immunoblotting

These were done as previously described [7].

Authors' contributions

KK carried out the streptavidin ChIP optimization experiments and drafted the manuscript. FP and EdB did the experiments involving the tandem biotin and V5 tags. SK did the Rad21 ChIP. FG and JS conceived of the study, participated in its design and coordination and co-authored the manuscript. All authors read and approved the final manuscript.

Additional material

Additional file 1

GATA-1 ChIP of the myb promoter. *A) Location of ChIP primers in the myb promoter. B) Comparison of different derivatives of immobilized streptavidin: NeutrAvidin, streptavidin agarose, streptavidin mutein and M280 Dynabeads. Relative enrichment is calculated over non-transfected BirA control cells. C) The effects of preclearing chromatin and using 1% FGEL in blocking the beads. Enrichment was calculated relative to negative primer set. D) The effect of omitting SDS from the sonication buffer. Relative enrichment is calculated over non-transfected BirA control cells.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2199-10-6-S1.pdf]

Additional file 2

GATA-1 ChIP of the GATA-2 gene locus. *A) Location of the ChIP primers in regulatory elements of the GATA-2 locus. B) GATA-1 binding using streptavidin M280 Dynabeads. C) The effect of blocking M280 Dynabeads with 1% FGEL. Enrichment was calculated relative to non-transfected BirA cells. Primer sequences are as published by Rodriguez et al. (ref. [8]).*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2199-10-6-S2.pdf]

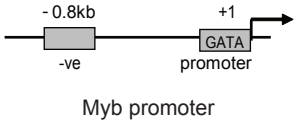
Acknowledgements

This work has been supported by grants from the Dutch Research Organization (NWO), the EU FP6 Networks of Excellence "Cells into Organs" and "The Epigenome", the EU FP6 Integrated Project "EUTRACC", the Cancer Genomics Center (NL) and the Center for Biomedical Genetics (NL).

References

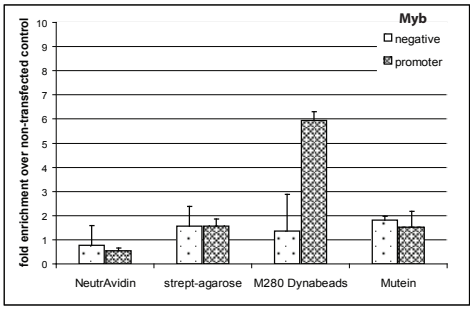
1. Metz B, Kersten GF, Hoogerhout P, Brugghe HF, Timmermans HA, de Jong A, Meiring H, ten Hove J, Hennink WE, Crommelin DJ, Jiskoot W: **Identification of formaldehyde-induced modifications in proteins: reactions with model peptides.** *The Journal of biological chemistry* 2004, **279:**6235-6243.
2. Orlando V, Strutt H, Paro R: **Analysis of chromatin structure by in vivo formaldehyde cross-linking.** *Methods (San Diego, Calif)* 1997, **11:**205-214.
3. Toth J, Biggin MD: **The specificity of protein-DNA crosslinking by formaldehyde: in vitro and in drosophila embryos.** *Nucleic acids research* 2000, **28:**e4.
4. Ford CF, Suominen I, Glatz CE: **Fusion tails for the recovery and purification of recombinant proteins.** *Protein expression and purification* 1991, **2:**95-107.
5. Diamandis EP, Christopoulos TK: **The biotin-(strept)avidin system: principles and applications in biotechnology.** *Clinical chemistry* 1991, **37:**625-636.
6. Laitinen OH, Nordlund HR, Hytonen VP, Kulomaa MS: **Brave new (strept)avidins in biotechnology.** *Trends in biotechnology* 2007, **25:**269-277.
7. de Boer E, Rodriguez P, Bonte E, Krijgsveld J, Katsantoni E, Heck A, Grosveld F, Strouboulis J: **Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice.** *Proceedings of the National Academy of Sciences of the United States of America* 2003, **100:**7480-7485.
8. Rodriguez P, Bonte E, Krijgsveld J, Kolodziej KE, Guyot B, Heck AJ, Vyas P, de Boer E, Grosveld F, Strouboulis J: **GATA-1 forms distinct activating and repressive complexes in erythroid cells.** *The EMBO journal* 2005, **24:**2354-2366.
9. Schatz PJ: **Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: a 13 residue consensus peptide specifies biotinylation in Escherichia coli.** *Bio/technology (Nature Publishing Company)* 1993, **11:**1138-1143.
10. Beckett D, Kovaleva E, Schatz PJ: **A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation.** *Protein Sci* 1999, **8:**921-929.
11. Cantor AB, Orkin SH: **Hematopoietic development: a balancing act.** *Current opinion in genetics & development* 2001, **11:**513-519.
12. Ferreira R, Ohneda K, Yamamoto M, Philipsen S: **GATA1 function, a paradigm for transcription factors in hematopoiesis.** *Molecular and cellular biology* 2005, **25:**1215-1227.
13. Horsman S, Moorhouse MJ, de Jager VC, Spek P van der, Grosveld F, Strouboulis J, Katsantoni EZ: **TF Target Mapper: a BLAST search tool for the identification of Transcription Factor target genes.** *BMC bioinformatics* 2006, **7:**120.
14. Kim J, Chu J, Shen X, Wang J, Orkin SH: **An extended transcriptional network for pluripotency of embryonic stem cells.** *Cell* 2008, **132:**1049-1061.
15. van Werven FJ, Timmers HT: **The use of biotin tagging in Saccharomyces cerevisiae improves the sensitivity of chromatin immunoprecipitation.** *Nucleic acids research* 2006, **34:**e33.
16. Viens A, Mechold U, Lehrmann H, Harel-Bellan A, Ogryzko V: **Use of protein biotinylation in vivo for chromatin immunoprecipitation.** *Analytical biochemistry* 2004, **325:**68-76.
17. Driegen S, Ferreira R, van Zon A, Strouboulis J, Jaegle M, Grosveld F, Philipsen S, Meijer D: **A generic tool for biotinylation of tagged proteins in transgenic mice.** *Transgenic research* 2005, **14:**477-482.
18. Sanchez C, Sanchez I, Demmers JA, Rodriguez P, Strouboulis J, Vidal M: **Proteomics analysis of Ring1B/Rnf2 interactors identifies a novel complex with the Fbxl10/Jhdm1B histone demethylase and the Bcl6 interacting corepressor.** *Mol Cell Proteomics* 2007, **6:**820-834.
19. Meier N, Krpic S, Rodriguez P, Strouboulis J, Monti M, Krijgsveld J, Gering M, Patient R, Hostert A, Grosveld F: **Novel binding partners of Ldb1 are required for haematopoietic development.** *Development (Cambridge, England)* 2006, **133:**4913-4923.
20. Kaur R, Dikshit KL, Raje M: **Optimization of immunogold labeling TEM: an ELISA-based method for evaluation of blocking agents for quantitative detection of antigen.** *J Histochem Cytochem* 2002, **50:**863-873.
21. Vogt RF Jr, Phillips DL, Henderson LO, Whitfield W, Spierto FW: **Quantitative differences among various proteins as blocking agents for ELISA microtiter plates.** *Journal of immunological methods* 1987, **101:**43-50.
22. Bayer EA, Wilchek M: **Application of avidin-biotin technology to affinity-based separations.** *Journal of chromatography* 1990, **510:**3-11.
23. Wilchek M, Bayer EA, Livnah O: **Essentials of biorecognition: the (strept)avidin-biotin system as a model for protein-protein and protein-ligand interaction.** *Immunology letters* 2006, **103:**27-32.
24. Lucast LJ, Batey RT, Doudna JA: **Large-scale purification of a stable form of recombinant tobacco etch virus protease.** *BioTechniques* 2001, **30:**544-546.
25. Kapust RB, Waugh DS: **Controlled intracellular processing of fusion proteins by TEV protease.** *Protein expression and purification* 2000, **19:**312-318.
26. Walker PA, Leong LE, Ng PW, Tan SH, Waller S, Murphy D, Porter AG: **Efficient and rapid affinity purification of proteins using recombinant fusion proteases.** *Bio/technology (Nature Publishing Company)* 1994, **12:**601-605.
27. Cordingley MG, Callahan PL, Sardana VV, Garsky VM, Colonno RJ: **Substrate requirements of human rhinovirus 3C protease for peptide cleavage in vitro.** *The Journal of biological chemistry* 1990, **265:**9062-9065.
28. Schwartz YB, Kahn TG, Pirrotta V: **Characteristic low density and shear sensitivity of cross-linked chromatin containing polycomb complexes.** *Molecular and cellular biology* 2005, **25:**432-439.
29. Solomon MJ, Larsen PL, Varshavsky A: **Mapping protein-DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene.** *Cell* 1988, **53:**937-947.
30. Zhou W, Dong L, Ginsburg D, Bouhassira EE, Tsai HM: **Enzymatically active ADAMTS13 variants are not inhibited by anti-ADAMTS13 autoantibodies: a novel therapeutic strategy?** *The Journal of biological chemistry* 2005, **280:**39934-39941.
31. Southern JA, Young DF, Heaney F, Baumgartner WK, Randall RE: **Identification of an epitope on the P and V proteins of simian virus 5 that distinguishes between two isolates with different biological characteristics.** *The Journal of general virology* 1991, **72(Pt 7):**1551-1557.
32. Solomon MJ, Varshavsky A: **Formaldehyde-mediated DNA-protein crosslinking: a probe for in vivo chromatin structures.** *Proceedings of the National Academy of Sciences of the United States of America* 1985, **82:**6470-6474.
33. Petrelli F, Moretti P, Paparelli M: **Intracellular distribution of biotin-14COOH in rat liver.** *Molecular biology reports* 1979, **4:**247-252.
34. Stanley JS, Griffin JB, Zempleni J: **Biotinylation of histones in human cells. Effects of cell proliferation.** *European journal of biochemistry/FEBS* 2001, **268:**5424-5429.
35. Antoniou M: **Induction of Erythroid-Specific Expression in Murine Erythroleukemia (MEL) Cell Lines.** In *Methods in Molecular Biology, Vol.7: Gene Transfer and Expression Protocols* Edited by: Murray EJ. The Humana Press Inc, Clifton, NJ; 1991.
36. Mito Y, Henikoff JG, Henikoff S: **Genome-scale profiling of histone H3.3 replacement patterns.** *Nature genetics* 2005, **37:**1090-1097.
37. Dolznig H, Boulme F, Stangl K, Deiner EM, Mikulits W, Beug H, Mullner EW: **Establishment of normal, terminally differentiating mouse erythroid progenitors: molecular characterization by cDNA arrays.** *Faseb J* 2001, **15:**1442-1444.
38. Litt MD, Simpson M, Recillas-Targa F, Prioleau MN, Felsenfeld G: **Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci.** *The EMBO journal* 2001, **20:**2224-2235.

A.

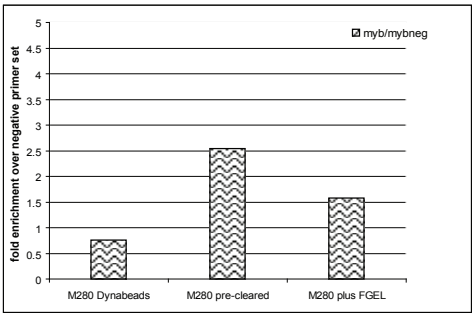


Myb promoter

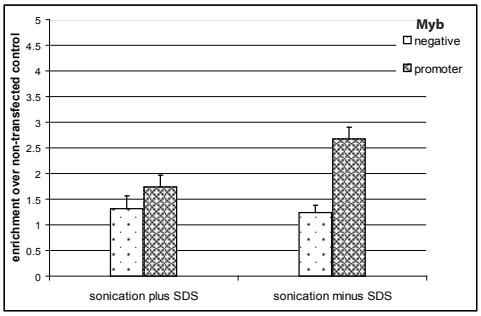
B.



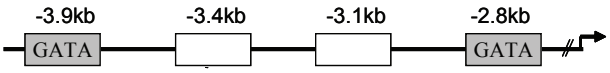
C.



D.

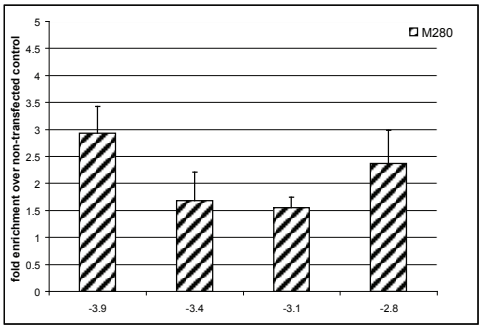


A.

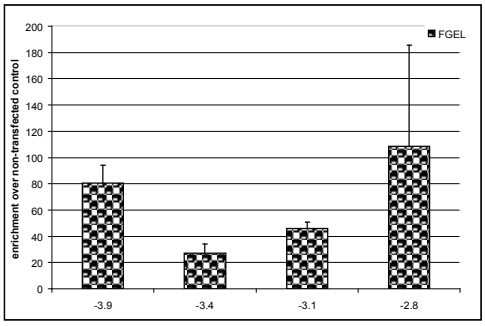

-3.9kb
-3.4kb
-3.1kb
-2.8kb
GATA
GATA

GATA-2

B.


M280
fold enrichment over non-transfected control
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
-3.9
-3.4
-3.1
-2.8

C.


FGEL
enrichment over non-transfected control
200
180
160
140
120
100
80
60
40
20
0
-3.9
-3.4
-3.1
-2.8

Chapter3

Isolation of a gene promoter in vivo: Identification of human γ-globin suppressor proteins

Farzin Pourfarzad[1], Ali Aghajanirefah[1], Thamar Bryn van Dijk[1], Nynke Gillemans[1], Ernie de Boer[1], Marieke von Lindern[2], Jeroen Demmers[3,4], Sjaak Philipsen[1,3] and Frank Grosveld[1,3,5].

[1]Department of Cell Biology, [2]Department of Hematology, [3]Netherlands Proteomics Centre, [4]Biomics Department and [5]Center for Biomedical Genetics, Erasmus MC, Dr. Molewaterplein 50, 3015GE, Rotterdam, The Netherlands.

Manuscript in preparation

Abstract

Reactivation of the human γ-globin ameliorates the effects of β thalassemia and sickle cell disease, however it is not known which (transcription) factors are bound to the γ-globin promoter to keep it suppressed in the adult stage. Here we show for the first time that promoter-bound transcription factors can be identified by purifying a gene promoter directly from mammalian cells *in vivo*. Using this method we identify transcription factors bound to the suppressed γ-globin gene promoter in the adult stage. We show that knockdown of these factors in human primary erythroid cells induces expression of the γ-globin genes. Since reactivation of the human γ-globin genes ameliorates the symptoms of β thalassemia and sickle cell disease, these factors provide novel targets for the development of drugs to treat these patients.

Introduction

The human α and β-globin gene loci produce hemoglobin tetramers which together with haem are responsible for the gas transport in the blood. The genes are developmentally regulated with one switch in the embryonic period for the α locus (from ζ to α) and β locus (from ε to γ) and a second switch in the β locus (from γ to β) around the time of birth. Conditions affecting the function of β-globin, known as β-thalassemias and sickle cell disease, are among the most frequent inherited single gene disorders in the human population. The most common treatment for these diseases is blood transfusion and iron chelation therapy or treatment with small molecules such as hydroxyurea or short chain fatty acids, which affect either red cell volume and/or lead to an increase of fetal γ globin gene expression (Bank, 2006a; Inusa, 2007; Mankidy et al., 2006; Stamatoyannopoulos, 2005) . However, these treatments are not very satisfactory since they do not lead to a normal quality of life nor prevent a relatively early death. At present, bone marrow transplantation is the only effective cure, a risky procedure that is not available to the large majority of patients.

The severity of β-thalassemia and sickle cell disease is greatly ameliorated by expression of γ-globin, which is normally expressed only during the fetal stage of development. Since almost all patients have normal γ-globin genes, reactivation of γ-globin expression in adults would provide a very elegant and attractive treatment of disease. In particular the treatment with short chain fatty acids aims to achieve such γ-globin gene activation, but is only partially successful (Perrine, 2008).

Despite intense research efforts by many laboratories, it is not known how the human γ-globin genes are normally suppressed around the time of birth when expression switches to the adult β-globin gene. Association studies on persons with relatively high γ-globin gene expression during adult life have indicated that a number of non-globin loci are involved in the suppression of the γ-globin genes (Jiang et al., 2006a; Lettre et al., 2008; Menzel et al., 2007; Thein and Menzel, 2009; Uda et al., 2008). The most promising of these, BCL11A, was recently shown to lead to elevated γ-globin gene expression when its activity is suppressed

(Sankaran et al., 2008). Interestingly the promoters of the γ-globin genes were previously identified as the regions responsible for their suppression (Yu M, 2006), whereas the BCL11A protein binds to a region downstream of the γ-globin genes. BCL11A interacts with the Gata1/Fog1/NuRD repressor complex, but whether this interaction is involved in γ-globin gene suppression is as yet not clear (Sankaran et al., 2008).

The importance of the promoter sequences has been well documented by genetic studies in mice (Yu M, 2006) and in a number of persons with point mutations in the γ-globin promoters leading to a condition known as hereditary persistence of fetal hemoglobin (HPFH). HPFH individuals sustain elevated γ-globin levels throughout adult life. The strongest known HPFH phenotype is caused by a G to A mutation at position –117 in the promoter, resulting in γ-globin expression at therapeutic levels (G Stamatoyannopoulos, 2001; Hardison et al., 2002). In order to gain more insight into the mechanism of suppression we developed a novel approach to directly identify the proteins bound to the γ-globin promoter. We purified the suppressed γ-globin promoter from erythroid cells using a protein tag that binds to the upstream promoter. This approach requires the use of a substantial amount of genetically modified cells that still regulate the γ-globin genes properly. Since the γ- to β-globin switch of the human β-globin locus and the effects of HPFH mutations can be phenotypically mimicked in the mouse (Berry et al., 1992; Li et al., 2001; Starck et al., 1994), we used erythroid cells derived from transgenic mice for these experiments. Here we show that this approach successfully identified a number of transcription factors whose suppression leads to the activation of the γ-globin genes in human erythroid cells. These factors provide novel targets for the development of drugs that release the suppression of γ globin genes in patients.

Results

Experimental design

We opted for an unbiased approach by introducing Tet operator (TetO) sequences in the upstream γ-globin promoter. Upon addition of doxycyclin, these sequences are bound by a tagged TetR protein that is used as a hook to fish out the γ-globin promoter region by affinity purification. Three basic parameters are important in such an approach. Firstly, the hook and its binding sites in the promoter should not disturb the normal expression pattern of the γ-globin gene. Secondly, since the γ-globin promoter fragment of interest constitutes less than 1-millionth of the genome, a large number of cells is needed to enable the purification of sufficient amounts of material required for further analysis. Thirdly, the control sample should be comparable to the experimental sample to allow the elimination of contaminants.

We used mouse transgenesis to generate a model system meeting these requirements. The basic design of the approach is schematically shown in Fig. 1. Binding sites for the bacterial tetracycline repressor protein (TetR) were inserted into a human β-globin minilocus (Grosveld et al., 1987) that contained the locus control region (LCR), the Aγ gene and the 3' hypersensitive site 1 (Fig. 1A). Previous experiments have shown that such a γ-globin

minilocus is regulated properly in the mouse (Dillon and Grosveld, 1991). The γ-globin gene is expressed in the embryo and early mouse fetal liver, and is silenced around E14-E16 in the fetal liver and remains silenced in the adult (Peterson et al., 1998; Strouboulis et al., 1992). Furthermore, the γ-globin gene remains active at the adult stage when the –117 G to A HPFH mutation is introduced in the promoter (Berry et al., 1992). The modified globin minilocus (*TetO-γ*) was introduced into transgenic mice that were bred in a p53 null background (*TetO-γ::p53null* mice) to facilitate the derivation of immortalized cell lines (Dolznig et al., 2001; Donehower et al., 1992; von Lindern, 2001). Next, we constructed a triple tag in the TetR protein (TetR3T, Fig. 1A) that enables sequential purification on HA and streptavidin affinity beads. The TetR3T cDNA was cloned in the hematopoietic expression vector pIE3.9IntpolyAA (Ohneda et al., 2002) and this construct (TetR3T) was introduced in transgenic mice. The *TetR3T* transgene was bred into mice expressing the *E. coli* BirA biotin ligase from the pEV locus(de Boer et al., 2003) . BirA will efficiently biotinylate the Bio tag of the TetR3T protein. Crossing of the *TetR3T::BirA* mice with the *TetO-γ::p53null* mice resulted in quadruple *TetR3T::BirA*::*TetO-γ::p53null* mice after several breeding and selection steps. The quadruple *TetR3T::BirA*::*TetO-γ::p53null* mice were used for the generation of fetal-liver derived erythroid progenitor cell lines (Dolznig et al., 2001; von Lindern, 2001) (Fig. 1B). One of these cell lines was used for the Aγ globin gene promoter purification after which the proteins bound to the promoter were identified by mass spectrometry (Fig 1C, D).

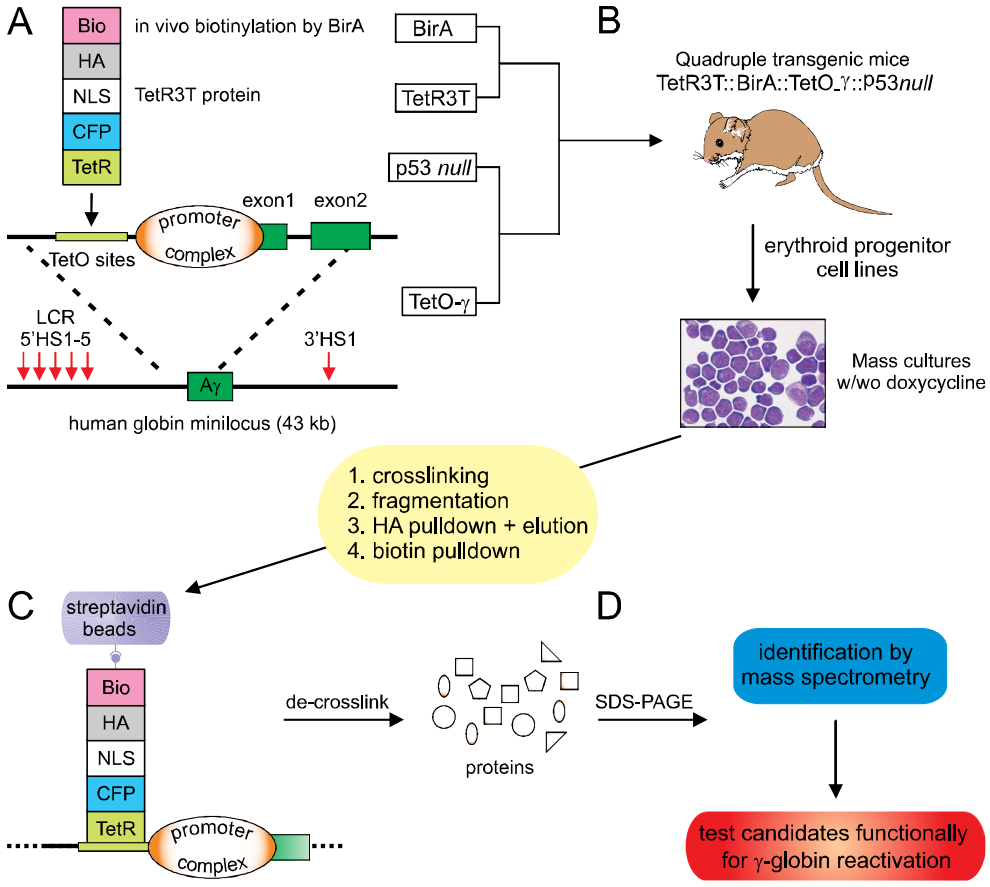


Figure 1. Scheme of the identification proteins bound to the suppressed γ globin promoter.
A) Depicts the generation of the TetO-γ and TetR3T-G1HRD constructs resulting in the two transgenic mice TetR3T and TetO-γ,that were crossed with BirA mice containing the BirA biotin ligase and $p53^{-/-}$ mice. **B)** Depicts the resulting *TetR3T::BirA::TetO-γ::p53*null mice that were used to derive the *TetR3T::BirA::TetO-γ::p53* cell line. **C)** depicts the purification of the γ globin promoter after crosslinking the *TetR3T::BirA::TetO-γ::p53* cell line chromatin, its fragmentation and HA/biotin pulldown. **D)** Depicts decrosslinking of the purified promoter proteins, their identification by mass spectrometry and functional testing for γ globin gene activation.

Development of the experimental system

First, the TetO site was tested in electrophoretic mobility shift assays with *TetR3T::BirA* fetal liver cell extracts. This showed strong binding of TetR3T to the TetO sequence, and some weak but specific Gata1 binding (not shown). We found no evidence for binding of other proteins. Next, seven copies of the TetO sequence were introduced into the StuI site 380 bp upstream of the transcription initiation site of the γ-globin gene in the γ-globin minilocus

(Fig. 1A). This *TetO-γ* locus was used for transgenesis, and several *TetO-γ* transgenic lines were selected for further analysis (Suppl. Fig. 1). The developmental expression pattern of γ-globin from the *TetO-γ* locus was analysed by quantitative S1 nuclease protection assays. The results obtained with line 05-23736-05, which contains two copies of the transgene, are shown in Figure 2.

From these data we conclude that the *TetO-γ* locus is expressed properly. It is active during the embryonic period, suppressed during the late fetal liver stage and it remains inactive in the adult.

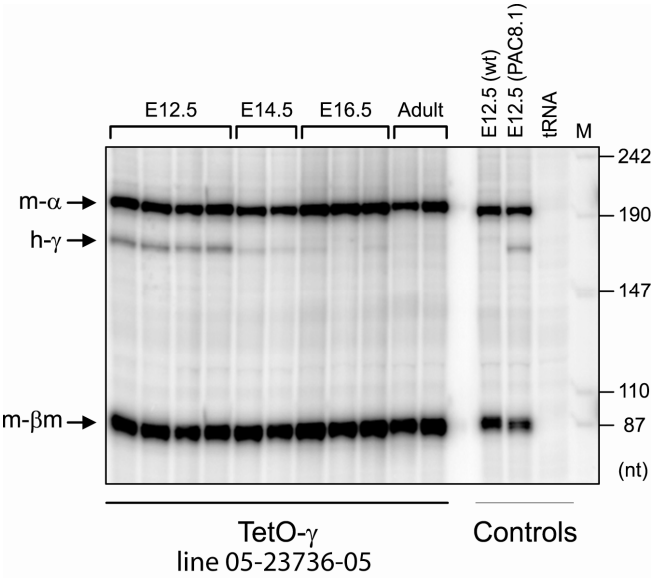


Figure 2. S1 nuclease protection analysis of globin expression in TetO-γ transgenic mice. RNA was prepared from fetal livers and adult blood of transgenic mice carrying two copies of the TetO-γ transgene (line 05-23736-05). Quantitative S1 nuclease protection assays were performed as described (Dillon and Grosveld, 1991). Controls are shown on the right. E12.5 wt: RNA from liver of wild type mouse E12.5 mouse embryo; E12.5 PAC8.1: RNA from liver of E12.5 mouse embryo carrying a complete human β-globin locus (Wai et al., 2003). The samples were separated on a 6% polyacrylamide urea gel; the positions of the protected fragments for mouse α-globin (m-α), βmajor globin (m-βm) and human γ-globin (h-γ) are indicated. Size marker: pUC18 cut with MspI.

In parallel, the TetR3T cDNA was cloned in the pIE3.9IntpolyAA vector (Ohneda et al., 2002) resulting in the TetR3T construct. This vector recapitulates the hematopoietic expression pattern of the mouse *Gata1* gene, driving expression at all developmental stages of erythroid cells, i.e also when the γ-globin gene is suppressed. The TetR3T protein is composed of the TetR DNA binding domain coupled to CFP, three copies of a nuclear localization signal, an HA tag and a biotinylation tag (Fig. 1A and Fig. 3A). The TetR3T-G1HRD construct was introduced into transgenic mice and these were bred to mice expressing the BirA biotin ligase from the *pEV* locus (*BirA*; Fig. 1A)(de Boer et al., 2003). Several mice were obtained expressing TetR3T protein in the nucleus of erythroid cells (Fig. 3B). The protein was specifically biotinylated by BirA, as shown by probing Western blots of crude nuclear extracts with HRP-streptavidin (Fig. 3C).

Interbreeding of the mouse lines resulted in quadruple *TetR3T::BirA*::*TetO-γ::p53null* mice that were used for the generation of fetal-liver derived erythroid progenitor cell lines (Fig. 1B).

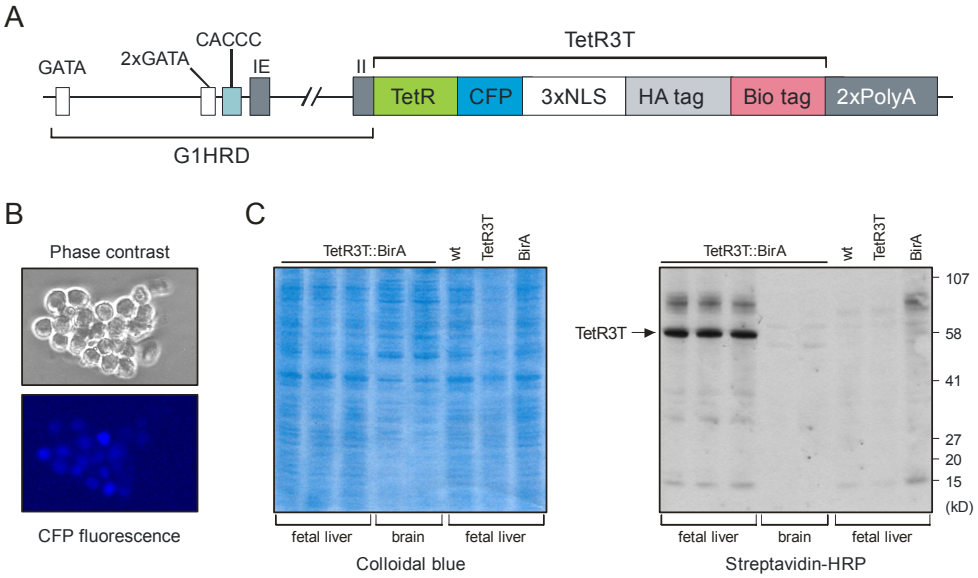


Figure 3. Streptavidin-HRP stained blot in fetal liver extracts of *TetR3T::BirA* compound transgenic embryos only.

Purification of the γ-globin promoter

Erythroid progenitor cell lines (von Lindern, 2001) were derived from *TetR3T::BirA*::*TetO-γ::p53null* fetal livers. The cell lines were cloned and further propagated in vitro. Out of these, a cell line was randomly selected that had repressed the human γ-globin gene, yet still expressed the endogenous α and β-globin genes at high levels (Suppl. Fig. 2). This cell line was expanded for the subsequent optimization of each step of the biochemical purification of the γ-globin gene promoter. First, we used chromatin immunoprecipitation (ChIP) to confirm that the γ-globin gene promoter was bound by the TetR3T protein, and that it was released upon the addition of doxycycline to the medium (Fig. 4A, B). We found that the γ-globin promoter could be efficiently pulled down by streptavidin beads (Fig. 4). Thus, the approach appears to work, however a single pulldown before mass spectrometry analysis would be insufficient to identify proteins bound to the γ-globin promoter. We therefore tested whether the chromatin captured with the HA beads could be eluted efficiently by the addition of HA peptide, allowing a subsequent pulldown step on the eluted chromatin. The result (Fig. 4B) shows that all of the bound material could be eluted efficiently by HA peptide.

Both pulldowns show an enrichment of the sequences immediately upstream and downstream of the TetO sites including all of the promoter sequences. Sequences further upstream (5'HS5) and downstream (exon 3) show no enrichment when compared to the controls (Fig. 4B).

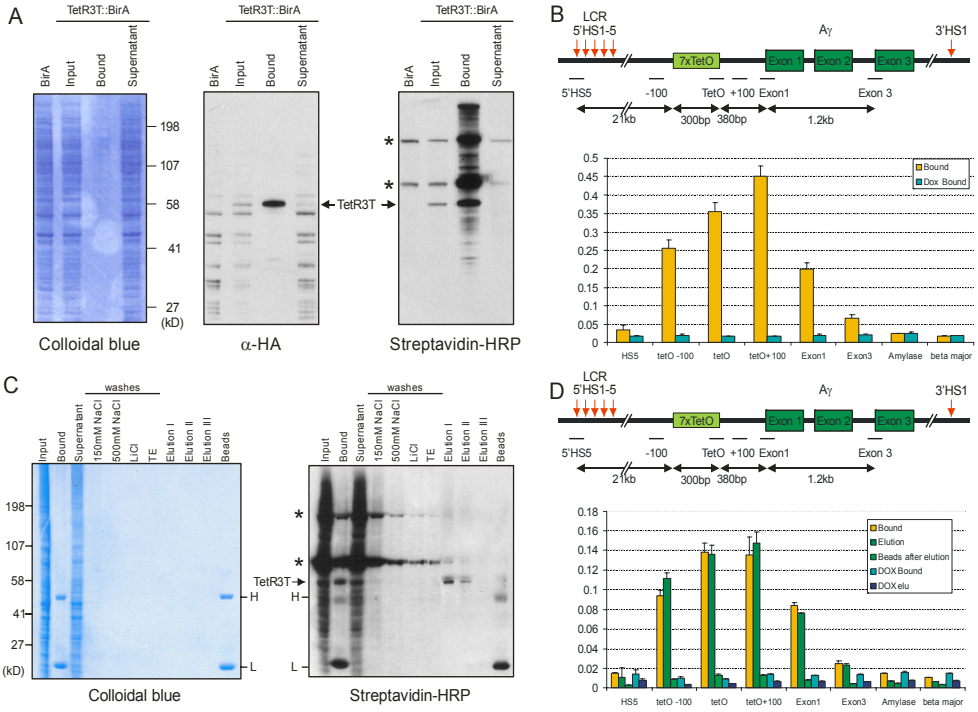


Figure 4. Affinity purification of the γ globin promoter from *TetR3T::BirA::TetO-γ::p53null* cells.
A) Left panel; gel electrophoresis of the nuclear extracts after colloidal blue staining. The first lane contains a control extract from an identical cell line not containing the TetR γ-promoter construct. Middle panel; detection of the TetR3T protein with an anti-HA antibody. Right panel; detection of the TetR3T protein with Streptavidin HRP. Bands indicated by a * are naturally biotinylated proteins (de Boer et al., 2003). **B)** Top shows a map of the γ globin locus and the distances of the various fragments from the TetO sites in the promoter. Bottom shows the chromatin precipitation on Streptavidin beads or in the presence or absence of doxycyclin. The mouse β-major and Amylase genes were used as negative controls. **C)** Left panel; gel electrophoresis of the nuclear extracts after colloidal blue staining after ChIP with an α-HA antibody. The lanes show the fractions bound, washed and eluted of the beads. H and L are the heavy and light chains of the α HA antibody on the beads. Right panel; detection of the TetR3T protein with Streptavidin HRP of the lanes shown in the left panel. Bands indicated by a * are naturally biotinylated proteins (de Boer et al., 2003). **D)** Top shows a map of the γ globin locus and the distances of the various PCR fragments from the TetO sites in the promoter. Bottom shows the PCR detection of the various fragments of the γ globin locus or the mouse β-major and Amylase control genes after binding on α-HA beads in the presence or absence of doxycyclin.

Subsequently all parameters for the purification of the promoter were optimized. This included fixation conditions, crosslinking, type of pull down beads, buffers, blocking agents to prevent non-specific binding, large scale culture and importantly the elution conditions (Suppl. Table 1). The optimization resulted in conditions that gave over 50-fold purification for both the HA- and the biotin pulldown (Fig. 4). In the final optimized protocol, the yield of the

purification steps measured as γ-globin promoter DNA recovery after the HA and biotin pulldowns was estimated to be ~2% of the starting material.

Proteomics analysis of the pulled down chromatin

Currently, the sensitivity of mass spectrometry for protein identification is in the order of 100 attomolar. Therefore, a minimum number of $12x10^9$ cells is required to recover sufficient amounts of γ-globin promoter complexes with the optimized procedure. For each experiment, we grew two cultures in parallel, one with and one without doxycycline. The cells were harvested and subjected to the promoter isolation procedure. After the final step, the chromatin was de-crosslinked and the proteins were separated by SDS-PAGE. The lanes were sliced and each slice was used for mass spectrometry analysis. We performed a total of three independent experiments in proliferative culture medium and one in differentiation medium condition. In the first experiment (575), we identified 534 proteins with a Mascot score >30. Of these proteins, 380 were also found in the samples of the doxycyclin-treated cells. Thus collectively, we found 154 unique candidate proteins bound to the γ-globin promoter (Suppl. Table 2 and 3). 89 of these were very unlikely candidates to be involved in the suppression of the γ-globin genes, such as DNA-directed RNA polymerase, glutathione S transferase and ribosomal RNA Proteins. These were relegated to the bottom of the list, leaving approximately 65 candidates. Amongst these were Gata1, Dnmt1, Cdc5l, Actl6a (Baf53A) and Chd4 (Mi-2β) that have been reported previously to bind to the γ-globin promoter or Locus Control Region in vitro (Harju-Baker et al., 2008; Mahajan et al., 2005; Olave et al., 2007). It also included a number of interesting novel proteins Apex1, ZBP-89 (Zfp148), Ctnnbl11, Fanci, Supt5, Wdr12 and others (suppl table 2 and 3).

Functional identification of suppressor proteins

We next checked for a number of the proteins whether they are functionally indeed suppressing the γ glpobin gene promoter and tested this immediately with the human homologues of these genes in human erythroid progenitor cells (HEP),(Leberbauer et al., 2005). Gata1 was not tested, because it is well known that its absence leads to erythroid cell death. However a number of the remaining proteins that could potentially function as repressor of γ-globin gene were functionally tested through shRNA- or siRNA-mediated knockdown assays. Several shRNA constructs for ZBP-89, CDC5L, APEX1, SUPT5h and CTNNBL1 and a pool of 4 siRNAs for FANCI were tested for their ability to lower the level of the respective proteins and whether such a knockdown would result in the upregulation of γ globin gene expression by S1 nuclease protection at messenger RNA level, and protein level by HPLC profiles of human hemoglobins.

ZBP-89

Five shRNAs were used to knock down ZBP-89 in HEP cells, of which only sh3 and sh5 resulted in a knockdown of ZBP-89 protein level (Fig. 5A). S1 nuclease protection analysis for detection of globin gene expression showed that ZBP-89 sh3 and sh5 result in γ-globin gene expression in adult HEP cells at mRNA level. At the protein level, γ-globin induction was measured by relative hemoglobin subtypes using high performance liquid chromatography (HPLC). In the same knock down experiment, HbF level was increased between 2 to 3 folds. Immunohistochemistry (IHC) showed the majority of the cells express HbF to varying level upon ZBP-89 depletion (Figure 5D).The knock down experiment reproduced the same results in HEP cells from 6 different healthy individuals (Fig. 5B and C).

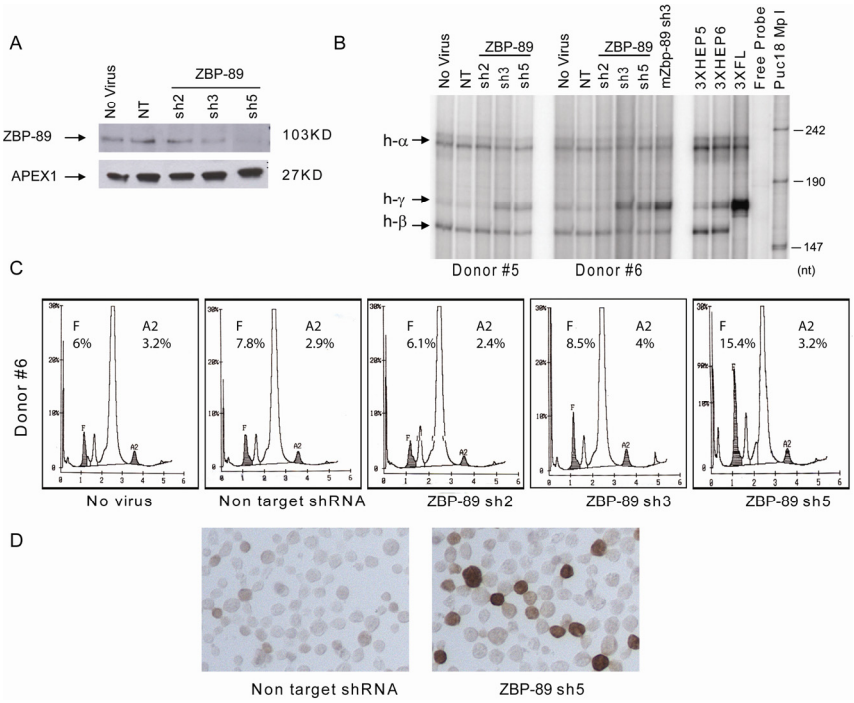


Figure 5. Lentiviral RNA interference-mediated gene silencing of ZBP-89.
A) Top panel immunoblot for ZBP-89 shRNA knock down in HEP cells. First lane no virus control; NT, Non-target shRNA lentiviral transduced HEP cells. Bottom panel, APEX1 staining of the same blot as loading control. **B)** S1 nuclease protection analysis of globin expression in ZBP-89 knock down HEP cells of two different donors 5 and 6 (donor#5 and 6). First lane no virus control, NT; Non-target shRNA, mZBP-89 sh3; Mouse ZBP-89 short hairpin RNA 3 also depletes human ZBP-89 and induces γ-globin gene expression. Extra controls in the right, 3XHEP 5 and 6; 3xRNA from no virus treated HEP 5 and HEP6 cells, 3xFL; Human fetal liver RNA, PUC18 MspI size marker. The position of human α, γ and β are indicated at the left of the panel. **C)** HPLC hemoglobin chain analysis of the ZBP-89 shRNA knock down in HEP6 (donor#6) cells. **D)** HbF Immunohistochemistry of the ZBP-89 shRNA knock down HEP6 cells.

CDC5L

Four shRNAs for CDC5L were transduced in HEP cells and all of them resulted in a knockdown of related protein (Fig. 6A). S1 nuclease protection shows that CDC5L sh2 and 4 induce γ-globin gene expression in adult HEP cells at mRNA level. The RNA qualities from other shRNAs were not good enough for an S1 nuclease protection assay; this could be due to the lethal effect of the knock down of CDC5L (Fig. 7B). γ-globin induction was also seen at the protein and cellular levels by HPLC and IHC, respectively (Fig. 7C and D). Fetal hemoglobin was increased from 4% to 16.1% with sh2 and to 9.6% with sh4. The knock down experiment produced the same results in HEP cells from 2 different healthy individuals.

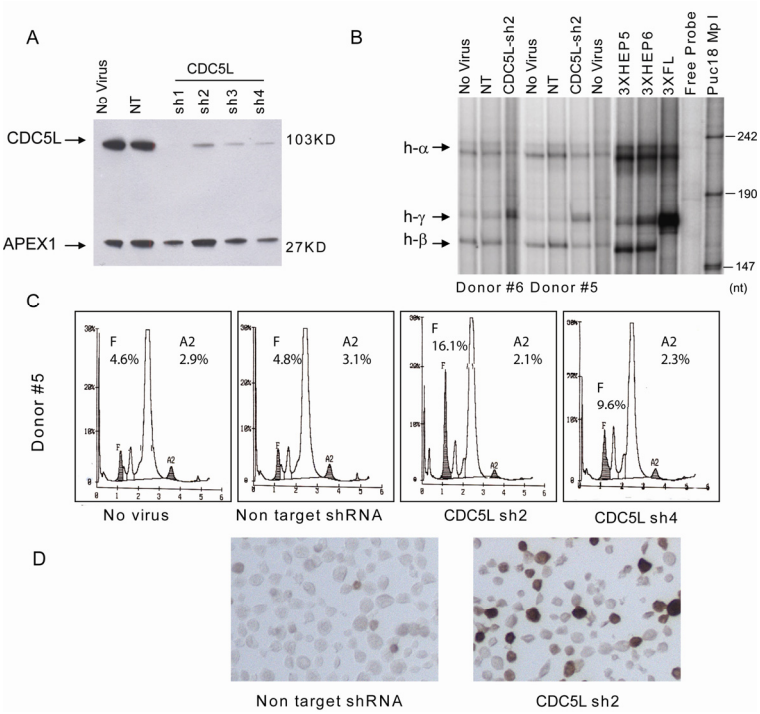


Figure 6. Lentiviral RNA interference-mediated gene silencing of CDC5L.
A) Top panel immunoblot for CDC5L short hairpin shRNA knock down in HEP cells. First lane no virus control; NT, Non-target shRNA lentiviral transduced cells. Bottom panel, APEX1 staining of the same blot as loading control. **B)** S1 nuclease protection analysis of globin expression in CDC5L sh2 knock down HEP cells of two different donors 5 and 6. First lane no virus control, NT; Non-target shRNA. Extra controls in the right, 3XHEP 5 and 6; 3xRNA from no virus treated HEP 5 and HEP6 cells, 3xFL; Human fetal liver RNA, PUC18 MspI size marker. The position of human α, γ and β are indicated at the left of the panel. **C)** HPLC hemoglobin chain analysis of the CDC5L shRNA knock down in HEP5(donor#5) cells. **D)** HbF Immunohistochemistry of the CDC5l shRNA knock down HEP5 cells.

APEX1

APEX1 was downregulated by four shRNAs in HEP cells and all of them resulted in a clear effect at protein level (Fig. 7A). The shRNA knock down of APEX1 was very lethal to the cells and only sh1 and sh3 transduced cells survived. S1 nuclease protection assay showed increased γ-globin levels after APEX1 knock down. HbF protein was also increase from 7% to 14 and 24% with sh1 and sh3, respectively (Fig. 7B and C). The increased level of HbF was also observed at the cellular level by HbF Immunohistochemistry after APEX1 knock down (Fig. 7D). The experiment was reproducible in 2 different donors.

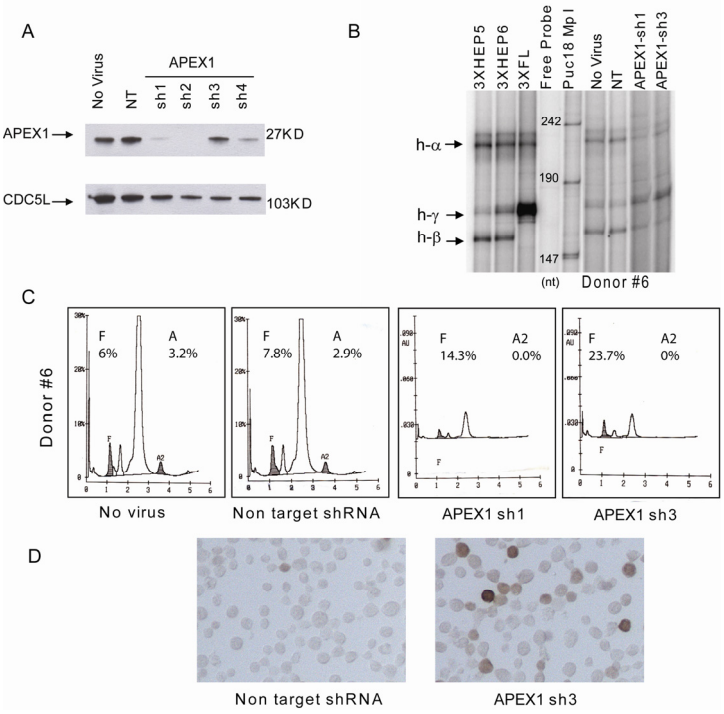


Figure 7. Lentiviral RNA interference-mediated gene silencing of APEX1.
A) Top panel immunoblot for APEX1 shRNA knock down in HEP cells. First lane no virus control; NT, Non-target shRNA lentiviral transduced cells. Bottom panel, CDC5L staining of the same blot as loading control. **B)** S1 nuclease protection analysis of globin expression in APEX1 knock down HEP cells. NT; Non-target shRNA. Extra controls in the left, 3XHEP 5 and 6; 3xRNA from no virus treated HEP 5 and HEP6 cells, 3xFL; Human fetal liver RNA, PUC18 MspI size marker. The position of human α, γ and β are indicated at the left of the panel. **C)** HPLC hemoglobin chain analysis of the APEX1 shRNA knock down in HEP6 (donor#6) cells. **D)** HbF Immunohistochemistry of the APEX1 shRNA knock down HEP6 cells.

FANCI was efficiently knocked down using a pool of 4 siRNAs and 4 fold γ-globin expression at RNA level was observed (data not shown). Supt5h and CTNNBL1 were also knocked down by different shRNA's but the decreased level of these proteins did not significantly change the level of γ-globin at mRNA or protein level (Suppl. Fig. 3 and 4).

Discussion

Target chromatin purification (TCP)

Here we show for the first time that transcription factors can be identified directly through the purification of the target promoter sequences (Targeted Chromatin Purification, TCP) *in vivo*. A direct approach has only been successfully carried out once in mammalian cells, using a hybridization technique (Proteomics of Isolated Chromatin, PICh) to identify the factors bound to telomeres (Déjardin and Kingston, 2009). Since each chromosome contains two telomeres the latter approach requires less material to be able to identify the proteins by mass spectrometry. Telomeres constitute approximately 0.01-0.07% of the genome (Déjardin and Kingston, 2009), whereas a unique promoter sequence of approximately 1kb constitutes a much smaller percentage of the genome. Although the results obtained with the two methods are not directly comparable they suggest that the TCP approach presented here may be more suitable to the isolation of unique sequences than PICh. The two approaches are fundamentally different and each has their advantages and disadvantages. Each method used formaldehyde crosslinking and a similar preparation of the chromatin. Each used sonication to fragment the chromatin, although restriction enzymes that still cleave formaldehyde fixed chromatin could be used to isolate a specific fragment of DNA from the chromatin. The PICh method involves a formaldehyde fixation and hybridization step to capture the target chromatin while TCP relies on the binding of the TetR protein to the TetO sequences, which is likely to be more efficient than hybridisation. TCP has the disadvantage that it requires the introduction of sequences that serve as a "hook" into the naturally occurring sequences and hence that controls should be carried out to ensure that the introduction of the hook has not changed the behavior of the promoter (or other target) sequence. The advantage is that once the sequences have been introduced multiple purification steps can be carried out with the same hook, using the appropriate tags and elution conditions. Even the biotinylation step could be provided with an "elution" step as the final purification either by using biotin analogues (Hirsch et al., 2002) or by adding a restriction site to the hook that would cleave the target sequence from the TetO sequences. In each approach it is essential to run a comparable negative control to allow the identification of candidate binding proteins. Dejardin and Kingston list approximately 200 proteins in the control PICh pulldown, probably due to mismatched hybridization and non specific sticking of irrelevant chromatin, while the TCP control showed more than 400 background proteins, probably due to non specific binding of TetR and also non specific sticking of irrelevant chromatin.

Here we have used a transgenic mouse approach and derived a cell line from it, because the human γ globin promoter presents a specific set of difficult problems. In many cases it should be possible to directly transfect the promoter or other sequence of interest directly into an appropriate cell line, however in the case of the human γ-globin genes this is extremely difficult, erythroid cells are very difficult to transfect (or homologously recombine) and γ-globin gene silencing is prone to position effects. The advantage of our method is that

one can easily generate mice that are p53 negative expressing both the TetR protein and the BirA ligase ubiquitously (Driegen et al., 2005; Katsantoni et al., 2003), which would allow the introduction of other genes expressed in erythroid cells.

γ globin gene suppression

TCP identified 154 proteins as possible candidates to bind to the suppressed γ globin gene promoter of which we classified 89 proteins as unlikely candidates to be involved in suppression. These are most likely abundant non-specific chromatin interacting proteins that are not seen in the control for as yet unknown reasons. We considered about 20 of the 65 candidates for their possible role in γ globin gene suppression by applying the following criteria; are they expressed in erythroid cells; are they known to be proteins bound to chromatin; are there good antibodies available? That left approximately 20 possible candidates for testing, including ZBP-89, Cdc5l, Apex1, Yb1, Ctnnbl1, Gata1, Hdac2, Prmt1, Supt5, Chd4 (mi2β) , Actl6a (Baf53), Nap1l1, Ehmt1, Son, Snw1 and FancI. We next asked whether the down regulation of these factors affect γ-globin expression in adult HEP cells. We used lentiviral transduction of multiple short hairpin RNA to reduce protein levels of ZBP-89, Cdc5l, Apex1, Yb1, Ctnnbl1, FancI, Supt5.

One of the most interesting factors was ZBP-89 which was only found in the promoter pull down experiment with the cells induced to terminal erythroid differentiation. ZBP-89 (BFCOL1, BERF-1, ZNF148 in human and Zfp148 in mouse) is a Krüppel-type, zinc-finger transcription factor that binds to a GC-rich region, and subsequently represses or activates known target genes. In the case of ZBP-89 acting as a repressor, it has been proposed that ZBP-89 and Sp1 compete for binding to the same or overlapping GC-rich sequences (Merchant et al., 1996). It is proposed that ZBP-89 represses vimentin expression via the specific recruitment of HDAC1 to the vimentin promoter (Wu et al., 2007). ZBP-89 interacts with GATA1, Mafk and it is involved in erythroid development and differentiation (Brand et al., 2004; Woo et al., 2008). Knock down of ZBP-89 in zebra fish results in bloodless phenotype by disrupting primitive and definitive erythropoiesis (Li et al., 2006). Binding of this factor is reported on GATA1 enhancer and also mouse β-globin LCR HS8 as well as other genes (Ohneda et al., 2009; Vernimmen et al., 2007). Mutations in -175 and -566 in GATA binding motifs results in HPFH condition (Chen et al., 2008; Liu et al., 2005). Moreover, it is well known that HDACs are involved in γ-globin silencing and use of HDAC inhibitors can partially activate γ-globin expression (Constantoulakis et al., 1989; Perrine et al., 1989; Perrine et al., 1987). Based on these observations together with our specific detection of ZBP-89 on the γ globin promoter in our pull-down assay we decided to study the effect of this factor in more detail. Knock down of ZBP-89 in human adult HEP cells increased γ-globin expression, only with specific shRNA's that resulted in depletion of ZBP-89, at mRNA, protein and cellular level in majority of the cells shown by γ-globin immunostaining (Fig. 5). The level of ZBP-89 knock

down seems to inversely correlate with γ-globin expression. We have also observed the same effect in mouse α and β globin locus embryonic genes (data not shown).

Unfortunately we have not yet been able to obtain reproducible data to show the binding of ZBP-89 to the γ-globin gene promoter in HEP cells to confirm this interaction. Nevertheless we are in the process of tagging this protein to study its DNA binding site with ChIP sequencing and also identify other binding protein partners (preliminary data presented in chapter 6).

Two of the other protein factors identified were Apex1 and Yb-1. Apurinic/apyrimidinic endonuclease 1 (Apex1)/Redox factor 1 (Ref-1) is a multifunctional protein involved in DNA repair activity, proofreading exonuclease activity and in modulating DNA binding activity of several transcription factors including AP-1, Creb, Nrf2 via reduction of a conserved cysteine in their DNA binding domains. Apex1/Ref-1 is involved during hemin (Iwasaki et al., 2006) mediated differentiation of K562 cells, regulates hematopoietic differentiation of embryonic stem cells, and knock down Apex1 gene expression in embryonic body (EB) cells results in a significant decrease in the frequency of hemangioblast formation and diminished formation of primitive and definitive hematopoietic colonies (Zou et al., 2007). Apex1, stably interacts with Y-box-binding protein 1 (Yb-1, Ybx1) and enhances its binding to the Y-box element (Chattopadhyay et al., 2008). Acetylation of Apex1 enhances its binding to distinct trans-acting complexes involved in activation or repression (Zou et al., 2007).

Y box-binbing protein (Yb-1) is important for late-stage embryonic development (Lu et al., 2005). It has been reported that Yb-1 is highly expressed in the spleen of GATA1 knockdown mice. GATA1 binds to the proximal GATA element in 5'-UTR region of Yb-1 gene and has a positive effect on its expression in K562 cells (Yokoyama et al., 2003b). Yb-1 binds to inverted CCAAT boxes (Spitkovsky et al., 1992), and it is highly expressed in erythroid cells of patients with refractory anemia (Yokoyama et al., 2003a). Yb-1 appears to function in the early stage of erythropoiesis, and aberrant expression of this protein may induce hematological diseases.

Since both of these proteins Apex1 and Yb-1 were present in our mass spectrometry results, and it is known that AP-1 and CREB also bind to human β-globin locus LCR HS2 (Johnson et al., 2002; Talbort and Grosveld, 1991), we decided to test whether these two factors are involved in γ-globin suppression. shRNA knock down of APEX1 with sh2 and 4 almost completely depleted APEX1 protein and was very lethal to the HEP cells. Short hairpin RNA number 1 and 3 efficiently decreased the APEX1 protein level and had a clear effect on induction of γ-globin expression from 2 to more than 3 fold (from 7% in the cells transduced with non targeting short hairpin RNA to 14% in sh1 and 23% in sh3) (Fig. 7).

Yb-1 shRNA knock down was also very efficient and also unfortunately lethal. However, the shRNA with a milder knock down effect showed a two fold increase of γ-globin expression (data not shown) consistent with the role of Yb-1 as a γ globin repressor.

CDC5L (cell division cycle 5-like) is a cell cycle regulator important for G2/M transition and has sequence specific DNA binding activity (CTCAGCG). Its DNA binding domain has similarities to Myb (Takashi and Kazuo, 1996), which is known to be involved in γ-globin repression (Jiang et al., 2006b; Kuroyanagi et al., 2006). In a sequence specific protein purification Cdc5l and Dnmt1 were identified to be able to bind the γ-globin promoter in a protein purification experiment using the human -198 Aγ HPFH mutation of γ-globin promoter (Olave et al., 2007). CDC5L also interacts with beta catenin like 1 protein (Ctnnbl1) also called nucleosome assembly protein (NAP) (Conticello et al., 2008).

These two proteins were also depleted using specific shRNAs and γ-globin expression levels were analyzed in HEP cells. Complete knock down of CDC5L was again lethal, but a 3-4 fold decreased level of CDC5L expression increased fetal hemoglobin between 2-4 fold (Fig. 6). A reduced level of CTNNBL1 however had no effect on γ-globin expression (Suppl. Fig. 3)..

FancI is another factor found in γ-globin promoter pull down mass spectrometry. FancI belongs to the Fanconi anemia complementation group (FANC). Its absence induces cytogenetic instability, hypersensitivity to DNA crosslinking agents, increased chromosomal breakage, and defective DNA repair. Fanconi anemia is associated with high HbF levels (Gumruk et al., 2008; Miniero, 1981). Expression profiling of erythroid progenitor cells grown from human fetal liver and adult blood shows that FANCI expression is 2 fold higher in adult erythroid progenitors (unpublished data). siRNA mediated knock down of FANCI induced γ-globin expression up to 4 fold in adult HEP cells (data not shown). However, the effect of FANCI depletion on DNA damage and genome instability may induce cellular stress and thereby indirectly contribute to the induction of γ-globin expression. Hence, FANCI presence in the γ-globin suppressor complex needs to be validated. A similar explanation may apply for the effect seen with other identified candidate protein factors with a role in cell cycle.

Supt5, also called spt5 is another factor detected in our experiment. P-TEFβ-mediated phosphorylation of hSpt5 is critical for transcription elongation (Yamada et al., 2006). DSIF (spt5/spt4) and NELF bind to RNA Pol II together and repress transcription elongation in the promoter-proximal region (Wada, 1998; Yamaguchi et al., 1999). However, knock down of SUPT5 in adult HEP cells did not have any effect on γ-globin expression (Suppl. Fig. 4).

Comparison between background proteins in our study and Déjardin *et al* (Déjardin and Kingston, 2009) reveals many shared proteins. In our knock down experiment we observed a γ-globin upregulating effect with 5 out of 7 tested factors. Moreover, identifying specific hits like Gata1, Dnmt1, Chd4 Cdc5l and few other that are already known to interact with γ-globin promoter, provide us with confidence regarding the validity of our approach in identifying factors that are specifically bound to the γ-globin promoter and regulate its activity.

It is possible that these protein factors are part of different complexes that interact with each other. For example ZBP-89 is known to interact with GATA1 (Woo et al., 2008) and GATA1 in turn interact with BCL11A (Sankaran et al., 2008) and both interact with the NuRD complex. Chd4, one of the NuRD complex subunits, is also present in our mass spectrometry hits. Our attempts to show reproducible binding of these factors by ChIP to the γ-globin

promoter have not yet been successful in convincingly demonstrating their *in vivo* binding to the suppressed γ-globin promoter in human HEP cells. This is likely due to technical issues. We are therefore actively pursuing tagging many of these factors followed by conventional ChIP experiments as well as genome-wide determination of their binding profiles using SOLEXA ChIP seq as well as identification of their binding partners using Imminoprecipitation coupled to mass spectrometry. We have already begun these studies with ZBP-89 and the preliminary results are presented in chapter 6.

Identifying γ-globin repressors is very important to understand the mechanism of developmental γ-globin gene switching and development of new therapeutic approaches. One such approach is the specific targeting of the assembly of γ-globin suppressor complex, e.g. designing short peptides preventing interaction of different subunits of the complex. Other therapeutic means such as blood transfusion, transplantation or pharmacological fetal hemoglobin inducers, each have their own complications and/or side effects. The screening of small molecule libraries to identify γ-globin inducers has not so far introduced any new compound which is more advantageous compare to the ones that already exist (Haley et al., 2003). The protein factors presented in this study, if they are confirmed to be part of the γ-globin repressor complex, could be potential targets for therapy.

Materials and methods

DNA constructs

A human γ globin gene ClaI/KpnI fragment from the γ globin minilocus (Dillon and Grosveld, 1991) was modified as a subclone by insertion of the tetO heptamer with the sequence TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAG (Gossen and Bujard, 1992) ligated to a short random sequence containing SwaI and PmeI restriction enzyme digestion sites (for the sequence see supplementary materials and methods) into the StuI site of the γ globin gene promoter 380 basepairs upstream of the transcription initiation site. A loxP sequence (CGATATAACTTCGTATAATGTATGCTATACGAAGTTATGGCGCGCCTT) was inserted in ClaI site (Fig. 1). The modified ClaI/KpnI fragment was cloned back into the minilocus by standard λ phage packaging (Stratagene) and transduction into E.Coli DH10B. DNA was isolated and the integrity of the modified minilocus established by cleavage with EcoRI and gel electrophoresis (Suppl. Fig. 1).

The 642 bp tetR binding domain (Gossen and Bujard, 1992) was cloned in frame onto CFP cDNA (pECFP-N1, Clontech). The insert was recloned into pEYFP-Nuc replacing eYFP (Clontech) thereby gaining the 3x NLS and SV40 polyA sequence. A biotinylation tag (de Boer et al., 2003) and HA tag (YPYDVPDYA) sepereated by few glycin residue as linker and restriction sites were PCR cloned in frame from legated overlapping oligonucleotides (see supplementary material and method for complete overlapping oligonucleotides sequence) in between the NLS and poly A resulting in TetR3T cDNA. The TetR3T fragment was excised by Eco RI/NotI and inserted into the NotI site of the G1HRD expression vector (Ohneda et al.,

2002). The resulting TetR3T-RG1HRD vector was tested for integrity (not shown) and standard transfection and expression in MEL cells which demonstated its localization to the nucleus by CFP fluorescence (not shown).

Transgenesis and cell line derivation

The minilocus DNA was subsequently cut with SalI, the insert fragment isolated by gel electrophoresis and injected into fertilized oocytes (Dillon and Grosveld, 1991) of p53-/- mice (Donehower et al., 1992).A number of γ minilocus transgenic mice was obtained, one of which contained only two copies of the minilocus (line 05-23736-05). Its integrity was was established by EcoRI digestion and Southern blots of genomic DNA (Supplementary Fig. 1). Similarly The Asp718/PvuI TetR3T restriction fragment was isolated from the vector by gel electrophoresis and injected into fertilized eggs obtained from BirA expressing mice (de Boer et al., 2003).The resulting mice, *TetR3T::BirA,* expressed the TetR3T protein in the erythroid lineage (Fig. 3). Finally the two mouse strains were bred together to yield the *TetR3T::BirA*::*TetO-*γ*::p53null* mice. These mice were bred together and 13.5d fetal liver cells were cultured (Dolznig et al., 2001) to obtain the *TetR3T::BirA*::*TetO-*γ*::p53null* cell line. The transgenic mouse breeding also yielded control mice and subsequently cell lines *TetR3T*::*TetO-*γ*::p53null*, *BirA*::*TetO-*γ*::p53null and TetR3T::BirA::p53null.*

Culturing of human proerythroblasts from blood (HEPs)

Culturing of undifferentiated cells

The culturing was essentially as published (Leberbauer et al., 2005). 40ml blood was sampled and buffy coat isolated by centrifugation. White cell were removed after Ficoll gradient purification from the inter phase and washed. For initial expansion, $5x10^6$ cells/ml were cultivated in serum-free medium (StemSpan; Stem Cell Technologies, Vancouver, BC, Canada) supplemented with Epo (2 U/mL EPREX Ortho-Biotech, Tilburg, The Netherlands), the synthetic glucocorticoid Dex (1 μM; Sigma, St Louis, MO) SCF (100 ng/mL; R&D Systems, Minneapolis, MN), and lipids (40 μg/mL cholesterol-rich lipid mix; Sigma) for about 5 days, the cells were fed every day by partial medium changes with the same medium. After 5 days the big proerythroblasts became visible and by day 7 they start to overgrow the culture. In order to purify proerythroblasts from lymphocytes and macrophages a percoll purification was carried out and the cells isolated from the interphase (followed by a wash). Homogeneous cultures of erythroid progenitors were kept in the same medium at $1.5\text{-}2x10^6$ cells/ml by daily partial medium changes. Proliferation kinetics and size distribution of the cells were monitored daily using an electronic cell counter (CASY-1; Schärfe-System, Reutlingen, Germany).

Induction of differentiation

To induce terminal differentiation, proliferating erythroblasts were washed in ice-cold PBS and reseeded at 1-1.5x10^6 cells/ml in StemSpan supplemented with Epo (10 U/ml), human transferring Holo (1 mg/mL; SCIPAC Ltd, UK) and lipids (40 μg/mL cholesterol-rich lipid mix; Sigma). Differentiating erythroblasts were maintained at 2–3x10^6 cells/ml by daily cell counts and partial medium changes.

γ globin promoter purification using chromatin pulldown

Part A. Cross linking and Chromatin preparation

The *TetR3T::BirA*::*TetO-γ::p53null cells* are grown in Stem-Pro 34 medium (Invitrogen)) supplemented with 2mM L-Glutamin (Invitrogen), Epo (1 U/mL EPREX Ortho-Biotech, Tilburg, The Netherlands), the synthetic glucocorticoid Dex (1 μM; Sigma, St Louis, MO) SCF (100 ng/mL; R&D Systems, Minneapolis, MN) (Dolznig et al., 2001) and harvested at the density of 2-3x10^6 cells/ml. The chromatin is crosslinked by adding formaldehyde directly to culture medium to a final concentration of 1%. It is incubated for 10 minutes at room temperature. The reaction is stopped by the addition of 1M glycin to final concentration 0.125M. The cells are spun at 1500 RPM for 3 min at 4°C (Eppendorf centrifuge 5810R), the supernatant is decanted and the cells are washed in ice cold PBS. The spin and wash are repeated twice. The cells are resuspended at a density of 2x10^8 cells/ml sonication buffer (10 mM Tris-HCl, 1 mM EDTA **,** 0.5 EGTA, PH 8.0) and 1x complete protease inhibitor (Roche), incubateed for 10min on ice. The cells are disrupted and the DNA fragmented by sonication at 0^0C (Sanyo, SoniPrep150 large probe, 14 times, amplitude 7; 15" on; 45"off). The sample is centrifuged for 10 minutes at 13,000 rpm at 4°C, and the supernatant aliquoted in 100μl per microcentrifuge tube (i.e 20x10^6 cells per tube). The pellet are discarded. The samples (chromatin) are snap frozen in liquid nitrogen and stored at -80°C.

(***Note:*** The sonication conditions may have to be optimized for different machines). The efficiency of the sonication is analysed by adding 8 μL 5 M NaCl to 200μl chromatin sample and the crosslinks reversed at 65°C for 4 hours. The DNA is recovered by phenol/chloroform extraction and run on an agarose gel to visualize the shearing efficiency.

Part B. Pulldown

Note: When proceeding to the PCR or Immunoblot protocol a portion of the diluted chromatin 2% (~20 μL) is kept to quantify the amount of DNA or Protein present in different samples at the PCR or Immunoblot protocol. This sample is considered to be the input/starting material.

Anti-HA-agarose beads Pulldown

100ul of the immunoprecipitating monoclonal Anti-HA-Agarose beads (Sigma Cat# A2095) are washed per sample with 1 ml PBS. The agarose is pelleted by brief centrifugation for 1min at 2000 RPM and the supernatant is discarded. This wash step is repeated once more. The beads are equilibrated with 1ml ChIP dilution buffer (0.01% SDS, 1.1% Triton X100, 1.2mM EDTA, 16.7mM tris-HCl, pH 8.1, 150mM NaCl) and the agarose pelleted by brief centrifugation for 1min at 2000 RPM. The supernatant is discarded. The equilibration step is repeated two more times.

The beads are blocked for 1 hour with 1% Fish skin gelatin (Sigma G7765), and 0.2mg/ml Chicken egg albumin (Sigma A-5503) and 40ul sonicated salmon sperm DNA (10mg/ml). The agarose is pelleted by brief centrifugation (1 min) at 2000 RPM the supernatant discarded.

Concurrent with preparing the beads a number of frozen chromatin tubes are thawed on ice. Each chromatin sample equal to $20x10^6$ cells (i.e. 100 μl) is taken up in 1 ml ChIP Dilution Buffer adding 1X complete protease inhibitors as above. The chromatin is precleared with 75 μL of Protein A Agarose/Salmon Sperm DNA (50% Slurry), (Upstate # 16-157C) for 30 minutes at 4°C with rotation, to reduce the nonspecific background. The agarose is pelleted by brief centrifugation and the supernatant collected. The precleared supernatant was added to the blocked beads and incubate at 4 °C overnight with rotation.

For a negative control, a no-antibody was performed by incubating the supernatant fraction with same amount of Protein A Agarose/Salmon Sperm DNA washed, equilibrated and blocked in the same way. A non-expressing TetR-HA-biotag fusion protein cell lysate or Doxycycline (1μg/ml sigma) treated cell lysate was also used as a negative control treated the same as the experimental sample.

The agarose was pelleted by centrifugation (1min at 2000 RPM) and the supernatant carefully removed (it contains the unbound, non-specific DNA or chromatin). The protein A agarose/antibody/ *TetR-HA-biotag* /DNA complex is washed for 3-5 minutes on a rotating platform with 1 mL of each of the buffers listed in the following order, all at 4^0C; a Low Salt Immune Complex Wash Buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, pH 8.1, 150 mM NaCl), one wash;a High Salt Immune Complex Wash Buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, pH 8.1, 500 mM NaCl), one wash; a LiCl Immune Complex Wash Buffer (0.25 M LiCl, 1% NP-40, 1% deoxycholic acid (sodium salt), 1 mM EDTA, 10 mM Tris, pH 8.1), one wash; TE Buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0), two washes.

The sample is at this stage a protein A/antibody/histone/DNA complex ready for either an Immunoprecipitation/Immunoblot assay (Section I) or Polymerase Chain Reaction (PCR) assay (Section II) or further purification.

Immunoprecipitation/Immunoblot protocol

Following washing of the beads the immunoprecipitated complex can be analyzed by immunoblot analysis. 25 µL of 1X Laemmli buffer is added per sample and boiled for 10 minutes. 20 µL is loaded per lane and the immunoblot procedure carried out as described per appropriate antibody.

PCR protocol to amplify DNA bound to the beads

500µl of freshly prepared elution buffer (1%SDS, 0.1M NaHCO3) is added per sample of the pelleted protein A agarose/antibody/ *TetR-HA-biotag/DNA* complexed beads. The sample is vortexed briefly to mix and then incubated at room temperature for 15 minutes with rotation. 20 µL 5 M NaCl is added to the elution (500 µL) and the protein-DNA crosslinks are reversed by heating at 65°C for more than 4 hours or over night. The sample can be stored overnight at 20°C.

Note: Include the input/starting material (*20* µL of the diluted chromatin above) plus 480 µl of elution buffer and 20 µl 5 M NaCl heating to 65 °C for 4 hours. 10 µL of 0.5 M EDTA 20 µL 1 M Tris-HCl, pH 6.5 is added plus 2 µl of 10 mg/mL Proteinase K to the eluates and incubated for one hour at 45°C. The DNA is cleaned by two phenol/chloroform and one chloroform extraction. 1µl of 20 µg/ml glycogen (Roche, 14267332) is added followed by 0.6 volumes of isopropanol. After 30min at -20°C the sample is centrifuged and the pellet washed with 70% ethanol and air dried. They are resuspended into water, the input into 400 µl and pull down sample into 200µl H2O for PCR. 4µl is used per PCR reaction.

Streptavidin paramagnetic beads Pulldown

For analysis of binding efficiency and γ-globin promoter by streptavidin paramagnetic beads via TetR3T protein, chromatin pull down was performed essentially similar to Anti-HA-agarose beads pulldown with slight modification. 50µl beads used per $20x10^6$ cells and beads were separated by magnet instead of centrifugation. In the PCR protocol beads were kept after elution and were discarded after first phenol chloroform. This is because of strong binding affinity of biotin- streptavidin that can not be eluted in elution buffer only.

RNA purification and Real- time RT-PCR analysis

Total RNA was extracted from cells using the TRI reagent (sigma) and used directly for S1 nuclease protection analysis of globin expression. For quantitative RT-PCR, cDNA was synthesized from 1 µg of total RNA using random hexamers and Superscript™ II RNase H-Reverse Transcriptase (Invitrogen). RNase free DNase (Invitrogen) was used to degrade possible DNA contaminant and primers were designed over an intron. PCR was performed as described below. Human *GAPDH and Actin* was used as an endogenous reference for normalization.

Quantitative PCR condition and primers

Quantitative real-time PCR (MyIQ, Bio Rad) was performed using 0.75 μl of SYBR Green I (sigma S9430) 1/2000 dilution in DMSO, Platinum taq kit (Invitrogen), 10 pmol of each primer, 4μl DNA sample under the following cycling conditions: 3 min. at 95°C followed by 40 cycles of 30 s at 95°C, 20 s at 56°C, 40 s at 60°C, 15 s at 75°C. Mouse βmajor and Amylase was used as an endogenous reference. Enrichment of specific sequences was calculated using the comparative C_T method (Livak and Schmittgen, 2001).

Human genomic βLCR HS5/F	CCCAAGCAAGGAAGTTGT
Human genomic βLCR HS5/R	CAGATGTCCTGTCCCTGTA
Human genomic tetO-100/F	AAAAGTCACAAAGAGTATATTCAAAAAG
Human genomic tetO-100/R	CAGGATTTTTGACGGGACAAA
tetO/F	AAGTCGAGCTCGGTACTACG
tetO/R	GACTTCTTTTGTCAGCCGTTTT
Human genomic tetO+100/F	CGGCTGACAAAAGAAGTCCT
Human genomic tetO+100/R	CCCAAGAGGATACTGCTGCT
Human genomic AγExonI/F	ACCCTTCAGCAGTTCCACAC
Human genomic AγExonI/R	CCCCACAGGCTTGTGATAGT
Human genomic AγExonIII/F	GACCGTTTTGGCAATCCATTTC
Human genomic AγExonIII/R	TTGTATTGCTTGCAGAATAAAGCC
Mouse genomic Amylase/F	CTCCTTGTACGGGTTGGT
Mouse genomic Amylase/R	AATGATGTGCACAGCTGAA
Mouse genomic βmajor/F	GGGAGAAATATGCTTGTCATC
Mouse genomic βmajor/R	CAACTGATCCTACCTCACCTT
Human cDNA FANCI/F	TGAGGAAGCTGGAACACTTAGG
Human cDNA FANCI/R	TGGTCCTGGAAAATGGTGAGC
Human cDNA GAPDH/F	GCCAAAAGGGTCATCATCTC
Human cDNA GAPDH/R	GGTGCTAAGCAGTTGGTGGT

HA elution and further purification by streptavidin pulldown

Note: for Mass spectrometry analysis the whole procedure was scaled up accordingly**.**

After binding to the HA beads and washing the sample was eluted in low salt wash buffer with 0.5mg/ml of HA peptide (SciLight Biotechnology, LLC) 3 times same volume as beads bed volume (i.e. 50μl) each time, shaking at 37C. Streptavidin paramagnetic beads (Dynabeads, M280, In Vitrogen) were blocked as described above for Anti-HA-agarose beads. The eluate was diluted 10x with ChIP dilution buffer plus protease inhibitor and bound to 50μl blocked Streptavidin paramagnetic beads overnight at 4^0C on a rotating wheel. Beads are separated by magnet, washed with 4 buffers 3-5 min each as described above (low salt

washing buffer, high salt washing buffer, LiCl washing buffer and 2x TE buffer) and boiled in 1X Laemmli sample buffer. The sample is run on a NuPAGE 4-12% gradient SDS acrylamide gel (Invitrogen).

Alternatively, for analysis of HA elution or bound DNA or protein fraction to streptavidin paramagnetic beads, DNA for PCR protocol or protein samples for Immunoblot protocol was prepared and analyzed as described.

Mass spectrometry

The mass spectrometry was carried out after gel electrophoresis. Each gel was cut into 15 slices and analysed by LC-MS/MS in an orbiTrap. The peptides were identified using the Mascot search engine.

RNAi and S1 nuclease protection

shRNA vectors were obtained from the TRC Mission human and mouse library from Sigma as bacterial glycerol stocks. Packaging were as described using puromycin selection (http://biomics.erasmusmc.nl/). HEP cells were transduced with lentiviruses containing shRNA's, puromycin was added at the concentration of 50 µg/ml for 24 hours and they were kept in culture for another 5 days before harvesting and further analysis. For FANCI siRNA was purchased (Darmacon, ON-TARGET plus SMART pools) and transfected as described (POLYplus, INTERFERin^{TM}). S1 nuclease was as described (Hanscombe et al., 1991).

Western blotting

Cell extracts after shRNA knock down were prepared, 200x10^5 cells in 100µl of RIPA buffer (10 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1% Nonidet P-40, 1% NaDOC, 0.1% SDS) and same volume of laemmli buffer was added and boild for 5 min. 20µl of each sample was resolved by 4-12% gradient SDS-PAGE (Invitrogen) , and transferred to 0.45 µm nitrocellulose membrane. The membranes were blocked for 1 hour in 1XTBS with 5% milk and 0.05% Tween20 and incubated in primary- (overnight) and secondary (1 hour) antibody. Primary antibodies: FANCI (KIA1794) abcam cat# ab15344; ZBP-89, abcam cat# ab69933; Supt5 santa cruze cat# sc-101158; CDC5L, abcam cat#ab51320; CTNNBL1, abcam cat#ab76243; APEX1 (Ref-1 C4)X cat# sc-17774 X, Yb-1 abcam cat#ab12148. Western blots were developed with the ECL detection kit (PIERCE).

HPLC Analysis of hemoglobin subtypes was measured by HPLC (BioRad).

Immunohistochemistry, cells were spotted on poly-prep slides (Sigma), fixed with 4% paraformaldehyde, permeabilized in 10 mM citric acid (pH 6.0), and blocked with 5% BSA.

Primary antibody incubation was performed in blocking solution for 16 hrs at 4°C, followed by peroxidase staining.

Acknowledgements

We are grateful to Ton de Wit and John Kong a San for the generation of transgenic mice, Ali Imam, and John Strouboulis for helpful advice, Kirsten van Lom for HPLC analysis of hemoglobin subtypes and Ortho-Biotech, Tilburg, The Netherlands for the donation of Epo. This work was supported by the NIH (USA), Landsteiner Foundation for Blood Transfusion Research (NL), the center for Biomedical genetics (NL) and the FP7 EUTRACC consortium (EU).

Supplementary Data

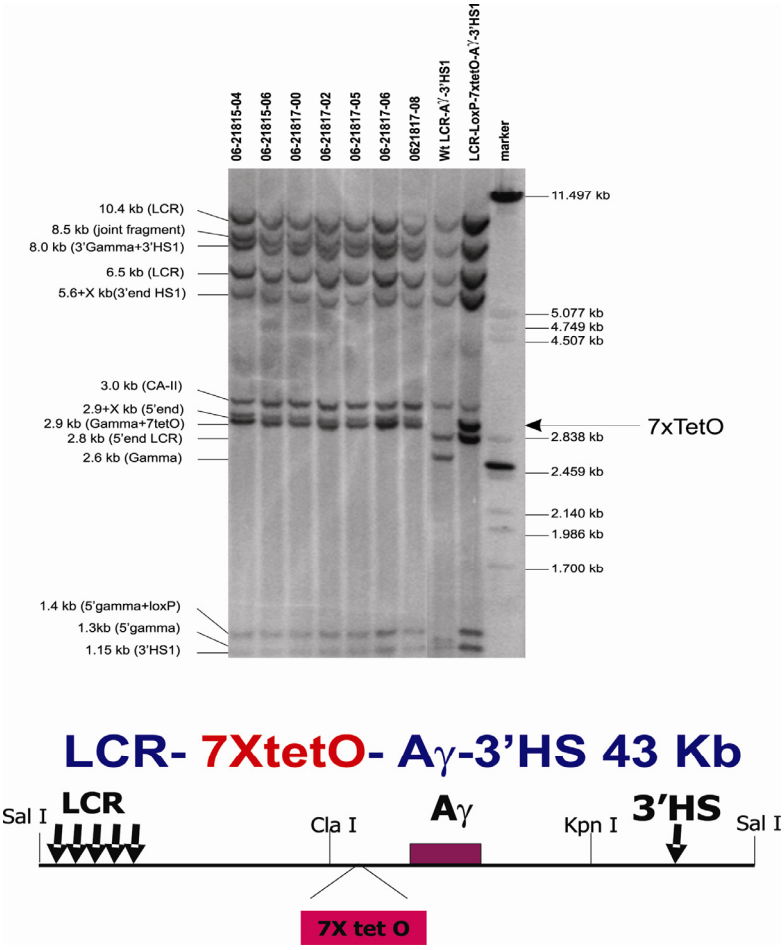
Overlapping oligos for HA-Bio tag Synthesis

HA-Bio/F1:	5'>GATCCTATCCTTACGATGTACCCGACTATGCATATTCTGCTGGTC
HA-Bio/F2:	5'>ATCACCACCACCATCACCATCACCATCACTTAGTTCCTCGTGGTTCTCCTG GTATTTCTG
HA-Bio/F3:	5'>GCGGAGGGGGTGGCGCGTCATCGTTGCGTCAGATCCTTGATAGTCAAAAAA TGGAGT
HA-Bio/F4:	5'>GGCG CAGCAACGCTGGGGGTTCTTAATAGC
His-Bio/R1:	5'>GGCCGCTATTAAGAACCCCCAGCGTTGC
His-Bio/R2:	5'>TGCGCCACTCCATTTTTTGACTATCAAGGATCTGACGCAACGATGACGCG CCACCCC
His-Bio/R3:	5'>CTCCGCCAGAAATACCAGGAGAACCACGAGGAACTAAGTGATG GTGATGGTGA TGGTGGT
HA-Bio/R4:	5'>GGTGATGACCAGCAGAATATGCATAGTCGGGTACATCGTAAGGATAGG

Overlapping oligos for 30 nucleotid spacer and restriction enzyme digestion sites followed by tetO repeats:

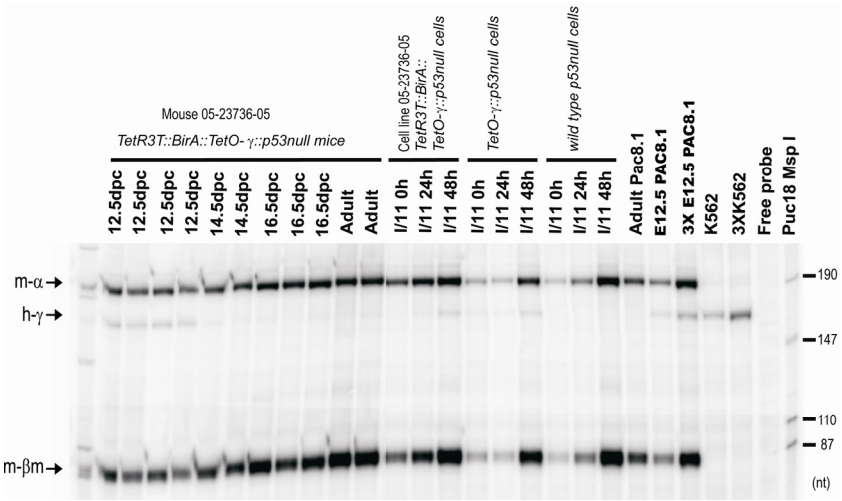
30nt_Pme/Swa/Stu/EcoRV/AspF
GTACTACGTAATACGACTCACTAGTGAGATGTATTTAAATAGCTTTGTTTAAACTGAGCGCCGGAGGCC TAGATATCG

30nt_Pme/Swa/Stu/EcoRV/AspR
GTACCGATATCTAGGCCTCCGGCGCTCAGTTTAAACAAAGCTATTTAAATACATCTCACTAGTGAGTCG TATTACGTA



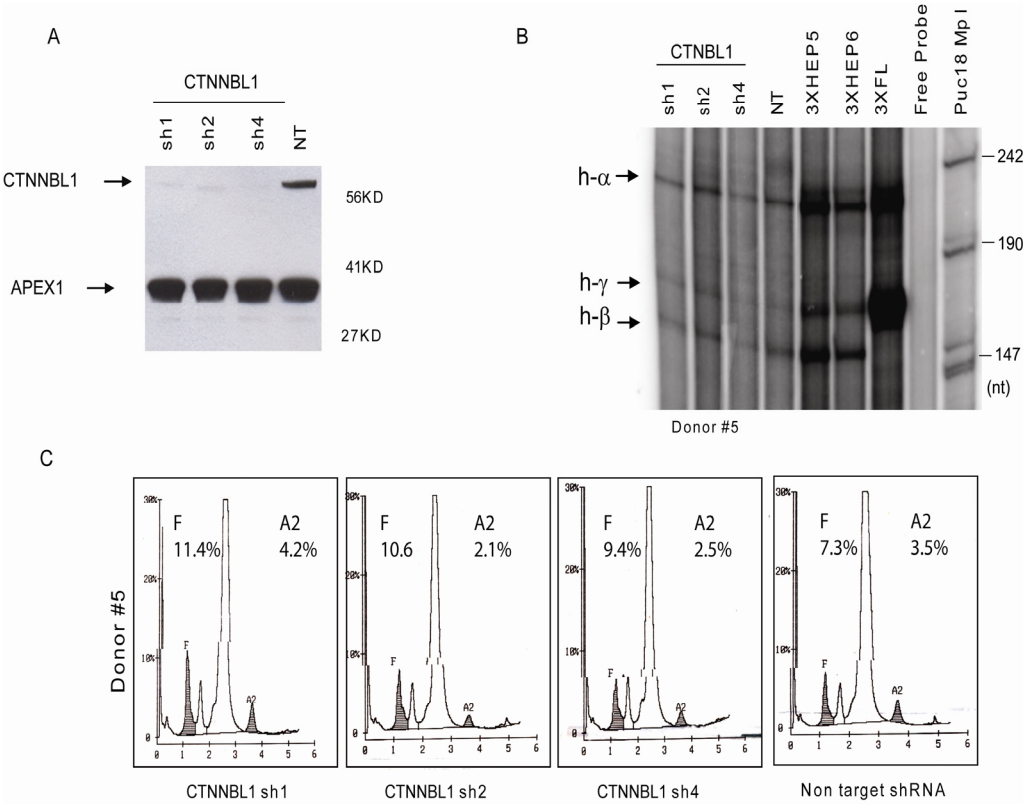
Supplementary Figure 1. The TetO modified γ globin minilocus

The drawing on the bottom shows the γ globin minilocus (Dillon and Grosveld, 1991) and the position of the insertion of 7 copies of the tetO operator sequences in the γ globin gene promoter. The top panel shows the EcoRI restriction digest of the normal γ globin minilocus (Wt LCR-Aγ-3'HS1) and the modified γ globin locus (LCR-LoxP-7xtetO-Aγ-3'HS1) cosmids. The fragment containg the TetO sequences is indicated. The lanes on the left show the Southern blot of several transgenic mice containing two copies of the locus and hybridized with the γ globin minilocus as a probe (Dillon and Grosveld, 1991) and a PstI fragment of carbonic anhydrase gene as loading control. The transgene is determined two head to tail copy using 5'HS5 and 3'HS1 probes (not shown). The identity of the bands is indicated on the left. The fragment sizes of the marker bands are indicated on the right.

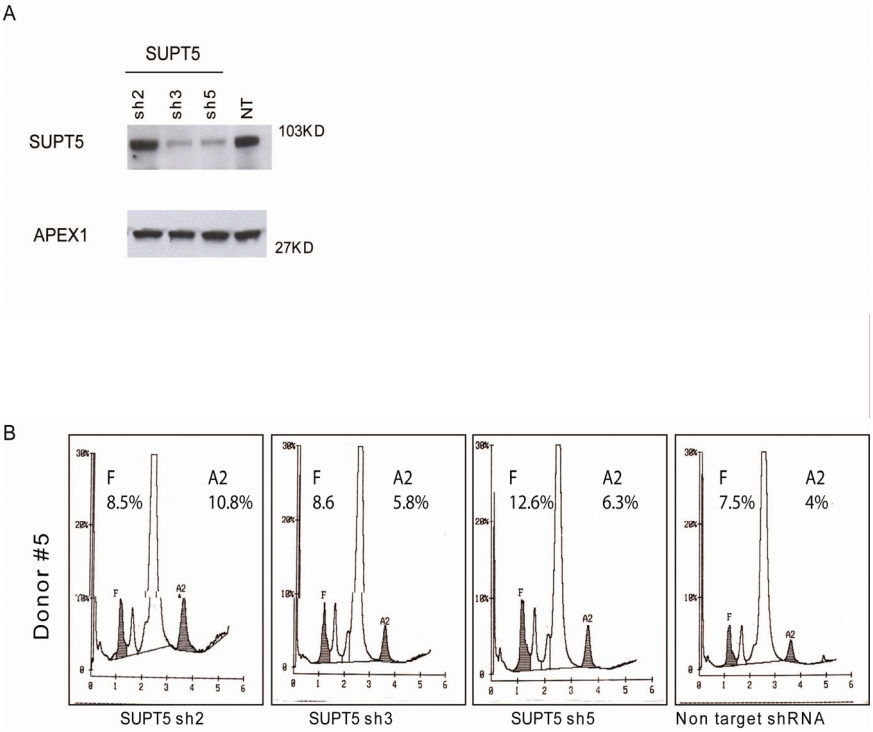


Supplementary Figure 2. S1 nuclease protection analysis of globin expression in cell line derived from day 13.5 fetal liver *TetR3T::BirA::TetO-γ::p53null* mice (05-23736-05).

RNA was prepared from different time point fetal livers and adult blood of transgenic mice derived from 05-23736-05 transgenic mouse line carrying two copies of the TetO-γ transgene, the BirA and TetR3T constructs (*TetR3T::BirA::TetO-γ::p53null*) and the cell line derived from the fetal liver of these mice that was used for the purification of the γ globin gene promoter. A control $p53^{-/-}$ cell line containing the modified γ globin construct (*TetO-γ::p53null*) and a wt $p53^{-/-}$ were used as controls. Three time points were measured for each cell line after induction for 0, 24 and 48 hrs. Quantitative S1 nuclease protection assays were performed as described(Dillon and Grosveld, 1991). Additional controls are shown on the right. Adult PAC8.1: RNA from adult blood of mouse carrying a complete human β-globin locus (PAC8.1) (Wai et al., 2003), E12.5 PAC8.1: RNA from liver of E12.5 PAC8.1 mouse embryo. 3x E12.5 PAC8.1: same as previous but 3 times the amount of RNA. K562 and 3xK562: RNA from the human γ globin expressing cell line K562 and 3 times the amount of that RNA. The samples were separated on a 6% polyacrylamide urea gel; the positions of the protected fragments for mouse α-globin (m-α), βmajor globin (m-βm) and human γ-globin (h-γ) are indicated. Size marker: pUC18 cut with MspI.



Supplementary Figure 3. Lentiviral RNA interference-mediated gene silencing of CTNNBL1.
A) immunoblot for CTNNBL1 shRNA knock down in HEP cells, Last lane Non-target shRNA lentiviral transduced cells; Loading control APEX1. **B)**S1 nuclease protection analysis of globin expression in CTNNBL1 knock down HEP cells. NT; Non-target shRNA. Extra controls in the right, 3XHEP 5 and 6; 3xRNA from no virus treated HEP 5 and HEP6 cells, 3xFL; Human fetal liver RNA, PUC18 MspI size marker. The position of human α, γ and β are indicated at the left of the panel. **C)** HPLC hemoglobin chain analysis of the CTNNBL1 shRNA knock down in HEP5 (donor#5) cells.



Supplementary Figure 4. Lentiviral RNA interference-mediated gene silencing of SUPT5.
A) immunoblot for SUPT5 shRNA knock down in HEP cells, Last lane Non-target shRNA lentiviral transduced cells; Loading control APEX1. **B)** HPLC hemoglobin chain analysis of the CTNNBL1 shRNA knock down in HEP5 (donor#5) cells.

• Titration of cross -linking agent	→ 1% formaldehyde
• Pre-clearing of the Chromatin	→ NO
• Blocking the beads	→YES
• Sonication buffer Vs. SDS lysis buffer	→ Sonication buffer
• Number of cells / 100 µl of HA beads	→ $2x10^7$ cells/100 µl beads
• Optimal streptavidin beads	→ M280
• HA peptide elution titration	→ 0.5mg/ml
• Cells grow in fermentor	→Good
• DNA shearing	→Sonication

Suppelmentry Table 1. Optimization of the purification protocol.
Each step that was optimized is indicated, details for each of the steps are available on request.

hit	Score	acc./id	mgigene	share 575 & 582	share 575 & 617	share 575 & pool	description
65	505	gi|6680229	Hmgb2		Hmgb2		high mobility group box 2 [Mus musculus]
143	299	gi|6680231	Hmgb3				high mobility group box 3 [Mus musculus]
149	290	gi|24496776	Bclaf1				BCL2-associated transcription
162	275	gi|55451	Ybx1				Y box-binbing protein [Mus musculus]
169	264	gi|4001805	Actl6a		Actl6a		BAF53a [Mus musculus]
173	258	gi|6753086	Apex1				apurinic/apyrimidinic endonuclease 1
188	236	gi|1405747	Sfrs2				PR264/SC35 [Mus musculus]
192	234	gi|18204699	Nol5a				Nucleolar protein 5A [Mus musculus]
193	232	gi|22779899	Cdc5l	Cdc5l	Cdc5l		cell division cycle 5-like [Mus musculus]
216	205	gi|39930425	Gtf2f2				general transcription factor IIF
222	197	gi|6755364	Sub1				RNA polymerase II transcrip.coactivator
223	197	gi|82950149	Dsp	Dsp			PREDICTED: desmoplakin isoform 2
250	167	gi|1765919	Dnmt1			Dnmt1	DNA methyltransferase 1 [Mus musculus]
252	166	gi|13096934	Fip1l1			Fip1l1	Fip1l1 protein [Mus musculus]
260	156	gi|2645205	Mybbp1a			Mybbp1a	p160 myb-binding protein [Mus musculus]
276	142	gi|6679947	Gata1			Gata1	GATA binding protein 1 [Mus musculus]
298	130	gi|7657357	Nap1l1				nucleosome assembly protein 1-like 1
304	124	gi|1389682	Jup	Jup			plakoglobin
310	122	gi|20373167	Luc7l2				LUC7-like 2 [Mus musculus]
342	99	gi|124358955	Son				Son cell proliferation protein
346	98	gi|86198327	Nolc1				nucleolar and coiled-body phosphoprotein
354	93	gi|47606347	Ctnnbl1		Ctnnbl1		Beta-catenin-like protein 1 (NAP)
359	91	gi|59798918	Snw1				SNW domain-containing protein 1
373	82	gi|31560213	Zfp830				zinc finger protein 830 [Mus musculus]
377	81	gi|17529977	Pak1ip1				p21-activated protein kinase-interacting
383	79	gi|22094123	Supt5h			Supt5h	suppressor of Ty 5 homolog
391	77	gi|13543768	Chd4			Chd4	Chd4 protein [Mus musculus]
392	77	gi|484530	H3F3A				H3.3 like histone MH921 - mouse
400	72	gi|7271905	Wdr12			Wdr12	nuclear protein Ytm1p [Mus musculus]
412	68	gi|68085626	Ccdc88b				Ccdc88b protein [Mus musculus]
414	67	gi|608528	Khdrbs1				p62 ras-GAP associated phosphoprotein
428	58	gi|6681069	Csrp1				cysteine and glycine-rich protein 1
429	58	gi|10946666	Gsdma1				gasdermin A1 [Mus musculus]
434	56	gi|33563236	Arhgdib				Rho, GDP dissociation inhibitor (GDI)
435	56	gi|4826962	RAC3				ras-related C3 botulinum toxin substrate 3
436	56	gi|6754816	Sept2				septin 2 [Mus musculus]
448	51	gi|151357840	Dnajc25				novel protein [Mus musculus]
459	47	gi|46909602	Samhd1				SAM domain- and HD domain-containing
466	46	gi|5353754	Ddb1			Ddb1	damage-specific DNA binding protein 1
467	46	gi|18204100	Pds5a			Pds5a	Pds5a protein [Mus musculus]
473	44	gi|487874	Dnm2				dynamin
474	44	gi|162138942	Gm381				hypothetical protein LOC214308
477	44	gi|189409138	Cand1			Cand1	TBP-interacting protein [Mus musculus]
478	44	gi|12839970	1700049E17Rik1				unnamed protein product [Mus musculus]
479	43	gi|17369057	Mbnl1				Muscleblind-like protein 1
480	43	gi|21312626	2400003C14Rik				hypothetical protein LOC71955
481	43	gi|126722700	Fanci				Fanconi anemia, complementation group I
482	43	gi|56404955	Trim42				Tripartite motif-containing protein 42
483	43	gi|21706613	Rrn3				Rrn3 protein [Mus musculus]
484	43	gi|124487275	2010204N08Rik	2010204N08Rik			hypothetical protein LOC69983
488	42	gi|22748609	Rgs9bp				regulator of G protein signaling 9-binding
489	42	gi|12861981	Myo1f				unnamed protein product [Mus musculus]
491	42	gi|12844270	Ppil2				unnamed protein product [Mus musculus]
496	40	gi|12834781	1110020G09Rik				unnamed protein product [Mus musculus]
498	39	gi|71725377	Fat2				FAT tumor suppressor homolog 2
501	38	gi|187957390	Esf1			Esf1	ESF1, nucleolar pre-rRNA processing
508	37	gi|157838004	Ehmt1	Ehmt1			euchromatic histone methyltransferase 1
509	36	gi|85702059	4931408C20Rik				hypothetical protein LOC210940
510	36	gi|148697999	E2f2			E2f2	E2F transcription factor 2
516	35	gi|74226354	Il13ra1	Il13ra1			unnamed protein product [Mus musculus]
520	35	gi|12860412	2610528E23Rik				unnamed protein product [Mus musculus]
526	33	gi|8394027	Ppp2r1a			Ppp2r1a	alpha isoform of regulatory A
530	32	gi|148683299					mCG17558 [Mus musculus]
533	32	gi|148675873					mCG1030867, isoform CRA_a
534	32	gi|40557580	Hivep3				zinc finger protein ZAS3 [Mus musculus]

Supplementary Table 2. Mass spectrometry hits in non induced *TetR3T::BirA::TetO-γ::p53null* cells. The hits present in *TetR3T::BirA::TetO-γ::p53null* cells treated with Doxycyclin for non-specific background controls are subtracted. Hits that were very unlikely to be associated with γ-globin suppression are removed. One of the typical experiments 575 is presented. The shared hits in the other experiments are listed in the next columns.

hit	Score	acc./id	mgigene	description
21	205	gi|28302223		Predicted gene, EG432987 [Mus musculus]
27	135	gi|13385998	Trap1	TNF receptor-associated protein 1 [Mus musculus]
37	108	gi|149255357	Hmgb1l	PREDICTED: similar to HMG-L6 [Mus musculus]
47	81	gi|149251801	4930422G04Rik	PREDICTED: similar to 4930422G04Rik protein [Mus musculus]
52	75	gi|167736365	Muc4	mucin 4 [Mus musculus]
53	75	gi|6680421	Il1rap	interleukin 1 receptor accessory protein isoform a [Mus musculus]
61	71	gi|27882581	D030074E01Rik	D030074E01Rik protein [Mus musculus]
67	68	gi|5932003	Naip2	neuronal apoptosis inhibitory protein-rs6 [Mus musculus]
68	63	gi|123210264		ryanodine receptor 3 [Mus musculus]
69	67	gi|29244180	Morc2b	microrchidia 2B [Mus musculus]
74	64	gi|13435022	Tlr3	toll-like receptor 3 [Mus musculus]
79	63	gi|39841025	Fchsd2	FCH and double SH3 domains 2 [Mus musculus]
81	58	gi|81868330	Ranbp2	E3 SUMO-protein ligase RanBP2 (Ran-binding protein 2)
82	63	gi|148692276	Nlrp9c	NACHT, LRR and PYD containing protein 9c [Mus musculus]
83	62	gi|27734194	Sox30	SRY-box containing gene 30 [Mus musculus]
84	61	gi|26331056	Chm	unnamed protein product [Mus musculus]
86	61	gi|148707900	Serpinb2	serine (or cysteine) peptidase inhibitor, clade B, member 2
90	58	gi|63561913	9130208D14Rik	PREDICTED: hypothetical protein LOC77700 [Mus musculus]
95	56	gi|149249923	Dcdc5	PREDICTED: doublecortin domain containing 5 [Mus musculus]
99	56	gi|7304903	Dbf4	activator of S phase kinase [Mus musculus]
100	56	gi|149260345	Zcwpw2	PREDICTED: similar to Zinc finger, CW type with PWWP domain 2
102	56	gi|1724124	Zfp148	transcription factor BFCOL1 [Mus musculus]
103	55	gi|149251592	Tchh	PREDICTED: similar to hCG1642996 [Mus musculus]
108	55	gi|4454797	Tcfeb	transcription factor TFEB [Mus musculus]
109	55	gi|71892410	Mmp27	matrix metalloproteinase 27 [Mus musculus]
117	53	gi|38016154	Nup50	nucleoporin 50 [Mus musculus]
118	53	gi|16169008	Serac1	serine active site containing 1 isoform 2 [Mus musculus]
126	51	gi|37537534	Rpap1	RNA polymerase II associated protein 1 [Mus musculus]
130	51	gi|26328953	Gse1	unnamed protein product [Mus musculus]
136	51	gi|74178013	Bckdk	unnamed protein product [Mus musculus]
144	50	gi|34147153	Zfp691	zinc finger protein 691 [Mus musculus]
145	50	gi|20270271	1190017O12Rik	hypothetical protein LOC68936 [Mus musculus]
147	50	gi|149270722	Ccdc147	PREDICTED: gene model 969, (NCBI) [Mus musculus]
149	47	gi|68534089		Pregnancy specific glycoprotein 19 [Mus musculus]
152	49	gi|26338359	D3Bwg0562e	unnamed protein product [Mus musculus]
153	44	gi|148705783	Wdr19	WD repeat domain 19, isoform CRA_a [Mus musculus]
156	49	gi|63079701	Spopl	speckle-type POZ protein-like [Mus musculus]
157	49	gi|148704531		mCG146314, isoform CRA_a [Mus musculus]
158	43	gi|148705925	Scfd2	sec1 family domain containing 2, isoform CRA_c [Mus musculus]
159	48	gi|74151988	Pkm2	unnamed protein product [Mus musculus]
161	48	gi|74181165	Smarca2	unnamed protein product [Mus musculus]
162	48	gi|148698432		mCG1040588, isoform CRA_c [Mus musculus]
163	48	gi|37359824	Pdcd11	mKIAA0185 protein [Mus musculus]
164	48	gi|37747414	Asz1	Ankyrin repeat, SAM and basic leucine zipper domain containing 1
167	48	gi|74219852	Hnrnpu	unnamed protein product [Mus musculus]
168	48	gi|74220725	Tatdn2	unnamed protein product [Mus musculus]
169	48	gi|210031395	Vash1	vasohibin 1 [Mus musculus]
173	47	gi|149240748	LOC100046802	PREDICTED: similar to Inhbb protein [Mus musculus]
176	47	gi|22122579	Fam13b	hypothetical protein LOC225358 [Mus musculus]
177	47	gi|148701361	Rtn3	reticulon 3, isoform CRA_c [Mus musculus]
178	47	gi|148704587	Lrrc9	leucine rich repeat containing 9, isoform CRA_d [Mus musculus]
179	47	gi|21450157	Lmcd1	LIM and cysteine-rich domains 1 [Mus musculus]
184	46	gi|149267752	EG622509	PREDICTED: similar to putative protein kinase [Mus musculus]
185	46	gi|148669182		mCG6318 [Mus musculus]
189	46	gi|123227600	Ubr4	zinc finger, UBR1 type 1 [Mus musculus]
194	46	gi|148697023		mCG1035397 [Mus musculus]
196	46	gi|148699387		mCG116173, isoform CRA_b [Mus musculus]
202	45	gi|38085189	Ccdc147	PREDICTED: hypothetical protein [Mus musculus]
203	45	gi|74213935	Znfx1	unnamed protein product [Mus musculus]
210	45	gi|50345972	Foxc1	forkhead box C1 [Mus musculus]
211	44	gi|149253812	100038949	PREDICTED: hypothetical protein [Mus musculus]
212	44	gi|28913562	Tcerg1l	Tcerg1l protein [Mus musculus]
213	44	gi|124430711	Ccdc76	coiled-coil domain containing 76 [Mus musculus]
221	44	gi|4521251	Nbn	Nibrin [Mus musculus]
223	44	gi|74196413	Magix	unnamed protein product [Mus musculus]
224	44	gi|34447211	Rad23a	RAD23a homolog [Mus musculus]
225	44	gi|94388538	4930485B16Rik	PREDICTED: hypothetical protein LOC380654 [Mus musculus]
228	44	gi|12853102	4921524L21Rik	unnamed protein product [Mus musculus]
230	43	gi|12839062		unnamed protein product [Mus musculus]
231	43	gi|149264597	LOC100047422	PREDICTED: hypothetical protein [Mus musculus]
234	39	gi|84490375		RAB GTPase activating protein 1-like isoform b [Mus musculus]
235	41	gi|148685519	Hirip3	HIRA interacting protein 3, isoform CRA_a [Mus musculus]
241	43	gi|9837264	1110032E23Rik	hypothetical basic protein I-19 [Mus musculus]
243	43	gi|148677742	Zfp532	zinc finger protein 532 [Mus musculus]

Supplementary Table 3. Mass spectrometry hits in induced *TetR3T::BirA::TetO-γ::p53null* cells. The hits present in *TetR3T::BirA::TetO-γ::p53null* cells treated with Doxycyclin for non-specific background controls are subtracted. Hits that were very unlikely to be associated with γ-globin suppression are removed. In this experiments (599) 244 hits were observed, 98 hits were present in Doxycyclin control. Out of 146 specific hits, 72 of them were unlikely to be associated with γ-globin suppression and this left us with 74 hits.

References

Bank, A. (2006). Regulation of human fetal hemoglobin: new players, new complexities Blood *107*, 435-443.

Berry, M., Grosveld, F., and Dillon, N. (1992). A single point mutation is the cause of the Greek form of hereditary persistence of fetal haemoglobin. Nature *358*, 499-502.

Brand, M., Ranish, J.A., Kummer, N.T., Hamilton, J., Igarashi, K., Francastel, C., Chi, T.H., Crabtree, G.R., Aebersold, R., and Groudine, M. (2004). Dynamic changes in transcription factor complexes during erythroid differentiation revealed by quantitative proteomics. Nat Struct Mol Biol *11*, 73-80.

Chattopadhyay, R., Das, S., Maiti, A.K., Boldogh, I., Xie, J., Hazra, T.K., Kohno, K., Mitra, S., and Bhakat, K.K. (2008). Regulatory Role of Human AP-Endonuclease (APE1/Ref-1) in YB-1-Mediated Activation of the Multidrug Resistance Gene MDR1. Mol Cell Biol *28*, 7066-7080.

Chen, Z., Luo, H.-Y., Basran, R.K., Hsu, T.-H., Mang, D.W.H., Nuntakarn, L., Rosenfield, C.G., Patrinos, G.P., Hardison, R.C., Steinberg, M.H.*, et al.* (2008). A T-to-G Transversion at Nucleotide -567 Upstream of HBG2 in a GATA-1 Binding Motif Is Associated with Elevated Hemoglobin F. Mol Cell Biol *28*, 4386-4393.

Constantoulakis, P., Knitter, G., and Stamatoyannopoulos, G. (1989). On the induction of fetal hemoglobin by butyrates: in vivo and in vitro studies with sodium butyrate and comparison of combination treatments with 5-AzaC and AraC. Blood *74*, 1963-1971.

Conticello, S.G., Ganesh, K., Xue, K., Lu, M., Rada, C., and Neuberger, M.S. (2008). Interaction between Antibody-Diversification Enzyme AID and Spliceosome-Associated Factor CTNNBL1. Molecular Cell *31*, 474-484.

de Boer, E., Rodriguez, P., Bonte, E., Krijgsveld, J., Katsantoni, E., Heck, A., Grosveld, F., and Strouboulis, J. (2003). Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. Nucl Acids Res *100*, 7480-7485.

Déjardin, J., and Kingston, R.E. (2009). Purification of Proteins Associated with Specific Genomic Loci. Cell *136*, 175-186.

Dillon, N., and Grosveld, F. (1991). Human gamma-globin genes silenced independently of other genes in the beta-globin locus. Nature *350*, 252-254.

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

Donehower, L.A., Harvey, M., Slagle, B.L., McArthur, M.J., Montgomery, C.A., Jr., Butel, J.S., and Bradley, A. (1992). Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. Nature *356*, 215-221.

Driegen, S., Ferreira, R., Van Zon, A., Strouboulis, J., Jaegle, M., Grosveld, F., Philipsen, S., and Meijer, D. (2005). A generic tool for biotinylation of tagged proteins in transgenic mice. Transgenic Res *14*, 477-482.

G Stamatoyannopoulos, F.G. (2001). Hemoglobin switching. In: G Stamatoyannopoulos, P Majerus, R Perlmutter and H Varmus, Editors, Molecular Basis of Blood Diseases (3rd ed), W Saunders Publishing Company, Philadelphia, 135–182.

Gossen, M., and Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc Natl Acad Sci U S A *89*, 5547-5551.

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

Gumruk, F., Tavil, B., Balta, G., Unal, S., and Gurgey, A. (2008). Significance of Fetal Hemoglobin Values in Detection of Heterozygotes in Fanconi Anemia: Reevaluation of Fetal Hemoglobin Values by a Sensitive Method. J Pediatr Hematol Oncol *30*, 896-899 810.1097/MPH.1090b1013e3181833835.

Haley, J.D., Smith, D.E., Schwedes, J., Brennan, R., Pearce, C., Moore, C., Wang, F., Petti, F., Grosveld, F., Jane, S.M.*, et al.* (2003). Identification and characterization of mechanistically distinct inducers of [gamma]-globin transcription. Biochemical Pharmacology *66*, 1755-1768.

Hanscombe, O., Whyatt, D., Fraser, P., Yannoutsos, N., Greaves, D., Dillon, N., and Grosveld, F. (1991). Importance of globin gene order for correct developmental expression. Genes Dev *5*, 1387-1394.

Hardison, R.C., Chui, D.H., Giardine, B., Riemer, C., Patrinos, G.P., Anagnou, N., Miller, W., and Wajcman, H. (2002). HbVar: A relational database of human hemoglobin variants and thalassemia mutations at the globin gene server. Hum Mutat *19*, 225-233.

Harju-Baker, S., Costa, F.C., Fedosyuk, H., Neades, R., and Peterson, K.R. (2008). Silencing of Agamma-globin gene expression during adult definitive erythropoiesis mediated by GATA-1-FOG-1-Mi2 complex binding at the -566 GATA site. Mol Cell Biol *28*, 3101-3113.

Hirsch, J.D., Eslamizar, L., Filanoski, B.J., Malekzadeh, N., Haugland, R.P., Beechem, J.M., and Haugland, R.P. (2002). Easily reversible desthiobiotin binding to streptavidin, avidin, and other biotin-binding proteins: uses for protein labeling, detection, and isolation. Analytical Biochemistry *308*, 343-357.

Inusa, E.C.a.B. (2007). Sickle Cell Anemia: Targeting the Role of Fetal Hemoglobin in Therapy Clinical Pediatrics *46*, 386 - 391.

Iwasaki, K., MacKenzie, E.L., Hailemariam, K., Sakamoto, K., and Tsuji, Y. (2006). Hemin-Mediated Regulation of an Antioxidant-Responsive Element of the Human Ferritin H Gene and Role of Ref-1 during Erythroid Differentiation of K562 Cells. Mol Cell Biol *26*, 2845-2856.

Jiang, J., Best, S., Menzel, S., Silver, N., Lai, M.I., Surdulescu, G.L., Spector, T.D., and Thein, S.L. (2006a). cMYB is involved in the regulation of fetal hemoglobin production in adults. Blood *108*, 1077-1083.

Jiang, J., Best, S., Menzel, S., Silver, N., Lai, M.I., Surdulescu, G.L., Spector, T.D., and Thein, S.L. (2006b). cMYB is involved in the regulation of fetal hemoglobin production in adults. Blood *108*, 1077-1083.

Johnson, K.D., Norton, J.E., and Bresnick, E.H. (2002). Requirements for utilization of CREB binding protein by hypersensitive site two of the beta-globin locus control region. Nucleic Acids Res *30*, 1522-1530.

Katsantoni, E.Z., Langeveld, A., Wai, A.W.K., Drabek, D., Grosveld, F., Anagnou, N.P., and Strouboulis, J. (2003). Persistent gamma-globin expression in adult transgenic mice is mediated by HPFH-2, HPFH-3, and HPFH-6 breakpoint sequences. Blood *102*, 3412-3419.

Kuroyanagi, Y., Kaneko, Y., Muta, K., Park, B.-S., Moi, P., Ausenda, S., Cappellini, M.D., and Ikuta, T. (2006). cAMP differentially regulates [gamma]-globin gene expression in erythroleukemic cells and primary erythroblasts through c-Myb expression. Biochemical and Biophysical Research Communications *344*, 1038-1047.

Leberbauer, C., Boulme, F., Unfried, G., Huber, J., Beug, H., and Mullner, E.W. (2005). Different steroids co-regulate long-term expansion versus terminal differentiation in primary human erythroid progenitors. Blood *105*, 85-94.

Lettre, G., Sankaran, V.G., Bezerra, M.A.C., Araújo, A.S., Uda, M., Sanna, S., Cao, A., Schlessinger, D., Costa, F.F., Hirschhorn, J.N.*, et al.* (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 U S A *105*, 11869-11874.

Li, Q., Duan, Z.J., and Stamatoyannopoulos, G. (2001). Analysis of the mechanism of action of non-deletion hereditary persistence of fetal hemoglobin mutants in transgenic mice. Embo J *20*, 157-164.

Li, X., Xiong, J.-W., Shelley, C.S., Park, H., and Arnaout, M.A. (2006). The transcription factor ZBP-89 controls generation of the hematopoietic lineage in zebrafish and mouse embryonic stem cells. Development *133*, 3641-3650.

Liu, L., R,, Du, Z., W,, Zhao, H., L,, Liu, X., L,, Huang, X., D,, Shen, J., Ju, L., M,, Fang, F., D,, and Zhang, J., W, (2005). T to C Substitution at -175 -173 of the Îbeta-Globin Promoter Affects GATA-1 and Oct-1 Binding in Vitro Differently but Can Independently Reproduce the Hereditary Persistence of Fetal Hemoglobin Phenotype in Transgenic Mice. J Biol Chem *280*, 7452-7459.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2-[Delta][Delta]CT Method. Methods *25*, 402-408.

Lu, Z.H., Books, J.T., and Ley, T.J. (2005). YB-1 Is Important for Late-Stage Embryonic Development, Optimal Cellular Stress Responses, and the Prevention of Premature Senescence. Mol Cell Biol *25*, 4625-4637.

Mahajan, M.C., Narlikar, G.J., Boyapaty, G., Kingston, R.E., and Weissman, S.M. (2005). Heterogeneous nuclear ribonucleoprotein C1/C2, MeCP1, and SWI/SNF form a chromatin remodeling complex at the beta-globin locus control region. Proc Natl Acad Sci U S A *102*, 15012-15017.

Mankidy, R., Faller, D.V., Mabaera, R., Lowrey, C.H., Boosalis, M.S., White, G.L., Castaneda, S.A., and Perrine, S.P. (2006). Short-chain fatty acids induce gamma-globin gene expression by displacement of a HDAC3-NCoR repressor complex. Blood *108*, 3179-3186.

Menzel, S., Garner, C., Gut, I., Matsuda, F., Yamaguchi, M., Heath, S., Foglio, M., Zelenika, D., Boland, A., Rooks, H.*, et al.* (2007). A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. Nat Genet *39*, 1197-1199.

Merchant, J.L., Iyer, G.R., Taylor, B.R., Kitchen, J.R., Mortensen, E.R., Wang, Z., Flintoft, R.J., Michel, J.B., and Bassel-Duby, R. (1996). ZBP-89, a Kruppel-like zinc finger protein, inhibits epidermal growth factor induction of the gastrin promoter. Mol Cell Biol *16*, 6644-6653.

Miniero, R., David, O., Saglio, G., Paschero, C., Nicola, P. (1981). The Hbf in Fanconi's anemia Pediatr Med Chir *3*, 167-170.

Ohneda, K., Ohmori, S.y., Ishijima, Y., Nakano, M., and Yamamoto, M. (2009). Characterization of a Functional ZBP-89 Binding Site That Mediates Gata1 Gene Expression during Hematopoietic Development. J Biol Chem *284*, 30187-30199.

Ohneda, K., Shimizu, R., Nishimura, S., Muraosa, Y., Takahashi, S., Engel, J.D., and Yamamoto, M. (2002). A minigene containing four discrete cis elements recapitulates GATA-1 gene expression in vivo. Genes Cells *7*, 1243-1254.

Olave, I.A., Doneanu, C., Fang, X., Stamatoyannopoulos, G., and Li, Q. (2007). Purification and Identification of Proteins That Bind to the Hereditary Persistence of Fetal Hemoglobin â€"198 Mutation in the Î³-Globin Gene Promoter. J Biol Chem *282*, 853-862.

Perrine, S.P. (2008). Fetal globin stimulant therapies in the beta-hemoglobinopathies: principles and current potential. Pediatr Ann *37*, 339-346.

Perrine, S.P., Miller, B.A., Faller, D.V., Cohen, R.A., Vichinsky, E.P., Hurst, D., Lubin, B.H., and Papayannopoulou, T. (1989). Sodium butyrate enhances fetal globin gene expression in erythroid progenitors of patients with Hb SS and beta thalassemia. Blood *74*, 454-459.

Perrine, S.P., Miller, B.A., Greene, M.F., Cohen, R.A., Cook, N., Shackleton, C., and Faller, D.V. (1987). Butryic acid analogues augment gamma globin gene expression in neonatal erythroid progenitors. Biochemical and Biophysical Research Communications *148*, 694-700.

Peterson, K.R., Navas, P.A., and Stamatoyannopoulos, G. (1998). beta-YAC transgenic mice for studying LCR function. Ann N Y Acad Sci *850*, 28-37.

Sankaran, V.G., Menne, T.F., Xu, J., Akie, T.E., Lettre, G., Van Handel, B., Mikkola, H.K.A., Hirschhorn, J.N., Cantor, A.B., and Orkin, S.H. (2008). Human Fetal Hemoglobin Expression Is Regulated by the Developmental Stage-Specific Repressor BCL11A. Science *322*, 1839-1842.

Spitkovsky, D.D., Royer-Pokora, B., Delius, H., Kisseljov, F., Jenkins, N.A., Gilbert, D.J., Copeland, N.G., and Royer, H.-D. (1992). Tissue restricted expression and chromosomal localization of the YB-1 gene encoding a 42kD nuclear CCAAT binding protein. Nucl Acids Res *20*, 797-803.

Stamatoyannopoulos, G. (2005). Prospects for developing a molecular cure for thalassemia. Hematology *10*, 255-257.

Starck, J., Sarkar, R., Romana, M., Bhargava, A., Scarpa, A.L., Tanaka, M., Chamberlain, J.W., Weissman, S.M., and Forget, B.G. (1994). Developmental regulation of human gamma- and beta-globin genes in the absence of the locus control region. Blood *84*, 1656-1665.

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

Takashi, H., and Kazuo, S. (1996). A cdc5+ homolog of a higher plant,â€‰Arabidopsisâ€‰thaliana. Proc Natl Acad Sci U S A *93*, 13371-13376.

Talbort, D., and Grosveld, F. (1991). The 5'HS2 of the globin locus control region enhances transcription through the interaction of a multimeric complex binding at two functionally distinct NF-E2 binding sites. EMBO J *10*, 1391-1398.

Thein, S.L., and Menzel, S. (2009). Discovering the genetics underlying foetal haemoglobin production in adults. Br J Haematol *145*, 455-467.

Uda, M., Galanello, R., Sanna, S., Lettre, G., Sankaran, V.G., Chen, W., Usala, G., Busonero, F., Maschio, A., Albai, G.*, et al.* (2008). Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of beta-thalassemia. Proc Natl Acad Sci U S A *105*, 1620-1625.

Vernimmen, D., Gobbi, M.D., Sloane-Stanley, J.A., Wood, W.G., and Higgs, D.R. (2007). Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. EMBO J *26*, 2041-2051.

Wada, T., Takagi, T., Yamaguchi, Y., Watanabe, D., Handa, H. (1998). Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription in vitro. EMBO J *17*, 7395-7403.

Wai, A.W.K., Gillemans, N., Raguz-Bolognesi, S., Pruzina, S., Zafarana, G., Meijer, D., Philipsen, S., and Grosveld, F. (2003). HS5 of the human beta-globin locus control region: a developmental stage-specific border in erythroid cells. EMBO J *22*, 4489-4500.

Woo, A.J., Moran, T.B., Schindler, Y.L., Choe, S.-K., Langer, N.B., Sullivan, M.R., Fujiwara, Y., Paw, B.H., and Cantor, A.B. (2008). Identification of ZBP-89 as a Novel GATA-1-Associated Transcription Factor Involved in Megakaryocytic and Erythroid Development. Mol Cell Biol *28*, 2675-2689.

Wu, Y., Zhang, X., Salmon, M., and Zehner, Z., E. (2007). The zinc finger repressor, ZBP-89, recruits histone deacetylase 1 to repress vimentin gene expression. Genes Cells *12*, 905-918.

Yamada, T., Yamaguchi, Y., Inukai, N., Okamoto, S., Mura, T., and Handa, H. (2006). P-TEFb-Mediated Phosphorylation of hSpt5 C-Terminal Repeats Is Critical for Processive Transcription Elongation. Molecular Cell *21*, 227-237.

Yamaguchi, Y., Takagi, T., Wada, T., Yano, K., Furuya, A., Sugimoto, S., Hasegawa, J., and Handa, H. (1999). NELF, a Multisubunit Complex Containing RD, Cooperates with DSIF to Repress RNA Polymerase II Elongation. Cell *97*, 41-51.

Yokoyama, H., Harigae, H., Takahashi, S., Kameoka, J., Miyamura, K., Ishizawa. K., Kaku, M., and Sasaki, T. (2003a). High expression of YB-1 gene in erythroid cells in patients with refractory anemia. Int J Hematol *3*, 213-218.

Yokoyama, H., Harigae, H., Takahashi, S., Takahashi, S., Furuyama, K., Kaku, M., Yamamoto, M., and Sasaki, T. (2003b). Regulation of YB-1 gene expression by GATA transcription factors. Biochemical and Biophysical Research Communications *303*, 140-145.

Yu M, H.H., Xiang P, Li Q, Stamatoyannopoulos G. (2006). Autonomous silencing as well as competition controls gamma-globin gene expression during development.

. Mol Cell Biol *26*, 4775-4781

Zou, G.-M., Luo, M.-H., Reed, A., Kelley, M.R., and Yoder, M.C. (2007). Ape1 regulates hematopoietic differentiation of embryonic stem cells through its redox functional domain. Blood *109*, 1917-1922.

Chapter4

Genetic factors influencing fetal hemoglobin expression and hydroxyurea response in β-thalassemia

Farzin Pourfarzad[1], Marieke von Lindern[2], Jun Hou[1], Sima Kheradmand Kia[3], Fatemehsadat Esteghamat[1,6], Azita Azarkeivan[4,5], Wilfred van IJcken[7], Sjaak Philipsen[1], Hossein Najmabadi[4,6] and Frank Grosveld[1,7].

[1]Department of cell Biology, Erasmus-MC, Rotterdam, The Netherlands [2]Department of Hematology, Erasmus-MC, Rotterdam, The Netherlands [3]Department of Biochemistry, Erasmus-MC, Rotterdam, The Netherlands [4]Kariminejad-Najmabadi Pathology and Genetics Center, Tehran, Iran [5]Blood Transfusion Organization, Thalassemia Clinic, Tehran, Iran [6]Genetic Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran [7]Center for Biomedical Genetics, NGI, The Netherlands.

Manuscript in preparation

Abstract

The β-thalassemia syndromes are caused by mutations in the β-globin gene locus that cause loss of β-globin expression. Hydroxyurea (HU) increases expression of fetal γ globin in postnatal life, and is therefore an appealing therapeutic approach to the β-thalassemias. Patients treated with HU fall into three categories: (i) 'Good responders' increase hemoglobin to therapeutic levels (ii) 'Moderate responders' increase hemoglobin levels but still need transfusions at longer intervals, and (iii) 'non-responders', who remain transfusion-dependent. The mechanisms underlying these differential responses remain unclear.

We generated RNA expression profiles of erythroblasts grown in the presence or absence of HU that were expanded from blood of 27 β-thalassemia patients being either absolute 'non-responders' or extremely good 'responders'. Many genes were upregulated in response to HU in erythroblasts from 'non-responders', including γ globin. In these cases, however, the γ globin starting level was extremely low and HU induced cell death. In contrast, treatment of erythroblasts from 'responders' with HU did not significantly change their gene expression profile. Their γ globin was upregulated, but the starting level of γ globin was already high. Interestingly, part of the gene program that was upregulated by HU in 'non-responder' erythroblasts was already highly expressed in the erythroblasts of 'responders' before HU treatment. We conclude from the gene expression profiles of 'responders' that the cells of these patients have adopted to constitutive stress conditions.

Introduction

Hemoglobin disorders, particularly β-thalassemia and sickle cell disease, are the most common single gene disorders, worldwide (Weatherall, 2001). They are caused by mutations in the β-globin locus resulting in abnormal or reduced rates of hemoglobin production (Stamatoyannopoulos and Grosveld, 2001). Clinical symptoms include anemia, α/β globin chain imbalance, infarction, bone marrow expansion, splenomegaly. Currently the patients have a life expectancy of about four decades.

During human development, the fetal γ-globin and adult β-globin are the main β-like globins. They associate with α-globin chains to produce HbF ($\alpha_2\gamma_2$) during fetal period and HbA ($\alpha_2\beta_2$) in adult life. The mechanism responsible for this developmentally regulated gene expression pattern, known as globin switching, has been the subject of intense research efforts during the last 30 years, mainly because (reactivation of) γ-globin expression is beneficial to β-hemoglobinopathy patients. In the case of β-thalassemia, γ-globin protein can reduce α-globin chain precipitation and also compensate for the absence of functional β-globin. In sickle cell patients, high γ-globin expression reduces hemoglobin polymerization in erythrocytes and reduces their deformability, thereby ameliorating the symptoms (de Paula et al., 2003; Hajjar and Pearson, 1994; Hoppe et al., 2000; Loukopoulos and Pavlides, 1998; Poillon, 1993; Zeng, 1995).

Several drugs can induce γ-globin gene expression and ameliorate the disease phenotype. Three well known HbF inducing agents are sodium butyrate (histon deacetylase inhibitor) (Constantoulakis et al., 1989; Perrine et al., 1989; Perrine et al., 1987), 5-azacytedine (DNA demethylating agent) (DeSimone, 1982; Dover, 1983; Ley 1983; Ley, 1982) and hydroxyurea (ribonucleotid reductase inhibitor) (Veith et al., 1985). Of these, hydroxyurea is currently the only FDA approved drug and it is widely used to treat patients with β-hemoglobinopathies. How hydroxyurea exactly induces γ-globin gene expression is poorly understood. Mechanisms proposed for the induction of HbF by HU include rapid erythroid regeneration, increased erythropoietin (EPO) production, apoptosis, nitric oxide (NO) production (Cokic et al., 2007), increased guanylate cyclase activity (Cokic et al., 2008), and activation of the p38 MAPK pathway (Park et al., 2001). Induction of HbF by HU in β-thalassemic patients was reported to be of similar magnitude as found in the cells of normal individuals (1.3- to 3.5-fold) and sickle cell patients (2- to 5-fold). HbF induction by HU was reported using erythroid cell cultures (Moi and Kan, 1990). In erythroid progenitor cells treated with HU *in vitro*, HbF induction was comparable to the increase of HbF in peripheral blood *in vivo* of sickle cell disease patients following HU therapy(Yang et al., 1997). HU increased RBC survival from 18.6 +/- 11 days to 70 +/- 21 days, as a result of decreased hemolysis (Ballas et al., 1999). This is presumably due to reduced intracellular HbS polymerization as a result of increased γ-globin expression.

The majority of patients increase HbF production upon HU treatment (Steinberg et al., 1997). Both the HbF baseline and the magnitude of the response vary widely among patients. The absolute response to HU and the HbF baseline may be dependent on genetic factors which modulate different regulatory pathways, including trans-acting factors involved in γ-globin production.

The ability to predict the HbF response to HU would be very useful for the selection of responders, and to prevent side effects of HU treatment. A correlation of SNPs with high HbF induction has been reported in several studies. The XmnI (G)γ polymorphism at -158 (C>T) is associated with high HbF response to HU in sickle cell disease and β-thalassemia in several studies(Alebouyeh, 2004; Bradai et al., 2003; Dixit et al., 2005; Karimi et al., 2005; Neishabury et al., 2008; Panigrahi et al., 2005; Verma et al., 2007; Yavarian et al., 2004). Furthermore, a SNP association study with the response to HU reported 17 and 20 SNPs significantly associated with the percentage of HbF and the response to HU, respectively (Ma et al., 2007).

Regulation of γ globin gene expression is complex and can be influenced by different regulatory pathways, genetic and environmental factors. The net outcome of these determines the response to HU (reviewed in Bank, 2006; Stamatoyannopoulos and Grosveld, 2001). We hypothesized that the activity of the regulatory mechanisms may be deduced from the comparison of the expression profiles of erythroid progenitor cells from HU 'responder' and 'non-responder' groups. In addition, data from such an approach may help to understand the mechanism by which HU induces γ-globin expression. It may also explain the difference

between 'responders' and 'non-responders' regarding base line levels of HbF in individuals, and factors that are involved in high γ-globin induction. Only a few of such studies have been reported. One study reported expression profiling of SCD patients with mild and severe phenotype (Jison et al., 2004). This study used peripheral blood for expression profiling that mainly contains erythrocytes. These are terminally differentiated cells without a nucleus and the program necessary for globin production has finalized. The expression profile of these cells may therefore not be very informative with respect to this process.

Here we studied two patient groups representing β-thalassemia patients that do not respond to HU treatment and remain fully dependent on regular blood transfusions and patients that respond well to HU and do not need treatment. Erythroblasts were expanded from peripheral blood, and erythroblast proliferation, hemoglobin production and gene expression were compared in the presence and absence of HU. Our data shows that cells derived from 'responders' express high levels of HbF compared to cells derived from 'non-responders'. Cells derived from 'non-responders' change the expression pattern of a large number of genes upon HU treatment, while cells derived from 'responders' displayed only minor changes after treatment. Differential gene expression profiles of these two groups indicated that high HbF was associated with a continuous stress level, and with genes that protect from stress-induced apoptosis. This suggest that HU is effective in 'responder' patients because their HbF level are relatively high and their erythroblasts already activated a stress response program which can protected them from HU's cytotoxic effects .

Results

HU induces hematopoietic differentiation and HbF production

Hematopoietic erythroid progenitor (HEP) cells were expanded from peripheral blood mononuclear cells as described (Leberbauer et al., 2005). We first titrated HU on cultured erythroblasts to determine the concentration dependent effect on cell survival and proliferation, and on accumulation of HbA and HbF. This pilot experiment involved two healthy donors. Erythroblasts were expanded from peripheral blood mononuclear cells for 10 days before HU was added in concentrations ranging from 0 to 400 μM and cell proliferation was monitored daily. Hemoglobin production, the percentage of different globin chains, and cell morphology were analyzed at day 16 when the experiment was terminated (Fig. 1).

Cell density was maintained between 1-2 million/ml by daily dilution and cumulative cell numbers were calculated. Increasing HU concentrations progressively decreased cell proliferation, but only 400 μM was immediately toxic to the cells (Fig. 1A). HU induced a concentration-dependent increase in total hemoglobin that reached a maximum at 100 μM HU (Fig. 1B). Analysis of hemoglobin subtypes by HPLC indicated that HU increases the HbF percentage from 1.7% in non HU treated to 6% in 100μM HU treated samples (Fig. 1C). Using total Hb levels, the absolute expression of the different hemoglobin chains can be calculated (Fig.1D). This shows that all hemoglobin chains were induced by HU. Finally, we analyzed

hemoglobinisation at the cellular level. Under proliferation conditions, addition of HU at concentrations from 0 to 200μM increased the number of hemoglobinized cells from 40% at 50μM HU up to 50% at 200μM HU(Fig. 1E).

Based on these observations and also others (Budzowska et al., 2004; Fibach et al., 1993; Nagai et al., 2003; Rodrigue et al., 2001), we conclude HU induces hemoglobinisation/erythroid differentiation in these cells. Since the best response with least cell toxicity was observed at 100μM, we decided to use 100μM HU for our study.

Selection and characterisation of the study population

We selected two groups of β-thalassemia patients based on the complete absence of a response to HU treatment (group I), or on a good response to HU treatment resulting in transfusion independence (group II). In addition, 4 patients were added to group II that are independent of both transfusion and HU treatment. The XmnI (G)γ polymorphism at (position -158C>T in the γ-globin locus) has been associated with increased γ-globin expression and was present in 29% of the non-responders and in 83% of the responders chromosomes (Table 1).

HU sensitivity of erythroblasts derived from responders and non-responders

First, we examined how HU sensitivity at the cellular level corresponds to the response of β-thalassemic patients. Erythroblasts were expanded from blood mononuclear cells of 27 β-thalassemic patients (Table 1). After 10 days of culture, cell samples were further expanded in the presence or absence of 100 μM HU. Cell proliferation was monitored daily and total hemoglobin production, ratio of hemoglobin chains and cell morphology analyzed 6 days after HU addition.

The first difference observed between cultures derived from group I and group II was the relatively poor growth rate of erythroblasts derived from group II, with the exception of #15 (figure 2A,B). Strikingly, addition of HU hardly affected the slow growth rate of 'responder' group II cultures; while 'non-responder' group I cultures were very sensitive to HU treatment. The increased sensitivity of group I cultures to HU treatments is shown in (Fig. 2C).

Total hemoglobin (in arbitrary units, a.u.) and the frequency of fetal versus adult hemoglobin expression was measured 6 days after the start of HU treatment (day 16 of erythroblast expansion), and the absolute distribution of different hemoglobin chains was calculated (figure 2D). Interestingly, HU treatment did not alter the expression of adult Hb but increased the expression of fetal globin (non-responders from 8 to 33 a.u., responders from 106 to 248 a.u.). Because the basal level of total hemoglobin is much higher in group II erythroblasts, only group II erythroblasts express high HbF levels upon HU treatment (Fig. 2D).

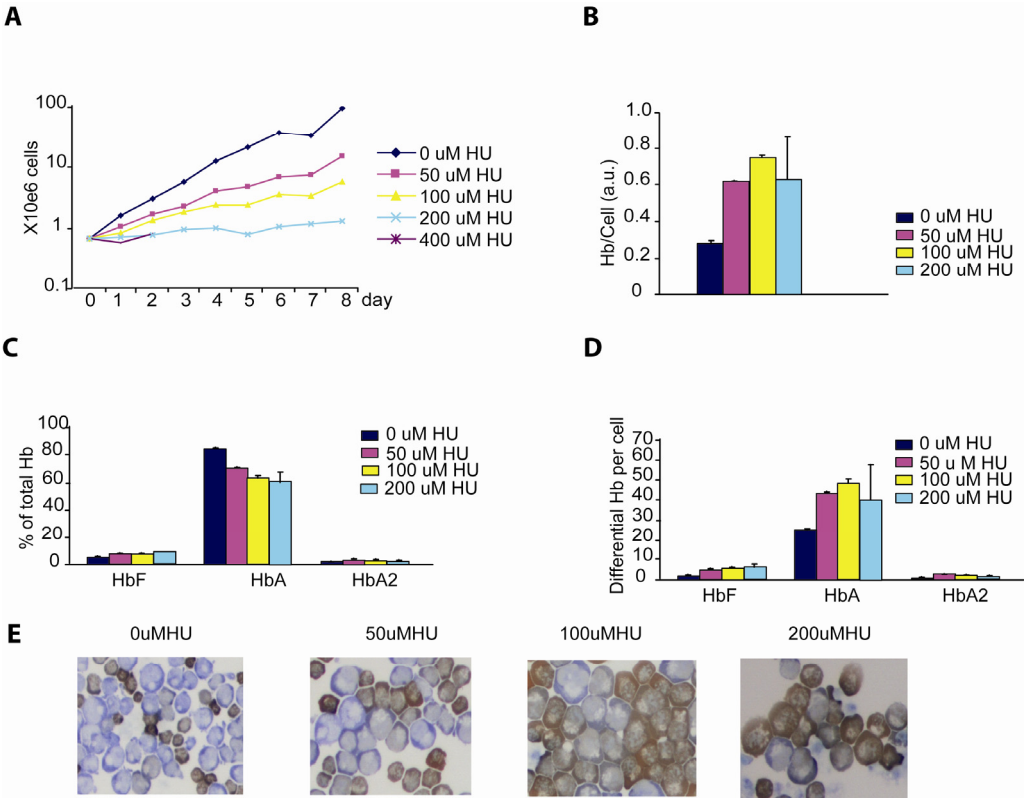


Figure 1. Hydroxyurea (HU) dose-response on healthy donor cells.
A) Human Erythroid progenitor (HEP) cells growth curves in different HU dose from 0 to 400 μM. **B)** The effect of different HU concentrations on total hemoglobin production after 6 days of treatment. Experiments are performed using cells from two healthy donors in triplicate. Hb levels are represented in arbitrary units (a.u.) **C)** HPLC globin chain analysis of samples in B. **D)** Representation of total hemoglobin chains by distributing total hemoglobin (A) in to different hemoglobin chain percentages (C) expressed in arbitrary units (a.u.). **E)** Representative cytospins of HEPs treated with different HU concentrations, stained with histological dyes and neutral benzidine. Hemoglobinized cells are stained brown.

Number	DOB	Sex	Thalassemia type	use of HU	Mutation	Xmn1
Non-Responders						
1	1983	Female	Major	-	IVSII-I/IVSI-110	-/-
2	1981	male	Major	-	IVSII-I/IVS II-I	+/-
3	1982	Female	Major	-	IVSII-I/IVSII-I	-/-
4	1992	Female	Major	-	IVSI-5/IVSI-5	-/-
5	1987	male	Major	-	C44/Cd 27	-/-
6	1977	Female	Major	-	IVSII-I/IVSII-I	+/+
7	1992	Female	Major	-	IVSI-I/IVSI-I	-/-
8	1984	male	Major	-	IVSI-25/IVSI-5	+/-
9	1986	male	Major	-	C39/C39	-/-
10	1978	Female	Major	-	IVSII-I/IVS II-I	+/-
11	1985	male	Major	-	C22/C30	-/-
12	1974	Female	Major	+	IVSII-I/IVSII-I	+/+
Number	DOB	Sex	Thalassemia type	use of HU	Mutation	Xmn1
Responders						
13	1967	male	Intermediate	+	IVS II-I/IVS II-I	+/+
14	1966	male	Major	+	IVSII-I/IVSII-I	+/+
15	1980	Female	Intermediate	+	IVS I-II0/ IVSI-II0	+/+
16	1977	Female	Major	+	C8/C8	+/+
17	1970	Female	Intermediate	+	IVSII-I/C22	+/+
18#	1999	Female	Intermediate	-	IVS II-I/IVS II-I	+/+
19	1985	male	Major	+	IVSII-I/IVSI-5	+/-
20	1985	male	Major	+	IVS1-110 /IVS1- 6	-/-
21	1996	Female	Major	+	IVSII-I/IVSI-128	+/+
22	1982	male	Intermediate	+	cd25-26/cd25-26	+/+
23*	1975	Female	Intermediate	-	cd25-26/cd25-26	+/+
25#	1986	male	Intermediate	-	IVSII-I/IVSII-I	+/+
26	1969	male	Intermediate	+	?/?	+/+
27	1977	Female	Intermediate	+	IVSII-I/IVSII-I	+/+
28	1982	male	Intermediate	-	IVSII-I/29bp deletion	-/-

Table 1. Thalassemic patients included in the study.
DOB, date of birth; Use of HU, + patient on drug, - no HU treatment; β-globin gene mutation and XmnI, -158 C>T Gγ XmnI polymorphism; # brother and sister ; * patient 24 did not present at sampling

Cell morphology

Morphological analysis of erythroblasts treated with HU for 3 days, and stained for hemoglobin in combination with histological dyes showed that group II erythroblast cultures accumulated more hemoglobinised cells, 25% in group II versus 10% in group I, after HU treatment, whereas group I cultures accumulated more pycnotic cells, 20% in group I versus 10% in group II (Fig. 2E).

In conclusion, erythroblasts from patients that do not respond to HU are very sensitive to HU. These cells have a higher propensity to succumb to cell death in response to HU. Furthermore, although the fold-change in HbF levels upon HU treatment is higher than that observed in cells from the 'responders', the very low starting level of HbF precludes induction of HbF to therapeutic levels.

Erythroblasts from HU 'responder' patients constitutively express a stress-program that is induced by HU in 'non-responder' patients

To carry out an expression profiling analysis, erythroblasts were expanded for 10 days (or longer if needed) and subsequently treated with 100μM HU or solvent for two days. Fresh medium plus HU was added after 1 day. RNA expression profiles were compared between 5 group I cultures (non-responders #3, #4, #7, #8, #11) and 5 group II cultures (#14, #15, #16, #18, #28).

Data were analysed as described in Materials and Methods. The expression profiles of different patient groups were compared using the SAM algorithm to identify HU response associated genes.

The first striking difference between group I and II erythroblast profiles is their response to HU at the gene expression level. HU induced many common changes in 'non-responder' derived group I erythroblasts (766 differentially expressed probe sets) whereas HU did not induce significant changes in group II erythroblasts (Fig. 3A).

Comparison of the expression profiles between group I and group II erythroblasts showed a moderate number differentially expressed genes in the absence of HU (145 genes) and the differences became smaller when the cells were treated with HU (100 genes) (Fig. 3A). The overlap between differentially expressed genes plus and minus HU consisted of 26 probe sets/genes.

In total, 245 probe sets were differentially expressed between the two groups of erythroblasts in absence of HU treatment. Expression data of these 245 probe sets were clustered (Fig. 3B). Clustering confirmed that the 'non-responder' group I samples are clearly separated from samples treated with HU. In contrast, the group II samples plus or minus HU are very similar and cluster independent of HU treatment. Clustering indicates that HU barely affected gene expression in samples #18 and #28, and only moderately in samples #14, #15 and #16.

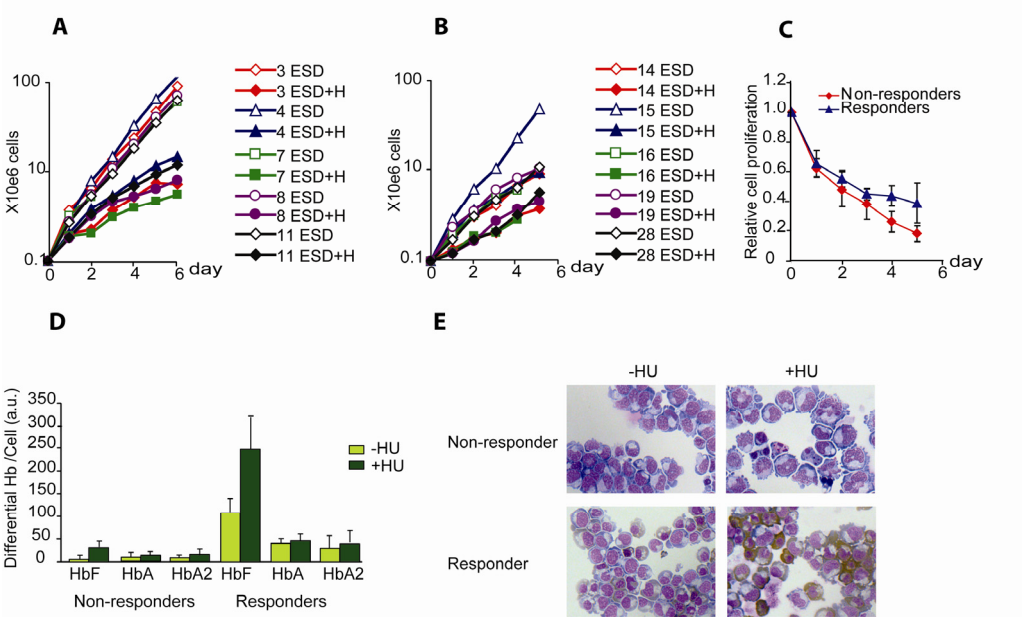


Figure 2. 'Responder' and 'non-responder' HEPs response to 100μM HU.
A) HEPs from 'non-responders' grow faster than those from 'responders' (E: Epo, S: SCF, D: Dex, ESD: Proliferation condition and H: HydroxyUrea treatment). **B)** in the absence of HU. After HU treatment, growth curves of 'non-responder' HEPs decline more than those of 'responders'. **C)** HU sensitivity curve in HEPs from 'responders' and 'non-responders', normalized by taking the ratio of cell proliferation of HU treated cells over non treated cells. **(D)** Hemoglobin chain induction by HU (6 days) in responders and non-responders. See legend to fig.1D for details. **(E)** Representative cytospins of Responder and non-responder HEPs treated with HU for 3 days stained with histological dyes and neutral benzidine. Hemoglobinised cells are stained brown.

Notably, some gene clusters are up- or downregulated with HU in non-responder group I samples, that are consistently expressed above or below average in all group II samples independent of treatment with HU. This suggests that cells from group II patients constitutively express a stress program that is otherwise only activated upon exposure to drugs such as HU, and that may be involved in induction of γ-globin.

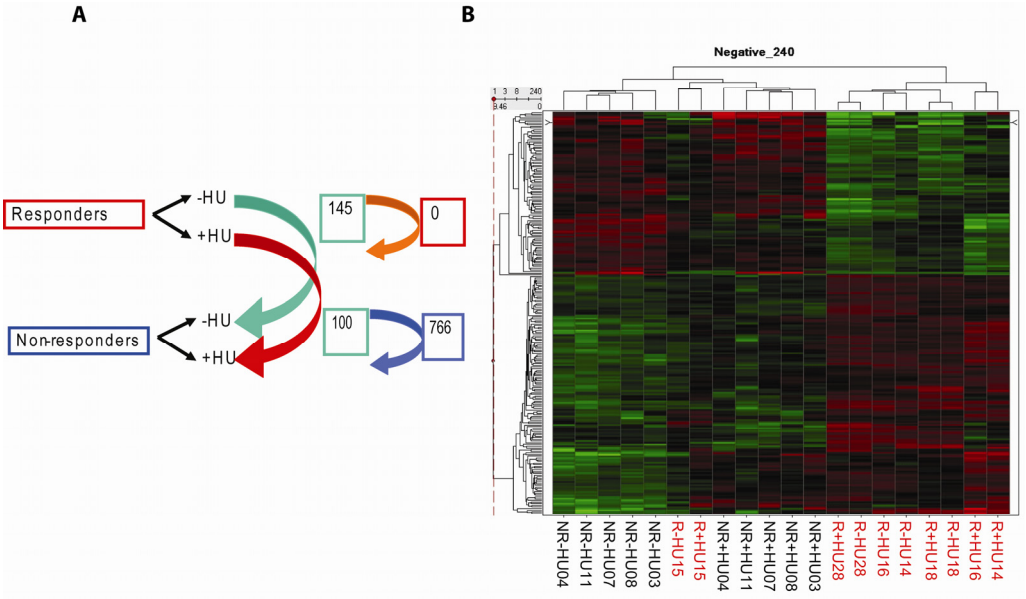


Figure 3. Expression profiling of patients.
A) The number of differentially expressed genes between Responders and Non- responders before and after HU treatment. **B)** Supervised clustering of differentially expressed genes. R: responders, NR: non-responders, +HU: HU treated, - HU: No HU treatment.

Regulation of genes involved in γ-globin expression and stress responses

The INK4b-ARF-INK4a locus

It was striking that the proliferation rate of erythroid cultures expanded from group II patients was much slower compared to those expanded from group I-derived cultures. The *INK4b*-ARF-INK4a locus is known to be involved in stress responses, but oligonucleotides on arrays cannot discriminate between the overlapping ORFs of $p14^{ARF}$ and $p16^{INK4a}$. Therefore, we analyzed expression of $p15^{INK4b}$, $p14^{ARF}$ and $p16^{INK4a}$ encoded by the *INK4b-ARF-INK4a locus* in more detail (Figure 4A). Expression of $p16^{INK4a}$ was increased upon HU treatment for both group I and group II. In contrast, expression of $p14^{ARF}$ was increased in group I compared to group II, and treatment with HU further increased expression in group I, but not in group II. Expression of $p14^{ARF}$ in group I erythroblast cultures treated with HU is on average 10-fold higher when compared to group II cultures treated with HU. Expression of $p15^{INK4b}$ is lower in group I compared to group II before HU treatment. However, HU significantly enhanced $p15^{INK4b}$ expression in group I cultures, but not in group II cultures. In conclusion, the failure of group II cultures to upregulate $p14^{ARF}$ and $p15^{INK4b}$ may contribute to their relative resistance to HU, however the differential expression of the CDKN2A locus does not explain the slow proliferation of group II cultures.

Stress response genes

From the array data we selected a number of genes with a role in the adaptation to stress responses that seemed consistently differentially expressed between cultures derived from group I or group II. Expression of these genes was analyzed using Q-PCR on independent cDNA samples. Forkhead box O3 (FOXO3) is a transcription factor inducing genes that enforce the oxidative firewall. Arginase 1 and 2 (ARG1, ARG2) compete with NO synthase (NOS) for the substrate L-arginine (Durante et al., 2007), and thereby protect against oxidative stress from NO. Homeodomain interacting protein kinase 2 (HIPK2), involved in appotosis, differentiation and also phosphorylative activation of C/EBP-p300(Steinmann et al., 2009; Yoshida and Kitabayashi, 2008). Arg2 and Hipk2 were shown to be Foxo3a target genes (Bakker et al., 2007). These genes were all consistently upregulated in group II compared to group I, and further upregulated upon treatment with HU (Fig. 4B). Notably, expression levels of ARG1 and ARG2 were higher in the untreated group II cultures than in HU-treated group I cultures and may play an important role in protection from HU-induced apoptosis in group II cultures (see discussion). Kruppel like factor 10 (KLF10) (Døsen-Dahl et al., 2008) was identified as a protein that protects ALL blasts and stroma cells against chemotherapy. KLF13 was demonstrated to upregulate the γ-globin gene through the CACCC promoter element (Asano et al., 1999). KLF10 and 13 were expressed at elevated levels in group II cultures compared to group I cultures and reached similar levels upon HU treatment.

Apoptosis genes

$BCLX_L$ protects erythroblasts from apoptosis. Both in the absence and presence of HU, $BCLX_L$ levels are higher in group II cultures. $BCLX_L$ is under control of STAT5, and also STAT5B is increased in group II cultures, although the difference is not as marked as $BCLX_L$ (Fig. 4C and not shown). Whether BCL6 is involved in apoptosis or in a different process is not clear. During VDJ rearrangement in B-cells it is responsible for methylation of the ATR gene and thus prevent the activation of DNA damage responses during the rearrangement process (Ranuncolo et al., 2007). Expression of BCL6 is not increased by HU, but it is expressed almost 10-fold higher in group II cultures (Figure 4C).

γ-globin regulation

Whereas erythropoietic stress induces γ-globin expression, several transcription factors and chromatin modifiers were also shown or suspected to modify γ-globin expression, among them BCL11A, SOX6, SOX4. Expression levels of SOX6 are approximately 3.5-fold higher in group II compared to group I cultures, whereas expression of SOX4 is slightly less. BCL11A is not different between both culture groups. Expression of the γ-globin gene itself at the transcript level is 4-fold higher in group II cultures compared to group I (Figure 4D).

Finally genes important to maintain expansion of erythroblasts such as MYB, BCL11A, PRMT5, IKAROS and STAT5B remained constant (data not shown).

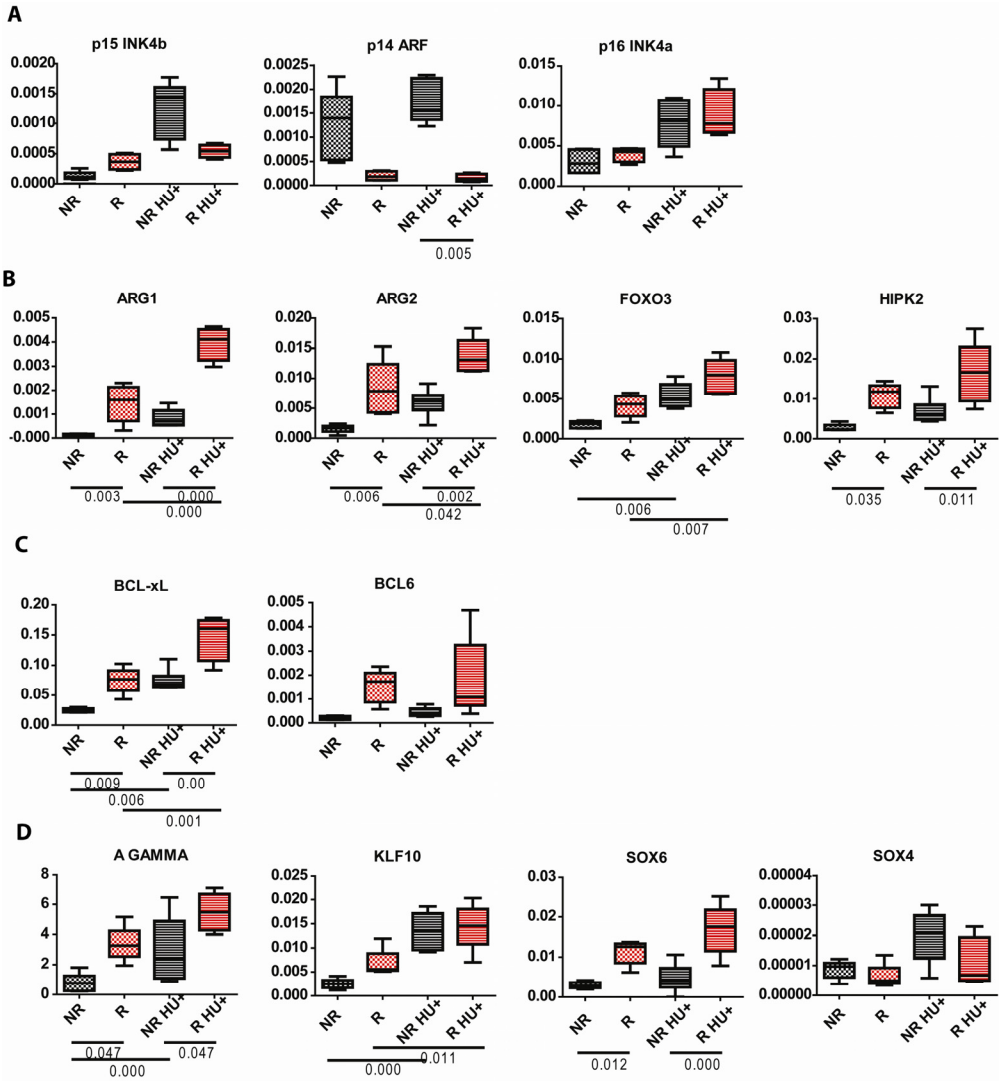


Figure4. qRT-PCR validation of selected genes (see Text)
A) INK4b-ARF-INK4a locus **B)** stress response genes **C)** Apoptosis response genes **D)** γ-globin regulation related genes. NR, 'non-responders' or group I; R, 'Responders' group II; HU+, 'HU treated'. Significant *p*-values > 0.050, are only indicated. P-values calculated by ANOVA, Bonferroni correction from Stata11.0 (Stata Corp, College Station, Texas USA), 5 or more patients were analyzed in each group (n≥5).

Discussion

Activation of γ-globin in patients suffering from β-thalassemia or sickle disease would be an effective cure. However repression of γ-globin after birth is well controlled and can not be reverted easily. Proliferative stress, such as induced by HU, can increase γ-globin expression, but not all patients respond with an increase in HbF that renders them less dependent on regular transfusions. We characterized the HU response of erythroblasts expanded from β-thalassemia patients that do or do not respond to HU, with respect to proliferation and differentiation kinetics, and the gene expression profiles. Erythroblasts expanded from patients that did not respond to HU (i.e. remained transfusion dependent) appeared to be very sensitive to HU, but the initial γ-globin levels are very low. HU upregulated γ-globin expression and that of many other genes, but this eventually resulted in cell death rather than in maturation to hemoglobinized cells; erythroblasts from patients that are or became independent of transfusion were characterized by high initial γ-globin levels, and a moderate to low response to HU in their gene expression profiles. This avoids cell death allowing maturation to hemoglobinized cells. The gene expression profile indicated that high HbF was associated with a gene expression signature of continuous stress, and with genes that protect from stress-induced apoptosis.

Selection of patients

The patients included in this study were selected from a large cohort of patients either because they remained fully transfusion dependent upon HU treatment (group I) or because they were essentially transfusion independent (group II). It can be argued that the basic differences between erythroblast cultures derived from group I and group II are due to the long term treatment of the patients with HU and selection for epigenetic changes that render the cells more resistant to HU. However, group I patients have also been treated with HU in the past, and more importantly, patient # 18, 23, 25 and 28 from group II have never been exposed to HU, they naturally express high HbF. The expression profile of cultures #18 and #28 are very similar to gene expression profiles of other group II cultures and hence the argument that HU might cause some epigenetic changes during years of HU treatment and adaptations of the cells of 'responders' is less likely. Furthermore, all patients stopped using the drug two weeks before blood sampling and the cells were cultured during at least 10 days from progenitors cycling in the periphery.

It is interesting that the culture derived from patient #15 (group II) had a proliferation rate that was comparable to group I cultures, but HbF levels comparable to group II cultures. Following clustering of expression profiling data, the transcriptome of culture #15 was also very much like those from group I cultures. However some genes were expressed at levels similar to those observed in group II cultures and we took advantage of this heterogeneity to pinpoint genes that might be directly involved in HbF response to HU. They include the HIPK2, FOXO3, CDH, ZNF711 and IGFBP7 genes (see below).

Genetic factors regulating γ-globin expression

The -158 XmnI polymorphism in the promoter of the γ-globin gene has been linked to HbF expression (Alebouyeh, 2004; Bradai et al., 2003; Dixit et al., 2005; Karimi et al., 2005; Neishabury et al., 2008; Panigrahi et al., 2005; Verma et al., 2007; Yavarian et al., 2004). In our patient cohort, 13/15 (83% of chromosomes) of group II carries a -158 XmnI C>T polymorphism, and 12 are homozygous for the Gγ polymorphism. In group I only 5/12 (29% of chromosomes) carries the polymorphism, 2 are homozygous for the mutation. Although there is a correlation between the presence of the mutation and HbF expression, this polymorphism does not solely determine HbF and HU response in this patient group.

Several factors were found to control γ-globin expression. *In vitro* cultures increase γ-globin expression in response to SCF and glucocorticoids (Gabbianelli et al., 2008), but the cultures of both groups were exposed to these conditions. Recently, expression of BCL11A was negatively associated with HbF expression. BCL11A is differentially expressed in pre- and postnatal erythroid cells and reduced expression of BCL11A increase γ-globin expression.

We found that group I- and group II-derived cultures express similar levels of BCL11A. However, some other transcription factors involved in γ-globin expression are increased in group II cultures. Q-PCR confirmed regulation of SOX6 (3.8-fold up in group II vs group I) which is in contradiction with other studies (Cohen-Barak et al., 2007; Yi et al., 2006). KLF10 (2,5-fold up in group II vs group I, 5-fold up in response to HU). Array data suggested more interesting genes differentially expressed in group II versus group I that are known to have an effect on the regulation of HbF or at least they might be related (Table 2).

The genes included in this table belong to one of the following categories: (A) directly involved in γ-globin regulation (B) indirectly might be related to γ-globin regulation based on literature or similarities with the genes in the same gene family (C) stress response or apoptosis genes (D) potentially interesting genes. Differential expression is calculated as fold changed comparing responders over non-responders. Several microarray data are confirmed by qRT-PCR.

	Gene Title	Gene Symbol	R vs. NR	qRT Confirmed	Reference
	Group A genes				
A	Hemoglobin, gamma A	HBG1	3.6		
A	Kruppel-like factor 13	KLF13	2.1		
A	SRY (sex determining region Y)-box 6	SOX6	2.1	3.8	(Sripichai et al., 2009)
A	inhibitor of DNA binding 1, dominant negative helix-loop-helix protein	ID1	2		(Scicchitano et al., 2003)
A	CCAAT/enhancer binding protein (C/EBP), beta	CEBPB	1.6		(Wall et al., 1996)
A	hematopoietically expressed homeobox	HHEX	1.6		(Scicchitano et al., 2003)
	Group B genes				
B	arginase, liver	ARG1	8	12.7	Ma et al. 2007
B	cytoplasmic polyadenylation element binding protein 4	CPEB4	3		
B	cAMP responsive element binding protein 3-like 1	CREB3L1	3		
B	interleukin 1 receptor, type I	IL1R1	3		
B	retinoid X receptor, gamma	RXRG	2.8		
B	argininosuccinate synthetase 1	ASS1	2.7		Ma et al. 2009
B	interleukin 6 signal transducer (gp130, oncostatin M receptor)	IL6ST	2.7		
B	SRY (sex determining region Y)-box 3	SOX3	2.5		
B	arginase, type II	ARG2	2.2	5.1	Ma et al. 2008
B	HMG-box transcription factor 1	HBP1	2.2		
B	Kruppel-like factor 10	KLF10	2.2	2.7	
B	tripartite motif-containing 10	TRIM10	2.2		(Blaybel et al., 2008)
B	interleukin 8	IL8	2.1		
B	nuclear factor (erythroid-derived 2), 45kDa	NFE2	2.1		
B	RAR-related orphan receptor A	RORA	2		
B	CKLF-like MARVEL transmembrane domain containing 8	CMTM8	-2		
B	interleukin 7	IL7	-2.1		
B	mitogen-activated protein kinase kinase 7	MAP2K7	-2.3		
B	SRY (sex determining region Y)-box 4	SOX4	-2.5	-1.4	
B	mitogen-activated protein kinase kinase 6	MAP2K6	-2.7		
B	v-myb myeloblastosis viral oncogene homolog (avian)-like 1	MYBL1	-2.8		(Scicchitano et al., 2003)
B	interleukin 7 receptor	IL7R	-4.4		
	Group C genes				
C	B-cell CLL/lymphoma 6 (zinc finger protein 51)	BCL6	5	7.3	
C	cyclin-dependent kinase inhibitor 2B (p15, inhibits CDK4)	CDKN2B	3.7	1	
C	Homeodomain interacting protein kinase 2	HIPK2	2.9	3.9	
C	BTG family, member 2	BTG2	2.2		
C	BCL2-like 1	BCL2L1	2	3	
C	forkhead box O3	FOXO3	2	2.3	
C	caspase 6, apoptosis-related cysteine peptidase	CASP6	-1.7		
C	homeodomain interacting protein kinase 1	HIPK1	-2		
C	cyclin-dependent kinase 6	CDK6	-2.1		
C	Mdm2, transformed 3T3 cell double minute 2, p53 binding protein (mouse) binding protein, 104kDa	MTBP	-2.1		
C	B-cell CLL/lymphoma 2	BCL2	-2.5		
C	DnaJ (Hsp40) homolog, subfamily C, member 12	DNAJC12	-3.8		
	Group D genes				
D	phosphodiesterase 4D interacting protein (myomegalin)	PDE4DIP	4.6		
D	mitogen-activated protein kinase 8 interacting protein 3	MAPK8IP3	2.8		
D	selenium binding protein 1	SELENBP1	2.6		
D	crumbs homolog 1 (Drosophila)	CRB1	2.4		
D	erythrocyte membrane protein band 4.9 (dematin)	EPB49	2.3		
D	zinc finger protein 711	ZNF711	-2.2		
D	microtubule-associated protein 7	MAP7	-2.5		
D	guanine nucleotide binding protein (G protein), beta 5	GNB5	-2.6		
D	cysteine-rich secretory protein 2	CRISP2	-2.7		
D	MAP7 domain containing 2	MAP7D2	-2.8		
D	Meis homeobox 1	MEIS1	-2.9		
D	sorbitol dehydrogenase	SORD	-3.4		

Table 2. Selected differentially expressed genes between non-responders (group I) and responders (group II). The genes included in this table belong to one of the following categories: **A)** directly involved in γ-globin regulation **B)** indirectly might be related to γ-globin regulation based on literature or similarities with the genes in the same gene family **C)** stress response or apoptosis genes **D)** potentially interesting genes. Differential expression is calculated as fold changed comparing responders over non-responders. Microarray data were confirmed in few cases by qRT-PCR.

The role of stress factors in γ-globin expression

The low proliferation rate of the group II-derived erythroblasts and their expression profiles suggest that these cells have adapted to permanent stress conditions, which is only reached in other cells upon HU treatment. The observation that HU did not alter gene expression in group II samples, while >700 probe sets were differentially expressed in response to HU in group I is likely to be an overestimation due to the relative homogeneity of the group I samples and the heterogeneity of the group II samples. However, comparison of profiles +/- HU particularly of group II patients 18 and 28 shows very few differences compared to the group I patients.

The most striking observation from the expression profiles is the expression of stress proteins in group II cultures. An example validated by Q-PCR is expression of FOXO3. FOXO3 is upregulated during erythroid differentiation (Bakker et al., 2004) and in response to various types of stress such as ROS and DNA damage. The consistent expression of MYB indicates that there is no difference in the differentiation stage of the cells (von Lindern et al., 1999). Hence the increased FOXO3 expression in group II compared to group I cultures and the further upregulation in response to HU indicates increased levels of cellular stress. Numerous FOXO3 target genes are also upregulated among which HIPK2, BTG1, and $p27^{KIP}$ (Bakker et al., 2007). Interestingly, the function of stress-induced FOXO3 protein is not to induce cell death, but to increase the potential of cells to prevent and repair oxidative damage and to slow down the cell cycle to allow for DNA repair before replication.

HU is a ribonucleotide reductase inhibitor that stalls cells in S phase (Szekeres et al., 1997). Stalled replication forks are potent inducers of senescence or apoptosis, but apparently also of HbF. Defects in the Fanconi anemia pathway result in increased replication fork stalling. It is therefore interesting to note that Fanconi anemia is associated with high HbF levels (Gumruk et al., 2008; Miniero, 1981). Importantly, a stalled replication fork activates ATR kinase, which is crucial to all downstream events. During B-cell development, BCL6 is required to methylate and silence the ATR locus to prevent ATR activation during VDJ recombination. It is interesting that increased tolerance to HU and high HbF levels are associated with high BCL6 levels. It is worthwhile to test the role of BCL6 in decreased cell death in presence of HU. Increased expression of $BCLX_L$ could also contribute to enhanced survival in presence of HU in group II cultures. In normal erythroid progenitors $BCLX_L$ is induced by EPO to maintain viability of erythroid cells during terminal maturation (Motoyama et al., 1999; Silva et al., 1996).

ARG 1 and ARG 2

HU also increases NO production through phosphorylation and activation of NOS (Cokic et al., 2006). Although NO has been implicated in upregulation of HbF through activation of γ-globin expression (Cokic et al., 2007; Lou et al., 2009), NO also inhibits growth of erythroid primary cells and colony cultures (Maciejewski et al., 1995). Arginase hydrolyzes L-arginine to

urea and L-ornithine in the urea cycle and inhibits nitric oxide (NO) production via competition with NOS for the substrate L-arginine (Durante et al., 2007). During erythroid differentiation NO levels decreased significantly (Kucukkaya et al., 2006). NO donors inhibit the hemoglobinization (Chénais et al., 1999). NO is inhibited more by fetal RBCs when compared to adult RBCs suggesting that fetal RBCs have a higher level of NO scavengers (Calatayud et al., 1998). On the other hand high NO concentrations promote apoptosis, while low NO concentrations result in resistance to apoptosis.

The Arginase 1, ARG1, Arginase 2, ARG2, and ASS1 (Argininosuccinate synthetase 1) genes are differentially expressed between group I and group II. Their expression is 8-, 2- and 2.7-fold increased in group II cultures compared to group I cultures. In a SNP association study ARG1 and ARG2, ASS1, NOS1 and NOS2A were also significantly associated with response to HU treatment (Ma et al., 2007). It could very well be that high expression of ARG1, ARG2 and ASS1 is necessary to protect hematopoietic progenitor cells against the excessive amount of NO after HU treatment to prevent them from apoptosis and scavenge the extra NO in the later stages of differentiation.

In conclusion, the biological and molecular analysis of erythroblast cultures of β-thalassemia patients that are fully dependent or independent of transfusion suggests that several mechanisms may be involved in high HbF expression and HU responsiveness. To be able to predict the HU response, it will be important to identify an expression fingerprint (predictor). This HU response predictor can be developed from the selected gene list presented in table 2 and further confirmed by analysis of a larger cohort of patients to establish reliability. Such a HU response predictor can subsequently be applied to a large number of patients to predict their response to HU.

Material and methods

Patients

27 β-thalassemic patients were selected from a large Iranian collection of more than 3000 β-thalassemic patients based on their clinical manifestation and response to hydroxyurea (HU).

After clinical consultation, measurement of globin levels and other haematological parameters and determination of their mutation in β-globin locus the patients were divided in two categories, namely group I, characterized based on complete absence of a HU response, 'non-responder', and group II 'responders' characterized by a good response to HU resulting in transfusion independence. In addition, 4 patients were added to group II that are independent of transfusion and HU treatment (Table 1 and data not shown). Patients in the responder group have a very mild thalassemic phenotype; they are transfusion independent and use HU. Patients in non-responder group have very severe phenotype, they are transfusion dependent and do not respond to HU. Age and sex were distributed similar

between the groups (group I: 5/12 male, 15-33 yrs median age 23.5; group II, 8/15male, 8-41 yrs median age 27).

Cell culture

The culturing of HEPs was essentially as published (Leberbauer et al., 2005). 40ml blood was collected per patient, and buffy coats isolated by centrifugation. White cell were removed after Ficoll gradient purification from the inter phase and washed. For initial expansion, $5x10^6$ cells/ml were cultivated at a density of $1\text{-}2x10^6$ cells/ml in serum-free medium (StemSpan; Stem Cell Technologies, Vancouver, BC, Canada) enriched with lipids (40 ng/ml cholesterol-rich lipid mix; Sigma) and supplemented with erythropoietin (2 U/ml kind gift of Orthobiotech, Tilburg, The Netherlands), dexamethasone (1 μM; Sigma, St Louis, MO) SCF (50 ng/ml, supernatant of CHO producer cells)(Leberbauer et al., 2005). The erythroblast culture was expanded until day 10 by daily partial medium changes and addition of fresh factors, keeping cell density between $1.5–2 \times 10^6$ cells/ml. Proliferation kinetics and size distribution of the cell populations were monitored daily using an electronic cell counter (CASY-1, Schärfe-System, Reutlingen, Germany). To induce terminal differentiation erythroblasts were washed and reseeded at $1.5\text{-}2 \times 10^6$ cells/ml in lipid-enriched StemSpan supplemented with Epo (5U/ml) and iron loaded transferrin (1 mg/ml; SCIPAC Ltd, UK) (Leberbauer et al., 2005). Differentiating erythroblasts were maintained at $2\text{-}3 \times 10^6$ cells/ml and harvested 48 hours after induction. After day 10 the cells from every patient were divided into 2 groups: proliferation and proliferation + 100ug/ml HU. The HU was refreshed every other day. RNA was isolated 48 hours after HU treatment, the remainder of the cells were kept in culture to determine total hemoglobin, HbF and growth rates.

Cell morphology

Cell morphology was analyzed in cytospins stained with histological dyes and neutral benzidine (Beug et al., 1982), using an OlympusBx40 microscope (40x objective, NA 0.65), an OlympusDp50 CCD camera and Viewfinder Lite 1.0 acquisition software.

Hemoglobin content

Small aliquots of the cultures were removed and analyzed for hemoglobin content by photometry as described (Bakker et al., 2004) and the relative ratios of globin chains were determined by HPLC (BioRad).

Globin locus mutation analysis

β-globin gene mutations in thalassemic patients were determined as described (Najmabadi et al., 2001).

The XmnI polymorphism (C>T substitution at position -158 upstream of the Gγ-globin gene) was detected by PCR and enzymatic digestion with Asp700, an isoschizomer of XmnI, followed by gel electrophoresis (Sutton et al., 1989).

RNA purification and Real-time RT-PCR analysis

Total RNA was extracted from cells using the TRI reagent (Sigma). For quantitative RT-PCR, cDNA was synthesized from 1 μg of total RNA using random hexamers and Superscript™ II RNase H-Reverse Transcriptase (Invitrogen). RNase-free DNase (Invitrogen) was used to degrade possible contaminating DNA and primers were designed over an intron. PCR was performed as described below.

Quantitative PCR condition and primers

Quantitative real-time PCR (MyIQ, Bio Rad) was performed using 0.75 μl of SYBR Green I (Sigma S9430) 1/2500 dilution in DMSO. Platinum Taq kit (Invitrogen), 10 pmol of each primer, 4μl cDNA sample under the following cycling conditions: 3 min. at 95°C followed by 40 cycles of 30 s at 95°C, 20 s at 56°C, 40 s at 60°C, 15 s at 75°C. Human *GAPDH and USP14* were used as endogenous references for normalization. Enrichment of specific sequences was calculated using the comparative C_T method (Livak and Schmittgen, 2001).

Oligo sequences used for qRT-PCR

BCL11A/F	5′-GTCTCGCCGCAAGCAAGG	BCLxl/F	5′-ACCTGAATGACCACCTAGAGC
BCL11A/R	5′-GCCGTGGTCTGGTTCATCATC	BCLxl/R	5′-CAGCGGTTGAAGCGTTCC
ARG1/F	5′-CAAGAAGAACGGAAGAATCAGC	STAT5B/F	5′-CATCCAGTACCAGGAGAGC
ARG1/R	5′-CCAGATGACTCCAAGATCAGG	STAT5B/R	5′-AGAGACACCTGCTTCTGC
IKZF1/F	5′-GGACCTCTCCACCACCTC	PRMT5/F	5′-CCATCAAAGCAGCCATTCTCC
IKZF1/R	5′-AATCCTCCGCACATTCTTCC	PRMT5/R	5′-TGGTGGTTGGTGCCTGTG
HIPK2/F	5′-GCCAGCCACGTCTCCAAGG	MYB/F	5′-CAGTGACGAGGATGATGAGG
HIPK2/R	5′-CACAGCCCAGGGACCACATG	MYB/R	5′-TGTTCCACCAGCTTCTTCAG
BCL6/F	5′-CTGAGGAGATGGGAGAGACC	Gamm/F	5′-AGGTGCTGACTTCCTTGGG
BCL6/R	5′-CAGCGTGTGCCTCTTGAG	Gamma/R	5′-GGGTGAATTCTTTGCCGAA
ARG2/F	5′-TGAGGTGGTTAGCAGAGC	P15 INK4b/F	5′-ATCACATGAGGTCAGGAGTTCG
ARG2/R	5′-AACCCAGACAACACAAAGG	P15 INK4b/R	5′-CCAGGTTCAAGCGAGTCTCC
FOXO3/F	5′-CGTTGCGTGCCCTACTTC	P14 ARF/F	5′-GGTTTTCGTGGTTCACATCC
FOXO3/F	5′-CTCTTGCCAGTTCCCTCATTC	P14 ARF/R	5′-CCTAGACGCTGGCTCCTC
SOX6/F	5′-CGAGACAACAGCAGCAACTTC	P16 INK4a/F	5′-CCCCTTGCCTGGAAAGATAC
SOX6/R	5′-GAGTCCGCTGGTCATGTGG	P16 INK4a/R	5′-AGCCCCTCCTCTTTCTTCCT
SOX4/F	5′-GTCCCACTCCTCCTCTTCC	GAPDH/F	5′-GCCAAAAGGGTCATCATCTC
SOX4/R	5′-CCGACGACGAACTGAAGC	GAPDH/R	5′-GGTGCTAAGCAGTTGGTGGT
KLF10/F	5′-ACCCAGGATGTGGCAAGAC	USP14/F	5′-AACGCTAAAGGATGATGATTGGG
KLF10/R	5′-TTCATCAGAACGGGCAAACC	USP14/R	5′-TTTGGCTGAGGGTTCTTCTGG

Affymetrix microarrays

Hematopoetic progenitor cells were lysed using TRIzol Reagent (Invitrogen, Carlsbad, CA), and then incubated at room temperature for 5 minutes before adding 0.2 µl of chloroform for each 1ml sample. After centrifuging at full speed (12000 rpm) for 20 minutes, the supernatant containing the RNA was precipitated with iso-propanol and centrifuged. The resultant RNA pellets were washed with 75% ethanol and dissolved in RNase-free water. If applicable, they were stored at –80 °C for further usage.

Assessment of RNA quality and concentration

The integrity if the isolated total RNA was verified on the Agilent 2100 BioAnalyzer (Agilent Technologies, Palo Alto, CA). Samples were kept for further processing if the 28s/18s ratio of its RNA was higher than 1.8. The concentrations of the RNAs were measured with a NanoDrop ND-111 UV-VIS spectrophotometer.

cRNA amplification and labelling

Double strand (ds) cDNA synthesis was performed according to the standardized protocol for One-Cycel cDNA synthesis from Affymetrix (Santa Clara, CA). Approximately 5 µg of total RNA was first converted to single strand cDNA in a 20 µl First-Strand Reaction Mix, containing poly-A control RNA, 100 µmol T7-Oligo Primer, 1x first strand buffer, 0.2 mol DTT 10 mmol dNTP mix and SuperScript II. In detail, the sample RNA, the poly-A control RNA and the T7-Oligo Primer were mixed and incubated for 10 min at 70 °C. Secondly, the first strand buffer, the DTT and the dNTP mix were added and incubated for 2 min at 42 °C, followed by adding SuperScript II and incubation of 1 hour at 42 °C. The ds cDNA was prepared from the resultant First-Strand Reaction Mix, mixed with 1x second strand reaction buffer, 30 mmol dNTP mix, E.coli DNA ligase, E.coli DNA Polymerase I and RNaseH. The mix was incubated for 2 hours at 16 °C, then supplemented with T4 DNA Polymerase, and incubated for another 5 minutes at 16 °C. The reaction was stopped by the addition of EDTA to a final concentration of 5 µM. The Sample Cleanup Module and GeneChip IVT Labeling Kit from Affymetrix were used to purify the synthesized ds cDNA, which was used to generate biotin-labeled cRNA, in the presence of 1x IVT Labeling buffer, IVT Labeling NTP Mix, IVT Labeling Enzyme Mix and RNase-free water in a total volume of 40 µl. After an incubation of 16 hours at 37 °C, the concentration and quality of the labelled cRNA were checked with NanoDrop ND-1000 UV-VIS spectrophotometer. An A_{260}/A_{280} ratio between 1.9 and 2.1 was considered acceptable. Approximately 20 µg cRNA per array was fragmented to an average size of 35-200 nucleotides by heating at 94 °C for 35 min, in the presence of a 1x Fragmentation Buffer in a total volume of 40 µl. The undiluted, fragmented samples were stored at –20 °C before being subjected to hybridization.

Hybridization

Hybridization was conducted following Affymetrix instructions for GeneChip® Human Genome U133 Plus 2.0 arrays. The GeneArray scanner 3000 (Affymetrix) was employed to detect the hybridization signals.

Preprocessing microarray data

Array Quality Control

Microarrays that did not pass the quality assessment were removed from further analyses. The quality metrics used to exclude microarrays was the statistics summary calculated by the GCOS algorithm during the processing of probe-level data. The primary inclusion criteria include: all arrays had to have comparable noise values (Raw Q, measurement for the pixel-to-pixel variation of probe cells on the chip); background values were within the range of 20 to 100; percent of present probe sets on the array should not be below 45%. The other criteria were: arrays with extremely high or low values for any of these parameters, e.g. values beyond the range of standard deviation ± median, were excluded; signal ratio of ≤3 of the 3′ / 5′ probe sets for GAPDH and Actin were used as a cut-off; labelling and hybridization were controlled by using standard spike-in controls according to the Affymetrix protocol; if global scaling was applied, the scaling factors for each array were within a three-fold range.

Array Data analysis

Quantile normalization

RMA (Robust Multi-Array average) is an integrated algorithm comprising background adjustment, quantile normalization, and expression summarization by median polish. The intensities of mismatch probes were entirely ignored due to their spurious estimation of non-specific binding. The intensities were background-corrected in such a way that all corrected values must be positive. The RMA algorithm utilized quantile normalization in which the signal value of individual probes was substituted by the average of all probes with the same rank of intensity on each chip/array. Finally Tukey's median polish algorithm was used to obtain the estimates of expression for normalized probe intensities.

Other transformations

Intensities of probe sets lower than 30 were reset to 30. The geometric mean for each probe set was calculated across all samples, or for each subgroup of samples firstly and then across all samples (OmniViz). The intensity values of individual probe sets in each sample were then displayed as the log 2 of the deviations to the calculated geometric means.

Probe sets filtering

Probe sets were involved in further analysis only if their expression levels deviated from the overall mean in at least one array by a minimum factor of 1, because the remaining data were unlikely to be informative. The result was that 36,471 probe sets were eliminated, and 18,204 probe sets remained for further analysis.

Unsupervised clustering and visualization of gene/sample similarity

Clustering was performed without taking into account any external information such as histology subtypes and tumor stages, with each of the selected 18,204 probe sets using the K-means algorithm (OmniViz). Similarities were measured by magnitude and shape (Euclidean distance). Pair-wised similarities between samples were sorted and visualized by the Pearson Correlation Matrix (OmniViz). The order of clusters and individual samples within each cluster was sorted according to the Pearson Correlation Coefficient.

Statistical analysis

The resulting 18,204 probe sets from the filtering step was the starting point for all supervised analyses which, for instance, correlated gene expression with the clinical variables such as the normal or disease, non-responder or responder. Class comparison analysis was performed by using Significance Analysis of Microarray (SAM), integrated in OminiViz version 5.1.

SAM discovered differentially expressed genes among different sample classes, e.g. between non-cancerous tissues and tumours or among different histology subtypes. Firstly this algorithm calculated the different expression for each gene between classes relative to the variation expected in the mean difference. To correct multiple testing, false discovery rate (FDR) was controlled by randomly permutating the classes of samples 100 times. Signature probe sets for assigned classes were selected by a change factor of 2 and a FDR of less than 1 percent.

Acknowledgments

We acknowledge the patients and their families for their contribution to this study, Dr. Yuri Mochkine and Dr. George Garinis for helpful advice and some microarray analysis, Kirsten van Lom for HPLC analysis of hemoglobin subtypes, Teus van Gent for statistical analysis and Ortho-Biotech, Tilburg, The Netherlands for the donation of EPREX. The work was supported by the NIH, NGI, NWO and Landsteiner foundation.

References

Alebouyeh, M., Moussavi, F., Haddad-Deylami, H., Vossough, P. (2004). Hydroxyurea in the treatment of major beta-thalassemia and importance of genetic screening. Ann Hematol *83*, 430-433.

Asano, H., Li, X.S., and Stamatoyannopoulos, G. (1999). FKLF, a Novel Kruppel-Like Factor That Activates Human Embryonic and Fetal beta -Like Globin Genes. Mol Cell Biol *19*, 3571-3579.

Bakker, W.J., Blazquez-Domingo, M., Kolbus, A., Besooyen, J., Steinlein, P., Beug, H., Coffer, P.J., Lowenberg, B., von Lindern, M., and van Dijk, T.B. (2004). FoxO3a regulates erythroid differentiation and induces BTG1, an activator of protein arginine methyl transferase 1. J Cell Biol *164*, 175-184.

Bakker, W.J., van Dijk, T.B., Parren-van Amelsvoort, M., Kolbus, A., Yamamoto, K., Steinlein, P., Verhaak, R.G.W., Mak, T.W., Beug, H., Lowenberg, B*., et al.* (2007). Differential Regulation of Foxo3a Target Genes in Erythropoiesis. Mol Cell Biol *27*, 3839-3854.

Ballas, S., Marcolina, M., Dover, G., and Barton, F. (1999). Erythropoietic activity in patients with sickle cell anaemia before and after treatment with hydroxyurea. Br J Haematol *105*, 491-496.

Bank, A. (2006). Regulation of human fetal hemoglobin: new players, new complexities. Blood *107*, 435-443.

Beug, H., Palmieri, S., Freudenstein, C., Zentgraf, H., and Graf, T. (1982). Hormone-dependent terminal differentiation in vitro of chicken erythroleukemia cells transformed by ts mutants of avian erythroblastosis virus. Cell *28*, 907-919.

Blaybel, R., Theoleyre, O., Douablin, A., and Baklouti, F. (2008). Downregulation of the Spi-1/PU.1 oncogene induces the expression of TRIM10/HERF1, a key factor required for terminal erythroid cell differentiation and survival. Cell Res *18*, 834-845.

Bradai, M., Abad, M.T., Pissard, S., Lamraoui, F., Skopinski, L., and de Montalembert, M. (2003). Hydroxyurea can eliminate transfusion requirements in children with severe beta-thalassemia. Blood *102*, 1529-1530.

Budzowska, M., Jaspers, I., Essers, J., de Waard, H., van Drunen, E., Hanada, K., Beverloo, B., Hendriks, R.W., de Klein, A., Kanaar, R*., et al.* (2004). Mutation of the mouse Rad17 gene leads to embryonic lethality and reveals a role in DNA damage-dependent recombination. EMBO J *23*, 3548-3558.

Calatayud, S., Beltrán, B., Brines, J., Moncada, S., and Esplugues, J.V. (1998). Foetal erythrocytes exhibit an increased ability to scavenge for nitric oxide. European Journal of Pharmacology *347*, 363-366.

Chénais, B., Molle, I., and Jeannesson, P. (1999). Inhibitory effect of nitric oxide on chemically induced differentiation of human leukemic K562 cells. Biochemical Pharmacology *58*, 773-778.

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. Experimental Hematology *35*, 358-367.

Cokic, V.P., Andric, S.A., Stojilkovic, S.S., Noguchi, C.T., and Schechter, A.N. (2008). Hydroxyurea nitrosylates and activates soluble guanylyl cyclase in human erythroid cells. Blood *111*, 1117-1123.

Cokic, V.P., Beleslin-Cokic, B.B., Noguchi, C.T., and Schechter, A.N. (2007). Hydroxyurea increases eNOS protein levels through inhibition of proteasome activity. Nitric Oxide *16*, 371-378.

Cokic, V.P., Beleslin-Cokic, B.B., Tomic, M., Stojilkovic, S.S., Noguchi, C.T., and Schechter, A.N. (2006). Hydroxyurea induces the eNOS-cGMP pathway in endothelial cells. Blood *108*, 184-191.

Constantoulakis, P., Knitter, G., and Stamatoyannopoulos, G. (1989). On the induction of fetal hemoglobin by butyrates: in vivo and in vitro studies with sodium butyrate and comparison of combination treatments with 5-AzaC and AraC. Blood *74*, 1963-1971.

de Paula, E.V., Lima, C.S., Arruda, V.R., Alberto, F.L., Saad , S.T., and Costa, F.F. (2003). Long-term hydroxyurea therapy in beta-thalassaemia patients. Eur J Haematol *70*, 151-155.

DeSimone, J., Heller, P., Hall, L., Zwiers, D. (1982). 5-Azacytidine stimulates fetal hemoglobin synthesis in anemic baboons. . Proc Natl Acad Sci U S A *79*, 4428-4431.

Dixit, A., Chatterjee, T.C., Mishra, P., Choudhry, D.R., Mahapatra, M., Tyagi, S., Kabra, M., Saxena, R., and Choudhry, V.P. (2005). Hydroxyurea in thalassemia intermedia—a promising therapy. Annals of Hematology *84*, 441-446.

Døsen-Dahl, G., Munthe, E., Nygren, M., K, Stubberud, H., Hystad, M., E, and Rian, E. (2008). Bone marrow stroma cells regulate TIEG1 expression in acute lymphoblastic leukemia cells: Role of TGFbeta/BMP-6 and TIEG1 in chemotherapy escape. Int J Cancer *123*, 2759-2766.

Dover, G.J., Charache, S.H., Boyer, S.H., Talbot, C.C. Jr, Smith, K.D. (1983). 5-Azacytidine increases fetal hemoglobin production in a patient with sickle cell disease. . Prog Clin Biol Res *134*, 475-488.

Durante, W., Johnson, F., K., and Johnson, R., A. (2007). ARGINASE: A CRITICAL REGULATOR OF NITRIC OXIDE SYNTHESIS AND VASCULAR FUNCTION. Clin Exp Pharmacol Physiol *34*, 906-911.

Fibach, E., Prasanna, P., Rodgers, G.P., and Samid, D. (1993). Enhanced fetal hemoglobin production by phenylacetate and 4- phenylbutyrate in erythroid precursors derived from normal donors and patients with sickle cell anemia and beta-thalassemia. Blood *82*, 2203-2209.

Gabbianelli, M., Morsilli, O., Massa, A., Pasquini, L., Cianciulli, P., Testa, U., and Peschle, C. (2008). Effective erythropoiesis and HbF reactivation induced by kit ligand in -thalassemia. Blood *111*, 421-429.

Gumruk, F., Tavil, B., Balta, G., Unal, S., and Gurgey, A. (2008). Significance of Fetal Hemoglobin Values in Detection of Heterozygotes in Fanconi Anemia: Reevaluation of Fetal Hemoglobin Values by a Sensitive Method. J Pediatr Hematol Oncol *30*, 896-899 810.1097/MPH.1090b1013e3181833835.

Hajjar, F.M., and Pearson, H.A. (1994). Pharmacologic treatment of thalassemia intermedia with hydroxyurea. The Journal of Pediatrics *125*, 490-492.

Hoppe, C., Vichinsky, E., Quirolo, K., Van Warmerdam, J., Allen, K., and Styles, L. (2000). Use of Hydroxyurea in Children Ages 2 to 5 Years With Sickle Cell Disease. J Pediatr Hematol Oncol *22*, 330-334.

Jison, M.L., Munson, P.J., Barb, J.J., Suffredini, A.F., Talwar, S., Logun, C., Raghavachari, N., Beigel, J.H., Shelhamer, J.H., Danner, R.L.*, et al.* (2004). Blood mononuclear cell gene expression profiles characterize the oxidant, hemolytic, and inflammatory stress of sickle cell disease. Blood *104*, 270-280.

Karimi, M., Darzi, H., and Yavarian, M. (2005). Hematologic and Clinical Responses of Thalassemia Intermedia Patients to Hydroxyurea During 6 Years of Therapy in Iran. J Pediatr Hematol Oncol *27*, 380-385.

Kucukkaya, B., Arslan, D.O., and Kan, B. (2006). Role of G proteins and ERK activation in hemin-induced erythroid differentiation of K562 cells. Life Sciences *78*, 1217-1224.

Leberbauer, C., Boulme, F., Unfried, G., Huber, J., Beug, H., and Mullner, E.W. (2005). Different steroids co-regulate long-term expansion versus terminal differentiation in primary human erythroid progenitors. Blood *105*, 85-94.

Ley , T.J., Anagnou, N.P., Noguchi, C.T., Schechter, A.N., DeSimone, J., Heller, P., Nienhuis, A.W. (1983). DNA methylation and globin gene expression in patients treated with 5-azacytidine. . Prog Clin Biol Res *134*, 457-474.

Ley, T.J., DeSimone, J., Anagnou, N.P., Keller, G.H., Humphries, R.K., Turne,r P.H., Young, N.S., Keller ,P., Nienhuis, A.W. (1982). 5-azacytidine selectively increases gamma-globin synthesis in a patient with beta thalassemia. N Engl J Med *307*, 1469-1475.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2-[Delta][Delta]CT Method. Methods *25*, 402-408.

Lou, T., Singh, M., Mackie, A., Li, W., and Pace, B. (2009). Hydroxyurea generates nitric oxide in human erythroid cells: mechanisms for gamma-globin gene activation. Exp Biol Med (Maywood) *234*, 1374-1382.

Loukopoulos, D., Voskaridou, E., Stamoulakatou, A., Papassotiriou, Y., Kalotychou, V., Loutradi, A., Cozma, G., Tsiarta, H.,, and Pavlides, N. (1998). Hydroxyurea Therapy in Thalassemia. Nucleic Acids Res *850*, 120-128.

Ma, Q., Wyszynski, D.F., Farrell, J.J., Kutlar, A., Farrer, L.A., Baldwin, C.T., and Steinberg, M.H. (2007). Fetal hemoglobin in sickle cell anemia: genetic determinants of response to hydroxyurea. Pharmacogenomics J *7*, 386-394.

Maciejewski, J.P., Selleri, C., Sato, T., Cho, H.J., Keefer, L.K., Nathan, C.F., and Young, N.S. (1995). Nitric oxide suppression of human hematopoiesis in vitro. Contribution to inhibitory action of interferon-gamma and tumor necrosis factor-alpha. The Journal of Clinical Investigation *96*, 1085-1092.

Miniero, R., David, O., Saglio, G., Paschero, C., Nicola, P. (1981). The Hbf in Fanconi's anemia Pediatr Med Chir *3*, 167-170.

Moi, P., and Kan, Y.W. (1990). Synergistic enhancement of globin gene expression by activator protein-1-like proteins. Proc Natl Acad Sci U S A *87*, 9000-9004.

Moitreyee, C.-K., Suraksha, A., and Swarup, A.S. (1998). Potential role of NF-kB and RXR beta like proteins in interferon induced HLA class I and beta globin gene transcription in K562 erythroleukaemia cells. Molecular and Cellular Biochemistry *178*, 103-112.

Motoyama, N., Kimura, T., Takahashi, T., Watanabe, T., and Nakano, T. (1999). bcl-x Prevents Apoptotic Cell Death of Both Primitive and Definitive Erythrocytes at the End of Maturation. J Exp Med *189*, 1691-1698.

Nagai, T., Tarumoto, T., Miyoshi, T., Ohmine, K., Muroi, K., Komatsu, N., Sassa, S., and Ozawa, K. (2003). Oxidative stress is involved in hydroxyurea-induced erythroid differentiation. Br J Haematol *121*, 657-661.

Najmabadi, H.,Karimi-Nejad, R., Sahebjam, S., Pourfarzad, F., Teimourian, S., Sahebjam, F., Amirizadeh, N. and Karimi-Nejad, M.H. (2001). The beta-thalassemia mutation spectrum in the Iranian population. hemoglobin *25(3)*, 285-296.

Neishabury, M., Azarkeivan, A., Oberkanins, C., Esteghamat, F., Amirizadeh, N., and Najmabadi, H. (2008). Molecular Mechanisms Underlying Thalassemia Intermedia in Iran. Genetic Testing *12*, 549-556.

Panigrahi, I., Dixit, A., Arora, S., Kabra, M., Mahapatra, M., Choudhry, V.P., and Saxena, R. (2005). Do alpha deletions influence hydroxyurea response in thalassemia intermedia? Hematology *10*, 61-63.

Park, J.-I., Choi, H.-S., Jeong, J.-S., Han, J.-Y., and Kim, I.-H. (2001). Involvement of p38 Kinase in Hydroxyurea-induced Differentiation of K562 Cells. Cell growth & Differentiation *12*, 481-486.

Perrine, S.P., Miller, B.A., Faller, D.V., Cohen, R.A., Vichinsky, E.P., Hurst, D., Lubin, B.H., and Papayannopoulou, T. (1989). Sodium butyrate enhances fetal globin gene expression in erythroid progenitors of patients with Hb SS and beta thalassemia. Blood *74*, 454-459.

Perrine, S.P., Miller, B.A., Greene, M.F., Cohen, R.A., Cook, N., Shackleton, C., and Faller, D.V. (1987). Butryic acid analogues augment gamma globin gene expression in neonatal erythroid progenitors. Biochemical and Biophysical Research Communications *148*, 694-700.

Poillon, W.N., Kim, B.C., Rodgers, G.P., Noguchi, C.T., Schechter, A.N. (1993). Sparing effect of hemoglobin F and hemoglobin A2 on the polymerization of hemoglobin S at physiologic ligand saturations. Proc Natl Acad Sci U S A *90*, 5039-5043.

Ranuncolo, S.M., Polo, J.M., Dierov, J., Singer, M., Kuo, T., Greally, J., Green, R., Carroll, M., and Melnick, A. (2007). Bcl-6 mediates the germinal center B cell phenotype and lymphomagenesis through transcriptional repression of the DNA-damage sensor ATR. Nat Immunol *8*, 705-714.

Rodrigue, C.M., Arous, N., Bachir, D., Smith-Ravin, J., Romeo, P.H., Galacteros, F., and Garel, M.C. (2001). Resveratrol, a natural dietary phytoalexin, possesses similar properties to hydroxyurea towards erythroid differentiation. Br J Haematol *113*, 500-507.

Scicchitano, M.S., McFarland, D.C., Tierney, L.A., Narayanan, P.K., and Schwartz, L.W. (2003). In vitro expansion of human cord blood CD36+ erythroid progenitors: Temporal changes in gene and protein expression. Experimental Hematology *31*, 760-769.

Silva, M., Grillot, D., Benito, A., Richard, C., Nunez, G., and Fernandez-Luna, J.L. (1996). Erythropoietin can promote erythroid progenitor survival by repressing apoptosis through Bcl-XL and Bcl-2. Blood *88*, 1576-1582.

Sripichai, O., Kiefer, C.M., Bhanu, N.V., Tanno, T., Noh, S.-J., Goh, S.-H., Russell, J.E., Rognerud, C.L., Ou, C.-N., Oneal, P.A.*, et al.* (2009). Cytokine-mediated increases in fetal hemoglobin are associated with globin gene histone modification and transcription factor reprogramming. Blood *114*, 2299-2306.

Stamatoyannopoulos, G., and Grosveld, F. (2001). Hemoglobin switching. In The molecular basis of blood diseases (3rd ed), G. Stamatoyannopoulos, Majeurs P., Perlmutter, P., Varmus H., ed. (Philadelphia, W.B. Saunders), pp. 135-182.

Steinberg, M.H., Lu, Z.-H., Barton, F.B., Terrin, M.L., Charache, S., Dover, G.J., and the Multicenter Study of, H. (1997). Fetal Hemoglobin in Sickle Cell Anemia: Determinants of Response to Hydroxyurea. Blood *89*, 1078-1088.

Steinmann, S., Schulte, K., Beck, K., Chachra, S., Bujnicki, T., and Klempnauer, K.H. (2009). v-Myc inhibits C/EBP[beta] activity by preventing C/EBP[beta]-induced phosphorylation of the co-activator p300. Oncogene *28*, 2446-2455.

Sutton, M., Bouhassira, E.E., and Nagel, R.L. (1989). Polymerase chain reaction amplification applied to the determination of beta-like globin gene cluster haplotypes. Am J Hematol *32*, 66-69.

Szekeres, T., Fritzer-Szekeres, M., Elford, H.L., and Jayaram, H.M. (1997). The Enzyme Ribonucleotide Reductase: Target for Antitumor and Anti-HIV Therapy. Crit Rev Clin Lab Sci *34*, 503-528.

Veith, R., Galanello, R., Papayannopoulou, T., and Stamatoyannopoulos, G. (1985). Stimulation of F-cell production in patients with sickle-cell anemia treated with cytarabine or hydroxyurea. N Engl J Med *313*, 1571-1575.

Verma, I.C., Kleanthous, M., Saxena, R., Fucharoen, S., Winichagoon, P., Raizuddin, S., Khan, S.N., Akbari, M.T., Izadyar, M., Kotea, N.*, et al.* (2007). Multicenter Study of the Molecular Basis of Thalassemia Intermedia in Different Ethnic Populations. Hemoglobin *31*, 439-452.

von Lindern, M., Zauner, W., Mellitzer, G., Steinlein, P., Fritsch, G., Huber, K., Lowenberg, B., and Beug, H. (1999). The Glucocorticoid Receptor Cooperates With the Erythropoietin Receptor and c-Kit to Enhance and Sustain Proliferation of Erythroid Progenitors In Vitro. Blood *94*, 550-559.

Weatherall, D.J., Clegg, J.B., Higgs, D.R., Wood, W.G. (2001). The hemoglobinopathies. In The Metabolic and Molecular Bases of Inherited Disease, C.R. Scriver, ed. (New York, NY, McGraw-Hill), pp. 4571–4636.

Yang, Y., M., Pace, B., Kitchens, D., Ballas, S., K., Shah, A., and Baliga, B.S. (1997). BFU-E colony growth in response to hydroxyurea: Correlation between in vitro and in vivo fetal hemoglobin induction. Am J Hematol *56*, 252-258.

Yavarian, M., Karimi, M., Bakker, E., Harteveld, C.L., and Giordano, P.C. (2004). Response to hydroxyurea treatment in Iranian transfusion-dependent beta-thalassemia patients. hematologica *89*, 1172-1178.

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.

Yoshida, H., and Kitabayashi, I. (2008). Chromatin regulation by AML1 complex. International Journal of Hematology *87*, 19-24.

Zeng, Y.T., Huang, S.Z., Re,n Z.R., Lu, .Z.H., Zeng, F.Y., Schechter, A.N., Rodgers, G.P. (1995). Hydroxyurea therapy in beta-thalassaemia intermedia: improvement in haematological parameters due to enhanced beta-globin synthesis. Br J Haematol *90*, 557-563.

Zhang, P., Basu, P., Redmond, L.C., Morris, P.E., Rupon, J.W., Ginder, G.D., and Lloyd, J.A. (2005). A functional screen for Krüppel-like factors that regulate the human gamma-globin gene through the CACCC promoter element. Blood Cells, Molecules, and Diseases *35*, 227-235.

Chapter 5

Increased γ-globin gene expression in β-thalassemia intermedia patients correlates with a mutation in 3'HS1

Adamantia Papachatzopoulou*, Polynikis Kaimakis*, **Farzin Pourfarzad**, Panagiotis G. Menounos, Panagiota Evangelakou, Panagoula Kollia, Frank Grosveld, George Patrinos.

*These authors contributed equally to this work.

Am J Hematol. 2007 Nov;82(11):1005-9.

Increased γ-globin gene expression in β-thalassemia intermedia patients correlates with a mutation in 3′HS1

Adamantia Papachatzopoulou,[1] Polynikis Kaimakis,[2] Farzin Pourfarzad,[2] Panagiotis G. Menounos,[3] Panagiota Evangelakou,[1] Panagoula Kollia,[4] Frank G. Grosveld,[2] and George P. Patrinos[2]*

[1] Department of General Biology, School of Medicine, University of Patras, Patras, Greece
[2] Faculty of Medicine and Health Sciences, MGC-Department of Cell Biology and Genetics, Erasmus MC, Rotterdam, The Netherlands
[3] Laboratory of Research, Nursing Military Academy, Athens, Greece
[4] Department of Biology, School of Medicine, University of Thessaly, Larissa, Greece

We report a novel set of genetic markers in the DNaseI hypersensitive sites comprising the human β-globin locus chromatin hub (CH), namely HS-111 and 3′HS1. The HS-111 (−21 G>A) and 3′HS1 (+179 C>T) transitions form CH haplotypes, which occur at different frequencies in β-thalassemia intermedia and major patients and normal (nonthalassemic) individuals. We also show that the 3′HS1 (+179 C>T) variation results in a GATA-1 binding site and correlates with increased fetal hemoglobin production in β-thalassemia intermedia patients. In contrast, the HS-111 (+126 G>A) transition, found in three normal chromosomes, is simply a rare polymorphism. We conclude that the CH haplotypes are useful genetic determinants for β-thalassemia major and intermedia patients, while the 3′HS1 (+179 C>T) mutation may have functional consequences in γ-globin genes expression. Am. J. Hematol. 82:1005–1009, 2007. © 2007 Wiley-Liss, Inc.

Introduction

The human β-globin locus contains five functional gene copies, arranged in their developmental order of expression (Fig. 1A). The most important control element is the locus control region (LCR), consisting of five DNaseI hypersensitive sites (HS1–5) [1]. Based on chromosome configuration studies in the mouse and human β-globin loci [2–4], HS5 of the LCR is thought to interact with two 5′ and 3′ distant HSs (HS-111 and 3′HS1) to form a chromatin hub (CH) through looping. Upon differentiation, the remainder of the LCR and the genes interact with the CH to form an active CH (ACH) [3,4]. Some β-thalassemia patients lack the LCR (but retaining at least one of the genes and HS-111), resulting in transcriptionally silent β-globin genes in an inactive chromatin configuration [5]. In contrast, a number of patients have large deletions at the 3′ end of the locus with or without deletion of 3′HS1. The deletions, leaving 3′HS1 but deleting the β-globin gene, result in β-thalassemia. Deletions, including 3′HS1, result in an increase of γ-globin transcription and are classified as hereditary persistence of fetal hemoglobin (HPFH). The common interpretation of this result is that the deletions including 3′HS1 result in juxtaposition of enhancers next to the γ-globin genes. However, it cannot be excluded that removing only 3′HS1 would be sufficient for the phenotype. 3′HS1 binds transcription factor CTCF, suggesting that 3′HS1 may exclude downstream enhancer interactions by forming the CH. Removal of CTCF binding to the mouse 3′HS1 does not change β-globin gene expression. However, the mouse β-globin locus does not contain a separate fetal-type globin gene that is only suppressed after birth. We therefore set out to determine whether 3′HS1 mutations may occur in patients with β-thalassemia intermedia who have increased γ-globin gene expression but no deletions in the 3′ end of the β-globin locus.

There are three types of β-thalassemia intermedia patients, those with mild mutations, i.e. β −101 C>T, those with coexisting α-thalassemia or HPFH, and those with high fetal hemoglobin (Hb F) levels in the absence of HPFH [6–9]. Genetic studies have shown that a number of the latter cases can be explained by sequence variations in cis, namely in the γ-globin gene proximal regulatory regions [10–13] or the LCR [14] or the Aγ-δ-globin intergenic region [15]. However, the phenotype of a number of patients cannot be explained by such sequence variations, and hence we looked in these patients whether we could find sequence variations in the distant HS-111 and, particularly, 3′HS1.

Results

Identification of CH haplotypes

Hematological and clinical data or our β-thalassemia intermedia patients are shown in Table I. The possible coexistence of α-thalassemia and α-globin genes triplication was investigated in all β-thalassemia intermedia patients by Southern blot analysis and by a PCR/restriction fragment length polymorphism analysis for the detection of α-globin gene point mutations, leading to α-thalassemia in the Greek population [16]. The results were negative in all cases. Also, the γ-globin genes promoter was HPFH mutation-free and the LCR contained no regulatory mutations in any of the patients. We therefore concluded that these β-thalassemia intermedia patients might be good candidates for possible mutations of the distant HSs.

Contract grant sponsor: Research Promotion Foundation, Cyprus; Contract grant number: ΠΔΕ06-02; Contract grant sponsor: NIH; Contract grant number: 5RO1HL073455-02; Contract grant sponsor: NWO, The Netherlands; Contract grant numbers: 912-02-054, 912-02-055.

*Correspondence to: Faculty of Medicine and Health Sciences, MGC-Department of Cell Biology and Genetics, Erasmus MC, PO Box 2040, 3000 CA, Rotterdam, The Netherlands. E-mail: g.patrinos@erasmusmc.nl

A. Papachatzopoulou and P. Kaimakis contirbuted equally to this work.

Received for publication 3 January 2007; Revised 15 March 2007; Accepted 4 April 2007

Am. J. Hematol. 82:1005–1009, 2007.

Published online 24 July 2007 in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/ajh.20979

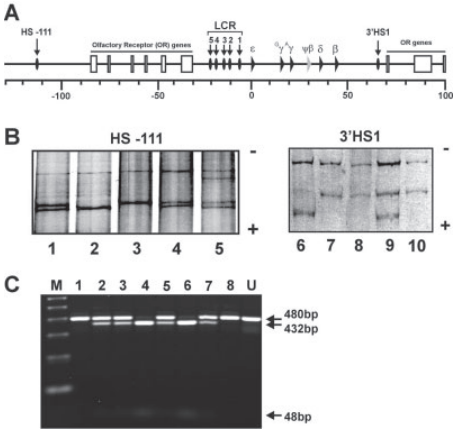


Figure 1. A. Schematic drawing of the human β-globin locus (not to scale), with the two distant HS, -111 and 3′HS1, forming the β-globin CH. B. Segments of the SSCP gel showing the different electrophoretic patterns corresponding to the nucleotide changes identified in HS-111 (−21 G>A and +126 G>A) and 3′ HS1 (+179 C>T) as indicated underneath. Lanes 1, 4: HS-111 (−21 G/A) heterozygote, Lane 2: HS-111 (−21 A/A) homozygote, Lane 3: HS-111 (−21 G/G) homozygote, Lane 5: HS-111 (−21 G/A) and HS-111 (+126 G>A) compound heterozygote, Lane 6: 3′ HS1 (+179 T/T) homozygote, Lanes 7, 8, 10: 3′ HS1 (+179 C/C) homozygotes, Lane 9: 3′ HS1 (+179 C/T) heterozygote. C. AlwNI restriction endonuclease analysis as a means of identification of the HS-111 (−21 G>A) variation. Lanes 1, 8: HS-111 (−21 G/G) homozygotes, Lanes 2, 3, 5, 7: HS-111 (−21 G/A) heterozygotes, Lanes 4, 6: HS-111 (−21 A/A) homozygotes, U: Undigested negative control, M: 123-bp size marker (New England Biolabs, Palo Alto, CA).

Single-strand conformation polymorphism (SSCP) analysis of PCR-amplified DNA from β-thalassemia intermedia patients indicated variant electrophoretic patterns in HS-111 and 3′HS1, encompassing the human β-globin locus (Fig. 1B), while no aberrant electrophoretic pattern was observed in PCR-amplified fragments containing HS5 (not shown). DNA sequence analysis revealed the presence of two novel sequence variations in the HS-111 (−21 G>A) and 3′HS1 (+179 C>T), respectively (+1: first base of HS core), which create four independent haplotypes, termed hereafter as CH haplotypes (Fig. 2A). HS-111 (−21 G>A) transition creates a new recognition site for AlwNI restriction endonuclease, allowing for its rapid detection in a larger population sample (Fig. 1C). By means of AlwNI restriction endonuclease and DNA sequence analysis, we also screened our β-thalassemia major patients and normal individuals, and established the distribution of the four different CH haplotypes (Fig. 2B) and allelic frequencies in our three subject groups. Calculation of allelic frequencies in the β-thalassemia major and intermedia patients and normal individuals indicated that the 3′HS1 (+179 T) variant allele was significantly more frequent in the β-thalassemia intermedia patient group (46%), compared with β-thalassemia major patients (15%) and normal subjects (16%; Fig. 2B). The HS-111 (−21 A) variant allele frequency in β-thalassemia intermedia patients (54.2%) was similar to that found in normal subjects (45.6%). Thus this association is not statistically significant for β-thalassemia intermedia. Interestingly, this variation was found less frequently in the β-thalassemia major patients (27.5%) and may be useful as a β-thalassemia major marker.

HS-111 and 3′HS1 cores have been shown to harbor transcription factor binding sites [18], and we therefore explored the possibility that the allelic changes described earlier affect protein binding. EMSA analysis showed that neither the HS-111 (−21 G) nor the HS-111 (−21 A) oligonucleotides have any protein binding properties in vitro (not shown). This was not unexpected, since the variant base lies outside the HS-111 core and does not contain any predicted transcription factor binding sites (http://www.biobase-international.com/pages/index.php?id=40). In contrast, the 3′HS1 (+179 T) sequence variation creates a putative GATA-1 binding site and indeed displayed a specific mobility shift in EMSA analysis (Fig. 3). Use of an anti-GATA-1 specific antibody resulted in a supershift of the protein/

TABLE I. Hematological, Clinical, and Molecular Data of the β-Thalassemia Intermedia Patients

Hematological and clinical data								Molecular data		
Sex/age	Hb (g/dL)	MCV (fL)	MCH (pg)	Hb A_2 (%)	Hb F (%)	Non-α/α[a]	Transfusion	Genotype	XmnI[b]	CH Haplotype
M/24	8.0	63.4	18.39	6.5	8.8	0.22	1–2/year	IVSI-1 (G>A)/Term +6 (C>G)	−/−	II/IV
M/36	7.5	81.3	21.2	4.3	50	0.38	2/year	IVSI-6 (T>C)/IVSI-110 (G>A)	−/−	I/IV (or II/III[c])
M/37	11.2	85.5	25.8	2.1	97.9	0.41	None	Cd39 (C>T)/FSC8 (ΔAA)	−/+	I/IV (or II/III[c])
M/45	8.3	80	23.7	2.7	30.0	0.38	2/year	IVSI-6 (T>C)/IVSI-6 (T>C)	−/+	II/IV
M/46	8.4	86.7	24.2	3.1	78.8	0.50	1 in 5 years	IVSI-6 (T>C)/Cd39 (C>T)	−/−	II/IV
M/56	8.2	80.7	24.7	1.8	93	0.39	1/year	IVSI-110 (G>A)/Cd39 (C>T)	+/+	I/I
F/33	9.1	75.6	23.2	3.4	64.6	0.39	None	IVSI-6 (T>C)/IVSII-1 (G>A)	−/+	II/IV
F/51	7.5	79.8	23.7	3.0	25.0	0.34	Occasionally	IVSI-110 (G>A)/Cd39 (C>T)	−/+	I/I
F/52	9.9	73.8	27	2.3	80.0	0.39	Occasionally	IVSI-6 (T>C)/Cd39 (C>T)	−/+	I/IV (or II/III[c])
F/70	7.7	95.6	25.8	5.0	22.0	0.33	2–3/year	IVSI-6 (T>C)/IVSI-6 (T>C)	−/−	III/IV
F/80	9.0	70	20	4.7	5.8	0.29	None	Cd39 (C>T)/+10 (ΔT)	−/−	I/II

[a]Biosynthetic ratios.
[b]G_γ-158 (C/T).
[c]Depending on whether the HS-111 (−21 A) and 3′HS1 (+179 T) variations are found or not to be linked in cis.

A

CH haplotypes	HS -111 -21 G/A	3'HS1 +179 C/T
I	G	C
II	A	C
III	G	T
IV	A	T

B

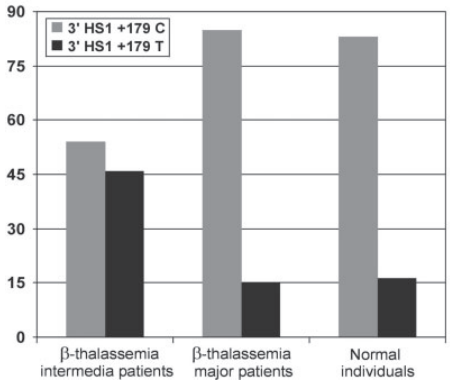


Figure 2. A. Depiction of the four CH haplotypes, determined by the HS-111 (−21 G/A) and 3′HS1 (+179 C/T) sequence variations. The coexistence of the variant bases in CH haplotype IV in cis is more likely to result from genetic recombination between haplotypes II and III, rather than to represent independent mutation events. In favor of this assumption is the existence of a 9-kb recombination hotspot, 5′ to the human β-globin gene [17]. B. Histogram showing the distribution of the 3′HS1 (+179 C/T) allelic frequencies in the three study groups, namely β-thalassemia intermedia, β-thalassemia major patients, and normal (nonthalassemic) individuals.

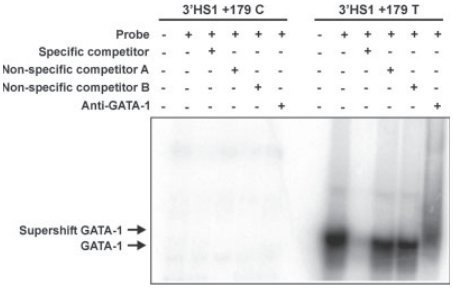


Figure 3. Electrophoretic mobility shift assays of the oligonucleotides containing either the 3′HS1 wild type (+179 C) or variant (+179 T) sequence using nuclear protein extracts prepared from DMSO-induced MEL cells. The oligonucleotide with the mutant 3′HS1 (+179 T) sequence contains an inverted GATA-1 binding site and yields a band shift identical to the reference (β-globin GATA-1) oligonucleotide (not shown). Competition with nonlabeled (specific) oligonucleotide results in the disappearance of the band shift, while competition with nonspecific [namely c-Myc (A) and EKLF (B)] results in restoration of the shifts at the previous position. Use of a GATA-1 antibody results in a supershifted complex at the expected, for the GATA-1, position. Use of nuclear protein extracts prepared from uninduced MEL cells yielded identical electrophoretic mobility shifts (not shown).

oligonucleotide complex confirming that a GATA-1 binding site has been generated.

A novel HS-111 (+126 G>A) variation is a rare polymorphism

An additional G>A transition was also identified 126 bp downstream of the starting point of the HS-111 core region (Fig. 1B) in three normal (nonthalassemic) chromosomes of Hellenic origin. EMSA analysis, using nuclear protein extracts from noninduced and induced MEL cells and both wild-type and variant probes at the latter position, revealed no mobility shifts (data not shown). These data, along with the fact that the HS-111 (+126) G>A variation is neither located in a phylogenetically conserved sequence nor on a known transcription factor binding site [18] strongly suggest that the latter substitution is a rare polymorphism without functional consequences.

Discussion

In this paper, we describe a new set of genetic markers in the human β-globin locus CH DNaseI HS-111 and 3′HS1. A number of association studies have been performed to correlate the mild symptoms of β-thalassemia intermedia patients with possible genetic determinants in the human β-globin locus [10–15]. Those studies indicated that several nucleotide variations, such as the Gγ −158 (C>T) polymorphism and the Aγ- and intergenic Aγ-δ haplotypes, are sufficient but not necessary for high Hb F production. This study adds a novel set of genetic markers for β-thalassemia, i.e., the CH haplotypes. Based on our EMSA analysis, CH haplotypes III and IV, characterized from the presence of the 3′HS1 (+179 T) variation, harbor an additional GATA-1 binding site inside 3′HS1 and are strongly associated with β-thalassemia intermedia, hitherto not observed even for the Gγ −158 C>T polymorphism. Haplotype analysis failed to identify a predominant haplotype linked either to the HS-111 (−21 G>A) or the 3′HS1 (+179 C>T) polymorphisms, although haplotype VII to be more frequently found linked with the latter polymorphism.

More than half of our β-thalassemia intermedia patients are homozygotes or compound heterozygotes for "mild" β-thalassemia mutations, which could potentially explain their mild symptoms. In particular, six patients carry the IVS I-6 (T>C) β-thalassemia mutation, occasionally leading to mild phenotype [9], in homozygous or compound heterozygous state, while two others are compound heterozygous for the +10 (ΔT) and Term +6 (C>G) mutations, respectively [19,20]. However, the mild symptoms observed in our β-thalassemia intermedia patients cannot be solely attributed to this fact, as there are many IVS I-6 T>C homozygote or compound heterozygote patients reported to be regularly transfused and present with β-thalassemia major symptoms [9,21,22 and unpublished observations]. It is noteworthy that in these patients, as well as in vast majority of β-thalassemia patients, the chromosomal background of the β-thalassemia mutation seems to be one of the compelling

factors related to the phenotypic outcome, rather then the β-thalassemia mutation itself [21,22 and unpublished observations]. Also, the +10 (ΔT) and Term +6 (C>G) mutations have been shown to result in drastically decreased steady state mRNA levels to 61.6% and 52–60%, using expression studies in stably transfected MEL cells and nuclear run-on experiments [23,24]. Of interest are also those β-thalassemia intermedia patients, whose Hb F levels were moderately to substantially high (M/56 = 93%, F/51 = 25%, and F/80 = 5.8%; Table I), despite the absence of the 3′HS1 (+179 T) variation. This can be possibly explained by the presence either of the XmnI polymorphism, previously demonstrated to correlate with high Hb F levels [10] or other yet unknown cis-regulatory elements contributing to the overall mild clinical symptoms.

It has been shown that HS-111 and 3′HS1 cause a spatial restriction by forming a loop containing the genes and the LCR and initiate folding of the human β-globin locus using an enzyme/substrate-like principle (but on a larger scale) [4]. This folding is thought to enable proper and precise interaction of the LCR with the globin genes at different developmental stages. It is at present not clear what the effect of the 3′HS1 (+179 C>T) variation would be, because it generates a GATA-1 binding site rather than destroying a known transcription factor binding site. Possibly it interferes with the participation of 3′HS1 in the ACH, which may bring in close proximity downstream enhancers in analogy to what is observed in the deletional HPFH cases. However in these patients, a competing δ- and β-globin gene would still be present to dampen the activation of the γ-globin genes when compared with the HPFH conditions [1].

Alternatively, the novel GATA-1 binding site may give 3′HS1 enhancer-like properties leading to mild γ-globin gene activation. In favor of this hypothesis is the absence of the 3′HS1 (+179 C>T) variation in β-thalassemia major patients with similar or even identical β-thalassemia genotype. However, the lack of γ-globin gene activation in normal individuals and β-thalassemia major patients with the same haplotypes suggests that in those β-thalassemia intermedia patients there must be additional sequence variations in the locus or interaction(s) with yet to be discovered or previously described modifier gene(s) [25–27], leading to the observed phenotype.

In essence, the HS-111 (−21 G>A) and 3′HS1 (+179 C>T) sequence variations represent useful genetic markers for β-thalassemia major and intermedia patients, respectively, while the 3′HS1 (+179 C>T) mutation may play a functional role in human γ-globin gene activation.

Materials and Methods

Subjects

Peripheral blood was collected with informed consent from 11 and 25 unrelated adult β-thalassemia intermedia and major patients, respectively, and 68 unrelated normal (nonthalassemic) male and female individuals of Hellenic origin, whose age ranged between 25 and 80 years. Hematological indices were measured with an automated coulter counter.

DNA analysis

Genomic DNA was extracted from peripheral blood leucocytes. Determination of human β-globin gene mutations and haplotype analysis was performed as previously described [12,28]. DNaseI HS-111, HS5, and 3′HS1, all of which form human β-globin locus CH, were amplified in three independent fragments, respectively (Fig. 1A), using 2.5 U of Taq DNA Polymerase (Invitrogen, Carlsbad, CA), 200 μM/dNTP, and 25 pmoles of each primer as follows: HS111-FW (5′-GCTTGGTGAAGTAGGAGATTC-3′), HS111-RV (5′-GAGAACCCTGTGAGTAAGGA-3′), HS5-FW (5′-ACCTTCTGGTTCACTCTGC-3′), HS5-RV (5′-GGAAACTGGAGAAACTGGGA-3′), 3′HS1-FW (5′-GCCTACTTCAGGTTTGTGG-3′), and 3′HS1-RV (5′-ACATTCCTATTTGCCAAGG-3′). PCR conditions were initial denaturation at 95°C for 3 min, followed by 30 cycles of 1-min denaturation at 95°C, 45-sec annealing at 55°C (fragments HS111 and 3′HS1) or 56°C (fragment HS5), and 45-sec elongation at 72°C. PCR products were then analyzed, using a nonradioactive (silver-staining) SSCP analysis, as previously described [29]. Restriction endonuclease analysis was performed at 37°C, using 300 ng of PCR-amplified DNA and 1 U of AlwNI restriction endonuclease (NEB, Palo Alto, CA). Finally, DNA sequence analysis was performed using automated fluorescent DNA sequencer (ABI PRISM® 310 Genetic Analyzer, Applied Biosystems, CA).

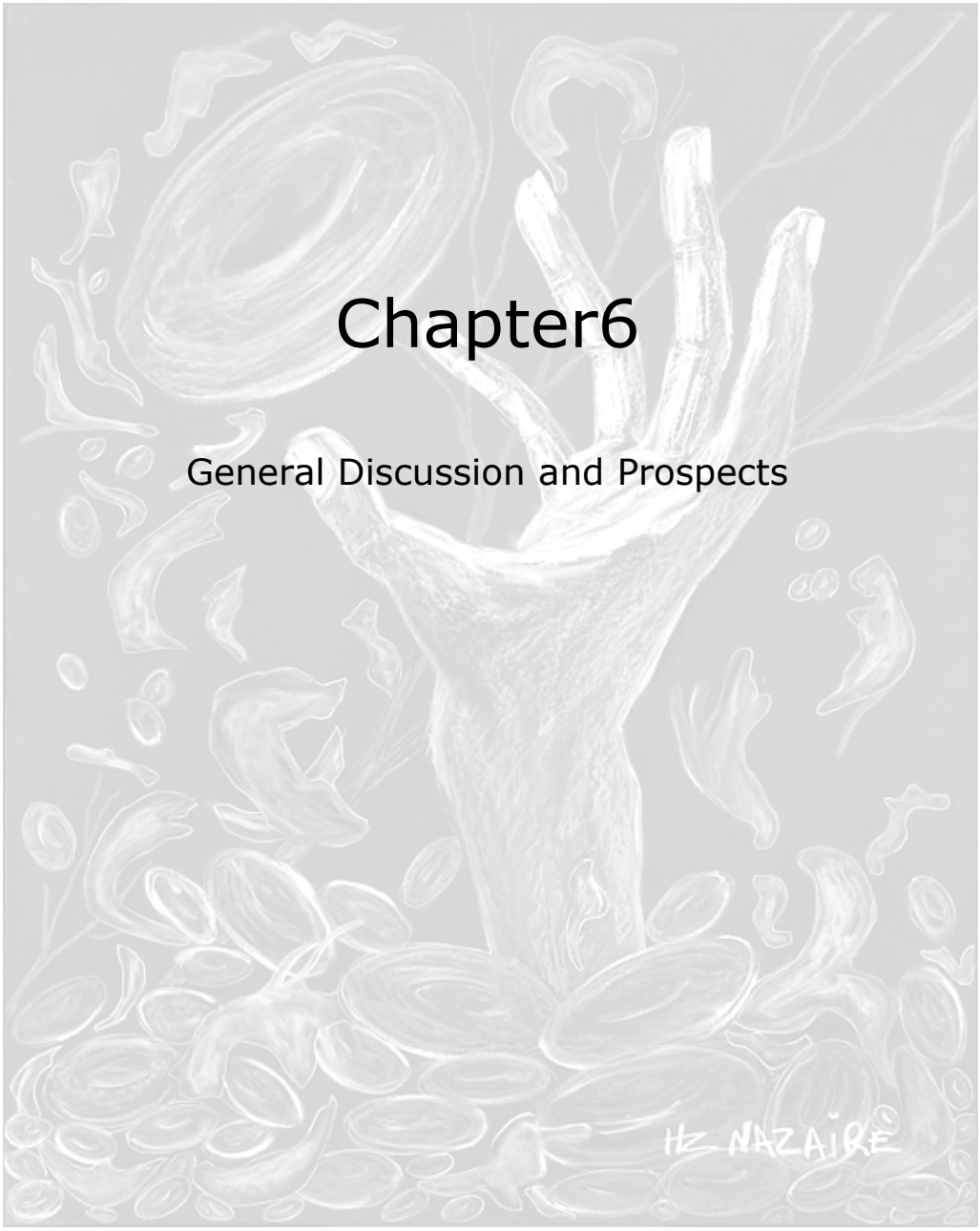
Electrophoretic mobility shift assays

Electrophoretic mobility shift assay (EMSA) analysis was performed using total nuclear protein extracts from MEL cells [30]. Five micrograms of nuclear extracts were used per reaction, and competitions were done using 100-fold molar excess of the indicated double-stranded oligonucleotides before addition of the nuclear extract. Super shift analysis was performed using a home-brew GATA-1 antibody. Nucleotide sequence of the oligonucleotides used in these studies is available upon request.

References

1. Stamatoyannopoulos G, Grosveld F. Hemoglobin switching. In: Stamatoyannopoulos G, Majerus PW, Perlmutter RM, Varmus H, editors. The Molecular Basis of Blood Diseases, 3rd ed. Philadelphia, PA: WB Saunders Company; 2001. pp 135–182.
2. Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W. Looping and interaction between hypersensitive sites in the active β-globin locus. Mol Cell 2002;10:1453–1465.
3. Palstra RJ, Tolhuis B, Splinter E, Nijmeijer R, Grosveld F, de Laat W. The β-globin nuclear compartment in development and erythroid differentiation. Nat Genet 2003;35:190–194.
4. Patrinos GP, de Krom M, de Boer E, et al. Multiple interactions between regulatory regions are required to stabilize an active chromatin hub. Genes Dev 2004;18:1495–1509.
5. Kioussis D, Vanin E, de Lange T, Flavell RA, Grosveld FG. β-Globin gene inactivation by DNA translocation in γβ-thalassaemia. Nature 1983;306:662–666.
6. Weatherall DJ, Clegg JB, Higgs DR, Wood WG. The hemoglobinopathies. In: Scriver CR, editor. The Metabolic and Molecular Bases of Inherited Disease. New York, NY: McGraw-Hill; 2001. pp 4571–4636.
7. Hardison RC, Chui DH, Giardine B, et al. HbVar: A relational database of human hemoglobin variants and thalassemia mutations at the globin gene server. Hum Mutat 2002;19:225–233.
8. Patrinos GP, Giardine B, Riemer C, et al. Improvements in the HbVar database of human hemoglobin variants and thalassemia mutations for population and sequence variation studies. Nucleic Acids Res 2004;32:D537–D541.
9. Cao A, Galanello R, Rosatelli C. Genotype-phenotype correlations in β-thalassaemias. Blood Rev 1994;8:1–12.
10. Gilman JG, Huisman TH. DNA sequence variation associated with elevated fetal G γ globin production. Blood 1985;66:783–787.
11. Lanclos KD, Oner C, Dimovski AJ, Gu YC, Huisman TH. Sequence variations in the 5′ flanking and IVS-II regions of the G γ- and A γ-globin genes of β S chromosomes with five different haplotypes. Blood 1991;77:2488–2496.
12. Patrinos GP, Kollia P, Papapanagiotou E, Loutradi-Anagnostou A, Loukopoulos D, Papadakis MN. Agamma-haplotypes: A new group of genetic markers for thalassemic mutations inside the 5′ regulatory region of the human Agamma-globin gene. Am J Hematol 2001;66:99–104.
13. Patrinos GP, Samperi P, lo Nigro L, Kollia P, Schiliro G, Papadakis MN. Evidence for the molecular heterogeneity of sickle cell anemia chromosomes bearing the βS/benin haplotype. Am J Hematol 2005;80:79–80.
14. Adekile AD, Dimovski AJ, Oner C, Lanclos KD, Huisman TH. Haplotype-specific sequence variations in the locus control region (5′ hypersensitive sites 2, 3, 4) of β S chromosomes. Hemoglobin 199317:475–478.
15. Papachatzopoulou A, Kourakli A, Markopoulou P, et al. Genotypic heterogeneity and correlation to intergenic haplotype within high HbF β-thalassemia intermedia. Eur J Haematol 2006;76:322–330.
16. Patrinos GP, van Baal S, Petersen MB, Papadakis MN. Hellenic National Mutation database: A prototype database for mutations leading to inherited disorders in the Hellenic population. Hum Mutat 2005;25:327–333.
17. Smith RA, Ho PJ, Clegg JB, Kidd JR, Thein SL. Recombination breakpoints in the human β-globin gene cluster. Blood 1998;92:4415–4421.
18. Bulger M, Schubeler D, Bender MA, et al. A complex chromatin landscape revealed by patterns of nuclease sensitivity and histone modification within the mouse β-globin locus. Mol Cell Biol 2003;23:5234–5244.
19. Athanassiadou A, Papachatzopoulou A, Zoumbos N, Maniatis GM, Gibbs R. A novel β-thalassaemia mutation in the 5′ untranslated region of the β-globin gene. Br J Haematol 1994;88:307–310.
20. Jankovic L, Dimovski AJ, Kollia P, Karageorga M, Loukopoulos D, Huisman TH. A C-G mutation at nt position 6 3′ to the terminating codon may be the cause of a silent β-thalassemia. Int J Hematol 1991;54:289–293.

21. Kattamis C, Hu H, Cheng G, Reese AL, et al. Molecular characterization of β-thalassaemia in 174 Greek patients with thalassaemia major. Br J Haematol 1990;74:342–346.
22. Ragusa A, Lombardo M, Beldjord C, et al. Genetic epidemiology of β-thalassemia in Sicily: Do sequences 5′ to the G γ gene and 5′ to the β gene interact to enhance HbF expression in β-thalassemia? Am J Hematol 1992;40: 199–206.
23. Sgourou A, Papachatzopoulou A, Psiouri L, et al. The β-globin C−>G mutation at 6 bp 3′ to the termination codon causes β-thalassaemia by decreasing the mRNA level. Br J Haematol 2002;118:671–676.
24. Sgourou A, Routledge S, Antoniou M, Papachatzopoulou A, Psiouri L, Athanassiadou A. Thalassaemia mutations within the 5′UTR of the human β-globin gene disrupt transcription. Br J Haematol 2004;124:828–835.
25. Miyoshi K, Kaneto Y, Kawai H, et al. X-Linked dominant control of F-cells in normal adult life: Characterization of the Swiss type as hereditary persistence of fetal hemoglobin regulated dominantly by gene(s) on X chromosome. Blood 1988;72:1854–1860.
26. Craig JE, Rochette J, Fisher CA, et al. Dissecting the loci controlling fetal haemoglobin production on chromosomes 11p and 6q by the regressive approach. Nat Genet 1996;12:58–64.
27. Dover GJ, Smith KD, Chang YC, et al. Fetal hemoglobin levels in sickle cell disease and normal individuals are partially controlled by an X-linked gene located at Xp22.2. Blood 1992;80:816–824.
28. Patrinos GP, Kollia P, Papadakis MN. Molecular diagnosis of inherited disorders: Lessons from hemoglobinopathies. Hum Mutat 2005;26:399–412.
29. Menounos P, Zervas C, Garinis G, et al. Molecular heterogeneity of the glucose-6-dehydrogenase deficiency in the Hellenic population. Hum Hered 2000;50:237–241.
30. Wall L, de Boer E, Grosveld F. The human β-globin gene 3′ enhancer contains multiple binding sites for an erythroid-specific protein. Genes Dev 1988;2:1089–1100.



Chapter6

General Discussion and Prospects

Isolation and identification of human γ-globin gene promoter suppressor proteins by Targeted Chromatin Purification (TCP)

β-thalassemia and Sickle Cell Disease (SCD) together comprise one of the most significant genetic health problems in the World. They are present in populations from Southeast Asia, Middle East, the Mediterranean regions and Africa (Mabaera et al., 2008). With the available treatments such as blood transfusion and hydroxyurea (HU) therapy, patients usually survive up to their third decade of life. While there are more advanced means of treatment under development (Bhatia and Walters, 2007; Shenoy, 2007); Cooley's Anemia Symposium 2010, New Nork), regardless of their efficiency and potential risks, they require extensive financial resources and sophisticated medical care, which is unavailable for the majority of patients in these regions. On the other hand, the current available treatments are risky, pleiotropic, reversible, and inconvenient combined with many side effects (reviewed in (Mabaera et al., 2008). Therefore, in addition to the traditional chelation methods of treatment, novel approaches are needed to develop more practical and effective treatment options.

The observations that increased fetal hemoglobin levels ameliorate the disease phenotype and reduces β-thalassemia and SCD complications led to the idea of γ-globin reactivation as a promising approach for the treatment of these diseases. Screening of small molecule libraries to identify γ-globin activators has unfortunately not led to the identification of new compounds that are more advantageous over the existing therapeutics (Haley et al., 2003; Mabaera et al., 2008). In fact, screening for small molecule activators may not result in the successful identification of effective candidates. γ-globin suppression is thought to involve protein-protein interactions. These types of interactions generally take place between large protein surfaces and small molecules are inherently unfavourable to inhibit such interactions (as e.g. opposed to enzyme active pockets or receptors). Hence, inhibitor design, or more specifically, peptide design may be required.

Understanding γ-globin regulation and its switching mechanism is the key for the development of such therapeutic approaches. Despite attempts made during last 30 years, the mechanism of γ-globin switching, or the silencing of γ-globin after birth, remains largely unknown. Possibilities include the recruitment of suppressor complexes or removal of activating factors, or a combination of both. During the last few years, different transcription factors and chromatin remodelers have been suggested to be involved in γ-globin switching. These include BCL11A (Sankaran et al., 2008; Sankaran et al., 2009), NF-E4 (Zhao et al., 2006), MBD2 (Rupon et al., 2006), TR2/TR4 (Tanabe et al., 2002) and IKAROS (O'Neill et al., 2000).

Although, all of these factors *per se* are presumably important for γ-globin gene repression, they have been largely identified via candidate approaches. It remains to be determined whether these factors form complexes with each other and how exactly they participate in the switching process. Do they co-operate with each other and/or other novel,

as yet unidentified, factors in γ-globin silencing? Are they in physical contact, as part of different heterogeneous complexes? Are they all necessary for γ-globin silencing at the same time? At what stage do they take part in the switching process?

To answer these questions we developed a novel method to isolate the suppressed γ-globin promoter chromatin in vivo, followed by proteomic analysis to identify previously known and novel components of the functional regulatory complexes on the γ-globin gene promoter in vivo, followed by functional validation.

Locus specific protein composition has been previously studied by different chromatin isolation strategies (Déjardin and Kingston, 2009; Griesenbeck et al., 2003; Jasinskas and Hamkalo, 1999). While each method could enrich the targeted region, none of them successfully yield sufficient amounts of single copy locus chromatin with enough purity to allow identification of the bound factors. Other methods including nucleic acid affinity capture (Olave et al., 2007) have also been used for purification of protein factors bound to specific DNA sequences. These strategies identify sequence-specific DNA-binding proteins using a ''bait'' DNA sequence outside of its endogenous context, or an in vitro capture approach. Although these studies have been useful, they do not provide information about protein composition of the chromatin *in vivo*.

Specific chromatin segments in yeast (Griesenbeck et al., 2003) and telomeres of mammalian cells have been isolated to be analyzed by Proteomics of Isolated Chromatin segments (PICh) (Déjardin and Kingston, 2009). PICh is the first direct approach that has been successfully carried out in mammalian cells, using a hybridization technique to identify factors bound to telomeres. Since each chromosome contains two telomeres, they constitute approximately 0.01-0.07% of the genome (Déjardin and Kingston, 2009), PICh requires less material for identification of the proteins by mass spectrometry. However a unique promoter sequence of approximately 1kb, as is the case for the γ-globin locus, our locus of interest, constitutes a much smaller percentage of the genome, and is therefore impossible to obtain sufficient quantity and purity using currently available techniques. Here we show for the first time that transcription factors can be directly isolated and identified by purification of the target promoter sequences (Targeted Chromatin Purification, TCP) *in vivo*. Identification of functional protein complexes by Mass Spectrometry has mainly relied on approaches using antibodies against candidate proteins and tagged proteins in cell lysates. The disadvantage of this approach is that one may be looking at heterogeneous complexes, which may be compositionally and functionally different at different target gene promoters. The TCP and PICh methods have the unbiased advantage of identifying unknown components of endogenous functional complexes on a specific DNA locus *in vivo*.

The TCP method has advantages and disadvantages over the PICh method. TCP involves the introduction of TetO repeat sequences to the locus of interest as a "hook." This requires an extra step of genetic engineering and could theoretically result in changes in the behavior of the locus. While the PICh method involves hybridization to capture the target chromatin, TCP relies on the binding of the TetR protein to the TetO sequences and is

therefore more efficient and specific than hybridization. Since the mammalian genome does not contain the TetO target sequence, the TetR protein is able to bind specifically and with high affinity to the TetO sequence (Kd 10^{-9}) (Kamionka et al., 2004). This enables several sequential purification and elution steps after cross linking resulting in less contaminating background proteins. The TetR-TetO interaction can be abolished by addition of doxycycline (tetracycline isomer) to the medium. This is of great importance in order to subtract nonspecific protein binding and identification of specific proteins in mass spectrometry. Thus, once the modified locus has been generated, the TCP approach may be more suitable than PICh for isolation of unique sequences, yielding cleaner chromatin prep by the use of different purification tags fused to TetR protein. The requirement for large quantities of starting material will likely be overcome in the near future given the increasing sensitivity of mass spectrometry.

The human γ-globin gene is prone to position effects and its developmental regulation can be easily disturbed. Moreover, hematopoietic cells transfection is extremely difficult. Therefore we used a transgenic mouse approach in order to identify protein factors bound *in vivo* to the suppressed γ-globin promoter. For isolation of the γ-globin promoter we inserted seven tetracycline operator repeat sequence, upstream (-380 StuI site) of γ-globin promoter *(tetO-γ).* According to the currently available information, no regulatory motif or important function has been reported for this site. The TetO repeats were inserted in the context of a human LCR γ-globin minilocus which is developmentally regulated in mouse, i.e. γ-globin switches off at mouse embryonic day 12.5 (Dillon and Grosveld, 1991). We then produced a transgenic mouse containing this modified construct in a p53 knock out background (p53null) (Dolznig et al., 2001) to be able to grow substantial amounts of hematopoietic erythroid progenitor cells derived from mouse fetal liver for our proteomic analysis. Insertion of TetO repeats did not disturb the developmental regulation of γ-globin. This mouse was then crossed to another double transgenic mouse expressing Bir A biotin ligase (de Boer et al., 2003) and a recombinant tetracycline repressor protein fused to Haemagglutinin (HA) and a Bio-tag (Bio) (TetR3T). Subsequent breeding of this double transgenic mouse (*TetR3T::BirA*) with *tetO-γ::p53null* mouse resulted in generation of *TetR3T::BirA::tetO-γ::p53null* mouse.

Binding of TetR-TetO, streptavidin-Biotin (Kd 10^{-15}) (Holmberg et al., 2005) and HA tag-HA antibody are among the most specific and strongest of their type. In addition, because the HA tag has no Histidine or lysine amino acid residues in its sequence and the Bio tag has only one lysine, they are preferred for purification processes involving formaldehyde cross linking (Kolodziej et al., 2009). The purification efficiency of TetR3T fusion protein by streptavidin and HA antibody beads is estimated at 20-50% for streptavidin beads and 20% for HA beads. Sequential HA and Bio promoter pull down efficiency using the optimized protocol was estimated 2% of the starting material with an enrichment of about 1000 fold over background. Currently the sensitivity of mass spectrometry is approximately 100 nmole, $6.02X10^{7}$ molecules ($[6.02X10^{23}]/10^{-16}$). Therefore we reasoned that $12X10^{9}$ cells would yield

enough material to be detected by mass spectrometry *(TetR3T::BirA::tetO-γ::p53null* cells can grow to the density of $3X10^6$ cells per millilitre).

Mass spectrometry of the γ-globin promoter pull down chromatin in a typical prep identified 534 proteins of which 380 were also present in the Doxycycline treated prep. From the remaining 154 proteins, 89 (such as ribosomal proteins, keratin, RNA binding proteins, nuclear membrane, etc.) were very unlikely to be involved in transcription or chromatin modifications and were therefore considered as non-specific background binding to the beads. We considered the remaining 65 proteins as potential candidates involved in γ-globin suppression. We selected a number of these candidates based on available information on their involvement in transcription regulation, chromatin remodeling, hematopoiesis and their potential role in γ-globin silencing, for further investigation. These included ZBP-89, Cdc5l, Apex1, Yb1, Ctnnbl1, Gata1, Hdac2, Prmt1, Supt5, Chd4 (mi2β), Actl6a (Baf53), Nap1l1, Ehmt1, Son, Snw1 and FancI. We investigated the role of some of these factors in the regulation of γ-globin expression using RNA interference-mediated knock down of the factors and examined their ability to suppress γ-globin expression. Reduced levels of Apex1, Yb1, Cdc5l, FancI and ZBP-89 increased γ-globin expression in adult human erythroid progenitors, suggestive of their role as direct transcriptional repressors of γ-globin, while down regulation of Ctnnbl1 and Supt5 had no affect on γ-globin expression.

PICh uses a specific nucleic acid probe and hybridization to isolate telomeric DNA and identify its associated telomere proteins (Déjardin and Kingston, 2009). While Background proteins in our study (TCP) and Déjardin et al (PICh) study were very similar, none of our specific targets were identified in the Déjardin et al PICh study and vice versa. We were specifically expecting transcription and chromatin modifier factors related to transcriptional suppression while Déjardin et al were looking for structural telomere proteins. Thus the target proteins were very different in the two studies and hence a different result is not unexpected. Nevertheless it partially confirms technical validity of our approach. Importantly, siRNA depletion of five out of seven candidate factors in our γ-globin promoter pull down identified by TCP resulted in increased γ-globin expression, thus suggestive of their role in γ-globin expression regulation. Moreover, several of the TCP-identified factors, including Cdc5l, Gata1, Hdac2, Prmt1, Chd4 (mi2β), Actl6a (Baf53) (Harju-Baker et al., 2008b; Mahajan et al., 2005a; Olave et al., 2007; Van Dijk et al., submitted), were previously reported to be involved in γ-globin regulation or interact with the γ-globin promoter, and therefore served as a positive control as to the specificity of our approach. Together these observations provide us with confidence about the relevance of protein factors identified by γ-globin promoter pull down in γ-globin regulation.

We chose several of the TCP-identified γ-globin interacting candidates to functionally characterize their potential role in the regulation of γ-globin expression. Apex1 (Apurinic/apyrimidinic endonuclease 1) and Yb-1 (Y box-binbing protein 1) are two of the factors identified in our γ-globin promoter pull down by TCP. Apex1 modulates DNA binding activity of several transcription factors including AP-1, CREB and Nrf2 (NF-E2-related factor

2); which are known to be bound to human β-globin locus LCR HS2 (Johnson et al., 2002; Talbort and Grosveld, 1991). APEX1 is involved in hemin (Iwasaki et al., 2006) mediated differentiation of K562 cells and hematopoietic differentiation of embryonic stem cells(Zou et al., 2007). Apex1, stably interacts with Yb-1 and enhances its binding to the Y-box element (inverted CCAAT boxes) (Chattopadhyay et al., 2008; Spitkovsky et al., 1992). Gata1 was also shown to induce Yb-1 expression by binding to it's 5'-UTR region (Yokoyama et al., 2003b). Yb-1 is implicated in embryonic development (Lu et al., 2005), early stage of erythropoiesis and hematological diseases (Yokoyama et al., 2003a). To examine their function in γ-globin expression, we depleted APEX1 and YB-1 from human adult progenitor cells (HEP) by shRNA. Consistent with their role as transcriptional repressors of γ-globin, YB-1 and APEX1 depletion resulted in 3 and 2 fold HbF induction, respectively.

Another interesting γ-globin repressor candidate identified in our TCP, Cdc5l (cell division cycle 5-like) is a cell cycle regulator with sequence specific DNA binding activity (CTCAGCG). Its DNA binding domain has similarities to Myb (Takashi and Kazuo, 1996) which is known to be involved in γ-globin repression (Jiang et al., 2006; Kuroyanagi et al., 2006). It is reported that Cdc5l and Dnmt1 bind *in vitro* to -198 Aγ HPFH mutation (Olave et al., 2007). CDC5L also interacts with beta catenin like 1 protein (Ctnnbl1), also called nucleosome assembly protein (NAP)(Conticello et al., 2008). These 3 proteins, Cdc5l, Dnmt1 and Ctnnbl1, were detected in two out of 3 TCP experiments (3 out of 3 for Cdc5l). To probe its potential role in γ-globin regulation, we depleted CDC5L using an shRNA approach. Although complete depletion of CDC5L by shRNA was lethal to the HEP cells, partial down regulation of CDC5L had a 2-4 fold increasing effect on γ-globin expression while reduced levels of CTNNBL1 had no effect on γ-globin expression.

Another TCP-identified candidate, FancI belongs to the Fanconi anemia complementation group (FANC). Its absence induces cytogenetic instability, hypersensitivity to DNA cross-linking agents and results in Fanconi anemia associated with high HbF levels (Gumruk et al., 2008; Miniero, 1981). We found that human hematopoietic progenitor cells express lower levels of FANCI in the fetal stage when compared to the adult stage (unpublished data). FANCI siRNA mediated knock down, induced γ-globin expression up to 4 fold in adult HEP cells, suggesting its involvement in repression of γ-globin, although we can not exclude if has a secondary effect due to genetic instability.

Another novel candidate in γ-globin promoter pull down by TCP is ZBP-89. ZBP-89 (BFCOL1, BERF-1, ZNF148 in human and Zfp148 in mouse), is a Krüppel-type, zinc-finger transcription factor. It has been suggested that ZBP-89 competes with Sp1 for the same binding site and subsequently represses or activates known target genes (Merchant et al., 1996; Wu et al., 2007). ZBP-89 interacts with several erythroid specific transcription factors such as Gata1, Mafk and has been implicated in erythroid development and differentiation. ZBP-89 has been shown to bind to Gata1 enhancer and the mouse β-globin LCR HS8 as well as other targets (Brand et al., 2004; Li et al., 2006; Ohneda et al., 2009; Vernimmen et al., 2007; Woo et al., 2008). It is proposed that ZBP-89 represses vimentin expression via the

specific recruitment of HDAC1 to the vimentin promoter (Wu et al., 2007). Moreover, it is well known that HDACs are involved in γ-globin silencing and use of HDAC inhibitors can partially activate γ-globin expression (Constantoulakis et al., 1989; Perrine et al., 1989; Perrine et al., 1987). Based on these observation we reasoned that ZBP-89 might be involved in γ-globin suppression by recruiting HDAC's and other transcription factors or chromatin remodelling complexes to the γ-globin gene promoters, as seen on other promoters.

shRNA mediated ZBP-89 knock down in human adult HEP cells increased γ-globin expression at the mRNA, protein as well as cellular level by γ-globin immunostaining. The level of ZBP-89 knock down seems to inversely correlate with γ-globin expression. We have also observed the same effect in mouse α and β globin locus embryonic genes. To study ZBP-89 in more detail, we generated a C-terminal tagged version of ZBP-89 fused to V5, Flag and Bio tag and expressed it in MEL cells expressing Bir A biotin ligase (de Boer et al., 2003), to identify ZBP-89 interacting proteins by mass spectrometry (Table 1) . We found that almost all of the Nucleosome Remodeling deacetylase Complex (NuRD) subunits interact with ZBP-89. We also found several hematopoietic specific transcription factors such as Runx1, Gfi1b and Fop in the immunoprecipitate. One of the ZBP-89 interacting partners identified in our mass spectrometry was Mbd3. Mbd3 proteomics data (Kransdorf et al., 2006 and unpublished data) indicates that it also interacts with Gata1, Baf53, Bcl11a, CCAAT displacement protein (CDP) and Fog1. All of these proteins have been implicated in γ-globin silencing (Chen et al., 2008; Harju-Baker et al., 2008a; Mahajan et al., 2005b; Sankaran et al., 2008; Sankaran et al., 2009; Skalnik et al., 1991). Our preliminary CoIP experiments confirmed many of these interactions (not shown). Thus these preliminary data suggests that ZBP-89 represses the γ-globin gene expression via recruitment of the repressive NuRD and MBD3 complexes.

Unfortunately we were unable to produce convincing and reproducible data indicating direct binding of ZBP-89 to the γ-globin promoter and its upstream sequences by conventional ChIP experiment in human hematopoietic erythroid progenitor (HEP) cells. This may be due to indirect or transient binding of the factors to the γ-globin promoter, binding of ZBP-89 to other regions of the γ-globin gene promoter not covered by our scanning primers. We made use of tagged version of ZBP-89 expressed in MEL cells to identify DNA binding sites of this factor in mouse genome by SOLEXA chip sequencing. This shows that ZBP-89 interacts with the promoter of many genes including Gata1, Bcl11a, Ldb1, Klf1 and many other crucial genes involved in hematopoietic development and differentiation. ZBP-89 also binds to 1,5 and 2 kb upstream region of mouse embryonic βh1 globin genes but not to the rest of the globin locus. This interaction is confirmed in mouse fetal liver I/11 type cells by ChIP experiments. The promoters of the human γ and mouse βh1 genes are very similar; however, the ZBP-89 interacting region in mouse is missing in human γ-globin promoter. We are actively pursuing the role of this factor in hematopoiesis and γ-globin gene regulation. Currently we are in the process of ZBP-89 SOLEXA chip sequencing in human erythroid progenitor cells to identify ZBP-89 binding sites in the human genome. We will also generate conditional ZBP-89 knock

out mouse containing a transgenic human β-globin locus to study erythropoiesis and more specifically mouse and human embryonic and fetal globins regulation.

However, an indirect role for ZBP-89 on γ-globin regulation via other factors such as Bcl11a, Gata1 is equally likely. Mutations in KLF1 results in HPFH phenotype in a Maltese family and it is suggested that KLF1 affects BCL11A expression (Borg J. manuscript submitted). This may also be the case for ZBP-89, as we detected binding of this factor to the promoter of the genes mentioned above.

There are several other interesting protein candidates identified in our study, including Chd4 (mi2β), Actl6a (Baf53), Ehmt1, Son, and Snw1 that we will further investigate their role in γ-gobin regulation.

Chromodomain helicase DNA binding protein 4 (**Chd4**/Mi2β) is a subunit of the Nucleosome Remodeling deacetylase Complex (NuRD) and **Actl6a** (Baf53) is part of SWI/SNF chromatin remodeling complex. It has been shown that NURD/MeCP1 and SWI/SNF complex interacts with human β-locus LCR HS2 core sequence (Mahajan et al., 2005b) and -566 Aγ globin GATA motif (Harju-Baker et al., 2008a). Chd4 also interacts with BCl11A (Sankaran et al., 2008) and it is required for multilineage differentiation in early hematopoiesis. Knock down of Chd4 results in activation of mouse embryonic globin genes (Yoshida et al., 2008).

Son is a negative regulatory element (NRE) binding protein. NRE is located immediately upstream of the core promoter and second enhancer of human hepatitis B virus (HBV) and represses their transcriptional activation function (Sun et al., 2001). This consensus motif (GA(G/T)AN(C/G)(A/G)CC) is also present upstream and down stream of γ-globin genes, in between human β locus LCR HS1 and HS2, and downstream of β-globin gene.

Ehmt1 is a transcriptionally repressive histone methyltransferase (Collins et al., 2008). **Snw1** (SKI interacting protein), a transcriptional co-repressor published to interact with other co-repressors (Hdac3, Prmt5, N-CoR and Sin3a) and mediates transcriptional repression by nuclear hormone receptors. Ski directly binds to MeCP2 and is required for the transcriptional repression mediated by MeCP2 (Kokura et al., 2001; Tabata et al., 2009). Thus, the available literature on these factors implicates them as repressors potentially involved in direct repression of γ-globin by binding to this locus. Future studies will focus on functional characterization of these factors in γ-globin regulation through their direct recruitment to the locus.

Genetic Factors influencing fetal Hemoglobin expression and hydroxyurea response in β-thalassemia

Repression of γ-globin after birth is strongly influenced by different cellular pathways such as the stressed hematopoiesis response, proliferative stress and signaling pathways such as those induced by Hydroxyurea (HU). In fact the majority (but not all) of β-thalassemia and sickle cell disease patients respond to HU with an increase in γ-globin expression. The magnitude of this response is largely dependent on HbF baseline of the patient (Chapter 4),

different genetic factors which in turn modulate different regulatory pathways, and trans-acting factors involved in γ-globin production (Bank, 2006; Ma et al., 2007; Stamatoyannopoulos and Grosveld, 2001; Steinberg et al., 1997).

To better understand these differences among HU 'responder' and 'non-responder' groups, we compared the transcriptome of several thalassemic patients in chapter 4. This approach can help to identify factors responsible for increased HbF baseline in different individuals, as well as γ-globin regulation pathways and mechanism of HbF induction by HU which are not fully understood. Prediction of the HbF response to HU would be very useful for drug administration to patients and can be used as a preventative measure against side effects of drug.

We found that the responders and the non-responders are strikingly different. Erythroblast cultures derived from non-responders grow faster compared to the responders, but they are very sensitive to HU treatment. Responders express significantly more hemoglobin, which consists mostly of fetal hemoglobin. Cultures derived from non-responders displayed differential expression of many genes upon HU treatment while the expression profiles of responder cells hardly changed.

The higher growth rate of 'non-responders' compare to 'responders' and their decline after HU treatment could be explained in part by differential expression of the INK4b-ARF-INK4a locus genes. This locus is known to be involved in stress responses (Gil and Peters, 2006). Expression of p16INK4a is increased upon HU treatment but in the absence and presence of HU it is comparable between 'non-responders' and 'responders'. In contrast, p14ARF expression is much higher in 'non-responders' and is further increased after HU treatment, while this does not hold true for 'responders'. p15INK4b expression is lower in 'non-responders' before HU treatment and it is significantly enhanced upon HU treatment while 'responders' express the same level of p15INK4b before and after the treatment. This failure of 'responder' cultures to up-regulate p14ARF and p15INK4b may contribute to their relative resistance to HU.

Differentiation of erythroblasts is not associated with a block of cell cycle progression. Instead, differentiation requires fast divisions with a short G1 transition and loss of size control to generate small erythrocytes. Hydroxyurea (HU), a ribonucleotid reductase inhibitor, inhibits S phase progression (Szekeres et al., 1997), induces apoptosis (Schrell et al., 1997) and enhances differentiation (Chapter 4). As a result, HU treatment is associated with large erythrocytes, because they failed to undergo the divisions coupled to terminal differentiation. Most likely it is the stalling of replication during S-phase that triggers a stress program leading to γ-globin expression. In β-thalassemia, progenitor cells commit to apoptosis during early differentiation (Perrine, 2005). The expression profiles of responders and non-responders clearly shows that high HbF expression is associated with stressed condition and also genes that protect the cells from stress induced apoptosis. An example of these genes is $BCLX_L$ which displays higher expression in non-responders. During erythroid differentiation, non-functional cells are eliminated by apoptosis. Therefore differentiation and apoptosis are

synchronous and share some factors. For example, $BCLX_L$ is down-regulated in differentiation induced apoptosis in erythroleukemia cells treated with chemical compounds (Benito et al., 1997; Terui et al., 1998), while in normal erythroid progenitors, its induced expression maintains cell viability (Motoyama et al., 1999; Silva et al., 1996).

Another example for adaptation of responders to proliferative stress is higher expression of FOXO3 and its many target genes such as HIPK2, BTG1, ARG1, ARG2 and p27KIP. The function of stress-induced FOXO3 protein is not to induce cell death, but to increase the potential of cells to prevent and repair oxidative damage and to slow down the cell cycle (Bakker et al., 2007). FOXO3 is upregulated in erythroid differentiation and also in response to stress. This protein is initially expressed higher in responder cultures and is further upregulated in response to HU compared to non-responder cultures. There are also many other genes showing higher expression in responders, which indicates that this group has adapted to permanent stress conditions. Thus, non-responders reach the same stress level as the responders only upon exposure to HU, as demonstrated by their gene expression profile.

HU increases NO (Nitric Oxide) production through eNOS (Nitric Oxide synthase), regulated by the mechanisms dependent on PKA and PI3K-induced stimulation of PKB/Akt (Cokic et al., 2006). NO increases cGMP and fetal hemoglobin levels in human erythroid progenitor cells (Cokic et al., 2007; Lou et al., 2009). Furthermore, soluble guanylate cyclase activators or its analogs are erythroid differentiation inducers and increase γ-globin gene expression in primary human erythroblasts (Ikuta et al., 2001; Osti et al., 1997). Exogenous NO inhibits growth of erythroid primary cells (Maciejewski et al., 1995). NO levels is significantly decreased in erythroid differentiation (Kucukkaya et al., 2006) and erythroid cell differentiation is inhibited by NO inducers such as butyric acid. High NO concentrations in most cases induce apoptosis while low NO concentrations can result in resistance to apoptosis. NO is inhibited more by fetal than by adult erythrocytes indicating more NO scavenging factors (Calatayud et al., 1998) in fetal compared to adult RBC's. Arginase hydrolyzes L-arginine to L-ornithine in the urea cycle. Arginase can modulate the production of nitric oxide by competition with NO synthase (NOS) for the same substrate, L-arginine (Durante et al., 2007). Arginase 1 (ARG1), Arginase 2 (ARG2) and Argininosuccinate synthetase 1 (ASS1) (also involved in urea cycle converting NO by product of citrulline to arginine) are differentially expressed between 'non-responders' and 'responders' in our data set. They are increased 8, 2 and 2.7 fold in 'responders' compared to 'non-responders', respectively. In a large multicenter SNPs association study, ARG1, ARG2, ASS1, NOS1 and NOS2A were significantly associated with response to HU treatment (Ma et al., 2007). Collectively, these data suggest that although NO is involved in HbF induction by HU and erythroid cell progenitors differentiation, high expression of ARG1, ARG2 and ASS1 is necessary to protect hematopoietic progenitor cells against the excessive amount of NO after HU treatment to scavenge the extra NO in the later stages of differentiation and prevent them from apoptosis.

These observations together with growth and differentiation differences between 'responders' and 'non-responders' suggest that because of these differentially expressed apoptotic, cell cycle and stress genes 'responders' are more resistant to the cytotoxic effects of HU. They can easily follow their differentiation program, survive longer and accumulate more HbF upon HU treatment. In contrast, 'non-responders' are very sensitive to HU; in response to HU treatment, the apoptotic, cell cycle arrest and stress response pathways are activated in the 'non-responders'.

In expression profiles of 'responders' and 'non-responders' a number of genes are differentially expressed between the two groups that are suspected to modify γ-globin expression. These genes are included in the list of genes presented in (Chapter 4; Table2) and may be involved in higher γ-globin expression baseline in 'responders'.

These differences can be potentially very useful to predict β-thalassemia and sickle cell disease patients responsiveness to HU treatment. To this end, using the gene list presented in Chapter 4, Table 2; it is possible to develop an expression fingerprint (predictor) to distinguish between HU responders and non-responders. The predictor has to be further confirmed by analysis of a larger cohort of patients.

Polymorphic markers associated with HbF expression

The basal HbF expression difference between the 'responder' and 'non-responder' groups may also be largely dependent on different genetic backgrounds. In fact there are a number of polymorphisms and genetic loci associated with high HbF, including BCL11A, HBS1L-MYB region on 6q23, Gγ-globin XmnI polymorphism (Lettre et al., 2008; Solovieff et al., 2009; Uda et al., 2008). One significant and clear genetic difference between 'non-responder' and 'responder' groups is the -158 XmnI Gγ polymorphism haplotype which is also frequently reported to be associated with high HbF production (Alebouyeh, 2004; Bradai et al., 2003; Dixit et al., 2005; Karimi et al., 2005; Neishabury et al., 2008; Panigrahi et al., 2005; Verma et al., 2007; Yavarian et al., 2004). In our patients, 83% of 'responder' chromosomes carry the -158 XmnI C>T polymorphism while in 'non-responders' only 29% carry this polymorphism.

The -158 XmnI Gγ polymorphism seems to be a very powerful marker for induced HbF expression but it is not the sole variation that one can rely on to screen β-thalassemia intermedia patients, as there are many cases carrying this haplotype which do not fall in this category. High HbF thalassemic patients (β-thalassemia intermedia) can be categorized in 3 different types. i) β-thalassemic patients with mild mutations, i.e. β -101 C>T; ii) β-thalassemia with coinheritance of α-thalassemia or HPFH, and iii) high fetal hemoglobin (HbF) levels in the absence of HPFH (Cao et al., 1994; Hardison et al., 2002; Patrinos et al., 2004b; Weatherall, 2001). Some of the High HbF patients in the absence of HPFH and coexisting α-thalassemia or mild mutation were explained by sequence variations in the γ-globin gene promoter (Gilman and Huisman, 1985; Lanclos et al., 1991; Patrinos et al., 2001; Patrinos et

al., 2005), the LCR (Adekile et al., 1993) or the Aγ-δ-globin intergenic region (Papachatzopoulou et al., 2006). However, the phenotype of a number of patients cannot be explained by such sequence variations. In Chapter 5 we examined several of these cases to find additional genetic variation that might explain the high HbF expression in these patients and introduce a new set of markers for thalassemia intermedia. We found a new genetic variation in 3'HS1 (+179 C>T) which is associated with β-thalassemia intermedia patient group (46%), compared to β-thalassemia major patients (15%) and normal subjects (16%). Electrophoretic mobility shift assay indicates that this variation creates a novel GATA1 binding site in 3'HS1. It has been shown that 3'HS1 is involved in active chromatin hub formation (ACH). The ACH brings the globin genes and regulatory sequences of the globin locus in a spatial conformation that appears to be required for proper developmental expression (Patrinos et al., 2004a). It is not clear what the effect of the 3'HS1 (+179 C>T) variation would be, and whether the new generated GATA1 binding site is functionally important for globin locus regulation. One possibility is that the new GATA1 binding site may interfere with participation of 3'HS1 in the ACH. This may bring the downstream enhancers into close proximity of γ-globin gene promoters, similar to what is observed in the deletional HPFH Cases (Arcasoy et al., 1997; Feingold and Forget, 1989; Katsantoni et al., 2003).

The 3'HS1 (+179 C>T) sequence variation represents a useful genetic marker for β-thalassemia intermedia patients. This marker can be studied in a larger patient and normal cohort. If the 3'HS1 (+179 C>T) is indeed significantly associated with high HbF, it can be used along with the -158 XmnI Gγ polymorphism for screening of β-thalassemia intermedia patients.

Conclusion

In conclusion, understanding the mechanism behind γ-globin gene switching and all the players involved in the process is critical to design new therapeutic approaches for β-thalassemia and sickle cell disease. Such insights will also be important for revealing the basic features of developmental gene regulation of the β globin locus. The development of different methods and approaches are necessary to overcome technical obstacles in understanding mechanisms of gene regulation. Here we have taken advantage of different disciplines to tackle this problem.

First we designed a new methodology (Targeted Chromatin Purification, TCP) to isolate and identify protein factors that are directly bound to the γ-globin gene promoter. This methodology can be optimized in the future by using different cross linkers, extra steps of purification, and more efficient Mass Spectrometry techniques. Although TCP is technically challenging, once it is proven to be effective for the proteomic characterization and composition of a single copy locus, its applicability can extend to other loci. Our study was a proof of principle for this technique. Using TCP, one can identify proteins bound to a regulatory element of interest based solely upon the presence of targeted DNA sequences, and this will allow the unbiased discovery of regulatory interactions on a given genomic locus.

Identification of regulatory complex composition would allow the development of new therapeutic means to disrupt the regulatory complex (e.g by inhibitory peptide or small molecule inhibition of a critical enzymatic component of the complex) for gene regulation therapy approaches.

At the same time while we are trying to understand factors that directly regulate the γ-globin gene regulation, the molecular connections to phenomena such as signalling pathways, stress erythropoiesis and alterations in cell cycle kinetics as well as polymorphic genetic alternation are also of great importance. To shed light on indirect γ-globin regulatory mechanisms we studied several thalassemic patients with high and low HbF levels and also their response to the γ-globin inducing compound, Hydroxyurea. We hypothesised that the activity of regulatory mechanisms can be deduced from the expression profiles of erythroid progenitor cells. We found that the transcriptome of these two groups of patients are very different. A number of genes were consistently differentially expressed between these two groups. These genes can be further analyzed for their role in γ-globin regulation and also used as a prognostic marker for Hydroxyurea responsiveness. Future studies will elucidate the complex signaling and regulatory networks that control gene expression in this important developmental locus and will have important implications for the treatment of diseases such β-thalassemia and sickle cell anemia.

hit number	Mascot score	symbol	accession	description (Entrez)
4	1647	Zfp148	IPI00123531	zinc finger protein 148
7	1357	Chd4	IPI00396802	chromodomain helicase DNA binding protein 4
10	962	Thrap3	IPI00556768	thyroid hormone receptor associated protein 3
13	781	Gatad2b	IPI00128615	GATA zinc finger domain containing 2B
17	720	Mta1	IPI00776055	metastasis associated 1
18	670	Mta1	IPI00624969	metastasis associated 1
19	669	Gatad2a	IPI00229784	GATA zinc finger domain containing 2A
22	591	Serbp1	IPI00471476	serpine1 mRNA binding protein 1
23	574	Zfp219	IPI00469594	zinc finger protein 219
27	556	Mta2	IPI00128230	metastasis-associated gene family, member 2
29	549	Rbbp4	IPI00122696	retinoblastoma binding protein 4
30	523	Rbbp7	IPI00122698	retinoblastoma binding protein 7
34	488	Top2a	IPI00122223	topoisomerase (DNA) II alpha
39	443	Ilf3	IPI00130591	interleukin enhancer binding factor 3
40	443	Pdcd11	IPI00551454	programmed cell death 11
43	418	Hdac2	IPI00137668	histone deacetylase 2
44	415	Hdac1	IPI00114232	histone deacetylase 1
48	396	Chd3	IPI00551435	chromodomain helicase DNA binding protein 3
52	379	Rcc2	IPI00222509	regulator of chromosome condensation 2
54	364	Ilf2	IPI00318550	interleukin enhancer binding factor 2
55	364	Sfrs1	IPI00420807	splicing factor, arginine/serine-rich 1 (ASF/SF2)
62	346	Mta3	IPI00221805	metastasis associated 3
77	289	Lig3	IPI00124272	ligase III, DNA, ATP-dependent
92	246	Dnmt1	IPI00469323	DNA methyltransferase (cytosine-5) 1
96	241	Dnmt1	IPI00474974	DNA methyltransferase (cytosine-5) 1
102	227	Smarca5	IPI00396739	SWI/SNF related, matrix associated
115	199	Bclaf1	IPI00169477	BCL2-associated transcription factor 1
116	198	Nop14	IPI00353010	NOP14 nucleolar protein homolog (yeast)
131	162	Chd8	IPI00858099	chromodomain helicase DNA binding protein 8
132	161	Ebna1bp2	IPI00111829	EBNA1 binding protein 2
136	158	Elavl1	IPI00108271	ELAV (embryonic lethal,Drosophila)-like 1
137	157	Numa1	IPI00263048	nuclear mitotic apparatus protein 1
138	157	Caprin1	IPI00121515	cell cycle associated protein 1
140	156	Zc3h18	IPI00648513	zinc finger CCCH-type containing 18
147	146	Rfc2	IPI00124744	replication factor C (activator 1) 2
148	145	Gtpbp4	IPI00117642	GTP binding protein 4
152	139	Mbd3	IPI00131067	methyl-CpG binding domain protein 3
154	136	Luc7l	IPI00410804	Luc7 homolog (S. cerevisiae)-like
155	136	Plk1	IPI00120767	polo-like kinase 1 (Drosophila)
158	134	Runx1	IPI00750028	runt related transcription factor 1
166	121	Sltm	IPI00229571	SAFB-like, transcription modulator
179	106	Fxr1	IPI00122521	fragile X mental retardation gene 1
181	104	Aplf	IPI00110815	aprataxin and PNKP like factor
183	103	Thoc4	IPI00114407	THO complex 4
184	102	Mov10	IPI00130328	Moloney leukemia virus 10
185	102	Sfrs5	IPI00314709	splicing factor, arginine/serine-rich 5
186	101	Cbx3	IPI00677454	chromobox homolog 3 (Drosophila HP1 gamma)
187	101	Hp1bp3	IPI00342766	heterochromatin protein 1, binding protein 3
188	100	Ssrp1	IPI00407571	structure specific recognition protein 1
193	96	Pelp1	IPI00321597	proline, glutamic acid and leucine rich protein 1
195	95	Mela	IPI00224370	melanoma antigen
196	95	Ybx1	IPI00120886	Y box protein 1
204	90	Gm6816	IPI00116049	predicted gene 6816
207	87	Cdk1	IPI00114491	cyclin-dependent kinase 1
208	87	Kif23	IPI00330356	kinesin family member 23
211	87	Pnn	IPI00317891	pinin
212	85	Supt16h	IPI00120344	suppressor of Ty 16 homolog (S. cerevisiae)
225	74	Fip1l1	IPI00110716	FIP1 like 1 (S. cerevisiae)
234	68	Rbbp4	IPI00881032	retinoblastoma binding protein 4
235	68	Acin1	IPI00121136	apoptotic chromatin condensation inducer 1
236	67	Cdk2ap2	IPI00132178	CDK2-associated protein 2
239	66	Ccar1	IPI00135207	cell division cycle and apoptosis regulator 1
247	64	Baz1a	IPI00461396	bromodomain adjacent to zinc finger domain 1A
252	61	Bag2	IPI00130304	BCL2-associated athanogene 2
266	54	Mbd2	IPI00131088	methyl-CpG binding domain protein 2
285	45	Cyfip1	IPI00330476	cytoplasmic FMR1 interacting protein 1

Table 1. Preliminary list of ZBP-89 interacting proteins in MEL cells identified by Biotin tagged ZBP-89 pull down followed by Mass Spectrometry.

References

Adekile, A., Dimovski, A., Oner, C., Lanclos, K., and Huisman, T. (1993). Haplotype-specific sequence variations in the locus control region (5' hypersensitive sites 2, 3, 4) of beta S chromosomes. Hemoglobin *17*, 475-478.

Alebouyeh, M., Moussavi, F., Haddad-Deylami, H., Vossough, P. (2004). Hydroxyurea in the treatment of major beta-thalassemia and importance of genetic screening. Ann Hematol *83*, 430-433.

Arcasoy, M.O., Romana, M., Fabry, M.E., Skarpidi, E., Nagel, R.L., and Forget, B.G. (1997). High levels of human gamma-globin gene expression in adult mice carrying a transgene of deletion-type hereditary persistence of fetal hemoglobin. Mol Cell Biol *17*, 2076-2089.

Bakker, W.J., van Dijk, T.B., Parren-van Amelsvoort, M., Kolbus, A., Yamamoto, K., Steinlein, P., Verhaak, R.G.W., Mak, T.W., Beug, H., Lowenberg, B*., et al.* (2007). Differential Regulation of Foxo3a Target Genes in Erythropoiesis. Mol Cell Biol *27*, 3839-3854.

Bank, A. (2006). Regulation of human fetal hemoglobin: new players, new complexities. Blood *107*, 435-443.

Benito, A., Lerga, A., Silva , M., Leon, J., and Fernandez-Luna, J. (1997). Apoptosis of human myeloid leukemia cells induced by an inhibitor of protein phosphatases (okadaic acid) is prevented by Bcl-2 and Bcl-X(L).

. Leukemia *11*, 940-944.

Bhatia, M., and Walters, M.C. (2007). Hematopoietic cell transplantation for thalassemia and sickle cell disease: past, present and future. Bone Marrow Transplant *41*, 109-117.

Bradai, M., Abad, M.T., Pissard, S., Lamraoui, F., Skopinski, L., and de Montalembert, M. (2003). Hydroxyurea can eliminate transfusion requirements in children with severe beta-thalassemia. Blood *102*, 1529-1530.

Brand, M., Ranish, J.A., Kummer, N.T., Hamilton, J., Igarashi, K., Francastel, C., Chi, T.H., Crabtree, G.R., Aebersold, R., and Groudine, M. (2004). Dynamic changes in transcription factor complexes during erythroid differentiation revealed by quantitative proteomics. Nat Struct Mol Biol *11*, 73-80.

Calatayud, S., Beltrán, B., Brines, J., Moncada, S., and Esplugues, J.V. (1998). Foetal erythrocytes exhibit an increased ability to scavenge for nitric oxide. European Journal of Pharmacology *347*, 363-366.

Cao, A., Galanello , R., and Rosatelli, C. (1994). Genotype-phenotype correlations in b-thalassaemias. Blood Reviews *8*, 1-12.

Chattopadhyay, R., Das, S., Maiti, A.K., Boldogh, I., Xie, J., Hazra, T.K., Kohno, K., Mitra, S., and Bhakat, K.K. (2008). Regulatory Role of Human AP-Endonuclease (APE1/Ref-1) in YB-1-Mediated Activation of the Multidrug Resistance Gene MDR1. Mol Cell Biol *28*, 7066-7080.

Chen, Z., Luo, H.-Y., Basran, R.K., Hsu, T.-H., Mang, D.W.H., Nuntakarn, L., Rosenfield, C.G., Patrinos, G.P., Hardison, R.C., Steinberg, M.H*., et al.* (2008). A T-to-G Transversion at Nucleotide -567 Upstream of HBG2 in a GATA-1 Binding Motif Is Associated with Elevated Hemoglobin F. Mol Cell Biol *28*, 4386-4393.

Cokic, V.P., Beleslin-Cokic, B.B., Noguchi, C.T., and Schechter, A.N. (2007). Hydroxyurea increases eNOS protein levels through inhibition of proteasome activity. Nitric Oxide *16*, 371-378.

Cokic, V.P., Beleslin-Cokic, B.B., Tomic, M., Stojilkovic, S.S., Noguchi, C.T., and Schechter, A.N. (2006). Hydroxyurea induces the eNOS-cGMP pathway in endothelial cells. Blood *108*, 184-191.

Collins, R.E., Northrop, J.P., Horton, J.R., Lee, D.Y., Zhang, X., Stallcup, M.R., and Cheng, X. (2008). The ankyrin repeats of G9a and GLP histone methyltransferases are mono- and dimethyllysine binding modules. Nat Struct Mol Biol *15*, 245-250.

Constantoulakis, P., Knitter, G., and Stamatoyannopoulos, G. (1989). On the induction of fetal hemoglobin by butyrates: in vivo and in vitro studies with sodium butyrate and comparison of combination treatments with 5-AzaC and AraC. Blood *74*, 1963-1971.

Conticello, S.G., Ganesh, K., Xue, K., Lu, M., Rada, C., and Neuberger, M.S. (2008). Interaction between Antibody-Diversification Enzyme AID and Spliceosome-Associated Factor CTNNBL1. Molecular Cell *31*, 474-484.

de Boer, E., Rodriguez, P., Bonte, E., Krijgsveld, J., Katsantoni, E., Heck, A., Grosveld, F., and Strouboulis, J. (2003). Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. Nucl Acids Res *100*, 7480-7485.

Déjardin, J., and Kingston, R.E. (2009). Purification of Proteins Associated with Specific Genomic Loci. Cell *136*, 175-186.

Dillon, N., and Grosveld, F. (1991). Human gamma-globin genes silenced independently of other genes in the beta-globin locus. Nature *350*, 252-254.

Dixit, A., Chatterjee, T.C., Mishra, P., Choudhry, D.R., Mahapatra, M., Tyagi, S., Kabra, M., Saxena, R., and Choudhry, V.P. (2005). Hydroxyurea in thalassemia intermedia—a promising therapy. Annals of Hematology *84*, 441-446.

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

Durante, W., Johnson, F., K., and Johnson, R., A. (2007). ARGINASE: A CRITICAL REGULATOR OF NITRIC OXIDE SYNTHESIS AND VASCULAR FUNCTION. Clin Exp Pharmacol Physiol *34*, 906-911.

Feingold, E.A., and Forget, B.G. (1989). The breakpoint of a large deletion causing hereditary persistence of fetal hemoglobin occurs within an erythroid DNA domain remote from the beta-globin gene cluster. Blood *74*, 2178-2186.

Gil, J., and Peters, G. (2006). Regulation of the INK4b-ARF-INK4a tumour suppressor locus: all for one or one for all. Nat Rev Mol Cell Biol *7*, 667-677.

Gilman, J.G., and Huisman, T.H. (1985). DNA sequence variation associated with elevated fetal G g globin production. . Blood *66*, 783-787.

Griesenbeck, J., Boeger, H., Strattan, J.S., and Kornberg, R.D. (2003). Affinity Purification of Specific Chromatin Segments from Chromosomal Loci in Yeast. Mol Cell Biol *23*, 9275-9282.

Gumruk, F., Tavil, B., Balta, G., Unal, S., and Gurgey, A. (2008). Significance of Fetal Hemoglobin Values in Detection of Heterozygotes in Fanconi Anemia: Reevaluation of Fetal Hemoglobin Values by a Sensitive Method. J Pediatr Hematol Oncol *30*, 896-899 810.1097/MPH.1090b1013e3181833835.

Haley, J.D., Smith, D.E., Schwedes, J., Brennan, R., Pearce, C., Moore, C., Wang, F., Petti, F., Grosveld, F., Jane, S.M*., et al.* (2003). Identification and characterization of mechanistically distinct inducers of [gamma]-globin transcription. Biochemical Pharmacology *66*, 1755-1768.

Hardison, R., C, Chui, D., H. K., Giardine, B., Riemer, C., Patrinos, G., P. , Anagnou, N., Miller, W., and Wajcman, H. (2002). HbVar: A relational database of human hemoglobin variants and thalassemia mutations at the globin gene server. Hum Mutat *19*, 225-233.

Harju-Baker, S., Costa, F.C., Fedosyuk, H., Neades, R., and Peterson, K.R. (2008a). Silencing of A{gamma}-Globin Gene Expression during Adult Definitive Erythropoiesis Mediated by GATA-1-FOG-1-Mi2 Complex Binding at the -566 GATA Site. Mol Cell Biol *28*, 3101-3113.

Harju-Baker, S., Costa, F.C., Fedosyuk, H., Neades, R., and Peterson, K.R. (2008b). Silencing of Agamma-globin gene expression during adult definitive erythropoiesis mediated by GATA-1-FOG-1-Mi2 complex binding at the -566 GATA site. Mol Cell Biol *28*, 3101-3113.

Holmberg, A., Blomstergren, A., Nord, O., Lukacs, M., Lundeberg, J., and Uhlén, M. (2005). The biotin-streptavidin interaction can be reversibly broken using water at elevated temperatures. Electrophoresis *26*, 501-510.

Ikuta, T., Ausenda, S., and Cappellini, M.D. (2001). Mechanism for fetal globin gene expression: Role of the soluble guanylate cyclaseâ€"cGMP-dependent protein kinase pathway. Proc Natl Acad Sci U S A *98*, 1847-1852.

Iwasaki, K., MacKenzie, E.L., Hailemariam, K., Sakamoto, K., and Tsuji, Y. (2006). Hemin-Mediated Regulation of an Antioxidant-Responsive Element of the Human Ferritin H Gene and Role of Ref-1 during Erythroid Differentiation of K562 Cells. Mol Cell Biol *26*, 2845-2856.

Jasinskas, A., and Hamkalo, B.A. (1999). Purification and Initial Characterization of Primate Satellite Chromatin. Chromosome Research *7*, 341-354.

Jiang, J., Best, S., Menzel, S., Silver, N., Lai, M.I., Surdulescu, G.L., Spector, T.D., and Thein, S.L. (2006). cMYB is involved in the regulation of fetal hemoglobin production in adults. Blood *108*, 1077-1083.

Johnson, K.D., Norton, J.E., and Bresnick, E.H. (2002). Requirements for utilization of CREB binding protein by hypersensitive site two of the beta-globin locus control region. Nucleic Acids Res *30*, 1522-1530.

Kamionka, A., Bogdanska-Urbaniak, J., Scholz, O., and Hillen, W. (2004). Two mutations in the tetracycline repressor change the inducer anhydrotetracycline to a corepressor. Nucl Acids Res *32*, 842-847.

Karimi, M., Darzi, H., and Yavarian, M. (2005). Hematologic and Clinical Responses of Thalassemia Intermedia Patients to Hydroxyurea During 6 Years of Therapy in Iran. J Pediatr Hematol Oncol *27*, 380-385.

Katsantoni, E.Z., Langeveld, A., Wai, A.W.K., Drabek, D., Grosveld, F., Anagnou, N.P., and Strouboulis, J. (2003). Persistent gamma-globin expression in adult transgenic mice is mediated by HPFH-2, HPFH-3, and HPFH-6 breakpoint sequences. Blood *102*, 3412-3419.

Kokura, K., Kaul, S.C., Wadhwa, R., Nomura, T., Khan, M.M., Shinagawa, T., Yasukawa, T., Colmenares, C., and Ishii, S. (2001). The Ski Protein Family Is Required for MeCP2-mediated Transcriptional Repression. J Biol Chem *276*, 34115-34121.

Kolodziej, K., Pourfarzad, F., de Boer, E., Krpic, S., Grosveld, F., and Strouboulis, J. (2009). Optimal use of tandem biotin and V5 tags in ChIP assays. BMC Mol Biol *10*, 6.

Kransdorf, E.P., Wang, S.Z., Zhu, S.Z., Langston, T.B., Rupon, J.W., and Ginder, G.D. (2006). MBD2 is a critical component of a methyl cytosine-binding protein complex isolated from primary erythroid cells. Blood *108*, 2836-2845.

Kucukkaya, B., Arslan, D.O., and Kan, B. (2006). Role of G proteins and ERK activation in hemin-induced erythroid differentiation of K562 cells. Life Sciences *78*, 1217-1224.

Kuroyanagi, Y., Kaneko, Y., Muta, K., Park, B.-S., Moi, P., Ausenda, S., Cappellini, M.D., and Ikuta, T. (2006). cAMP differentially regulates [gamma]-globin gene expression in erythroleukemic cells and primary erythroblasts through c-Myb expression. Biochemical and Biophysical Research Communications *344*, 1038-1047.

Lanclos, K.D., Oner, C., Dimovski, A.J., Gu, Y.C., and Huisman, T.H. (1991). Sequence variations in the 5' flanking and IVS-II regions of the G gamma- and A gamma-globin genes of beta S chromosomes with five different haplotypes. Blood *77*, 2488-2496.

Lettre, G., Sankaran, V.G., Bezerra, M.A.C., Araújo, A.S., Uda, M., Sanna, S., Cao, A., Schlessinger, D., Costa, F.F., Hirschhorn, J.N.*, et al.* (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 U S A *105*, 11869-11874.

Li, X., Xiong, J.-W., Shelley, C.S., Park, H., and Arnaout, M.A. (2006). The transcription factor ZBP-89 controls generation of the hematopoietic lineage in zebrafish and mouse embryonic stem cells. Development *133*, 3641-3650.

Lou, T., Singh, M., Mackie, A., Li, W., and Pace, B. (2009). Hydroxyurea generates nitric oxide in human erythroid cells: mechanisms for gamma-globin gene activation. Exp Biol Med (Maywood) *234*, 1374-1382.

Lu, Z.H., Books, J.T., and Ley, T.J. (2005). YB-1 Is Important for Late-Stage Embryonic Development, Optimal Cellular Stress Responses, and the Prevention of Premature Senescence. Mol Cell Biol *25*, 4625-4637.

Ma, Q., Wyszynski, D.F., Farrell, J.J., Kutlar, A., Farrer, L.A., Baldwin, C.T., and Steinberg, M.H. (2007). Fetal hemoglobin in sickle cell anemia: genetic determinants of response to hydroxyurea. Pharmacogenomics J *7*, 386-394.

Mabaera, R., West, R.J., Conine, S.J., Macari, E.R., Boyd, C.D., Engman, C.A., and Lowrey, C.H. (2008). A cell stress signaling model of fetal hemoglobin induction: what doesn't kill red blood cells may make them stronger. Experimental Hematology *36*, 1057-1072.

Maciejewski, J.P., Selleri, C., Sato, T., Cho, H.J., Keefer, L.K., Nathan, C.F., and Young, N.S. (1995). Nitric oxide suppression of human hematopoiesis in vitro. Contribution to inhibitory action of interferon-gamma and tumor necrosis factor-alpha. The Journal of Clinical Investigation *96*, 1085-1092.

Mahajan, M.C., Narlikar, G.J., Boyapaty, G., Kingston, R.E., and Weissman, S.M. (2005a). Heterogeneous nuclear ribonucleoprotein C1/C2, MeCP1, and SWI/SNF form a chromatin remodeling complex at the beta-globin locus control region. Proc Natl Acad Sci U S A *102*, 15012-15017.

Mahajan, M.C., Narlikar, G.J., Boyapaty, G., Kingston, R.E., and Weissman, S.M. (2005b). Heterogeneous nuclear ribonucleoprotein C1/C2, MeCP1, and SWI/SNF form a chromatin remodeling complex at the beta-globin locus control region. Proc Natl Acad Sci U S A *102*, 15012-15017.

Merchant, J.L., Iyer, G.R., Taylor, B.R., Kitchen, J.R., Mortensen, E.R., Wang, Z., Flintoft, R.J., Michel, J.B., and Bassel-Duby, R. (1996). ZBP-89, a Kruppel-like zinc finger protein, inhibits epidermal growth factor induction of the gastrin promoter. Mol Cell Biol *16*, 6644-6653.

Miniero, R., David, O., Saglio, G., Paschero, C., Nicola, P. (1981). The Hbf in Fanconi's anemia Pediatr Med Chir *3*, 167-170.

Motoyama, N., Kimura, T., Takahashi, T., Watanabe, T., and Nakano, T. (1999). bcl-x Prevents Apoptotic Cell Death of Both Primitive and Definitive Erythrocytes at the End of Maturation. J Exp Med *189*, 1691-1698.

Neishabury, M., Azarkeivan, A., Oberkanins, C., Esteghamat, F., Amirizadeh, N., and Najmabadi, H. (2008). Molecular Mechanisms Underlying Thalassemia Intermedia in Iran. Genetic Testing *12*, 549-556.

O'Neill, D.W., Schoetz, S.S., Lopez, R.A., Castle, M., Rabinowitz, L., Shor, E., Krawchuk, D., Goll, M.G., Renz, M., Seelig, H.-P.*, et al.* (2000). An Ikaros-Containing Chromatin-Remodeling Complex in Adult-Type Erythroid Cells. Mol Cell Biol *20*, 7572-7582.

Ohneda, K., Ohmori, S.y., Ishijima, Y., Nakano, M., and Yamamoto, M. (2009). Characterization of a Functional ZBP-89 Binding Site That Mediates Gata1 Gene Expression during Hematopoietic Development. J Biol Chem *284*, 30187-30199.

Olave, I.A., Doneanu, C., Fang, X., Stamatoyannopoulos, G., and Li, Q. (2007). Purification and Identification of Proteins That Bind to the Hereditary Persistence of Fetal Hemoglobin â€"198 Mutation in the Î³-Globin Gene Promoter. J Biol Chem *282*, 853-862.

Osti, F., Corradini, F.G., Hanau, S., Matteuzzi, M., and Gambari, R. (1997). Human leukemia K562 cells: induction to erythroid differentiation by guanine, guanosine and guanine nucleotides. Hematologica *82*, 395-401.

Panigrahi, I., Dixit, A., Arora, S., Kabra, M., Mahapatra, M., Choudhry, V.P., and Saxena, R. (2005). Do alpha deletions influence hydroxyurea response in thalassemia intermedia? Hematology *10*, 61-63.

Papachatzopoulou, A., Kourakli, A., Markopoulou, P., Kakagianne, T., Sgourou, A., Papadakis, M., and Athanassiadou, A. (2006). Genotypic heterogeneity and correlation to intergenic haplotype within high HbF b-thalassemia

intermedia. . Eur J Haematol *76*, 322-330.

Patrinos, G., Kollia, P., Papapanagiotou, E., Loutradi-Anagnostou, A., Loukopoulos, D., and Papadakis, M. (2001). Agamma-haplotypes: A new group of genetic markers for thalassemic mutations inside the 50 regulatory region of the human Agamma-globin gene. Am J Hematol *66*, 99-104.

Patrinos, G., Samperi, P., lo Nigro, L., Kollia, P., Schiliro, G., and Papadakis, M. (2005). Evidence for the molecular heterogeneity of sickle cell anemia chromosomes

bearing the bS/benin haplotype. Am J Hematol *80*, 79-80.

Patrinos, G.P., de Krom, M., de Boer, E., Langeveld, A., Imam, A.M.A., Strouboulis, J., de Laat, W., and Grosveld, F.G. (2004a). Multiple interactions between regulatory regions are required to stabilize an active chromatin hub. Genes Dev *18*, 1495-1509.

Patrinos, G.P., Giardine, B., Riemer, C., Miller, W., Chui, D.H.K., Anagnou, N.P., Wajcman, H., and Hardison, R.C. (2004b). Improvements in the HbVar database of human hemoglobin variants and thalassemia mutations for population and sequence variation studies. Nucleic Acids Res *32*, D537-541.

Perrine, S.P. (2005). Fetal Globin Induction--Can It Cure beta- Thalassemia? Hematology Am Soc Hematol Educ Program *2005*, 38-44.

Perrine, S.P., Miller, B.A., Faller, D.V., Cohen, R.A., Vichinsky, E.P., Hurst, D., Lubin, B.H., and Papayannopoulou, T. (1989). Sodium butyrate enhances fetal globin gene expression in erythroid progenitors of patients with Hb SS and beta thalassemia. Blood *74*, 454-459.

Perrine, S.P., Miller, B.A., Greene, M.F., Cohen, R.A., Cook, N., Shackleton, C., and Faller, D.V. (1987). Butryic acid analogues augment gamma globin gene expression in neonatal erythroid progenitors. Biochemical and Biophysical Research Communications *148*, 694-700.

Rupon, J.W., Wang, S.Z., Gaensler, K., Lloyd, J., and Ginder, G.D. (2006). Methyl binding domain protein 2 mediates gamma-globin gene silencing in adult human betaYAC transgenic mice. Proc Natl Acad Sci U S A *103*, 6617-6622.

Sankaran, V.G., Menne, T.F., Xu, J., Akie, T.E., Lettre, G., Van Handel, B., Mikkola, H.K.A., Hirschhorn, J.N., Cantor, A.B., and Orkin, S.H. (2008). Human Fetal Hemoglobin Expression Is Regulated by the Developmental Stage-Specific Repressor BCL11A. Science *322*, 1839-1842.

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

Schrell, U.M.H., Rittig, M.G., Anders, M., Kiesewetter, F., Marschalek, R., Koch, U.H., and Fahlbusch, R. (1997). Hydroxyurea for treatment of unresectable and recurrent meningiomas. I. Inhibition of primary human meningioma cells in culture and in meningioma transplants by induction of the apoptotic pathway. J Neurosurg *86*, 845-852.

Shenoy, S. (2007). Has stem cell transplantation come of age in the treatment of sickle cell disease? Bone Marrow Transplant *40*, 813-821.

Silva, M., Grillot, D., Benito, A., Richard, C., Nunez, G., and Fernandez-Luna, J.L. (1996). Erythropoietin can promote erythroid progenitor survival by repressing apoptosis through Bcl-XL and Bcl-2. Blood *88*, 1576-1582.

Skalnik, D.G., Strauss, E.C., and Orkin, S.H. (1991). CCAAT displacement protein as a repressor of the myelomonocytic-specific gp91-phox gene promoter. J Biol Chem *266*, 16736-16744.

Solovieff, N., Milton, J.N., Hartley, S.W., Sherva, R., Sebastiani, P., Dworkis, D.A., Klings, E.S., Farrer, L.A., Garrett, M.E., Ashley-Koch, A*., et al.* (2009). Fetal hemoglobin in sickle cell anemia: genome-

wide association studies suggest a regulatory region in the 5' olfactory receptor gene cluster. Blood, blood-2009-2008-239517.

Spitkovsky, D.D., Royer-Pokora, B., Delius, H., Kisseljov, F., Jenkins, N.A., Gilbert, D.J., Copeland, N.G., and Royer, H.-D. (1992). Tissue restricted expression and chromosomal localization of the YB-1 gene encoding a 42kD nuclear CCAAT binding protein. Nucl Acids Res *20*, 797-803.

Stamatoyannopoulos, G., and Grosveld, F. (2001). Hemoglobin switching. In The molecular basis of blood diseases (3rd ed), G. Stamatoyannopoulos, Majeurs P., Perlmutter, P., Varmus H., ed. (Philadelphia, W.B. Saunders), pp. 135-182.

Steinberg, M.H., Lu, Z.-H., Barton, F.B., Terrin, M.L., Charache, S., Dover, G.J., and the Multicenter Study of, H. (1997). Fetal Hemoglobin in Sickle Cell Anemia: Determinants of Response to Hydroxyurea. Blood *89*, 1078-1088.

Sun, C.-T., Lo, W.-Y., Wang, I.-H., Lo, Y.-H., Shiou, S.-R., Lai, C.-K., and Ting, L.-P. (2001). Transcription Repression of Human Hepatitis B Virus Genes by Negative Regulatory Element-binding Protein/SON. J Biol Chem *276*, 24059-24067.

Szekeres, T., Fritzer-Szekeres, M., Elford, H.L., and Jayaram, H.M. (1997). The Enzyme Ribonucleotide Reductase: Target for Antitumor and Anti-HIV Therapy. Crit Rev Clin Lab Sci *34*, 503-528.

Tabata, T., Kokura, K., Dijke, P.t., and Ishii, S. (2009). Ski co-repressor complexes maintain the basal repressed state of the TGF-β target gene, SMAD7, via HDAC3 and PRMT5. Genes Cells *14*, 17-28.

Takashi, H., and Kazuo, S. (1996). A cdc5+ homolog of a higher plant,â€‰Arabidopsisâ€‰thaliana. Proc Natl Acad Sci U S A *93*, 13371-13376.

Talbort, D., and Grosveld, F. (1991). The 5'HS2 of the globin locus control region enhances transcription through the interaction of a multimeric complex binding at two functionally distinct NF-E2 binding sites. EMBO J *10*, 1391-1398.

Tanabe, O., Katsuoka, F., Campbell, A.D., Song, W., Yamamoto, M., Tanimoto, K., and Engel, J.D. (2002). An embryonic/fetal beta-type globin gene repressor contains a nuclear receptor TR2/TR4 heterodimer. EMBO J *21*, 3434-3442.

Terui, Y., Furukawa, Y., Kikuchi, J., Iwase, S., Hatake, K., and Miura, Y. (1998). Bcl-x is a regulatory factor of apoptosis and differentiation in megakaryocytic lineage cells. Exp Hematol *26*, 236-244.

Uda, M., Galanello, R., Sanna, S., Lettre, G., Sankaran, V.G., Chen, W., Usala, G., Busonero, F., Maschio, A., Albai, G.*, et al.* (2008). Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of beta-thalassemia. Proc Natl Acad Sci U S A *105*, 1620-1625.

Verma, I.C., Kleanthous, M., Saxena, R., Fucharoen, S., Winichagoon, P., Raizuddin, S., Khan, S.N., Akbari, M.T., Izadyar, M., Kotea, N.*, et al.* (2007). Multicenter Study of the Molecular Basis of Thalassemia Intermedia in Different Ethnic Populations. Hemoglobin *31*, 439-452.

Vernimmen, D., Gobbi, M.D., Sloane-Stanley, J.A., Wood, W.G., and Higgs, D.R. (2007). Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. EMBO J *26*, 2041-2051.

Weatherall, D.J., Clegg, J.B., Higgs, D.R., Wood, W.G. (2001). The hemoglobinopathies. In The Metabolic and Molecular Bases of Inherited Disease, C.R. Scriver, ed. (New York, NY, McGraw-Hill), pp. 4571–4636.

Woo, A.J., Moran, T.B., Schindler, Y.L., Choe, S.-K., Langer, N.B., Sullivan, M.R., Fujiwara, Y., Paw, B.H., and Cantor, A.B. (2008). Identification of ZBP-89 as a Novel GATA-1-Associated Transcription Factor Involved in Megakaryocytic and Erythroid Development. Mol Cell Biol *28*, 2675-2689.

Wu, Y., Zhang, X., Salmon, M., and Zehner, Z., E. (2007). The zinc finger repressor, ZBP-89, recruits histone deacetylase 1 to repress vimentin gene expression. Genes Cells *12*, 905-918.

Yavarian, M., Karimi, M., Bakker, E., Harteveld, C.L., and Giordano, P.C. (2004). Response to hydroxyurea treatment in Iranian transfusion-dependent beta-thalassemia patients. hematologica *89*, 1172-1178.

Yokoyama, H., Harigae, H., Takahashi, S., Kameoka, J., Miyamura, K., Ishizawa. K., Kaku, M., and Sasaki, T. (2003a). High expression of YB-1 gene in erythroid cells in patients with refractory anemia. Int J Hematol *3*, 213-218.

Yokoyama, H., Harigae, H., Takahashi, S., Takahashi, S., Furuyama, K., Kaku, M., Yamamoto, M., and Sasaki, T. (2003b). Regulation of YB-1 gene expression by GATA transcription factors. Biochemical and Biophysical Research Communications *303*, 140-145.

Yoshida, T., Hazan, I., Zhang, J., Ng, S.Y., Naito, T., Snippert, H.J., Heller, E.J., Qi, X., Lawton, L.N., Williams, C.J., *et al.* (2008). The role of the chromatin remodeler Mi-2Î² in hematopoietic stem cell self-renewal and multilineage differentiation. *22*, 1174-1189.

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

Zou, G.-M., Luo, M.-H., Reed, A., Kelley, M.R., and Yoder, M.C. (2007). Ape1 regulates hematopoietic differentiation of embryonic stem cells through its redox functional domain. Blood *109*, 1917-1922.

Summary

Thalassemia and sickle cell disease (SCD) together comprise one of the most significant genetic health problems in the world. The available treatments are risky, pleiotropic, reversible, inconvenient and have many side effects. Therefore, novel approaches are needed to develop better treatments. The observation that increased fetal hemoglobin levels ameliorate these diseases has led to the idea of γ-globin reactivation as a promising approach for the treatment. Despite attempts made during last 30 years, the exact mechanism of γ-globin silencing after birth remains largely unknown. However important progress has been made in recent years. Several factors have been described with the potential to release γ-globin suppression. The most important of these is BCL11A, although it is presently not clear with what factors it interacts to exert its function on the γ-globin gene promoter. It is therefore important to identify the protein complexes bound to suppressed γ-globin gene promoter *in vivo*. To this end we have developed novel methodology, Targeted Chromatin Purification (TCP), for the purification of proteins bound to a single copy (γ-globin) locus followed by proteomic analysis.

In Chapter 2, we have optimized protein epitope tagging for cross-linked chromatin protein purification, which can be used for Chromatin Immunoprecipitation (ChIP) with or without massive parallel sequencing (ChIP-seq) and Targeted Chromatin Purification (TCP). The optimized protocol led to the genome wide identification of binding sites of chromatin associated proteins by ChIP sequencing. The epitope tagging approach overcomes the usual problems of antibody based ChIP protocols. The results obtained in Chapter 2 served as the basis for the development of Targeted Chromatin Purification (TCP), (Chapter3).

In Chapter 3, we show that promoter-bound transcription factors can be identified by purifying a gene promoter directly from mammalian cells *in vivo*. We have identified several transcription factors bound to the suppressed γ-globin gene promoter in adult cells. We show that knockdown of these factors in human primary erythroid cells indeed induces the expression of the γ-globin genes. These factors therefore provide novel targets for the development of drugs to treat β-thalassemia or sickle cell disease patients. One of these factors is ZBP-89 which we show in Chapter 6 to interact with NuRD/MeCP1 suppressor complexes, Gata-1, Bcl11a and several other novel proteins. Genome wide ChIP-seq of Bio tagged ZBP-89 shows that ZBP-89 interacts with mouse embryonic βh1 gene promoter and several other genes implicated in hematopoiesis development and differentiation such as Gata-1 and Bcl11a.

In Chapter 4, we compared the transcriptome of several thalassemic patients with differential response to Hydroxyurea (HU) treatment for increased fetal Hemoglobin (HbF). We show that high HbF is associated with a continuous stress level, and with genes that protect from stress-induced apoptosis. This suggest that HU is effective in 'responder' patients because their HbF levels are relatively high and their erythroblasts have already activated a stress response program which protects them from the cytotoxic effects of HU. We also

identify a number of genes that are consistently differentially expressed between high and low expressing HbF groups. We suggest that this list aids in the identification of factors responsible for increased HbF baseline in different individuals, factors involved in regulatory pathways and the mechanism of HbF induction by HU. This gene list may also allow a prediction of the HbF response to HU in patients.

In Chapter 5, we examined several β-thalassemia intermedia patients with high HbF levels. The already known causes of high HbF were ruled out and we investigated novel genetic polymorphism associated with the high HbF in these patients. We found an additional genetic variation in the 3' hypersensitive site (3'HS1) of the β-globin locus (+179 C>T). It creates a novel GATA-1 binding site which may explain the high HbF expression in these patients by changing the spatial organization of active chromatin hub.

Samenvatting

Thalassemie en Sikkelcelziekte behoren tot de meest frequente genetische aandoeningen in de wereld. Op dit moment zijn de ziektes niet te genezen en heeft de behandeling veel bijwerkingen. Het is daarom belangrijk dat er nieuwe behandelingen worden ontwikkeld. Onderzoek heeft aangetoond dat de aanwezigheid van foetaal hemoglobine (γ-globine) de ziekteverschijnselen kan verminderen. Om deze reden is het reactiveren van het foetale globine een veel belovende aanpak. Ondanks 30 jaar onderzoek is er nog steeds weinig bekend van het mechanisme hoe het γ-globine gen wordt uitgezet na de geboorte. Pas zeer recent is er belangrijke vooruitgang geboekt met de identificatie van een aantal factoren die de expressie van γ-globine beïnvloeden. De belangrijkste hiervan is BCL11A, een eiwit dat achter het γ-globine gen aan het DNA bindt, maar waar niet van bekend is met welke factoren in de γ-globine promoter het interacties aangaat. Het is daarom belangrijk om de eiwitcomplexen te identificeren die *in vivo* aan de foetale promoter binden. Om dit te realiseren hebben wij een nieuwe methode ontwikkeld: Targeted Chromatin Purification (TCP).Met deze techniek wordt de γ-globine promoter met de daaraan gebonden eiwitten geïsoleerd, waarna de eiwitten met behulp van massaspectrometrie worden geïdentificeerd.

In hoofdstuk 2 wordt de optimalisatie beschreven van een protocol om eiwitten met een zogenaamde 'epitope-tag' te zuiveren nadat het eiwit complex gecrosslinked is met het DNA. Dit chromatine immunoprecipitatie protocol kan samen met parallelle sequencing (ChIP-sequencing) en TCP worden gebruikt. Met behulp van dit protocol zijn de bindingsplaatsen van een aantal chromatine geassocieerde eiwitten in het genoom geïdentificeerd. De epitope-tag zorgt er voor dat problemen die normaal optreden bij eiwit zuivering voor ChIP-seq niet meer optreden. De resultaten van hoofdstuk 2 vormen de basis voor de in hoofdstuk 3 beschreven ontwikkeling van TCP.

In hoofdstuk 3 wordt de identificatie beschreven van een aantal promoter-gebonden eiwit complexen na zuivering van de γ-globine promoter uit erythroïde cellen. Een aantal van deze transcriptie factoren zijn betrokken bij het uitschakelen van het γ-globine gen in volwassen cellen, want als het niveau van deze eiwitten verlaagd wordt gaat de expressie van γ-globine omhoog. Deze factoren lenen zich daarmee voor de ontwikkeling van farmacologische interventie voor de behandeling van Thalassemia of Sikkelcelziekte. Eén van de eiwitten, ZBP-89, bindt aan het NuRD/MeCP1 complex, Gata-1, Bcl11a en verschillende andere nieuwe geïdentificeerde eiwitten zoals beschreven in hoofdstuk 6. Genoom brede ChIP-sequencing van 'Bio-tagged' ZBP-89 laat zien dat ZBP-89 bindt aan de promoter van het embryonale βh1-globine gen. ZBP-89 bindt ook meerdere genen waarvan bekend is dat ze betrokken zijn bij hematopoietische ontwikkeling, zoals Gata-1 en Bcl11a.

In hoofdstuk 4 wordt het transcriptoom bestudeerd van thalassemie patiënten. Deze patiënten zijn behandeld met Hydroxyurea (HU) om het foetaal hemoglobine (HbF) te verhogen. Een deel van de patiënten heeft baat bij deze behandeling (responders), anderen niet (non-responders). De data laten zien dat verhoging van HbF samen gaat met verhoogde

cellulaire stress en verhoging van de expressie van genen die betrokken zijn bij de bescherming tegen apoptose. HU lijkt een effectief middel dat alleen voor de patiënten die al een verhoogd niveau aan HbF hebben. Deze groep heeft een geactiveerd cellulair stress programma en daardoor zijn de erythroblasten beschermd tegen de cytotoxische effecten van HU. Ook is een aantal genen geïdentificeerd die verschillen bij patiënten uit de groep met een hoog HbF of een laag HbF niveau. Deze lijst zal van nut zijn om genen te identificeren die verantwoordelijk zijn voor de verhoging van HbF niveaus in patiënten als wel van factoren die betrokken zijn bij de regulatie van HbF inductie. De lijst is ook van prognostische waarde om de respons te voorspellen van patiënten op HU behandeling.

In hoofdstuk 5 worden thalassemie patiënten met hoge HbF niveaus onderzocht, in aanvulling op de al bekende oorzaken van een hoog HbF niveau. Hierbij is een nieuwe genetische variatie in de 3' hypersensitive site (3'HS1) van de β-globin locus (+179 C>T) gevonden. Deze verandering creëert een additionele GATA-1 bindingsplaats op het DNA die mogelijk de hoge HbF expressie van deze patiënten verklaard door de ruimtelijke organisatie van de zogenaamde 'active chromatin hub' te veranderen.

Curriculum Vitae

Personal Details

Name: Farzin Pourfarzad
Birth date: 16-07-1970
Birthplace: Tehran, Iran

Education and Research experience

1992-1996 B.Sc. Biology, major of Microbiology, Azad University, Zanjan-Iran.

1996-1998 M.Sc. Human Genetics, School of public Health, Tehran University of Medical Sciences & Health Services, Tehran-Iran.

1997- 2001 Prenatal diagnosis of common hereditary diseases in Iran, KarimiNejad and Najmabadi Genetic and Pathology Center, Tehran- Iran.

1999-2001 Instructor of molecular genetic laboratory, Genetic Research Center Welfare and Rehabilitation University, Tehran-Iran.

2000-2001 Azad University, Zanjan-Iran. Lecturer in basic genetics & Biology.

2001-2002 Post-graduate program in Human Genetics, Department of Human Genetics, Leiden University, The Netherlands.

2003-2009 PhD, Department of Cell Biology, Medical genetic Cluster, Erasmus University Medical Centre, Rotterdam, The Netherlands (Prof.Dr. F.G. Grosveld).

2009-Present Postdoctoral fellow, Department of Cell biology, Erasmus University Medical Centre, Rotterdam, The Netherlands (Prof. Dr. F. Grosveld).

List of publications

1. **EZH2-dependent chromatin looping controls INK4a and INK4b, but not ARF, during human progenitor cell differentiation and cellular senescence.** Kheradmand Kia S, Solaimani Kartalaei P, Farahbakhshian E, Pourfarzad F, von Lindern M, Verrijzer CP. *Epigenetics Chromatin. 2009 Dec 2;2(1):16.*

2. **Optimal use of tandem biotin and V5 tags in ChIP assays.** Kolodziej KE, Pourfarzad F, de Boer E, Krpic S, Grosveld F, Strouboulis J. *BMC Mol Biol. 2009 Feb 5;10:6.*

3. **The Hellenic type of nondeletional hereditary persistence of fetal hemoglobin results from a novel mutation (g.-109G>T) in the HBG2 gene promoter.** Chassanidis C, Kalamaras A, Phylactides M, Pourfarzad F, Likousi S, Maroulis V, Papadakis MN, Vamvakopoulos NK, Aleporou-Marinou V, Patrinos GP, Kollia P. *Ann Hematol. 2009 Jun;88(6):549-55. Epub 2008 Dec 3.*

4. **Screening of Iranian thalassemic families for the most common deletions of the beta-globin gene cluster.** Esteghamat F, Imanian H, Azarkeivan A, Pourfarzad F, Almadani N, Najmabadi H. *Hemoglobin. 2007;31(4):463-9.*

5. **Increased gamma-globin gene expression in beta-thalassemia intermedia patients correlates with a mutation in 3'HS1.** Papachatzopoulou A, Kaimakis P, Pourfarzad F, Menounos PG, Evangelakou P, Kollia P, Grosveld FG, Patrinos GP. *Am J Hematol. 2007 Nov;82(11):1005-9.*

6. **The cypriot and Iranian National Mutation Frequency Databases.** Kleanthous M, Patsalis PC, Drousiotou A, Motazacker M, Christodoulou K, Cariolou M, Baysal E, Khrizi K, Moghimi B, Pourfarzad F, van Baal S, Deltas C, Najmabadi H, Patrinos GP. *Hum Mutat. 2006 Jun;27(6):598-9.*

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

8. **The beta-thalassemia mutation spectrum in the Iranian population.** Najmabadi H, Karimi-Nejad R, Sahebjam S, Pourfarzad F, Teimourian S, Sahebjam F, Amirizadeh N, Karimi-Nejad MH. *Hemoglobin. 2001 Aug;25(3):285-96.*

9. **Fetal globin expression is regulated by Friend of Prmt1.** van Dijk T, Gillemans N, Pourfarzad F, van Lom K, von Lindern M, Grosveld F, and Philipsen S. *Submitted to Blood.*

10. **Functional analysis of the role of TPMT gene promoter VNTR polymorphism in TPMT gene transcription.** Zukic B, Radmilovic M, Stojiljkovic M, Tosic N, Pourfarzad F, Dokmanovic L, Janic D, Colovic N, Philipsen S, Patrinos GP, Pavlovic S. *Submitted to Pharmacogenomics.*

Summary of PhD training and teaching activities

Name PhD student: Farzin Pourfarzad Erasmus MC Department: Medical Genetic center (MGC) graduate school	PhD period:01-01-2003 until 29-12-2009 Promotor: F.G. Grosveld

1. PhD training	
	Year
General academic skills	
- Course on Laboratory Animal Science	2002
- Molecular Medicine	2004
In-depth courses (e.g. Research school, Medical Training)	
- CELERA Genome database browser	2002
- Transgenesis, gene Targeting and Gene Therapy	2003
- Comparative Gene Function Analysis	2003
- Epigenetics and chromatine	2004
- Browsing genes and genome with ensemble	2008
Presentations	
- 19th MGC symposium, September, Rotterdam	2009
- Winter School of the International Graduiertenkolleg "Transcriptional Control in Developmental processes". Kleinwalsertal	2006-7
- 3rd Iranian Congress of Genetic disorders and disabilities	2004
- Netherland Consortium for Systems Biology	2009
International conferences	
- 3rd Iranian Congress of Genetic disorders and disabilities	2004
- 15th Hemoglobin Switching	2006
- 16th Hemoglobin Switching	2008
Seminars and workshops	
- MGC Promovendi Workshops	2003-6
- Hematology department Promovendi seminar	2005
2. Teaching activities	
Supervising PhD student	
- Role of ZBP-89 in γ-globin regulation	2008-9

There are a number of people who were involved in the completion of this thesis. I would like to show my appreciation for their help and effort in the next couple of pages. I will not forget their contribution during the past several years to my work and life. With everyone there are many good memories that I would liked to mention, but for that, I would need to write another book!!!

Of course Frank comes first. Frank, thank you very much for accepting me into your lab, you saved me from a very difficult situation and gave me a project that was really my interest. Thanks for being so supportive and also critical during all these years. I was always amazed how you manage to do all of what you do and still have time for me with an open mind every time I needed your input.

Dear Sjaak, thank you for your critical suggestions and support on my project during all these years. My project benefitted greatly from your input. Thank you for organizing the Kleinwalsertal meetings, which was always great.

Dear Marieke, it was a great pleasure to work with you. You were always so helpful and encouraging that I had no other choice but to become interested in much further developed versions of foggy ideas I had in mind before talking to you.

Dear Dies, thank you for joining my committee, for reading my thesis and for all your comments. Dear Elaine and Ruud thanks for accepting to be in my committee.

Dear Ali Imam, George and John I am really grateful for all the different molecular biology techniques you have taught me, and for helping me to find my way in the lab in the beginning. Special thanks to Ali for all kinds of discussions and comments on my thesis. Dear George, thanks for taking the effort to come over and join my committee.

Petros, as you also mentioned in your thesis, we shared a lot and I have nothing more to add to that. Unfortunately, it appears we are close to an end; I already missed you during the last few months. I wish you the best in Amsterdam. Dear Ali, your collaboration in the last year of my PhD was very helpful. I am very happy that you joined me and hopefully we will continue to work on our interesting findings.

Dear Eric and Alireza, my paranimfs, thank you very much for all your help during my promotion. Alireza, I wish you the best for the rest of your PhD. Eric and Charlotte; thank you very much for the very nice French dishes from all your different roots and fun we had together. Thanks to the rest of the present and former members of lab 702; Robert-Jan, Wouter, Bas, Arnie, Sanja, Atina, Boet (also for great helps outside of the lab), Maureen, Mariette, Daan, Eric, Jorgen, Petra, Xiao, Ralph, Mary and Dorota.

I had a very good time during the years I have been working in lab 710. Thanks to Dubi, Alex, Ton, Rik, Rick, Rien, Ruud, Irem, Sylvia and John for creating a cheerful atmosphere in the lab. Dear Ton and John, thanks for the generation of the transgenic mice in one go. Ton, Alex and Rick were also helping me to deal with my personal paper works! Thank you. Thanks to dear collogues of labs 706 and 716; Nynke and Thamar (helpful technical advice and correction of smenvatting), Pavlos, Luara, Harald, Jun, Raymond, Umut, Debbie, Tiago, Erik, Ernie (always positive and cheerful, special thanks for the big chromatin preps), Katy, Patrick, Miyata, Vesna, Edgar and Javier. Special thanks to Sahar for her great help in many different ways.

Many Thanks to Ton and Pim (from Pc-support), Melle and Arthur, Marieke and Jasperina (You both are a wonderful help), Beb, Sanja, Leo, Mouse house, kitchen ladies, who always did a great job when I needed their help. There are a number of other people in the cluster 15 who made this time fruitful; Peter, Joost, Gert, Niels, Rini, Martine, Robert Jan Galjaard, Lobke (you have been a great help, thanks), Prashanth, Noorie, Ekim, Anna, Stavrus, Jeroen, Yuri, Jan, Adone, Willem (appendix to cluster 15!), Frank Sleutels, Kris, Siska, Daniel, Jeffry, Ksenija, Eva, Wendy, Cristina, Eugin, Karl, George Garinis, Eskeatnaf, Polynikis and many other people that I probably missed.

Dear Hossein and Marjan, I learned the basics of molecular biology in your lab and always enjoyed your friendship and working with you. I am really happy that you join us for my promotion, and are in my committee. THANK YOU.

Dear Rob and Tokameh, I am really grateful that you convinced me to change labs and my situation. It did work out, and it was all because of your encouragement and helpful advice. Well, that was only in the beginning of our friendship and I still benefit and enjoy your friendship. THANK YOU.

I would like to remember and acknowledge the efforts of my University teachers and colleagues in Iran, Dr. Rahnema, Prof. Farhoud, Prof. Mehdipour, Dr. Azarkeyvan, Prof. Karimi-nejad and all the colleagues in his laboratory and also in the Welfare and Rehabilitation University in Tehran which were fantastic places to work and learn. I am also grateful to the friends and colleagues in the department of human genetics in Leiden; Prof. Rune Frants, Payman, Corina, Carlos, Ko, Leny, Arja, Aart, Ivonne, Gery, Martijn, Ludo and Vilo.

Anke and Max, thank you very much for your great hospitality during my first 3 month in the Netherlands, for introducing me to Dutch culture, the Sinterklaas cadeautjes and poems, which we receive every year.

Maziar, I really enjoyed your visits during the last few years. Thanks for coming over for my promotion. I hope we can go for our next camping destination very soon, with your Goldwing of course. My best Friends, Farzin and Hengameh (we missed you during all these years), Amir, Sanaz, Mehdi, Hossein Deilami, Hossein Hojabr and his family.

Media, Maryam and Babak we had such a nice time all together during our first years in Leiden, thanks for all your helps and keeping in touch since then. Gelareh, Negin, Matt and Stefan thank you for all the nice times you spend with us, all your helps and friendships. I would also like to thank your parents Khaleh Zari, Dr. Arastafar and family Saidavi who made us feel part of their family. Thanks to the Iranian community, who made this time wonderful, Payman, Vilo, Maryam (Atakhorami), Mehrnaz and Saieed (my technical support!), Ronak, Ghazal, Farhad, Azadeh, Sevda, Parisa, Amir Hossein, Elnaz, Omid (thanks on behalf of the whole family!) and Parham specially for the lay out.

Dear Mona, I really appreciate your patience and your art for designing the cover of my thesis. My dear cousins, Sayeh, Mehrdad, Melina, Sahar and Ali you were always very encouraging, thanks for coming to my promotion.

This thesis would have not been possible without the love and support of my family. My mother and step father (Saieed), I don't know how I can say thank you; your help and your supports are always endless. My dear father, I really enjoyed your visit to us in the Netherlands, it was a great opportunity to spend some good time with you. My lovely grandma, you know how much I love and miss you. Hossein, Mohsen, Jamileh, Nahid, Maryam, Farhad and Farshad; I always wish you the best. All of you are the driving power for me and you were involved with your love and wonderful support.

Sima's family, Baba, Maman, Mohammad, Simin, Bahman, Saeeideh, Shirin, Parastoo, Saman and Arian, you also participated in all of this the same way as my own family, thank you. Special thanks to Simin for spending a lot of time and effort while we were away from Iran. I know how frustrating the task we asked you to do was, although you did it with joy.

And now the person who participated in every single step of this thesis together with me, Sima. You understood all my up and downs, although you had to deal with your own! The first 3 months without you in the Netherlands was enough for me to realize that I can't do the rest without you. My little Jasmin, you grow up with this thesis and add an additional taste to our life. I enjoy it every time you call me Babish. I love you both.

Farzin

5-Aza	5-Azacytidine
a.u.	arbitrary units
ACH	Active Chromatin Hub
Apex1	Apurinic/apyrimidinic endonuclease 1
ARG1, ARG2	Arginase 1 and 2
ASS1	Argininosuccinate Synthetase 1
Bio tag	Biotin tag
BirA	Biotin ligase
C/EBP	CCAAT/Enhancer-Binding Protein
CBF	CCAAT-Binding Factor
CBP	CREB Binding Protein
CDC5L	Cell Division Cycle 5-like
CDP	CCAAT Displacement Protein
CH	Chromatin Hub
ChIP	Chromatin Immunoprecipitation
Ctnnbl1	Beta catenin like 1 protein
DNMT	DNA methyltransferase
DNMT3A	DNA Methyltransferase
DR-1	Direct Repeats
DRED	Direct Repeat Erythroid Definitive
EB	Embryonic Body
EKLF	Erythroid Kruppel Like Factor
EPO	Erythropoietin
FANC	Fanconi Anemia Complementation group
FDA	Food and Drug Administration
FDR	False Discovery Rate
FOXO3	Forkhead box O3
GC	Guanylate Cyclase
H3ac	Histone 3 acetylated
H3K4me3	Histone 3 lysine 4 methylated
HA tag	Hemagglutinin tag
HAT	Histone Acetyltransferase
Hb	Hemoglobin
HbA	Adult Hemoglobin
HbF	Fetal Hemoglobin
HbS	Sickle Hemoglobin
HBV	Human Hepatitis B Virus
HDAC	Histone deacetylase
HEP	Human Erythroid Progenitor cells
HIPK2	Homeodomain Interacting Protein Kinase 2
HPFH	Hereditary Persistence Fetal Hemoglobin
HPLC	High Performance Liquid Chromatography
HS	DNase I Hypersensitive Sites
HU	Hydroxyurea
IHC	Immunohistochemistry
iPS	induced Pluripotent Stem
LARC	LCR Associated Remodeling Complex
LCR	Locus Control Region
MBD2	Methyl-CpG Binding Domain 2 protein
MeCP1	Methyl-CpG Binding Protein Repressor complex 1

NAP	Nucleosome Assembly Protein
NO	Nitric oxide
NOS	NO synthase
NRE	Negative Regulatory Element
NuRD	Nucleosome Remodeling and Deacetylating
PICh	Proteomics of Isolated Chromatin segments
PYR	PYR complex
QTL	Quantitative Trait Loci
RBC	Red Blood Cells
RMA	Robust Multi-Array average
SCD	Sickle Cell Disease
SCF	Stem Cell Factor
SSP	Stage Selector Protein
TCP	Target Chromatin Purification
TetO	Tetracycline Operator sequence
TetO-γ	TetO-modified γ-globin minilocus
TetR	Tetracycline Repressor protein
TetR3T	Triple Tag TetR protein
Yb-1	Y-box-binding protein 1