Characterization of the molecular functions of the transcription factor GATA-1 in hematopoiesis

Χαρακτηρισμός των μοριακών μηχαισμών δράσης του μεταγραφικού παράγοντα GATA-1 στην αιμοποίηση

PhD Thesis

Elena Karkoulia

Supervised by:

Prof. Achilleas Gravanis

Prof. George Mavrothalassitis

Dr. John Strouboulis

To my uncle Bernhard Oesch

CONTENTS

Chapter 1

Introduction

Hematopoiesis is a dynamic process where initiation and maintenance of hematopoietic stem cells, as well as their differentiation into erythroid, myeloid and lymphoid lineages, are tightly regulated by a network of transcription factors. Understanding the genetic controls of hematopoiesis is crucial as perturbations in hematopoiesis lead to diseases such as anemia, porphyria, thrombocytopenia, or malignancies, including leukemias and lymphomas. Animal models, conventional and conditional knockout mice, have been of particular use towards our understanding of the genetic controls of hematopoiesis. However, knockout mice for many of the key hematopoietic transcription factors, including GATA-1, are embryonic lethal, thus precluding the analysis of their roles during hematopoiesis. The long term scope of this work is to study in a much more systematic way GATA-1 functions in regulating the differentiation of distinct hematopoietic lineages. To these ends, we use the biotinylation tagging of GATA-1 as a uniform platform for the proteomic and global gene target analysis by establishing a system for the expression of physiological levels of biotin-tagged GATA-1 in all hematopoietic lineages where it is active by a knock-in approach. Herein, we describe the biotin tag knock-in approach we followed and the data we obtained from the characterization of the molecular functions of GATA-1 in the murine erythroid lineage. However, the study of the erythroid compartment cannot be separated from the hematopoietic system as a whole and therefore this introduction presents the hematopoietic system, with a focus on the erythroid lineage and GATA-1 functions.

1.1 Hematopoiesis

Hematopoiesis is the process of generation of all the mature blood cells from a rare pool of hematopoietic stem cells (HSCs). Mature blood cells have a limited lifespan and need to be constantly replenished in order to maintain homeostasis of the organism (Zon, 2001). There are at least eight types of blood cells, varying in their appearance and function (Figure 1, Ferreira et al., 2005). The most abundant cells in the blood are the red blood cells or erythrocytes, occupying 45% of its volume, whereas the rest of cell types -white blood cells- occupy about 1% of the blood volume and include the platelets, the granulocytes (neutrophils, eosinophils and basophils), the monocytes and the lymphocytes.

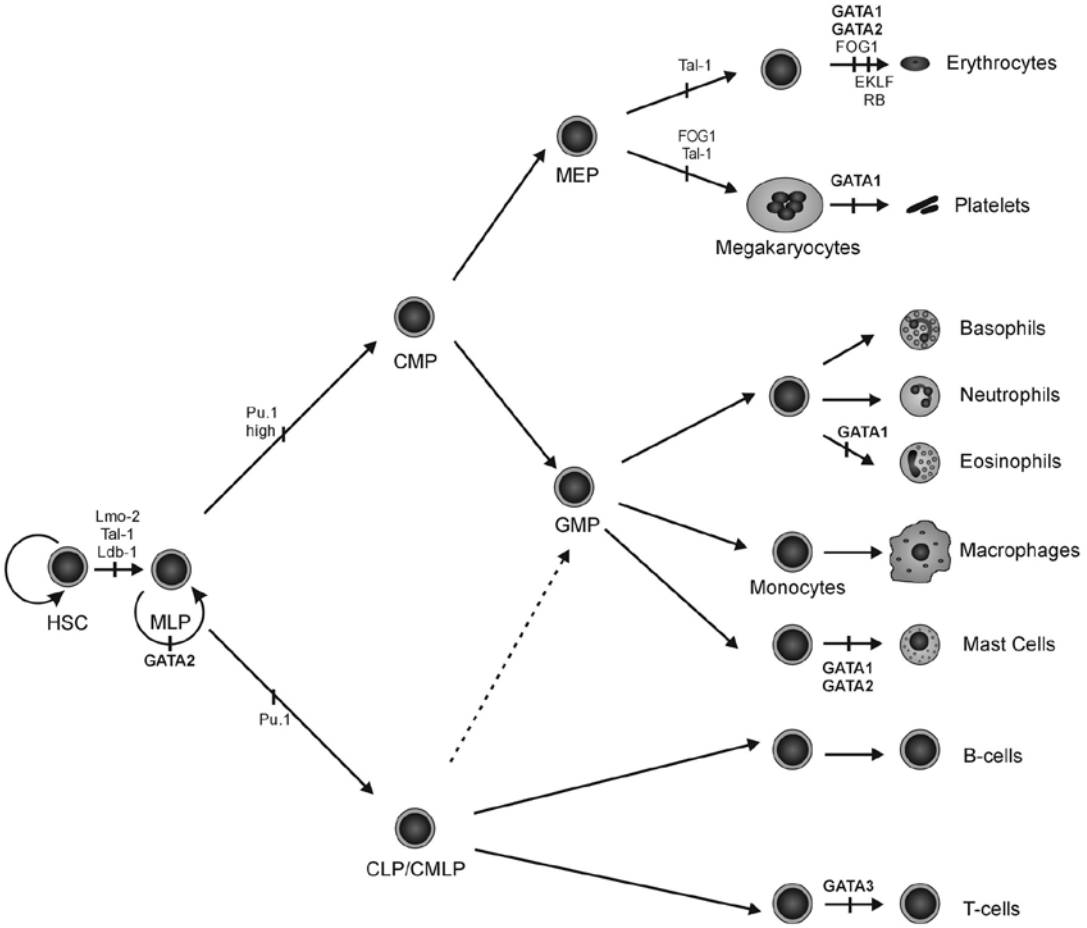


Figure 1: The hematopoietic tree.

Schematic representation of the main lineage commitment steps in hematopoiesis. The hematopoietic stem cell (HSC) is the basis of the hematopoietic hierarchy and gives rise to multilineage progenitors (MLP), which can differentiate into all the hematopoietic lineages. MLPs become lineage restricted to the lymphoid and myeloid lineages in the common lymphoid progenitor (CLP) and common myeloid progenitor (CMP), respectively. CLPs can give rise exclusively to B and T cells, while CMPs can give rise to megakaryocyte-erythrocyte progenitors (MEP) and granulocyte-monocyte progenitors (GMP). Alternatively, it is also believed that the first lineage commitment separates myeloid and erythroid potential, in the CMP, from myeloid-lymphoid potential, in the common myeloid lymphoid progenitor (CMLP). CMLPs can then further differentiate in B cells, T cells, and GMPs (dashed line). Hematopoietic GATA factors and GATA1 cofactors relevant for the development of particular hematopoietic lineages are indicated (Ferreira et al., 2005).

Megakaryocytes and erythroid cells are produced in the bone marrow and their final products of differentiation, platelets and red blood cells respectively, are released in the blood stream. Platelets serve in wound healing and vessel fissuring by clot formation and also inflammatory responses. Red blood cells are specialized cells that lose their nucleus to maximize space for gas transportation.

Lymphocytes are part of the innate and adaptive immune system that confers protective function to the organism. B and T cells are produced in the bone marrow and complete their maturation in the spleen, thymus, lymph nodes and lymph tissue. Upon recognition of an antigen, B cells are activated to produce antibodies against it. There are two types of T cells, helper T cells (Th) that interact with a) activated B cells and nurse their differentiation and antibody production and b) phagocytes in the elimination of pathogens or dendritic cells in the antigen presenting process and cytotoxic T cells (Tc) that are involved in the destruction of infected host cells. Another type of cells, related to T cells, are the natural killer cells (NK), which are quite unspecifically reactive to tumour cells or virus-infected cells and induce cytotoxicity to the target cells.

The maturation and function of B and T cells is closely related to mononuclear phagocytes, which include the dendritic cells and the macrophages. They are both produced in the bone marrow and migrate to different tissues to exert their functions. Dendritic cells endocytose antigens and present them to B and Th cells in the lymph nodes or spleen. Macrophages are phagocytic cells that manage the cell waste in the tissue of residency. They remove senescent erythrocytes, leukocytes, and megakaryocytes by phagocytosis and digestion.

Neutrophils and eosinophils also possess phagocytotic properties. They are both formed in the bone marrow and circulate in the blood. Neutrophils contain lysosomes that are very effective in the destruction of circulating bacteria. Eosinophils can also phagocytose and destroy ingested microorganisms, however their main function is triggering of the response against parasites that are not phagocytosed by specialized cells due to their large size. Eosinophils are recruited to the surface of the parasite, via chemotaxis, where they release their granules that contain cytotoxic agents. Moreover, they release substances that counteract the stimuli originating from mast cells and basophils.

Mast cells and basophils are also granulocytes and exert similar function, differing in that basophils circulate in the blood stream whereas mast cells are immobilized in the tissues. Upon activation, they release mediator histamines, which attract eosinophils (Zon, 2001).

Remarkably, all classes of blood cells derive from a single common multipotent progenitor, the so-called hematopoietic stem cell (HSC) (Lemischka et al., 1986), which was recently shown to emerge from the hemangioblast. In murine embryos, a common precursor, the hemangioblast, gives rise to endothelial, blood and vascular smooth muscle cells. Recent advances in live cell imaging allowed time-lapse visualization of mammalian HSCs, emerging directly from endothelial cells (Eilken et al., 2009).

HSCs have two main features: a) self-renewal, as they possess a high proliferative potential and b) pluripotency, as they can differentiate to any blood cell type through a hierarchical series of lineage decisions via different progenitor cells (Lemischka et al., 1986). Although some other multipotent hematopoietic progenitor cells can also generate daughters with the same lineage potential for limited periods of time, it is only the HSCs that can self-renew for long enough so as to support the hematopoietic system for a lifetime (reviewed by Schroeder, 2010).

During vertebrate ontogeny, hematopoiesis occurs in successive but overlapping waves at distinct anatomical niches (Figure 2, Costa et al., 2012). Overall, the hematopoietic process is distinguished in primitive and definitive hematopoiesis based on the type of blood cells generated. Primitive hematopoiesis takes place during embryogenesis, is transient, and results in the generation of unipotent blood cells directly from the mesoderm. It occurs in the extraembryonic yolk sac blood islands and results in the formation of primitive macrophages -the first line of defense of the organism- and primitive erythrocytes, which transport oxygen. Definitive hematopoiesis occurs in two distinct waves and results in the generation of multipotent blood cells with diverse lineage potential to support blood cell development throughout the lifespan of the organism. The first wave of definitive hematopoiesis produces a transient population of cells, the erythroid-myeloid progenitors (EMPs) in the yolk sac and the fetal liver. The second wave of definitive hematopoiesis occurs at the hemogenic epithelium of the embryo, which includes the aorta-gonad-mesonephros (AGM) region, the yolk sac and the placenta and produces the HSCs, which enter circulation and migrate to the fetal liver to support embryonic hematopoiesis. Soon before birth and throughout adulthood, hematopoiesis occurs at the bone marrow, where HSCs differentiate to lineage-committed progenitors and self-renew (reviewed by Sood & Liu, 2012; Costa et al., 2012, Baron et al., 2012).

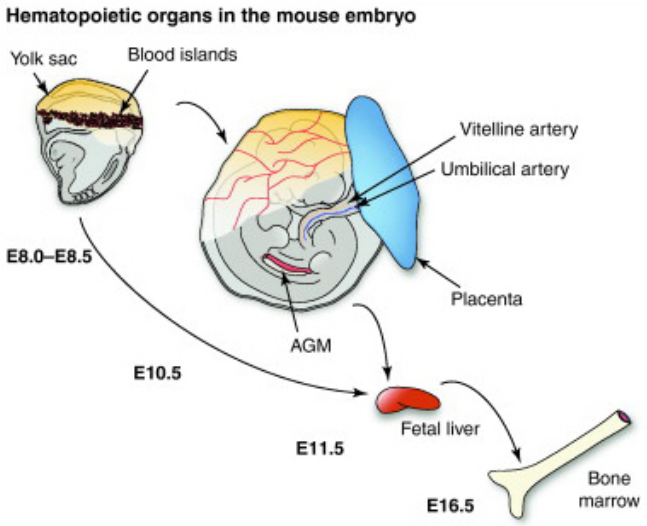


Figure 2: Embryonic sites of blood development.
In murine embryos, hematopoiesis takes place in several tissues where blood cells are generated and/or undergo maturation. The first hematopoietic progenitors are found extraembryonically at E8.0–E8.5, in the yolk sac blood islands in close proximity to emerging endothelial cells. Once circulation is established, blood cells colonize other developing hematopoietic organs. Around E10.5, the AGM, placenta, umbilical artery and vitelline artery initiate the generation of blood precursors that, together with yolk sac cells, colonize to the fetal liver rudiment around E11.5. The fetal liver is the major hematopoietic site where blood progenitors expand and/or mature. Finally, the bone marrow is colonized by precursors from the fetal liver before birth and remains the main hematopoietic niche throughout adult life (Costa et al., 2012).

1.2 Erythropoiesis

Erythropoiesis is the generation of mature red blood cells from the HSCs through a multistep, tightly regulated process. HSCs and multipotent progenitors (MLP) generate the common myeloid progenitors (CMP), which progress into megakaryocytic/erythroid progenitors (MEP). MEP cells, upon the action of growth factors and the combinatorial action of certain transcription factors, are differentiated into erythropoietin (Epo)-responsive burst forming unit-erythroid (BFU-E) and colony forming unit-erythroid (CFU-E). The CFU-Es are then differentiated into orthochromatic normoblasts and finally via enucleation yield reticulocytes and then red blood cells (Figures 1, 3). As erythropoiesis progresses, erythroid cells gradually lose their proliferative potential, mature and eventually extrude their nuclei. Mature erythrocytes are biconcave disks without mitochondria or organelles, full of hemoglobin for maximal capacity to transport oxygen to the tissues.

Erythropoiesis occurs in multiple waves and at distinct anatomical locations. The first wave occurs in blood islands in the yolk sac in both mouse and humans and produces primarily primitive erythroid nucleated cells (EryP). The second wave is also transient and arises in the yolk sac but is definitive, composing enucleated erythroid lineages. The third wave emerges from HSCs in the fetal liver, includes enucleated erythroid cells, and during ontogeny moves from the liver to the spleen and finally to the bone marrow (Fraser et al., 2007; McGrath et al., 2011; reviewed by Baron et al., 2012).

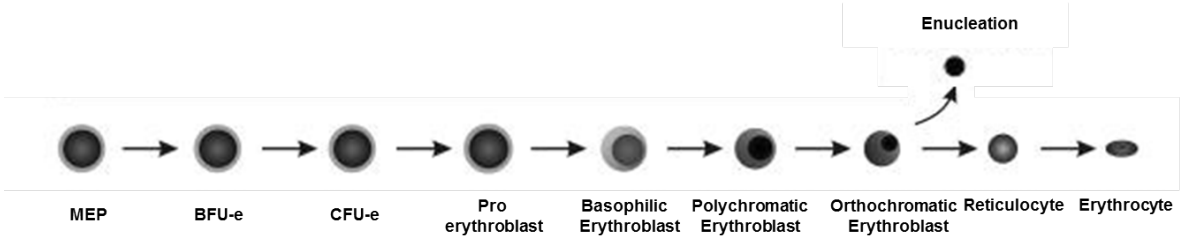


Figure 3: Schematic representation of the different stages in erythroid differentiation.
Megakaryocytic/Erythroid progenitors (MEP) give rise to burst forming unit-erythroid (BFU-E) and colony forming unit-erythroid (CFU-E), which are erythroid restricted progenitors. Terminal erythroid differentiation starts with the proerythroblast and terminates upon enucleation and the formation of erythrocytes (adapted from Ferreira et al., 2005).

The proerythroblast is the earliest morphologically identifiable erythroid progenitor, with large size (14-19μm and 11-13μm in diameter in human and mouse, respectively) and large nucleus (occupies 80% of the cell volume) with visible nucleoli, surrounded by basophilic cytoplasm. Basophilic erythroblasts are slightly smaller cells (12-17μm and 10-11μm in diameter in human and mouse, respectively), where condensation of chromatin begins to occur, giving the nucleus a granular appearance. The cytoplasm is highly basophilic due to the presence of large numbers of ribosomes. Hemoglobinization starts at the polychromatic stage, in which stage the cells (12-15μm and 9-10μm in diameter in human and mouse, respectively) and their nuclei are visibly smaller. Full hemoglobinization occurs at the orthochromatic cells, which are the smallest nucleated erythroid cells (8-12μm and 7-8μm in diameter in human and mouse, respectively) with very condensed chromatin and shrunk nuclei. Finally the nucleus is lost at the reticulocyte stage. Reticulocytes are slightly larger than the erythrocytes (7-8μm and 3.5-4.5μm in diameter in human and mouse, respectively), have irregular shapes and contain certain cytoplasmic organelles. The overall process, from the proerythroblast to the reticulocyte stage, takes 48-72 hours to complete and it is after another 48 hours that the cytoplasm becomes devoid or organelles and the cell acquires the discoid shape of the erythrocyte (Allen & Dexter, 1982) (Figure 3).

Definitive erythropoiesis occurs in the erythroblastic islands (blood islands). As observed by Bessis in the late 1950s, erythroblastic islands, which are composed of a central macrophage surrounded by erythroblasts, represent the primary erythroid niche within the developing fetus and in the adult bone marrow (Bessis, 1958). Electron microscopy defined the architecture and distribution of cells within the erythroblastic islands and provided evidence for the absolute association of the erythroid cells and the macrophages (Figure 4, Bessis et al., 1978). The more immature erythroid precursors are located closer to the center and as they mature they move away from the body of the macrophage. Extensive macrophage-erythroblast and erythroblast-erythroblast adhesive interactions are necessary for a thriving definitive erythropoietic compartment (Palis, 2004).

One obvious function of the macrophage is phagocytosis of the nuclei expelled by the erythroid precursors during differentiation. However, emerging evidence indicates that erythropoiesis is regulated by the balance of positive and negative feedback regulatory mechanisms within the erythroblastic islands, which are mediated via the macrophage-erythroblasts interactions (reviewed by Tsiftsoglou et al., 2009)

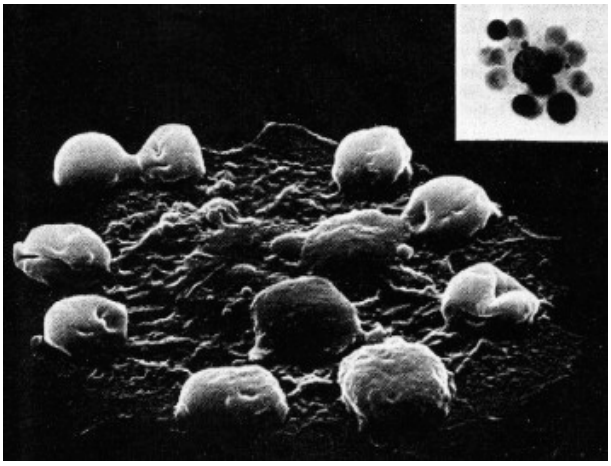


Figure 4: Scanning electron micrography of an erythroblastic island from rat bone marrow. Erythroblastic island after 2 hours *in vitro* culture examined by scanning electron microscopy. Top left insert: the same island seen with light microscopy. Note the two nuclear extrusions, one early (bottom) and the other almost complete (top left) (Bessis et al., 1978).

Whether maturation of the primitive erythroblasts also occurs in proximity to macrophages in the yolk sac or whether they migrate soon after generation to the definitive erythroblastic islands for assistance with their enucleation, has been for long under investigation. Recent work demonstrated that circulating primitive erythroblasts can also interact with erythroblastic islands in the fetal liver during maturation and enucleation (Isern et al., 2008). Moreover, no erythroblastic islands have been identified in the yolk sac so far and should not be confused with the primitive erythroid niche, the yolk sac blood islands. These islands originate from the hemangioblast and are clusters of primitive erythroblasts surrounded by an endothelial covering, nested between the endoderm and the mesoderm of the yolk sac (reviewed by Ferkowicz & Yoder, 2005; Manwani & Bieker, 2008).

A combination of signals and transcription factors define the transition from the primitive to the definitive erythropoietic wave. It is well established that it is hypoxia that primarily promotes erythropoiesis via an increase of erythropoietin (Epo) production. During development of the embryo, all necessary nutrients and oxygen are supplied by the mother through the placenta under highly hypoxic conditions. The first red blood cells appear in the circulation of the embryo to contribute towards the adaptation of the fetus to the changing needs for oxygen. There is a linear correlation between the production of Epo and the increase of hypoxia. EpoR knockout mice showed normal primitive erythropoiesis, whereas definitive was arrested at the CFU-E stage (Lin et al., 1996). During primitive erythropoiesis, erythroblasts are produced in a synchronized manner to support the needs of the embryo in nutrients and oxygen. As development progresses, hypoxia levels increase and so do the levels of Epo produced, resulting in a switch from primitive to definitive erythropoiesis, which will remain the main repository of erythroid cells from that moment on and after birth throughout adulthood. The levels of hypoxia in the fetus remain high, as are the levels of Epo and this is why the erythroid cells are highly proliferative in the embryo. Epo responsiveness remains high in definitive erythropoiesis and affects the CFU-E precursors (Palis & Segel, 1998). Later through erythroid differentiation the levels of Epo decrease due to GATA-2 and NF-kB upregulation, which bind to upstream enhancers of the Epo gene, inhibiting its transcription (reviewed by Chateauvieux et al., 2011).

Moreover, the transition from primitive to definitive erythropoiesis coincides with a switch in globin expression. Globin chains and heme form the hemoglobin tetramers that bind oxygen or carbon dioxide. Hemoglobins are classified into embryonic, fetal and adult ones and differ in their affinity to bind the gas molecules. Primitive erythrocytes in human express the embryonic hemoglobins [Gower 1 ($\zeta 2\varepsilon 2$), Gower 2 ($\alpha 2\varepsilon 2$) and Portland ($\zeta 2\gamma 2$)], definitive erythrocytes of fetal origin express the fetal hemoglobin ($\alpha 2\gamma 2$) and definitive erythrocytes in the adult express hemoglobin which is consisted of two α and two β globins ($\alpha 2\beta 2$). It should be noted though that globin switching is not lineage dependent. It has been reported that primitive erythroid cells, in addition to embryonic chains, produce fetal and adult globins and that definitive erythroid cells, in addition to fetal and adult globins, produce embryonic globins (Stamatoyannopoulos et al., 1987).

Transcription factor function during erythropoiesis involves not only the activation of the erythroid specific genes but also the suppression of the alternative lineage transcriptional programs of the HSCs and the multipotential hematopoietic progenitors as well as the suppression of cell division. Thus, transcription factors and cofactors serve as activators as well as repressors in directing erythroid differentiation. The activities of a number of key transcription factors and cofactors in erythropoiesis will be discussed below.

Models to study erythropoiesis

The study of erythropoiesis can be approached either directly in an animal model (*in vivo*), or in cultured primary cells isolated from an animal (*ex vivo*) or in cell lines that resemble the *in vivo* system (*in vitro*). Various cell lines and animal models have been employed to study erythropoiesis, however for the purpose of this thesis, emphasis will be given in the MEL, G1E and ES cells and the mouse model utilized herein.

Cell lines

Permanent cell lines grown in culture maintain their neoplastic phenotype and behave like a rather homogeneous pool of cells. Upon addition of certain growth factors or suitable inducers they can differentiate to cells resembling normal counterparts, such as orthochromatic normoblasts.

Murine erythroleukaemia (MEL) cells are Friend leukemia virus-transformed spleen derived cells arrested in the proerythroblast stage, that start differentiating and hemoglobinizing upon stimulation with several agents. Cells committed to differentiation undergo loss of proliferation, selective expression/repression of sets of genes and production of vast amounts of hemoglobin. However, MEL cells do not respond to Epo and terminally differentiate into reticulocytes, with only a small portion of the population becoming enucleated (Antoniou, 1991).

G1E cells are GATA-1 null erythroblasts derived from gene-targeted embryonic stem cells. These cells exogenously express the human antiapoptotic Bcl-2 gene and continuously proliferate when in culture supplied with stem cell factor (SCF) and low amounts of Epo. They start differentiating in a fashion that largely recapitulates normal erythropoiesis, after inducible activation of a GATA-1-ER fusion protein and under higher amounts of Epo (Weiss et al., 1997).

Mouse ES cells have been extensively used for *ex vivo in vitro* differentiation into hematopoietic lineages (reviewed by Olsen et al., 2006). In general, two experimental approaches are widely applied. In the first approach pioneered by Keller and colleagues, hematopoietic cells are derived from disaggregated embryoid bodies (EB) cultured in the presence of specific cytokines and growth factor combinations that promote growth and differentiation of specific hematopoietic lineages (Wiles and Keller, 1991; Keller, 1993). In the second approach pioneered by Nakano and colleagues, ES cells are cultured in the presence of the OP9 stromal cell line which promotes hematopoietic differentiation by providing a supportive hematopoietic environment (Nakano et al., 1996; Kitajima et al., 2003). It should be noted that in ES cell derived protocols (particularly in the OP9 based system), hematopoietic differentiation proceeds synchronously. Protocols have been described for the *in vitro* differentiation of ES cells into all of the hematopoietic lineages that GATA-1 has been functionally associated with, including dendritic cells (Olsen et al., 2006 and references therein). Primitive erythroid cells are distinguishable by virtue of their large nucleated morphology and the expression of embryonic globin genes, such as βh1. Definitive erythroid cells are smaller and express adult stage globin genes, such as βmajor. Addition of dexamethasone in the culture, which mimics a stress erythropoiesis response, leads to an expansion of erythroid progenitors and improved yields of erythroid cells (Carotta et al., 2004).

Mice

Various animal models, including xenopus, zebrafish, avian systems (primarily chicken) and mice, have been used to study hematopoiesis/erythropoiesis. The production of red blood cells in mice occurs in two successive waves and includes several stages that highly resemble human erythropoiesis and as such, mice are a very good system for *in vivo* studies of erythropoiesis.

In mice, the first erythroid cells appear at day 7.5 of gestation in the blood islands in the yolk sac, are nucleated and complete their maturation, including enucleation, in the blood stream. Circulating, enucleated primitive red blood cells can be detected up to day 15 of gestation (Palis & Yoder, 2001, Fraser et al., 2007). Definitive erythroid cells appear at day 10.5 of gestation in the fetal liver and by 12.5 dpc definitive enucleated erythrocytes can be detected in the blood. During ontogeny, definitive erythropoiesis moves from the liver to the spleen and finally to the bone marrow (15 dpc).

Primitive erythrocytes in mice express the embryonic haemoglobins, which consist of ζ- and βh1-globins at E7.5, followed by α1-, α2- and εy globins as they mature into reticulocytes by E15.5. Definitive erythroid cells express the adult haemoglobins, which consist of α1-, α2- globins and βmajor-, βminor- globins (reviewed by Tsiftsoglou et al., 2009; Sankaran & Orkin, 2010). Of note, there is some expression of the definitive erythroid globin genes within the primitive erythroid lineage, but this expression is at low levels and the extent of expression increases dramatically with robust production of red cells from the definitive lineage (Kingsley et al, 2006).

Transcriptional control of erythropoiesis

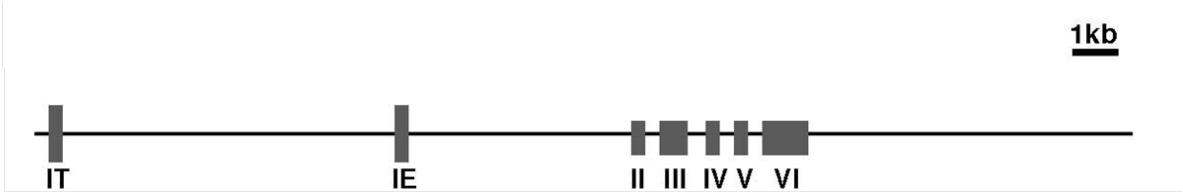
1.3 GATA-1

GATA-1 (also known as Eryf-1, NF-E1, NF-1 and GF-1) is a critical transcription factor for erythroid differentiation. It is the founding member of the GATA family of proteins, which consists of six transcription factors, GATA-1 to GATA-6 (reviewed by Morceau et al., 2004; Lowry & Mackay, 2006 and references therein). They all bind to the DNA consensus sequence (A/T)GATA(A/G) by two characteristic C_4 (Cys-X_2-Cys-X_{17}-Cys-X_2-Cys) zinc-finger motifs, referred to as the GATA fingers. All members of the GATA family share sequence homology in the zinc finger regions only; outside those, the conservation between GATA factors is low. GATA-1 is expressed in the hematopoietic system mainly; GATA-2 and GATA-3 are expressed in various tissues, including the hematopoietic system, whereas GATA-4 to GATA-6, are expressed in other mesodermal and endodermal tissues.
GATA-1 was isolated as a protein with binding specificity to the 3' enhancer of the β-globin locus and was cloned from a MEL cell line cDNA library. The human homologue was cloned soon after and assigned to the X-chromosome at position Xp21-11. The murine homologue is also located on the X-chromosome.

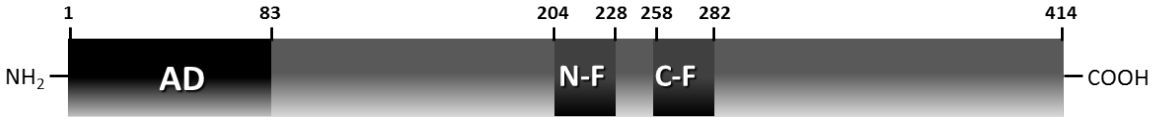
GATA-1 is expressed in erythroid cells, megakaryocytes, mast cells, eosinophils and dendritic cells in hematopoiesis and in the Sertoli cells in the testis. The promoter of the murine GATA-1 gene has a CACC box and a double palindromic GATA site, which suggests an autoregulatory role for GATA-1. The net effect of this autoregulation, whether it is positive or negative, remains controversial as it depends on the cell type and/or the stage of differentiation of the cells. For example, a targeted deletion of the palindromic GATA site results in loss of the eosinophil lineage only (Yu et al., 2002) whereas overexpression of GATA-1 in the erythroid lineage causes downregulation of the endogenous GATA-1 (Whyatt et al., 2000). The GATA-1 gene has two tissue specific promoters: a distal promoter, which contains the first exon and is specifically transcribed in the Sertoli cells and a proximal promoter, which encodes a shorter transcript than the one transcribed in the testis and is found in the erythroid cells only (Figure 5).

Three functional domains have been described for GATA-1 protein: an N-terminal domain reported to act as a transcriptional activation domain in transient transcription assays (Shimizu et al., 2001) and two zinc finger domains located toward the C-terminus of the protein (Mackay et al., 1998). The C-terminal-most, zinc finger is responsible for binding to DNA, whereas the N-terminal zinc finger modulates and stabilizes DNA binding, for example, to more complex palindromic GATA motifs. Both zinc fingers are involved in protein-protein interactions.

GATA-1 possesses an alternative translation initiation site located at methionine 84. Translation from this alternative site results in a 40kD protein, termed as GATA-1s, which lacks the 83 N-terminal most amino acids that correspond to the transactivation domain. This protein is found in hematopoietic tissues, where GATA-1 is active, albeit at much lower levels than GATA-1 (Calligaris et al., 1995).



A. GATA-1 allele



B. GATA-1 protein

Figure 5: GATA-1 gene (A) and protein structure (B).
A: Schematic representation of the murine GATA-1 gene. IT, First exon, testis; IE, first exon erythroid.
B: The murine GATA-1 protein domain structure. AD, activation domain; N-F, N-terminal most zinc-finger; C-F, C-terminal most zinc finger.

GATA-1 mice

GATA-1 is essential for erythroid differentiation and various GATA-1 -low, -null or -high mice have been generated to study the GATA-1 phenotype in erythropoiesis.

GATA-1 null mice are embryonic lethal with embryos dying around E10.5-11.5 due to severe anemia. Primitive erythropoiesis appears normal, however defective proliferation of the primitive erythrocytes was observed in blood (Fujiwara et al., 1996). GATA-1 null heterozygous females are born at expected rates, however suffer from mild anemia, from which they recover during adulthood (Pevny et al., 1991; Fujiwara et al., 1996). Detailed genetic analysis showed that GATA-1 null cells, primitive and definitive, fail to undergo complete differentiation, are arrested at the proerythroblast stage and subsequently die due to apoptosis (Weiss et al., 1994; Pevny et al., 1995). Colony assays showed that GATA-1 knockout cells were able to differentiate to macrophages and neutrophils. GATA-1 null megakaryocytes were also formed, however at lower rates, whereas mast cells were normally produced. It should be noted though, that the latter two lineages in addition to GATA-1 express GATA-2, which compensates for the absence of GATA-1 and rescues the GATA-1 null phenotype (Pevny et al., 1995).

GATA-1 levels fluctuate during hematopoiesis and erythropoiesis: they are basal in the hematopoietic progenitors, increase in erythroid committed cells (peak at the CFU-E stage) and decrease during terminal differentiation (Ferreira et al., 2005). Various mouse models have been generated to study alterations in the levels of GATA-1. GATA1.05 mice carry a disruption in the erythroid promoter of GATA-1 (Figure 5) resulting in a less than 5% GATA-1 expression. Male mice die at E12.5, whereas heterozygous female mice are born, have a shorter lifespan, display anemia, thrombocytopenia and splenomegaly (Takahashi et al., 1997; Suwabe et al., 1998; Takahashi et al., 1998) and gradually develop leukaemia (Shimizu et al., 2004). *In vitro* assays showed that the GATA1.05 derived cells escape apoptosis, however remain blocked at the proerythroblast stage (Suwabe et al., 1998). It appears thus that as little as 5% of GATA-1 expression, is sufficient to prevent cells from apoptosis, yet not enough to drive differentiation.

Decreased expression of GATA-1 was also studied in mice with disrupted hypersensitive site I (HS I) in the promoter of GATA-1 due to insertion of a neo cassette (Shivdasani et al., 1997). These mice express 4-5 folds less GATA-1 than the wild type counterparts and die at E10.5-11.5. Nevertheless, 5% of male mice are born, half of which die within two days after birth and the rest survive to adulthood suffering from severe anemia and reduced number of platelets (McDevitt et al., 1997; Shivdasani et al., 1997; Vyas et al., 1999). Heterozygous females have thrombocytopenia and splenomegaly, with spleens containing higher numbers of BFU-Es and CFU-Es than wildtypes.

Using chromosome-X-inactivation of a GATA-1 transgene under the regulation of the LCR and the β-globin promoter, mice that overexpressed GATA-1 were generated. Heterozygous females are able to survive adulthood, whilst transgenic males die between E12.5 to E13.5 from severe anemia. Overexpression of GATA-1 in the erythroid lineage resulted in a block in differentiation and subsequently to lethal anemia (Whyatt et al., 2000).

GATA-1 target genes

Chromatin immunoprecipitation (ChIP) coupled to massive parallel sequencing (ChIP-seq) has provided the tools to obtain a genome wide comprehensive view of how transcription factors interact to drive tissue-specific differentiation. ChIP sequencing studies on GATA-1 using various, mostly erythroid cells, hematopoietic cell lines have revealed numerous known and novel GATA-1 direct gene targets (reviewed by Kerenyi & Orkin, 2010). An overview of known GATA-1 gene targets is given below, whereas in chapter 3 of this thesis data from the ChIP-sequencing experiments conducted to identify novel *in vivo* GATA-1 targets in the erythroid lineage are presented.

GATA-1 autoregulation

GATA-1 levels vary from basal in the hematopoietic progenitors to maximal in erythroid committed cells and low during terminal differentiation (Ferreira et al., 2005).

In vivo studies have shown that the levels of GATA-1 protein are critical for erythroid differentiation to occur, as low amounts of GATA-1 lead to apoptosis and overexpression of GATA-1 inhibits differentiation (Weiss et al., 1994; Pevny et al., 1995; Whyatt et al., 2000). Nevertheless, GATA-1 overexpressing erythroid cells overcome impaired differentiation *in vivo* in the presence of wildtype cells. A homeostatic model was proposed whereby mature erythroid cells generate a signal to the more immature cells to degrade GATA-1 and accomplish the last stages of differentiation (Gutierrez et al., 2004).

Rescue experiments in GATA-1 null embryonic stem cells have demonstrated the requirement of an intact GATA element in the distal promoter, suggesting autoregulatory control of GATA-1 transcription. Further characterization provided evidence that GATA-1 expression in these cells does not involve additional regulatory factors and is maintained entirely by an autoregulatory loop (Baron & Farrington, 1994). In another study, deletion analysis identified a 317 base pair cis-acting regulatory element, sufficient to activate reporter gene expression in primitive and definitive erythroid cells and megakaryocytes in transgenic mice. HS I contains multiple potential consensus binding sites for regulatory factors including a composite GATA-E-box motif that binds a multiprotein complex including GATA-1, SCL/tal-1,E2A, Lmo2, and Ldb-1 (Wadman et al., 1997; Vyas et al., 1999).

GATA-1 and anti-apoptotic genes

Lineage committed cells need to activate a number of anti-apoptotic pathways in order to counteract apoptotic signals which are intrinsic to cell lineage differentiation induction (Nagata, 1997; Ashkenazi & Dixit, 1998). The hematopoietic system needs to tightly regulate the production of cells of each lineage so that no surplus of a particular lineage occurs and thus lineage differentiation is balanced between survival and apoptosis (Testa, 2004).

GATA-1 null cells, primitive and definitive, fail to undergo complete differentiation, are arrested at the proerythroblast stage and subsequently die due to apoptosis (Weiss et al., 1994; Pevny et al., 1995). However, in the presence of exogenously expressed human Bcl-2, an anti-apoptotic factor, they survive and remain proliferative (Weiss et al., 1997). GATA-1 directly regulates the transcription of Bcl-2 and Bcl-xL, both members of the Bcl family of anti-apoptotic factors (Weiss et al., 1997; Gregory et al., 1999; Tanaka et al., 2000). In addition, Bcl-xL is also regulated upon EpoR activation, which in turn is induced by GATA-1 (Zon et al., 1991; Komatsu et al., 1997). Bcl-xL null ES cells fail to terminate erythroid differentiation suggesting that the anti-apoptotic activity of Bcl-xL is an indispensable part of the erythroid program (Motoyama et al., 1999). Overexpression of Bcl-2 increases the viability and proliferative rate of HSCs in the embryo (Orelio et al., 2004).

GATA-1 and globins

GATA-1 was isolated as a protein with binding specificity to the 3' enhancer of the β-globin locus and was soon shown to bind to multiple regulatory regions in both α- and β-globin loci (Wall et al., 1988; Evans et al., 1988; Martin & Orkin, 1990). GATA-1 null erythroid cells produce hemoglobin despite the

absence of GATA-1 and it was thus proposed that perhaps GATA-1 does not play a crucial role in their transcription (Weiss et al., 1994). Most probably though, it is the activity of GATA-2 that is upregulated and replaces GATA-1 in these cells that drives the expression of the globins since it was shown that GATA binding sites in the LCR region of β-globin are necessary for its expression (Philipsen et al., 1993). Significantly, GATA-1 and its cofactor Friend of GATA, FOG1, were shown by 3C (Chromosomal Conformation Capture) approach to mediate spatial interactions between the β-globin gene and the LCR (Vakoc et al., 2005).

GATA-1 and erythroid genes

Erythropoiesis is stimulated by the major growth factor of erythropoiesis, erythropoietin (Epo), via the promotion of proliferation, differentiation and survival of erythroid precursors (BFU-E, CFU-E) (reviewed by Lacombe & Mayeux, 1999). Upon Epo binding to its receptor EpoR, the receptor undergoes conformational changes that, through the EpoR/Jak2/Stat5 signaling axis, trigger activation of several signal transduction pathways (reviewed by Richmond et al., 2005). The EpoR gene has a GATA site at -30bp and binding of GATA-1 and transactivation of the gene has been reported (Zon et al., 1991; Komatsu et al., 1997). In a second report it was shown that GATA-1 is expressed prior to the EpoR but its expression is strongly enhanced by EpoR-mediated signals (Chiba et al., 1993).

Enforced ectopic GATA-1 expression studies have provided evidence not only for an essential role but also for an instructive role for GATA-1 in erythroid differentiation. For example, ectopic GATA-1 expression in highly purified murine progenitor cells (myeloid or lymphoid) ‘‘instructed’’ their differentiation toward the erythroid and megakaryocytic lineages that GATA-1 normally regulates (Heyworth et al., 2002; Iwasaki et al., 2003). Thus, GATA-1 is capable of imposing a transcription program that promotes erythroid differentiation in cells that otherwise would have the potential to give rise to different cell types (Graf, 2002). This regulation involves both upregulation of the erythroid genes and suppression of the genes involved in multipotentiality or alternative hematopoietic programs, such as GATA-2 or the critical myeloid transcription factor PU.1 (Rodriguez et al., 2005). All erythroid genes described include GATA binding sites in their promoters, including gata-1 itself, gata-2, klf1 and scl/tal1 loci (reviewed by Tsiftsoglou et al., 2009).

TAL-1 and EKLF also play important roles in the erythropoietic program. The promoter regions of both transcription factors include GATA sites. GATA-1 was shown to bind to the promoter of tal1 in cooperation with the ubiquitously expressed SP1 transcription factor and to synergistically *in vitro* modulate the erythroid-specific expression of the tal1 gene (Lecointe et al., 1994). The motif GATA/E-box/GATA present in an upstream enhancer element was shown to be essential for the *in vivo* expression of eklf (Anderson et al., 1998). Rodriguez et al. (2005) showed GATA-1 binding in association with the TAL1/LDB1 complex to the upstream enhancer of the eklf gene in differentiated erythroid cells.

GATA-1 and cell cycle genes

Overexpression of GATA-1 in MEL cells reversed DMSO-induced terminal erythroid differentiation. Hence, transcription of the globin genes was not induced and the cells retained their proliferative status, without arresting in the G1 phase. Moreover, cyclin E-dependent kinase activity was not reduced as normally expected for differentiating cells. These observations suggested a role for GATA-1 in the regulation of genes involved in the cell cycle and more precisely in the G1 to S-phase transition (Whyatt et al., 1997).

Rylski et al. (2003) used a complementation assay based on the synchronous inducible rescue of GATA-1^{-} G1E erythroblasts to investigate how GATA-1 expression regulates the cell cycle in differentiating erythroid precursors. Microarray data showed that a number of cell cycle genes, such as $p27^{kip1}$, were activated, while others, such as cyclin D2, were repressed. Overall, it was revealed that GATA-1 not only drives differentiation but also results in rapid synchronous cell cycle arrest and inactivation of cyclin-dependent kinases (Cdks). In addition, it was shown that the G1 cell cycle arrest required during differentiation is fully dependent on the GATA-1-mediated repression of *c-myc*, by direct binding of GATA-1 to the promoter of the *myc* proto-oncogene.

Additional data from another study revealed that GATA-1, in a tricomplex with retinoblastoma protein (pRb) and E2F transcription factors, regulated cell cycle arrest and altered the G1 to S-phase progression of erythroid precursors forcing them to stop proliferating in order to differentiate.

However, the particular study did not address whether the GATA-1/pRb/E2F-2 tricomplex possesses specific transcriptional activity in addition to the mere sequestration of E2F-2. Preliminary ChIP assays though, indicated GATA-1 binding to known E2F targets genes such as Cdc6 (Kadri et al., 2009).

GATA-1 and microRNAs

MicroRNAs (miRNAs) are small (21-25nt) noncoding RNAs that repress gene expression at the posttranscriptional level by binding to the 3'-UTRs of mRNAs and, depending on the extent of sequence complementarity, inducing mRNA degradation or translational inhibition (reviewed by Chang & Mendell, 2007). The target range of a single miRNA can be broad as it can bind to several mRNAs through imperfect complementarity and hence a small number of miRNAs can regulate expression of numerous mRNAs, in various tissues, including the hematopoietic system. Binding of miRNA to mRNA within the RNA-induced silencing complex (RISC) leads to either translational inhibition or to destruction of the target mRNA. Both of these functions are executed by Argonaute 2 (Ago2). Hematopoietic specific conditional knockout of the Ago2 gene in mice resulted in defective erythropoiesis (O'Carroll et al., 2007). A number of studies have highlighted the importance of miR-451 and miR-144 in erythropoiesis (Rathjen et al., 2006; Bruchova et al., 2007; Zhan et al., 2007; Masaki et al., 2007; Lu et al., 2008). MiR-451 was found to be highly upregulated during erythroid maturation, which is consistent with the findings of Dore et al., (2008) that miR-451 gene locus is under the transcriptional control of GATA-1. ChIP experiments validated direct binding of GATA-1 to an upstream cis-regulatory region to activate RNA polymerase II-mediated transcription of a single common precursor RNA encoding both mature miR-451 and miR-144 miRNAs. Importantly, miR-451 knockdowns in zebrafish blocked erythroid maturation (Dore et al., 2008), whereas in another study GATA-2 mRNA was found to be among the miR-451 target sequences thus implicating miR-451 in the GATA-2 downregulation that is required for terminal erythroid maturation to proceed (Pase et al., 2009).

<u>GATA-1 protein-protein interactions</u>

Transcriptional regulation of erythropoiesis by GATA-1 is partly accomplished by protein interactions with other transcription factors or co-factors. GATA-1 has been reported to interact with many transcription factors, such as FOG-1, EKLF, TAL-1/SCL, PU.1 and cofactors such as CBP/p300, Brg1, Med1, MeCP1/NuRD, and others (reviewed by Cantor & Orkin, 2002; Ferreira et al., 2005; Lowry & Mackay, 2006). An overview of the key complexes GATA-1 participates in, based on previous work from Rodriguez et al. (2005) and others is given below, whereas in chapter 3 of this thesis data from the mass spectrometry analysis we followed to identify novel *in vivo* GATA-1 partners in the erythroid lineage are presented.

Using an *in vivo* biotinylation tagging approach coupled with mass spectrometry, Rodriguez et al. (2005) isolated and characterized nuclear GATA-1 protein complexes in differentiated MEL cells. This work showed that GATA-1 forms two independent complexes with FOG-1, with and without the repressive MeCP1/NuRD chromatin remodeling/histone deacetylase complex. In addition, GATA-1 forms distinct complexes with the hematopoietic transcription factors TAL-1/SCL (and associated partners) or Gfi-1b and with the chromatin remodeling complex ACF/WCRF (Figure 6).

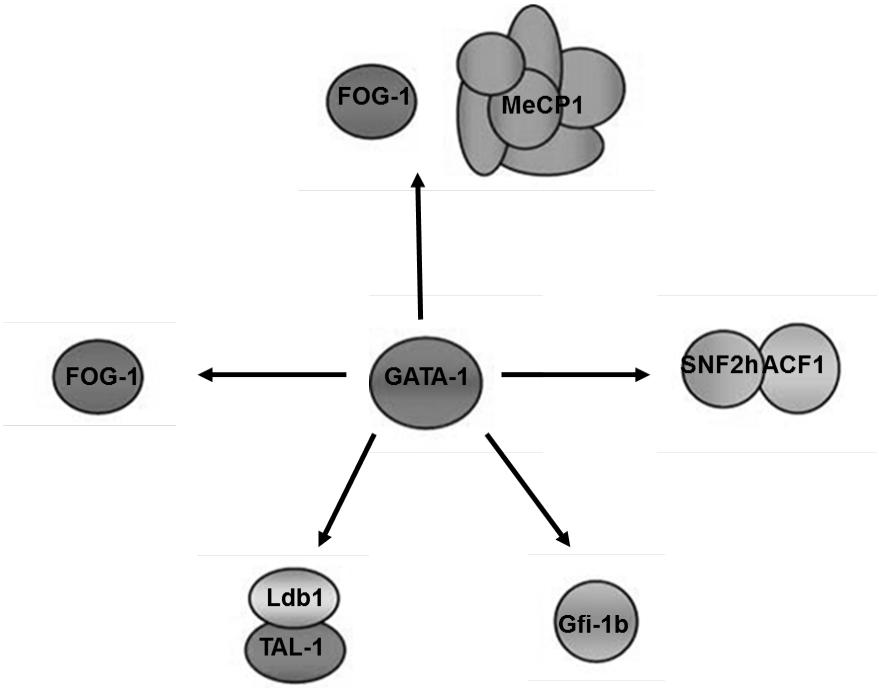


Figure 6: Model for the distinct GATA-1 complexes and their role in erythropoiesis.
(Adapted from Rodriguez et al., 2006).

GATA-1

In vitro studies have shown that GATA-1 can self-associate using its zinc fingers, both in solution and when bound to DNA (Crossley et al., 1995). These interactions enable GATA-1 to work as a powerful transcriptional activator, since mutations of particular residues in the zinc fingers reduces GATA-1 transactivation potential in reporter assays and diminishes GATA-1 positive autoregulation in zebrafish (Mackay et al., 1998; Nishikawa et al., 2003).

In vivo studies of the contribution of GATA-1 self-association to the regulation of erythropoiesis was studied with transgenic complementation rescue experiments using GATA-1 mutants that lack the self-association ability. A combination of three lysine mutations in the zinc fingers resulted in reduced expression of the transferrin receptor and heme biosynthesis enzymes, indicating that self-association is indispensable for GATA-1 to direct proper erythroid development (Shimizu et al., 2007).

FOG-1

Friend of GATA-1 (FOG-1), a protein containing nine widely-spread zinc fingers, was identified in a yeast two-hybrid screen in MEL cells as a GATA-1 co-factor (Tsang et al., 1997). It binds to the N-terminal finger of GATA-1 mainly via zinc finger 6 although fingers 1, 5 and 9 also contribute to the binding (Fox et al., 1998, 1999). FOG-1 is expressed in the megakaryocytic and erythroid lineages in a tissue specific pattern that resembles GATA-1. FOG-1 null mice are embryonic lethal at day E10.5-11.5 due to defects in primitive and definitive erythropoiesis, arrest of erythroid cells in the proerythroblast stage and subsequently severe anemia (Tsang et al., 1997). Analysis of the megakaryocytic lineage revealed complete absence of megakaryocytes in FOG-1 null mice due to an early block in the megakaryocytic differentiation (Tsang et al., 1998). From the comparison of the GATA-1 and FOG-1 null phenotypes it appears that the two factors have overlapping as well as distinct roles in hematopoiesis and erythropoiesis. FOG-1 null mice suffer from a complete loss of megakaryocytes, whereas GATA-1 mice have a block in megakaryocytic differentiation at mid-maturation, suggesting a broader role for FOG-1 in the megakaryocytic lineage. In contrast in the erythroid lineage, despite both knockouts showing arrest of cells at the proerythoblast stage, FOG-1 null erythroid precursors live longer than those obtained from GATA-1 null mice, suggesting FOG-1 independent GATA-1 functions (Tsang et al., 1998).

The importance of the interaction between GATA-1 and FOG-1 has been highlighted by mutation assays, where a single amino-acid change in the N-terminal zinc finger of GATA-1 abolishing interaction with FOG-1 (Crispino et al, 1999), resulted in lethality in mice due to severe anemia (Chang et al, 2002) and is associated with dyserythropoietic anemia in patients (Nichols et al, 2000).

Rodriguez et al. (2005) showed that the overlapping functions of GATA-1 and FOG-1 in erythropoiesis occur in the context of two distinct complexes, a GATA-1/FOG-1/MeCP1 complex and a GATA-1/FOG-1 complex. The association of GATA-1 and FOG-1 with the MeCP1 complex represses early hematopoietic and alternative lineage transcription programs, whilst the GATA-1/FOG-1 complex upregulates the erythroid transcription program. (Crispino et al, 1999; Fox et al, 1999; Letting et al, 2004; Pal et al, 2004). The fact that only one of the three isoforms of FOG-1 interacts with the GATA-1/MeCP1 complex, may provide a mechanism for the selective formation of the GATA-1/FOG-1/MeCP1 complex.

Gfi-1b

Growth factor-independent 1b (Gfi-1b) is an essential hematopoietic factor that contains six C-terminal C_2H_2 zinc-fingers, which bind a defined DNA consensus sequence, and an N-terminal SNAG domain which is associated with repression (Duan and Horwitz, 2003; Doan et al, 2004). Gfi-1b knockout resembles that of GATA-1, with embryonic lethality at E15 due to impaired erythroid and megakaryocytic differentiation in the fetal liver (Saleque et al, 2002). GATA-1/Gfi-1b complex binds to the promoters of the mitogenic genes myc and myb and suppresses cell division (Rodriguez et al., 2005). Huang et al. (2005) showed that Gfi-1b expression is upregulated by GATA-1 in early erythroid maturation with expression levels declining in later stages, possibly mediated through Gfi-1b repressing its own promoter in a negative feedback loop. Despite the similarities between the FOG-1 and Gfi-1b knockout phenotypes, no FOG-1 and Gfi-1b interaction was observed in differentiating MEL cells (Rodriguez et al., 2005). Possibly, the two factors regulate common gene targets through distinct complexes and binding sites or the functions of GATA-1 with FOG-1 and Gfi-1b are separate and for example, GATA-1/FOG-1 complex regulates differentiation versus GATA-1/Gfi-1b that controls proliferation arrest, with each function being essential for erythropoiesis.

Rb

The tumor suppressor protein retinoblastoma (RB) plays important roles in many stages of the differentiation process, including regulation of progenitor proliferation, terminal cell cycle exit, induction of tissue-specific gene expression and protection from apoptosis (reviewed by Du & Searle, 2009; Poznic, 2009; Gordon & Du, 2011). The phosphorylation state of RB determines the progression of cell cycle from G1 to S-phase. In G1, RB exists in an active, hypophosphorylated state, in complex with E2F, preventing it from activating the transcription of genes necessary for entrance to S-phase and cell cycle progression. The Rb null mutation leads to embryonic lethality characterized by neuronal defects and anemia due to erythropoietic impairment (Clarke et al., 1992; Clark et al., 2004). Overexpression of GATA-1 in MEL cells alleviated DMSO-induced terminal erythroid differentiation by direct interaction with RB and suppression of the genes involved in the G1 to S-phase transition of the cell cycle (Whyatt et al., 1997; Kadri et al., 2009).

PU.1

PU.1, a member of the E-twenty-six family of transcription factors, is the master regulator of the myeloid blood cell lineage. Inactivation of the PU.1 gene causes defects in the development of myeloid lineages and B and T lymphocytes (Scott et al, 1994; McKercher et al., 1996). GATA-1 and PU.1 cross-antagonize for determining lineage commitment (erythroid versus myeloid) during hematopoiesis, with GATA-1 repressing PU.1 activity during erythropoiesis and PU.1 repressing GATA-1 function during myelopoiesis. Several independent studies have revealed that the two proteins are able to physically interact both *in vitro* and in a cellular context (Rekhtman et al., 1999; Nerlov et al., 2000; Zhang et al., 2000). NMR titration data and mutagenesis experiments revealed that the PU.1-Ets domain and the GATA-1 C-terminal zinc finger (CF) form a low affinity interaction in which specific regions of each protein are implicated (Liew et al., 2006). Whereas the repression of GATA-1 activity by PU.1 involves the creation of repressive chromatin structures, inhibition of PU.1 activity by GATA-1 appears to involve the transcription factor c-Jun (Zhang et al., 1999).

EKLF

The erythroid Krüppel-like factor (EKLF) is an erythroid-specific zinc finger transcription factor that is critical for globin gene regulation and erythropoiesis in general (reviewed by Bieker, 2005). It contains

three conserved Cys_2His_2 type zinc-fingers, located at the C-terminal end of the protein that mediate DNA binding to CACC box motifs present in the β-globin promoter and in the promoters of many erythroid genes (Gregory et al., 1996). The significance of the CACC box in the human β-globin promoter is underscored by the fact that naturally occurring mutations in this element are associated with thalassemia (Feng et al., 1994). Merika and Orkin (1995) showed synergistic activity between EKLF and GATA-1 at the transcriptional level and provided evidence for physical interaction of the zinc finger DNA-binding domain of GATA-1 with the Krüppel zinc finger domains, which serves to facilitate or stabilize chromosomal loops between distant locus control region (LCR) and promoter elements. EKLF null mice are embryonic lethal at E14-15 due to defective definitive erythropoiesis (Nuez et al., 1995; Perkins et al., 1995). Analysis of the EKLF knockout mice showed a failure in β-globin gene activation in the fetal liver stage accompanied by a ''closed'' chromatin structure at the β-globin gene promoter and loss of DNase I hypersensitivity at HS3 of the LCR, thus suggesting important roles for EKLF in the organization of an active chromatin domain in the β-globin gene locus (Wijgerde et al., 1996). Further characterization of the mice revealed impaired expression of genes involved in hemoglobin biosynthesis pathways (Drissen et al., 2005) and perturbation of genes involved in cell cycle regulation, resulting in a delay in the G1 to S-phase transition (Pilon et al., 2008).

SCL/TAL-1/LMO2/LDB1/E2A complex

TAL-1/SCL is a basic helix-loop-helix (bHLH) transcription factor that binds to a short consensus DNA motif (CANNTG) called the E-box. TAL-1/SCL expression essentially mirrors that of GATA-1 as it is expressed in erythroid cells, megakaryocytes, and mast cells (reviewed by Cantor & Orkin, 2002; Kim & Bresnick, 2007). TAL-1/SCL null mice are embryonic lethal at day E9.5, displaying a complete absence of recognizable hematopoiesis in the yolk sac. Early organogenesis appeared to be otherwise normal. Transcripts of master regulators of hematopoiesis such as GATA-1 and PU.1 transcription factors were absent from RNA from scl-/- yolk sacs and embryos (Robb et al., 1995; Shivdasani et al., 1995). Further analysis by *in vitro* differentiation assays and *in vivo* studies in chimeric mice revealed that in the absence of TAL-1/SCL, hematopoiesis, including the generation of red cells, myeloid cells, megakaryocytes, mast cells, and both T and B lymphoid cells, is undetectable, suggesting that SCL/tal-1 functions very early in hematopoietic development (Porcher et al., 1996). In erythroid cells, TAL-1/SCL forms a complex with the ubiquitous E47/E2A bHLH partner and also with the hematopoietic LMO2 and ubiquitous Ldb1 LIM domain containing cofactors (Wadman et al., 1994; Schuh et al., 2005; Goardon et al., 2006). This complex interacts with GATA-1 to form a so-called pentameric complex that binds to composite Ebox/GATA-1 DNA motifs that are spaced 9–11 nucleotides apart. Such motifs are present in many erythroid genes and in the regulatory elements of key transcription factor genes including the TAL-1/SCL and GATA-1 genes themselves and the EKLF gene (Wadman et al., 1997; Cohen-Kaminsky et al., 1998; Anderson et al., 1998). Several lines of evidence support a critical role for the pentameric complex in erythropoiesis. Significantly, ectopic coexpression of TAL-1/SCL, LMO2, and GATA-1 in Xenopus embryos promoted erythropoiesis (Mead et al., 2001). Recent evidence for novel protein interacting partners of the pentameric complex emerges, including the ETO-2 repressor protein (Schuh et al., 2005; Goardon et al., 2006, Meier et al., 2006).

BRG1

SWI/SNF complexes are involved in both activation and repression of transcription by selectively regulating chromatin remodeling through direct interaction with specific DNA-binding proteins. Kadam et al. (2000) proposed that one mechanism by which mammalian SWI/SNF regulates specific subsets of genes is by interacting with distinct zinc finger proteins, including GATA-1, through the BRG1 and BAF155 or BAF170 subunits, and targeting the native complex to nucleosomal sites. This results in extended chromatin accessibility facilitating the active domain of the protein to recruit other components of the transcription apparatus to promote initiation. Involvement of BRG1 in erythropoiesis has been demonstrated by recruitment in the β-globin locus, where once recruited, BRG1 is required for the LCR to adopt an open chromatin structure (Armstrong et al., 1998). Mutant mouse embryos expressing lower levels of brg1 suffer from anemia due to a block in erythroid maturation, thus providing further evidence for a role for Brg1 in erythropoiesis (Bultman et al., 2005). Xu et al. (2003) showed that a DNA-binding complex containing TAL-1/SCL, E47, GATA-1, LMO2 and Ldb1 stimulates P4.2 (protein 4.2) transcription in erythroid progenitors via two E box–GATA elements in the gene's proximal promoter. BRG1 also associates with this complex and both the E box and GATA DNA-binding sites in these elements are required for BRG1 recruitment. In this case,

expression of Brg1 actively represses P4.2 gene expression, and most likely other genes, by histone deacetylation until the terminal differentiation of erythroid progenitors (Xu et al., 2006).

Post-translational modifications

GATA-1 can be acetylated *in vivo* and *in vitro* by the ubiquitously expressed acetyltransferases P300 and CREB-Binding Protein (CBP) (Boyes et al., 1998; Hung et al., 1999). Acetylation occurs in two highly conserved lysine residues present in the C-terminal part of each zinc-finger. Boyes et al. (1998) reported that GATA-1 acetylation increased the factor's ability to bind DNA in *in vivo* studies with chicken models, a phenotype which was not confirmed in *in vivo* studies with mouse models (Hung et al., 1999). Acetylation of GATA-1 is however required to trigger the *in vitro* differentiation of G1E cells (Hung et al., 1999). In another study, Lamonica et al. (2006) showed that acetylation of GATA-1 increases its ability to occupy chromatin in erythroid cells. In addition, acetylation of GATA-1 marks the factor for degradation via ubiquitination (Hernandez-Hernandez et al., 2006).

GATA-1 can also be phosphorylated at seven serine residues. Six of these residues, located at the N-terminus of the protein, are phosphorylated in non-differentiated MEL cells. The seventh serine (S310), which is in proximity to the DNA binding domain, becomes phosphorylated upon induction with DMSO. This could imply a potential role for phosphorylation in the activity of GATA-1, albeit substitution of serines to alanines in *in vitro* studies did not affect the activity or DNA binding of GATA-1 (Crossley & Orkin, 1994). Yu et al. (2005) used the interleukin 3-dependent hematopoietic cell line Ba/F3 to show that upon stimulation with interleukin-3, GATA-1 is strongly phosphorylated at residue serine 26 by a MAPK-dependent pathway. Phosphorylation of GATA-1 increased GATA-1-mediated transcription of the E4bp4 survival gene without significant alterations in the DNA-binding affinity of GATA-1. Phosphorylation site mutants revealed that GATA-1 phosphorylation at serine 26 is required for the antiapoptotic function of GATA-1, which is mediated by the up-regulation of the pro-survival gene Bcl-X(L). Moreover, experiments in both primary fetal liver erythroid progenitors and in cultured erythroid cells showed that Epo-induced maturation was dependent on GATA-1 phosphorylation at serine 310 (S310). Further studies revealed that GATA-1 phosphorylation at S310 is triggered by EpoR via the PI3-kinase/AKT signaling pathway.

GATA-1 was shown to be sumoylated both *in vivo* and *in vitro* at lysine 137, which is located at the transcriptional activation domain. Further characterization of the pathway revealed that GATA-1 sumoylation is promoted by the RING finger protein PIASy and results in a dramatic repress of the GATA-1 transcriptional activity (Collavin et al., 2004). In another study, by Lee et al. (2009) it was demonstrated that sumoylation of GATA-1 at lysine 137 promotes transcriptional activation, however only at GATA-1 targets that are regulated by GATA-1 and FOG-1. Mutations in K137 residue or in V205G residue, which abolishes GATA-1 binding to FOG-1, yielded the same phenotype even though sumoylation of GATA-1 was FOG-1 independent and GATA-1 sumoylation was dispensable for FOG-1 binding.

Recent work provided evidence for the ubiquitination of GATA-1 by the chaperone HSP27. It was shown that in two different models of erythroid differentiation, human erythroleukemic K562 cells and CD34 human cells in *ex vivo* culture, induced to differentiate, the ablation of HSP27 resulted in increased levels of GATA-1 and impaired maturation of the cells. It was demonstrated that in the late stages of erythropoiesis, HSP27 becomes phosphorylated, enters the nucleus, binds to acetylated GATA-1 and promotes its ubiquitination and proteasomal degradation (de Thonel et al., 2010). An observation which confirmed previous work from Hernandez-Hernandez et al. (2006), proposed that the acetylation of GATA-1 marks the factor for degradation via ubiquitination.

Clinical cases of GATA-1 mutations

As would be expected, mutations in GATA-1 have subsequently been found to have important clinical significance, and are directly linked to deregulated formation of certain blood cell lineages. There are five human diseases caused by GATA-1 mutations: X-linked thrombocytopenia (XLT), X-linked thrombocytopenia with thalassemia (XLTT), congenital erythropoietic porphyria (CEP), transient myeloproliferative disorder (TMD) and acute megarakaryoblastic leukemia (AMKL) associated with Trisomy 21, and, lastly, a particular subtype of anemia associated with the production of GATA-1s, a shortened, mutant isoform of the wild-type GATA-1. The locations of the GATA-1 mutations appear to have critical phenotypic consequences. For example, the crucial N-finger domain harbors all known

GATA-1 mutations causing the GATA-1 related cytopenias. In Trisomy 21, acquired somatic mutations lead to the production of GATA-1s, which is associated with TMD and AMKL (Mundschau and Crispino, 2006). The different phenotypic expressions associated with GATA-1 mutations illustrate the function of the transcription factor in overall body homeostasis (reviewed by Ciovacco et al., 2008).

1.4 Scope of this thesis

My project is concerned with the molecular basis of transcription factor functions in cellular differentiation. Transcription factors play key roles in regulating "batteries" of genes (termed transcription programs) by turning them on or off during differentiation. It is the identity and activities of these genes that determine the cell's terminally differentiated and fully functional phenotype. Gene knockout studies have been invaluable in revealing the functional association of individual transcription factors with the differentiation of specific cell types; however these approaches do not provide insight into the molecular basis of transcription factor function or into the direct gene targets that these factors regulate in shaping the cell's identity. I propose to address these questions by focusing on the function of key hematopoietic transcription factors in red blood (erythroid) cell differentiation. A number of transcription factors including GATA-1, FOG-1, Gfi-1b and TAL-1 are essential for erythroid cell differentiation. The gene knockouts of these factors in mice result in the arrest of erythroid differentiation at the proerythroblast stage (the first distinct stage of erythroid differentiation) causing embryonic lethality due to anemia. All of these factors have been reported to interact physically and/or genetically in erythroid cells; however, it is not known how their interactions are integrated in regulating gene targets. I intend to address this question by identifying the gene target networks coordinately regulated by these factors in red blood cell differentiation. To these ends I will focus my study on GATA-1, which is a master regulator of the erythroid cell differentiation and will use the biotinylation tagging of GATA-1 as a uniform platform for the proteomic and global gene target analysis. This way I will work with physiological levels of biotin-tagged GATA-1 in all hematopoietic lineages where it is active and will attempt to characterize in a much more systematic way the different protein protein interactions GATA-1 undergoes and the distinct genetic programs GATA-1 regulates in the various hematopoietic lineages.

Chapter 2

Materials and Methods

Constructs
2.95kb, 2.1kb and 2.3kb regions of the *mus musculus* GATA-1 gene locus were cloned by PCR from the mouse fosmid clone WIBR1694F07 purchased from the Children's Hospital Oakland Research Institute (CHORI, Figure 7). Primer sequences are listed in the Appendix, Table 5. GATA-1 gene fragments were recloned into the *SpeI* site, the *XbaI* and the *EagI* sites of the pBluescript II SK (Invitrogen) vector, respectively. The neomycin resistance gene was cloned into the *EcoRV* site of the 2.95kb GATA-1 fragment, the 2.1kb fragment was N-terminally tagged by introducing into the *NcoI* site overlapping the start codon an oligonucleotide linker with *NcoI* overhangs coding for the 23-aa biotin tag (Schatz, 1993). Lastly, the thymidine kinase gene was cloned into the *SacII* site of the 2.3kb fragment (Figure 7).

Embryonic Stem (ES) cell clones
S129, E14 and a knock-in S129 ES cell line (named iB10p) that ubiquitously expresses the BirA biotin ligase from the Rosa26 locus (Driegen et al., 2005) were electroporated with *AflII* linearized bioGATA-1/pBluescript plasmid. Clones were double selected for ganciclovir and neomycin (Invitrogen).

100µg of targeting construct/ electroporation were digested with *AflII*. Complete linearization of the construct was confirmed by loading an aliquot on a 0.8% agarose gel. Digested DNA was precipitated with 2.5 volumes of 100% ethanol and pellet was resuspended in 100µl PBS under sterile conditions.

ES cells were cultured in 10cm plates coated with filter sterilized 0.1% gelatin in PBS in KO DMEM (Invitrogen) supplemented with 15% FBS, 1% L-glutamine, 1% Penicillin/Streptomycin, 10^{-4} M β-mercaptoethanol, 1,000 units/ml LIF (ESGRO).

Day 1: 3 x 10cm ~70% confluent ES cells were used for 2 electroporations. Medium was aspirated and cells were washed 2 times with PBS. Cells were incubated with 0.25% Trypsin-EDTA for 2min at 37ºC and pipetted up and down to obtain a single cell suspension. 10ml medium/dish was added to remove the cells from the dish and extra 10ml medium was used to wash the dishes and extract all the cells. Pooled cells (40ml) were centrifuged at 1,000 rpm for 5min at room temperature and pellet was resuspended in 40ml PBS. 2ml cell suspension was added straight to a 10cm plate with medium to be used as the no electroporation control plate. Cells were counted and centrifuged at 1,000 rpm for 5min at room temperature. This last wash step was repeated once and the cells were resuspended in 600µl PBS/electroporation (i.e 1.2ml if 3 x 10cm dishes were used). 100µg of targeting construct/electroporation were used and the 700µl of the cells+DNA mix were transferred in a 0.4cm electrocuvette and electroporated at 250V, 500µF and for a time constant of ~7sec. Cells were incubated at room temperature for 10min prior to resuspending them in 9ml medium. 1ml of cells was added to each of the 10ml medium containing dishes.

Day 2: Medium was changed accordingly:

	G418 (0.3mg/ml)	Ganciclovir (2µM)
A) 1 plate non electroporated	+	+
B) 1 plate electroporated	+	-
C) 1 plate electroporated	-	-
All test plates	+	+

A: control plate to verify that the antibiotics kill all the cells that do not express the neo cassette
B: control plate to verify that ganciclovir is efficient and (get an estimate of the reduction in the non-homologous recombination events).
C: control plate to verify that the electroporation itself did not kill the cells (this plate should be confluent within 3-4 days).

Medium was changed every 1-2 days, depending on the rate of cell growth and death.

Day 8-9: Colonies were visible and were picked under the microscope.

For each single plate: medium was removed and dish was washed with 10ml PBS. 12ml PBS were added, 5ml of which were removed (so cells ended up being covered in 7ml of PBS). A P20 pipette was set to 15µl and colonies were loosened by pressing against them with the tip of the pipette and were transferred to a V-shaped 96-well plate, by sucking up all the 15µl with the pipette. 20µl of

0.25% trypsin-EDTA were added to each well, the plate was incubated for 4min at 37ºC and cells were disaggregated to obtain single cell suspensions. 200µl of medium/Ganciclovir/G418 were added to each well and cells were incubated for 2-3 days at 37ºC to reach confluence.

After ~2-3 days -when most of the cells approached confluence- cells were split to 2 plates for DNA extraction and 1 plate for freezing.

2 x 96-well plates were gelatin coated and filled with 180µl medium/Ganciclovir/G418 for every single 96-well plate of cells and stored in the 37ºC cell incubator. 12ml of filter sterilized Quench medium, containing 50% medium/Ganciclovir/G418-50% FBS and 10ml of 2xFreezing medium, containing 20% DMSO-80% medium/Ganciclovir/G418 were prepared per plate and incubated at room temperature and on ice, respectively. Medium was aspirated from all wells and cells were washed twice with 150µl PBS/well. 20µl of 0.25% trypsin-EDTA were added to each well, the plate was incubated for 4min at 37ºC and cells were disaggregated to obtain single cell suspensions. A further 2min incubation step at 37ºC followed, cells were examined under the microscope to ensure that the wells contained single cell suspensions and 110µl of Quench medium per well were added, pipetting up and down. 20µl of cells were added to each well in each of the two gelatin coated 96-well plates (stored in the incubator) and plates were transferred at 37ºC until wells were fully confluent being suitable for DNA extraction. Extra care was taken to ensure that well A1 in the new dishes derived indeed from well A1 of the original plate.

The remaining 80µl of cells were incubated on ice for 5min and ice cold 2xFreezing medium was added to them, pipetting up and down and the plate was stored at -80ºC.

Cells were ready for DNA extraction when the gelatin treated plates turned fluorescent yellow. Medium was removed from all wells with growing cells, washed once with 150µl PBS and 100µl of lysis buffer/well were added. Lysis buffer was prepared with 100mM Tris-HCl pH 8.5, 5mM EDTA, 0.2% SDS, 200mM NaCl and 200µg/ml proteinase K. Plates were incubated overnight at 37ºC and lysates were transferred in eppendorfs with 20µl 3M NaAc, 80µl MilliQ and 200µl Phenol:Chloroform. A 5min centrifugation step followed and DNA was precipitated with 0.6V of cold isopropanol and 1µl of glycogen carrier (20mg/ml). Precipitates were washed with 70% ethanol and resuspended in 50µl MilliQ+ 1µl RNase (10mg/ml).

Genotyping and recombination analysis

Targeted ES cells and mice were screened by Southern blot analysis. The probes used included: a) a 5' probe located outside the 5' homology sequence of the construct, which hybridizes to a 15.8kb *HindIII* fragment in the wild type GATA-1 locus or to a 9.9kb *HindIII* fragment in the correctly targeted locus, b) a 3'probe located outside the 3' homology sequence of the construct, which hybridizes to a 9.9kb *BamHI* fragment in the wild type GATA-1 locus or to a 7.5kb *BamHI* fragment in the correctly targeted locus, c) the neo probe which hybridizes to the 9.9kb *HindIII* fragment in the correctly targeted locus and d) the biotin tag probe which hybridizes to a sequence of >1kb in the presence of the biotin tag and <1kb in the absence of the tag.

Day 1: 10-20µg of genomic DNA were digested overnight in a final volume of 70µl at 37^{o}C and the efficiency of the digestion was tested on a 0.7% agarose gel. Fully digested samples were loaded on a 0.7% agarose gel in 2x TAE, run at 70V during the day. Gel was photographed adjacent to a ruler and then inverted in a large tank. DNA on the gel was depurinated by shaking in 0.2M HCl for 20min and a step to denature the gel followed, with 0.5M NaOH/1.5M NaCl (denaturing solution) with shaking for 1hr. In the meantime, Nylon membrane (Amersham's Hybond N^{+}) was cut to size and rehydrated by wetting in dH_2O. Two sheets of Whatman 3MM paper were also cut to size. A piece of Saran wrap was fully stretched and taped down on the bench and gel was faced down on the wrap. The Nylon membrane was briefly floated in denaturing buffer and placed on top of the gel smoothing out any air bubbles. All 4 ends of the gel and the membrane were sealed with some parafilm. Similarly, the two sheets of 3MM paper were wet in denaturing buffer and placed over the filter, again smoothing out any air bubbles. A pack of paper towels was assembled on top of the gel and a weight was placed on top of it. Transfer was carried out overnight.

Day 2: The gel and the filter were inverted and the lanes on the filter were marked with a pencil. The gel was checked in the DNA gel imager to verify complete transfer and was discarded. The filter was

neutralized by brief (i.e. 1min.) floating of the membrane in neutralizing buffer. While still wet, the DNA was cross-linked with UV in the DNA gel imager, for 2min.

Hybridizing Buffer
For 250ml: 50µl 0.5M EDTA, 112.5ml SDS 20% (or 18gr SDS), 4.3gr sodium dihydrogen phosphate monohydrate, 16.5gr sodium phosphate dibasic dihydrate

Neutralizing buffer
For 1lt: 60.55gr Tris, 175.32 NaCl

A waterbath was set at 37^{o}C, a rotating incubator at 65^{o}C and a heat block at 100^{o}C. The hybridizing buffer was preheated at 65^{o}C and used to prehybridize the membrane at 65^{o}C for a couple of hours (approximately 20ml of hybridizing solution without meshes or 25ml with meshes for a big 20x20cm filter).

100ng of probe DNA + 11,5µl 2xLS + 1µl BSA (10mg/ml) + 5µl random primers (500µg/ml) were boiled for 5min, briefly centrifuged and placed on ice. 3µl of ^{32}PdCTP and 0.5µl of Klenow fragment (10units/µl) were added to the eppendorf and the tube was incubated for 3hrs at 37^{o}C.

A G-50 column (Bio-Rad columns) was prepared by filling it in with G-50 and brief centrifuging to discard excess of H_2O. The labeled probe was passed through the column, centrifuged at 2,000rpm for 3min and labeling efficiency was checked with a Geiger counter. The labeled probe was boiled for 5min, briefly centrifuged and placed on ice for 3-5min prior to mixing it with the hybridizing buffer (the buffer that the membrane is prehybridizing in) in a falcon. The membrane was hybridized in the hybridizing buffer +probe mix at 65^{o}C overnight.

2xLS
25µl 1M Hepes pH 6.6, 25µl DTM, 7µl OL

DTM
0,1mM dGTP, 0,1mM dCTP, 0,1mM dTTP in TM or
0,1mM dGTP, 0,1mM dATP, 0,1mM dTTP in TM
Depends on the radioactivity available, for example, if there is ^{32}PdATP a DTM that included all dNTPs except dATP was prepared.

TM
250mM Tris-HCL pH 8.0, 25mM $MgCl_2$, 50mM mercaptoethanol (14.1 M)

OL
1mM Tris-HCl pH 7.5, 1mM EDTA pH 7.5

Day 3: The membrane was rinsed a couple of times in 0.3xSSC/0.1% SDS (50-100ml for big membranes) at room temperature and then washed 2x in 50-100ml of 0.3xSSC/0.1% SDS at 65^{o}C for 1 hour per wash. The membrane was then checked for background with a Geiger counter and once verified that no further washing with higher stringency was necessary, a 0.1xSSC/0.1% SDS wash followed. The membrane was then placed in a cassette for 3 days at -80^{o}C.

The membrane was stripped with 5mM Tris-HCl pH 8.0, 0.1% SDS. Tris-HCl buffer (without the SDS) was boiled, then the SDS was added and the membrane was incubated for 1hr with shaking at room temperature. The last step was repeated once and radioactivity was checked with a Geiger counter.

Transgenic mice
Chimeric mice were generated by injecting clones from S129 ES cells generated as above into C57BL/6 blastocysts. Chimeras were then crossed with C57/BL6 females and germ-line transmission was examined by coat color. Screening of mice for transgenics was carried out by Southern blots using as probes the neomycin gene and two genomic regions 5' and 3' of the GATA-1 gene (Figure 7).

Cell cultures
Fetal liver cells from E12.5 embryos were expanded in serum free medium (StemPro34, Gibco) supplemented with StemPro34 supplement, 1% Penicillin/Streptomycin, 1% L-glutamine, 0.5 units/ml Erythropoietin (Epo), 100ng/ml stem cell factor (SCF) and 10^{-6}µM dexamethasone at a concentration of 10^6 cells/ml for 3 days. Early proerythroblasts (negative for the Ter119 cell surface marker), and mature erythroid cells (positive for Ter119) were purified as previously described (von Lindern et al. 2001; Schuh et al. 2005). Differentiation was induced by withdrawal of dexamethasone and the addition of 10units/ml of erythropoietin and 1mg/ml transferrin.

C88 MEL cells were grown in DMEM-10% FBS-1% penicillin/streptomycin and induced to differentiate with 2% DMSO as previously described (Antoniou, 1991).

HEK293T cells were grown in DMEM-10% FBS-1% penicillin/streptomycin.

G1E-ER cells (Rylski et al., 2003) were maintained in Modified Dulbecco's Medium (IMDM) supplemented with 15% FBS, 1% penicillin-streptomycin, 2 u/ml erythropoietin, 0.14µM mono-thioglycerol (MTG), and 50ng/ml kit-ligand. Expression of ER-GATA-1 fusion protein was induced with 1 µM β-estradiol.

Lentivirus production and transduction
Day 0: 2-2.5x10^6 HEK293T cells were seeded on 10cm^2 plates.

Day 1: Medium was replaced with 8ml DMEM-1%FBS (no antibiotics) prior to transfection. For the transfection of each plate, we used 6µg pMD2_VSVG, (envelope plasmid, Didier Trono lab http://tronolab.epfl.ch/), 15µg psPAX2, (packaging plasmid, http://tronolab.epfl.ch/) and 20µg of transfer vector (plko.1 or 3548 clone, TRC mission human and mouse library from Sigma, Moffat et al., 2006). 1ml of calcium-phosphate precipitate/plate was prepared by mixing the vector, packaging and envelope plasmids and H_2O up to 0.5ml and by adding dropwise 0.5ml 2xHBS. No additional mixing was required and the mix was incubated at room temperature for 20min. Poly-ethyleneimine-Na_2HPO_4-DNA precipitate was added to the cells dropwise and after 3-4hrs the medium was replaced with DMEM-10% FBS-1% penicillin/streptomycin (protocol described in Trono lab webpage http://lentiweb.com/protocols_lentivectors.php).

Day 2: Supernatant was collected and stored at 4°C.

Day 3: Supernatant was collected and stored at 4°C.

Day 4: Supernatant was collected and pooled with day 2 and day 3 supernatant media. Harvested viral supernatant was filter sterilized and ultracentrifuged at 20k for 4hrs, 4°C for fetal liver cells transduction or centrifuged at 3k for 10min, 4°C for MEL cells transduction.

2 x HBS (for 500ml)
8g NaCl, 0.38g KCl, 0.1g Na2HPO4, 5g Hepes, 1g Glucose. pH was brought to 7.05, 2.5M $CaCl_2$

Non-differentiated MEL cells were transduced with viral supernatants from HEK293T cells (day 0) and 20hrs later media were replaced (day 1). On day 2, puromycin was added at a concentration of 3.5µg/ml, as determined by a puromycin "kill curve" conducted in both MEL and fetal liver cells. Differentiation was induced after 3 days of selection (day 5) and cells were expanded for another 3 days. Cells were collected on day 8 for cytospin preparations, RNA isolation, nuclear extracts and flow cytometry.

Western blotting, Immunoprecipitations, Streptavidin pull-downs, Activity Assay and MS
Nuclear protein extracts were prepared using the NUN extraction buffer as described (Lavery & Schibler, 1993) and 50µg of protein were loaded per lane.

Cells were centrifuged at 1,000rpm for 5min at room temperature and washed twice with PBS and protease inhibitors at 4°C. Cells were lysed in 1ml of ice cold NP-40 Lysis Buffer for 10min on ice and centrifuged at 1,000rpm for 5min at 4°C. Supernatant was collected, 10% glycerol was added and

tube was snap frozen in liquid nitrogen and stored at -80^{o}C as the cytoplasmic extracts. Pellet was washed with NP-40 Lysis Buffer and the crude nuclear pellet was resuspended in 1 volume of Nuclear Storage Buffer (NSB) so as to have the nuclear slurry and after resuspension by gentle agitation, nine volumes of 1.1x NUN Buffer were added. Tube was placed in a rotating wheel for 1hr at 4^{o}C and centrifuged at maximum speed for 40min at 4^{o}C. 10% glycerol was added to the supernatant and tube was snap frozen in liquid nitrogen and stored at -80^{o}C.

NP-40 Lysis Buffer: 10mM Tris-HCl pH 7.4, 10mM NaCl, 3mM $MgCl_2$, 0.5% NP-40 and 1% protease inhibitors

Nuclear Storage Buffer NSB: 20mM Tris-HCl pH 7.9, 75mM NaCl, 0.5mM EDTA, 0,85mM DTT, 50% glycerol and 1% protease inhibitors

1.1X NUN: 1.1M Urea, 0.33M NaCl, 1.1% NP-40, 27.5mM Hepes pH 7.6, 1.1mM DTT and 1% protease inhibitors

Protease inhibitors were added fresh prior to using the buffers.

Filters were incubated with anti-GATA1 antibody (N6 and M20, Santa Cruz Biotechnology; 11852 and 11963, Abcam), anti-nucleophosmin anti-B23 antibody (kind gift from Pui K. Chan, Baylor College), anti-GATA2 antibody (H116, Santa Cruz Biotechnology), anti-NRF2 antibody (H300, Santa Cruz Biotechnology), anti-DNMT1 antibody (BioAcademia 70-201), anti-DNMT3a antibody (ab23565, Abcam) and anti-FOG1 antibody (M20, Santa Cruz Biotechnology). Streptavidin pull-downs and protein immunoprecipitations were performed with streptavidin, anti-GATA1 antibody (N6, Santa Cruz Biotechnology) and anti-DNMT1 antibody (BioAcademia 70-201) respectively, as described by Rodriguez et al., 2006. Pulled down proteins were processed and analysed by mass spectrometry (LC-MSMS, MS) as described (Rodriguez et al., 2006). DNMT methylase activity was assayed using the Active Motif DNMT Activity/Inhibition Assay Kit cat. no. 55006.

To detect GATA-1 post-translational modifications 10^9 proliferating and 10^9 differentiated E12.5 fetal liver cells were lysed in a 3:1 mix of RIPA (25 mM Tris pH 8.2, 50 mM NaCl, 0.5% NP40, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulphate) and lysis buffer (5% sodium dodecyl sulphate, 0.15 M Tris pH 6.7, 30% glycerol) at a concentration of $2x10^4$ cells/µl in the presence of protease inhibitors (Complete Inhibitor Cocktail, Roche), 2.5 µM of the proteasome inhibitor, MG132, 5 mM sodium phosphate, 1 mM potassium fluoride, 1 mM sodium orthovanadate, pH 8.2 to prevent dephosphorylation, 10mM butyrate and 50mM nicotinamide (NAM) to inhibit deacetylation and 1mM N-ethyl maleimide (NEM) to prevent desumoylation. Samples were sonicated, boiled for 5min and used for streptavidin pull-downs as described by Rodriguez et al., 2006. Pulled down proteins were processed and analysed by mass spectrometry (LC-MSMS, MS) as described (Rodriguez et al., 2006).

Mass spectrometry

Following binding, streptavidin beads were resuspended in 50 mM ammonium bicarbonate and treated with trypsin (sequencing grade; Promega) to approximately 60 ng/mg of total protein, followed by overnight incubation at 37°C (64). The supernatant containing the trypsin-treated peptides was then recovered by magnetically removing the beads. Peptides released by in-gel or on-bead trypsinization were analyzed by nano-LC-MS/MS performed on either a CapLC system (Waters, Manchester, United Kingdom) coupled to a Q-ToF Ultima mass spectrometer (Waters), operating in positive mode and equipped with a Z-spray source, or on a 1100 series capillary LC system (Agilent Technologies) coupled to an LTQ-Orbitrap or LTQ-FT-MS mass spectrometer (both from Thermo Scientific) operating in positive mode and equipped with a nanospray source. Peptides were trapped and separated on a Jupiter C18 reversed-phase column (Phenomenex) using a linear gradient from 0 to 80% medium B (where medium A 0.1 M acetic acid and medium B 80% [vol/vol] acetonitrile, 0.1 M acetic acid) using a splitter. The column eluate was sprayed directly into the electrospray ionization source of the mass spectrometer. Mass spectra were acquired in continuum mode; fragmentation of the peptides was performed in data-dependent mode.

Streptavidin Pull-down
Day 1: 50µl of Dynabeads M280 (Invitrogen) were used per 1mg of protein. Beads were rinsed 3x in 1ml PBS at room temperature, using a magnetic rack to concentrate the beads during the rinses. Beads were blocked with 1ml HENG/ chicken egg albumin 200µg/ml for 1hr at room temperature in a rotating wheel. 1/10 of the extracts were transferred in a new tube to be used as the input sample. The salt concentration of the rest of the nuclear extracts used was adjusted to 150mM KCl (essential for efficient binding) by adding 2 volumes of ice cold HENG Buffer. Diluted extracts were added to the blocked beads and incubated in a rotating wheel overnight at 4°C.

Day 2: The magnetic rack was used to concentrate the beads and supernatant was transferred in a new tube. The beads were washed 5x 10min with HENG Wash Buffer in a rotating wheel at 4°C and then rinsed in PBS and transferred to new tubes. Protein was eluted by resuspension in 1x Laemmli buffer and incubation for 10min at 80°C. Supernatant equal to 3x the volume of the eluted protein was used in order to load the same amount of protein from the input, the supernatant and the pull-down samples and the appropriate amount of Laemmli buffer was added. Input and supernatant samples were also incubated for 10min at 80°C prior to loading them on an acrylamide gel as input, pull-down and supernatant fractions.

HENG Buffer: 10mM HEPES-KOH pH 9, 1.5mM $MgCl_2$, 0.25mM EDTA, 20% Glycerol and 1% protease inhibitors

HENG Wash Buffer: 250mM KCl and 0.3% NP40 in HENG Buffer and 1% protease inhibitors

Protease inhibitors were added fresh prior to using the buffers.

Immunoprecipitation
Day 1: 150µl of Protein Sepharose beads –in two separate tubes, 100µl for the IgG control sample and 50µl for the IP sample- per 500µg of protein were equilibrated by washing them 3x with 1x PBS at room temperature. Both tubes were washed 2x with 1ml 100mM sodium citrate pH 5.0. 10µg of antibody and 10µg of IgG, each diluted in 1ml HENG buffer were added to the beads of the IP tube and the IgG tube, respectively. Tubes were incubated for 2hrs at room temperature in a rotating wheel. Beads were washed 2x with 1ml 100mM sodium citrate pH 5.0, followed by a wash with 1ml 200mM Triethanolamine Hydrochloride pH 8.2. 1ml of freshly made ice cold DMP Buffer was added and samples were incubated for 30min at room temperature with rotation. Beads were washed with 1ml 50mM Tris pH 7.5 and incubated in 1ml 50mM Tris pH 7.5 for 15min at room temperature with rotation. Beads were washed 3x with PBS-0.001% Tween20. At this point different procedures were followed for the two tubes: the IP beads were incubated on ice in the last wash buffer (PBS-0.001% Tween20) for 1hr while the IgG beads were blocked with 1ml HENG and 200 µg/ml chicken egg albumin for 1hr at room temperature with rotation. The IP beads were then blocked with 1ml HENG and 200 µg/ml chicken egg albumin for 1hr at room temperature with rotation and in parallel the protein extracts were diluted 1:3 in ice cold HENG Buffer to adjust the salt concentration to 150mM KCl (essential for efficient binding) and were added to the blocked IgG beads. The IgG tube was incubated for 1hr at 4°C in a rotating wheel. The IP tube supernatant was discarded and the IgG supernatant, i.e the extracts, was transferred to the IP tube for overnight incubation at 4°C with rotation. The IgG beads were washed 5x 5min with HENG Wash buffer at 4°C with rotation, rinsed 2x with PBS and the protein bound to the beads was eluted by resuspension in 1x Laemmli buffer and incubation for 10min at 80°C. Tube was stored at -20°C as the IgG control sample.

Day 2: The IP tube supernatant was transferred to a separate tube in order to be used as the supernatant sample. Laemli buffer was added and the tube was incubated for 10min at 80°C. The IP beads were washed 5x 5min with HENG Wash buffer at 4°C with rotation, rinsed 2x with PBS and the protein bound to the beads was eluted by resuspension in 1x Laemmli buffer and incubation for 10min at 80°C.

HENG Buffer: 10mM HEPES-KOH pH 9, 1.5mM $MgCl_2$, 0.25mM EDTA, 20% Glycerol and 1% protease inhibitors

HENG Wash Buffer: 250mM KCl and 0.3% NP40 in HENG Buffer and 1% protease inhibitors

Protease inhibitors were added fresh prior to using the buffers.

DMP Buffer: 0.01037g DMP powder in 2ml of Triethanolamine Hydrochloride pH 8.2

GATA-1 ChIP

10^7 $Ter119^+$ or $Ter119^-$ cells were cross-linked with 1% formaldehyde for 10min at room temperature. The reaction was quenched with a final concentration of 0.125M glycine; cell pellets were further processed as described (Schuh et al. 2005). Immunoprecipitations of triplicates per condition were performed using anti-GATA1 antibody purchased from Abcam (cat no. 11852). See Figure 20A for a comparison of 5 commercially available anti-GATA1 antibodies.

Day 1: 10^7 cells per sample were cross-linked with 1% formaldehyde for 10min with shaking at room temperature. The reaction was quenched with 0.125M glycine for 10min with shaking at room temperature. Cells were centrifuged at 3,000rpm for 5min at 4°C. Cell pellet was resuspended in 1ml ice cold PBS and 1% proteinase inhibitors, transferred in eppendorfs and centrifuged at 3,000rpm for 5min at 4°C. Cells were lysed with 200µl ChIP Lysis buffer/ChIP reaction on ice for 10min. Chromatin was sonicated in a Diagenode bioruptor using high amplitude, 30sec on and 30sec off settings, for 18min to break up the DNA into fragments of maximum 500bp length. Tubes were centrifuged at maximum speed for 10min at 4°C. Supernatant was transferred in FACS tubes and 1/10 of the volume, i.e 20µl were transferred in a separate tube, mixed with 280µl of ChIP Dilution buffer and stored in the -20°C to be used as the input sample. The 180µl of the lysate were mixed with 1.8ml of ChIP Dilution buffer and the samples were pre-cleared by adding 40µl of salmon sperm blocked protein A or G beads (Millipore)/ 2ml sample for 1hr with rotation at 4°C. This pre-clearing step was performed in a master mix tube for both the antibody and the IgG/ or no antibody samples. Tubes were centrifuged at 3,000rpm for 5min at 4°C. Supernatant was aliquoted into FACS tubes (2ml in each) and 5µg of the antibody or 5µl of the ChIP Dilution buffer (no antibody control sample) were added for 4hrs with rotation at 4°C. 60µl of protein A or G beads were added to the samples for an overnight incubation with rotation at 4°C.

Day 2: FACS tubes were centrifuged at 3,000rpm for 5min at 4°C and supernatant was discarded. Beads were resuspended in 1ml of Low Salt ChIP buffer and transferred to eppendorfs for 10min incubation with rotation at 4°C. Beads were concentrated by centrifuging at 1,000rpm for 1min at 4°C. Supernatant was discarded. Beads were washed with 1ml of High Salt buffer for 10min with rotation at 4°C and centrifuged at 1,000rpm for 1min at 4°C. Supernatant was discarded. Beads were washed with 1ml of $LiCl_2$ buffer for 10min with rotation at 4°C and centrifuged at 1,000rpm for 1min at 4°C. Supernatant was discarded. Beads were washed 2x with 1ml of TE buffer for 10min with rotation at 4°C and centrifuged at 1,000rpm for 1min at 4°C. Supernatant was discarded. Samples were briefly centrifuged at maximum speed and all of the supernatant was discarded. Complexes were eluted by adding 150µl of freshly made Elution buffer for 15min with rotation at room temperature. Beads were centrifuged at 3,000rpm for 5min at room temperature and supernatant was transferred in new eppendorfs. The elution step was repeated by adding 150µl of freshly made Elution buffer for 15min with rotation at room temperature. Beads were centrifuged at 3,000rpm for 5min at room temperature and supernatant was added to the same eppendorf. For the reverse cross-linking of the chromatin 12µl of 5M NaCl was added to each of the ChIP samples as well as in the input samples (stored at -20°C) for 4hrs at 65°C. DNA was cleaned up by the addition of 6µl 0.5M EDTA, 12µl 1M Tris-HCl pH 6.5, 1.2µl 10mg/ml proteinase K for 1hr at 45°C. 500µl of phenol/chloroform was added in each sample and tubes were centrifuged at 13,000rpm for 5min at room temperature. Supernatant was transferred in new eppendorfs and 1ml of 100% ethanol, 50µl 3M NaAc and 1µl glycogen carrier were added and tubes were stored overnight at -20°C.

Day 3: DNA was centrifuged at 13,000rpm for 10min at 4°C and pellet was washed with 70% ethanol. DNA pellet was air-dried after last step and resuspended in 200µl of MilliQ. Input samples were diluted 1:10 and 3µl of ChIP DNA or diluted input DNA per real time PCR were used.

ChIP Lysis Buffer: 50mM Tris-HCl pH 8.1, 10mM EDTA, 1% SDS and 1% protease inhibitors

ChIP Dilution Buffer: 0.01% SDS, 1.1% Triton X-100, 1.2mM EDTA, 16.7mM Tris-HCl pH 8.1, 167mM NaCl and 1% protease inhibitors

Low Salt Buffer: 20mM Tris-HCl pH 8.1, 2mM EDTA, 150mM NaCl, 1% Triton X-100, 0.1% SDS and 1% protease inhibitors

High Salt Buffer: 20mM Tris-HCl pH 8.1, 2mM EDTA, 500mM NaCl, 1% Triton X-100, 0.1% SDS and 1% protease inhibitors

LiCl Buffer: 10mM Tris-HCl pH 8.1, 0.5M LiCl, 1mM EDTA, 1% NP-40, 1% DOC and 1% protease inhibitors

TE Buffer: 10mM Tris-HCl pH 8.0, 1mM EDTA and 1% protease inhibitors

Elution Buffer: 0.1M NaHCO3 and 1% SDS

Protease inhibitors were added fresh prior to using the buffers.

Streptavidin ChIP

$5x10^7$ $Ter119^-$ and $5x10^7$ $Ter119^+$ ES cells were cross-linked with 1% formaldehyde for 10min at room temperature. The reaction was quenched with a final concentration of 0.125M glycine and lysed cells were sonicated to obtain fragments with average length of 150-200bp. Sheared chromatin was pre-cleared with protein A-magnetic beads blocked in salmon sperm and 1%gelatin fish skin in ChIP dilution buffer. Supernatant lysates were incubated overnight with blocked streptavidin- and protein A-magnetic beads. Beads were washed two times with 2% SDS and further processed as described by Schuh et al. 2005.

Day 1: $5x10^7$ cells per sample were cross-linked with 1% formaldehyde for 10min with shaking at room temperature. The reaction was quenched with 0.125M glycine for 10min with shaking at room temperature. Cells were centrifuged at 3,000rpm for 5min at 4°C. Cell pellet was resuspended in 1ml ice cold PBS and 1% proteinase inhibitors, transferred in eppendorfs and centrifuged at 3,000rpm for 5min at 4°C. Cells were lysed with 200µl ChIP Lysis buffer/ChIP reaction on ice for 10min. Chromatin was sonicated in a Diagenode bioruptor using high amplitude, 30sec on and 30sec off settings, for 18min to break up the DNA into fragments of maximum 500bp length. Tubes were centrifuged at maximum speed for 10min at 4°C. Supernatant was transferred in FACS tubes and 1/10 of the volume, i.e 20µl were transferred in a separate tube, mixed with 280µl of ChIP Dilution buffer and stored in the -20°C to be used as the input sample. The 180µl of the lysate were mixed with 1.8ml of ChIP Dilution buffer and the samples were pre-cleared by adding 50µl of blocked protein A or G Dynabeads (Invitrogen)/ 2ml sample for 1hr with rotation at 4°C.

In order to block the beads: 3x washes with PBS, 5min each with rotation at room temperature were followed by an 1hr incubation in 1ml ChIP Dilution Buffer with 1mg sonicated salmon sperm DNA and 1% fish skin gelatin with rotation at room temperature.

The magnetic rack was used to concentrate the beads and supernatant was transferred in FACS tubes, whereas the beads were discarded. 50µl of blocked M280 Dynabeads (Invitrogen) and 1% protease inhibitors were added to the tubes and samples were incubated overnight with rotation at 4°C.

Day 2: Beads were concentrated using a magnetic rack and supernatant was discarded. Beads were resuspended in 1ml of 2% SDS and transferred to eppendorfs for 10min incubation with rotation at 4°C. Beads were concentrated using a magnetic rack and supernatant was discarded. Beads were washed with 1ml of Low Salt buffer for 10min with rotation at 4°C and concentrated using a magnetic rack. Supernatant was discarded. Beads were washed with 1ml of High Salt buffer for 10min with rotation at 4°C and concentrated using a magnetic rack. Supernatant was discarded. Beads were washed with 1ml of $LiCl_2$ buffer for 10min with rotation at 4°C and concentrated using a magnetic rack. Supernatant was discarded. Beads were washed 2x with 1ml of TE buffer for 10min with rotation at 4°C and concentrated using a magnetic rack. Supernatant was discarded. Samples were briefly centrifuged at maximum speed and all of the supernatant was discarded. Complexes were eluted by adding 150µl of freshly made Elution buffer for 15min with rotation at room temperature. Beads were concentrated using a magnetic rack and supernatant was transferred in new eppendorfs. The elution step was repeated by adding 150µl of freshly made Elution buffer for 15min with rotation

at room temperature. Beads were concentrated using a magnetic rack and supernatant was added to the same eppendorf. For the reverse cross-linking of the chromatin 12µl of 5M NaCl was added to each of the ChIP samples as well as in the input samples (stored at -20°C) for 4hrs at 65°C. DNA was cleaned up by the addition of 6µl 0.5M EDTA, 12µl 1M Tris-HCl pH 6.5, 1.2µl 10mg/ml proteinase K for 1hr at 45°C. 500µl of phenol/chloroform was added in each sample and tubes were centrifuged at 13,000rpm for 5min at room temperature. Supernatant was transferred in new eppendorfs and 1ml of 100% ethanol, 50µl 3M NaAc and 1µl glycogen carrier were added and tubes were stored overnight at -20°C.

Day 3: DNA was centrifuged at 13,000rpm for 10min at 4°C and pellet was washed with 70% ethanol. DNA pellet was air-dried after last step and resuspended in 200µl of MilliQ. Input samples were diluted 1:10 and 3µl of ChIP DNA or diluted input DNA per real time PCR were used.

ChIP Lysis Buffer: 50mM Tris-HCl pH 8.1, 10mM EDTA, 1% SDS and 1% protease inhibitors

ChIP Dilution Buffer: 0.01% SDS, 1.1% Triton X-100, 1.2mM EDTA, 16.7mM Tris-HCl pH 8.1, 167mM NaCl and 1% protease inhibitors

Low Salt Buffer: 20mM Tris-HCl pH 8.1, 2mM EDTA, 150mM NaCl, 1% Triton X-100, 0.1% SDS and 1% protease inhibitors

High Salt Buffer: 20mM Tris-HCl pH 8.1, 2mM EDTA, 500mM NaCl, 1% Triton X-100, 0.1% SDS and 1% protease inhibitors

LiCl Buffer: 10mM Tris-HCl pH 8.1, 0.5M LiCl, 1mM EDTA, 1% NP-40, 1% DOC and 1% protease inhibitors

TE Buffer: 10mM Tris-HCl pH 8.0, 1mM EDTA and 1% protease inhibitors

Elution Buffer: 0.1M NaHCO3 and 1% SDS

Protease inhibitors were added fresh prior to using the buffers.

ChIP-Seq Libraries and Data Processing

Anti-GATA1 duplicates and 'no Antibody' ChIP DNA were processed for Illumina high throughput sequencing at the Genomics Group of the Wellcome Trust Centre for Human Genetics (Oxford, UK). ChIP-Seq libary construction and analysis were performed according to the Illumina protocol.

Real-Time PCR

Primers for *gata-1*, *gata-2*, *c-myc*, *pu.1*, *nrf2, klf, bach1*, *ho-1*, *nfe2l2*, *myc*, *mad*, *max* and *mxd4* were designed using Primer Express. SYBR Green-based quantitative qPCR (ABI SYBR Green PCR master mix, ABI) was performed on input, GATA1 immunoprecipitated and IgG control material, or on input and streptavidin ChIP material, for three independent $Ter119^-$ and $Ter119^+$ samples. Samples were analyzed in triplicates using an ABI StepOne™Real-Time PCR System (ABI, Foster city, CA). Enrichment was determined by dividing the amount of this in the immunoprecipitated fraction by the amount of target sequence in input DNA (Litt et al., 2001). Results were analyzed relative to a control sequence in the *gapdh* locus. IgG ChIP enrichment values were subtracted from their corresponding GATA-1 immunoprecipitated ChIP enrichment values. Primers for sequences that were negative for GATA-1 binding were tested for all the gene loci assayed. Primer sequences are listed in Table 5, Appendix.

Expression profiling

Expression profiling was performed using Sentrix Mouse-6 Expression BeadChip arrays from Illumina® (Illumina Inc., San Diego, CA) on three independent $Ter119^-$ and $Ter119^+$ purified erythroid progenitor populations. RNA was extracted using RNAqueous (Ambion, Austin, TX) and assessed for integrity using the Agilent Bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA). All samples presented RNA integrity (RIN) scores above 9.5. Samples were then processed for array hybridization

and data accumulation at the Genomics Group, Wellcome Trust Centre for Human Genetics. Briefly, amplification was performed using the Illumina Total Prep RNA Amplification kit (Ambion) according to the manufacturer's instructions. Amplified cRNA was hybridized to the BeadChip arrays according to the manufacturer's guidelines and detected with Fluorolink Streptavidin-Cy3 (Amersham Biosciences). The raw intensity values obtained for the scanned array images were compiled using Illumina Bead Studio. The data was filtered so that any probe with a detection score <0.95 across all samples was removed from the analysis prior to log transformation (base2) and quantile normalization. Gene expression was normalized using the RMA method from R and a p-value cut-off of 0.01 was used for the two tailed t-test. Genes with a 2-fold change of expression were defined as differentially expressed.

Expression profiling was performed using the Mouse Genome 430 2.0 Affymetrix GeneChip Arrays. RNA was isolated from 10^7 cells, using the RNAeasy Mini Kit from QIAgen, cat. no. 74104, following the manufacturer's protocol. Samples were DNase Q treated and assessed for integrity using the Agilent Bioanalyzer. Samples with RNA integrity (RIN) scores above 8.0 were processed for array hybridization and data accumulation at the Expression Profiling Unit at the BSRC "Alexander Fleming" Institute.

Cytospin preparations

10^5 cells were cytocentrifuged and stained with May-Grünwald/ Giemsa stain. Slides were incubated for 3min in undiluted May-Grünwald solution, rinsed and further incubated for 15min in 1:20 diluted Giemsa stain.

For the benzidine staining: 60 mg of o-diansidine (Sigma D-9143) were diluted in 29.7ml H_2O and 0.5ml glacial acetic acid and incubated at 37°C for 30-60 minutes, until fully dissolved. Reagent was stored in a light-proof container at 4°C for several months. Prior to staining, fresh benzidine reagent was prepared by mixing one part of hydrogen peroxide to 10 parts of benzidine reagent. In order to quantitate hemoglobinization, one part of the above preparation was added to 10 parts of the media containing cells. Cells were incubated for 1-2 minutes at room temperature and slides were prepared. Positive cells appeared brown and negative cells were clear.

FACS analysis

Measuring ROS: 10^6 cells were incubated with 10μM hemin for 4hrs or 100μM H_2O_2 for 6hrs and stained with 5μM 2',7'-dichlorodihydrofluorescein diacetate (H2DCFDA) (Invitrogen) for 30min. Cells were analyzed in a FACS Calibur flow cytometer (BD Biosciences) using the CellQuest software (BD Biosciences).

Measuring apoptosis/ cell death: 10^6 cells were stained with Annexin V FITC (sc-4252, Santa Cruz Biotechnology) according to the manufacturer's protocol. Samples were incubated with Propidium Iodide (PI) 10min in the dark prior to FACS analysis.

Measuring DNA content: 10^6 cells were fixed with EtOH and stained with PI/RNase A for 30min. Cells were analyzed in a FACS Calibur flow cytometer (BD Biosciences) using the CellQuest software (BD Biosciences).

Chapter 3

Results

3.1 A biotin tag knock-in approach for studying GATA-1 functions in distinct hematopoietic lineages.

GATA-1 is a DNA binding zinc-finger transcription factor that is essential for the differentiation of erythroid, megakaryocytic, eosinophilic and mast cell lineages in hematopoiesis and of dendritic cells in the bone marrow, as evidenced primarily by the analysis of GATA-1 gene knockout mice (Ferreira et al., 2005; Gutierrez et al., 2004). The essential roles of GATA-1 as a key regulator of these hematopoietic lineages has also been clearly demonstrated by the ectopic expression of GATA-1 in avian and mammalian hematopoietic progenitor cells (or even cells committed to the myeloid lineage) which re-programs them towards hematopoietic lineages normally regulated by GATA-1, e.g. erythroid, megakaryocytic and eosinophilic (Kulessa et al., 1995; Heyworth et al., 2002; Hirasawa et al., 2002; Iwasaki et al., 2003). It is of note that the different hematopoietic lineages that GATA-1 regulates represent distinct cellular differentiation pathways (though erythroid and megakaryocytic cells share a common differentiation origin). Thus, an important question arises as to how GATA-1 carries out its key regulatory functions in distinct hematopoietic lineages?

A number of possibilities may address this question. First, GATA-1 interacts with different protein partners in distinct lineages, e.g. with FOG-1 in erythroid and megakaryocytic cells or C/EBPα in eosinophils (Tsang et al., 1997; McNagny & Graf, 2002). Second, GATA-1 protein levels in multipotential progenitor cells may play a role in deciding cell fate (McNagny and Graf, 2002; Kulessa et al., 1995). There is also evidence that domains of the GATA-1 protein may serve distinct functions in different hematopoietic lineages (Hirasawa et al., 2002). Lastly, DNA sequence may also play a role in the genetic programs that GATA-1 regulates in distinct lineages. For example, a GATA-E-box motif that binds the GATA-1/TAL-1 complex is frequently associated with erythroid genes (Wadman et al., 1997) whereas a high affinity palindromic double GATA site is frequently associated with eosinophilic genes (Du et al., 2002; Yu et al., 2002). Despite all this evidence, work addressing the basis of GATA-1 differential regulatory functions in distinct hematopoietic lineages remains largely fragmented.

The long term aim of the work described here is to characterize in a much more systematic way GATA-1 functions in regulating the differentiation of distinct hematopoietic lineages. This translates into two aims: (i) the characterization of the different protein interactions that GATA-1 undergoes and (ii) the distinct (and overlapping?) genetic programs that GATA-1 regulates in different hematopoietic lineages. To these ends, we will use the biotinylation tagging of GATA-1 as a uniform platform for the proteomic and global gene target analysis, by establishing a system for the expression by a knock-in approach of physiological levels of biotin-tagged GATA-1 in all hematopoietic lineages where it is active.

Biotinylation tagging involves the fusion of a short (14-23aa) peptide tag to the protein of interest. The tag is recognized by the E. coli BirA protein biotin ligase which biotinylates it (Cull & Schatz, 2000). Thus, co-expression in cells of the tagged protein together with BirA results in the in vivo biotinylation of the tagged protein, which is then amenable to various manipulations by using the very high affinity biotin-streptavidin system. The Strouboulis lab has applied this approach in biotinylating GATA-1 and using streptavidin binding of nuclear extracts in characterizing GATA-1 protein complexes in cultured erythroid cells (de Boer et al., 2003; Rodriguez et al., 2005). Biotinylation tagging has also been used successfully in place of antibodies in ChIP assays, with obvious advantages due το the very high affinities of streptavidin binding allowing high stringencies to be employed (Kim et al., 2008; van Werven & Timmers, 2006; Viens et al., 2004; de Boer et al., 2003). Recent work has led to the further optimization of conditions for streptavidin-ChIP that result in optimal enrichment with lower background and improved ChIP sequencing results (Kolodziej et al., 2009). Soler et al. (2011) showed that the high affinity reagents available for the biotin tags allow high enrichments in ChIP experiments even for factors expressed at low levels in the cells and also for TFs that do not directly bind DNA (and thus could be more difficult to crosslink to their cognate binding sites) like chromatin modifying enzymes and transcription co-factors. This observation is of particular importance since although GATA-1 is highly expressed in erythroid cells, its expression in eosinophils or mast cells is significantly lower. Moreover, there has been no evidence to-date showing that biotinylation affects the physiological properties of tagged proteins. Thus, biotinylation tagging of GATA-1 can be used as a uniform platform for the characterization of protein complexes and gene target networks even in lineages where it is expressed in low levels.

I. Targeting strategy

In order to set up a system for expressing physiological levels of biotin-tagged GATA-1 protein in all tissues where it is normally active, we knocked-in a short biotin tag into the coding region of the X-linked mouse GATA-1 gene in male ES cells. For the purpose of this experiment three different embryonic stem (ES) cell lines were used; two strains of wild type ES cells -S129 and E14- and a knock-in S129 ES cell line (named iB10p) that ubiquitously expresses the BirA biotin ligase as a knock-in in the Rosa26 locus (Driegen et al., 2005). For positive selection of the correctly targeted clones, a neo cassette, flanked by loxP sites, was inserted upstream of the biotin tag; for negative selection, a thymidine kinase cassette was cloned downstream of the 3' homology region (Figure 7).

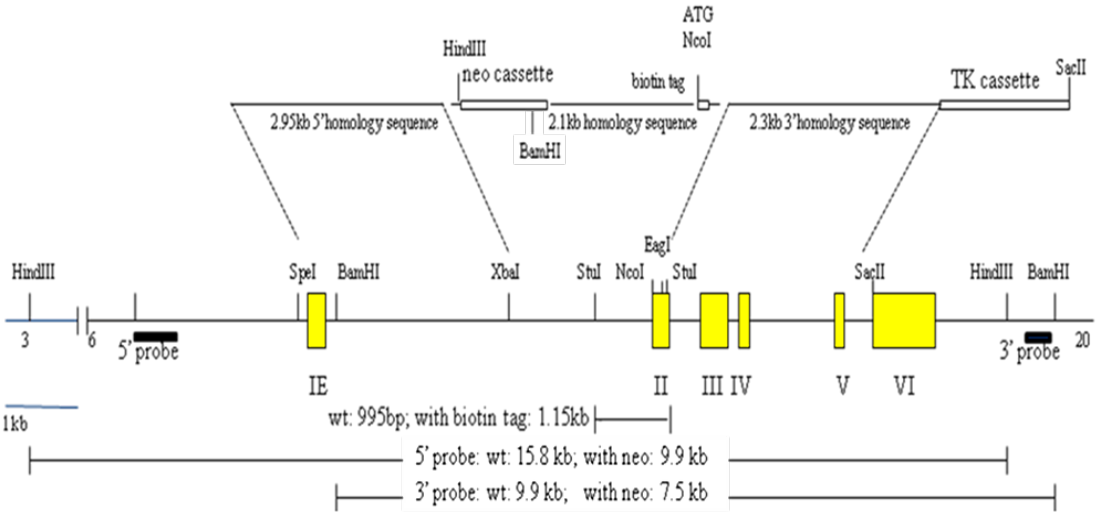


Figure 7: Cloning strategy for the bioGATA-1 knock-in.

The biotin tag has been inserted in the GATA-1 ATG start-codon in exon II, followed by the TEV protease cleavage site. A neo cassette, flanked by loxP sites, has been cloned 2kb upstream of the tag and a thymidine kinase (TK) cassette has been inserted 3' of the 3' homology arm, at around 3kb 3' of the tag for counterselection. The targeting construct includes two regions of homology, a 2.95kb 5' homology sequence and a 2.3kb 3' homology sequence. IE: Exon I-erythroid. Also shown are the restriction digests, the 5' and 3' probes used for screening recombinant clones and the sizes of the DNA fragments detected.

II. Screening of ES clones

In total, 317 targeted ES cell clones were screened by Southern blots, 57 of which were identified as positive for the knock-in. More specifically, 43 clones out of 127 iB10p clones, 5 out of 52 E14 clones and 9 out of 138 S129 clones were identified as being correctly targeted. All targeted clones were further screened by Southern blots, using probes for both 5' and 3' arms, as well as for the neo gene (Figure 8).

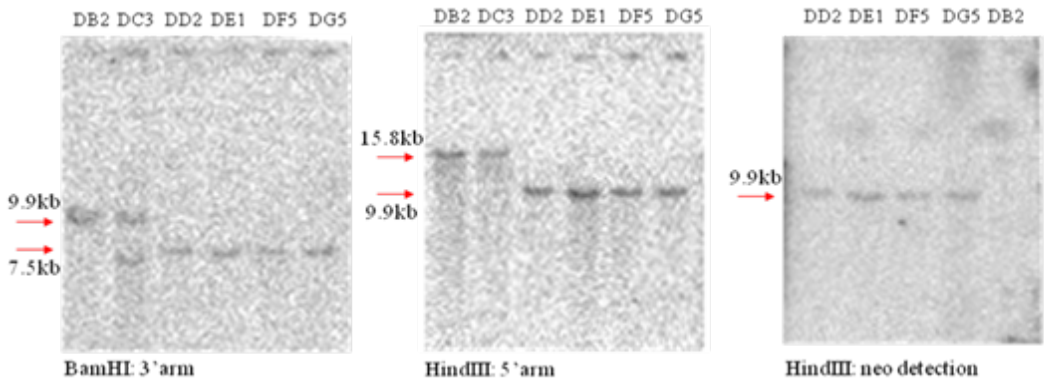


Figure 8: Screening the S129 clones by Southern blots.

Four positive and two negative clones digested with BamHI for the 3' targeting screening and HindIII for the 5' targeting screening and neo detection (see Figure 7).

The screening process was completed with a karyotype analysis of the S129 correctly targeted clones in order to select two clones for injections into blastocysts to generate chimeric mice. The same methodology, screening by Southern blots and karyotyping, was also followed for eight out of the 43 correctly targeted iB10p cell line clones, before proceeding to the neo cassette excision by Cre recombinase.

III. Screening of bioGATA-1 mice

After karyotyping, S129 targeted clones DE1 and DF5 (Figure 8) were selected for blastocyst injections in pseudopregnant C57/DBA2J female mice. In total, 13 chimeric male mice were born and crossed with C57/BL6 wild-type female mice for the generation of heterozygous female bioGATA-1 knock-in mice (Figure 9). Only agouti F1 female mice were screened in the first round of matings since GATA-1 is X-linked and male mice inherit the X-chromosome from the wild-type mother and the Y-chromosome from the chimeric father.

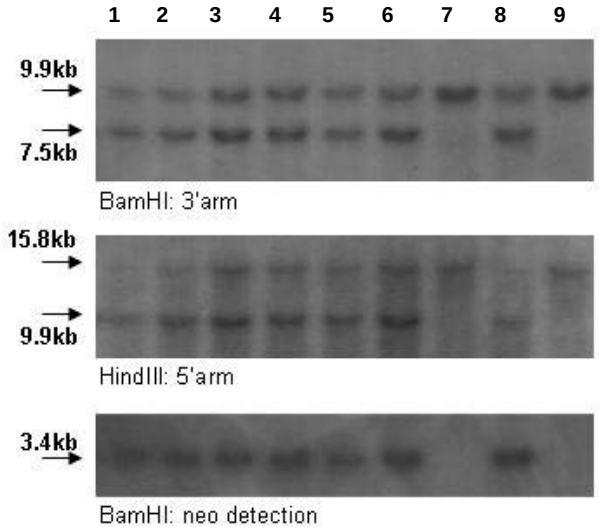


Figure 9: Screening the bioGATA-1 F1 mice by Southern blots.
Genomic DNA from tail biopsies of nine F1 female mice were digested with HindIII for the 5' targeting screening and BamHI for the 3' targeting screening and for neo detection (see Figure 7). Seven mice appear heterozygous for the biotin tag GATA-1 knock-in (lanes 1-6, 8) whereas the remaining two mice are wild-type (lanes 7, 9). Lower panel shows the BamHI blot re-probed with the neo cassette probe.

Both DF5 and DE1 clones resulted in fertile chimeric offspring and heterozygous F1 female mice, however we decided to carry on with a single mouse line (DF5) and freeze the second line (DE1). The next round of matings included the F1 heterozygous females crossed with the BALB/c-TgN(CMV-Cre)1Cgn male mice (Schwenk et al., 1995), in order to excise the neo selection cassette. These mice express the Cre recombinase ubiquitously under the CMV promoter hence, expression of the Cre recombinase in the germline will result in the excision of the neo cassette, which is flanked by loxP sites (see Figure 7). Both male and female F2 bioGATA-1 CMV-Cre mice were born (Figure 10) and intercrossed to obtain F3 homozygous female bioGATA-1 CMV-Cre mice (Figure 11).

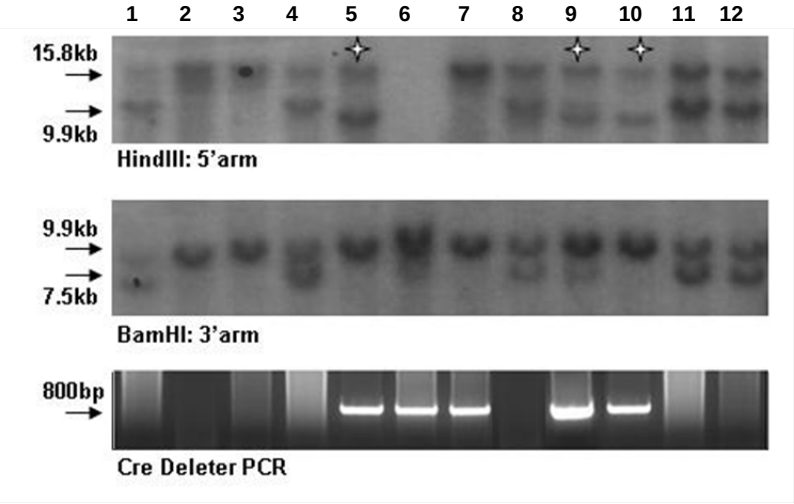


Figure 10: Screening the bioGATA-1 F2 mice by Southern blots and PCRs.
Genomic DNA tail biopsies of 12 F2 mice were digested with HindIII for the 5' probe and BamHI for the 3' probe (see Figure 7) and were also screened by PCR for detection of the Cre recombinase. Top panel: eight mice appear heterozygous for the bioGATA-1 knock-in (lanes 1, 4, 5, 8-12). Mice denoted with an asterisk are heterozygous female F2 bioGATA-1/CMV-Cre mice (lanes 5, 9, 10), in which the 9.9kb band has been replaced by a 8.9kb band, due to the neo gene excision. The mouse in lane 9 appears to have undergone partial Cre-mediated recombination. Neo excision is confirmed in the middle panel: after neo excision, the internal BamHI site disappears, resulting in a single wild type-like 9.9kb band (e.g. lanes 5 and 10). Bottom panel: Cre recombinase detection by PCR (lanes 5-10).

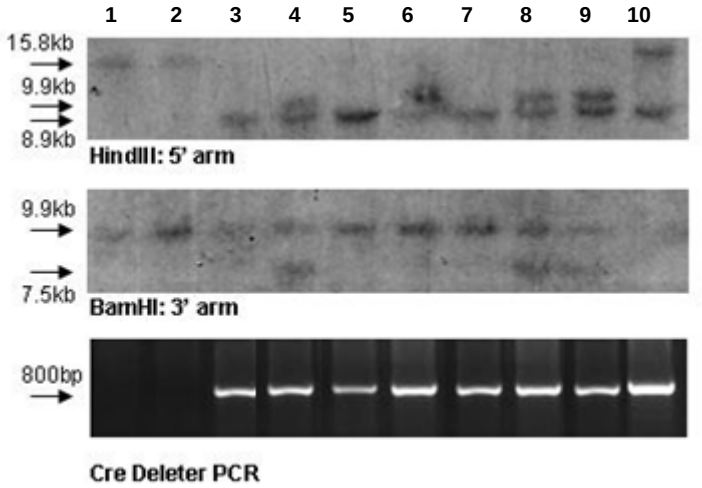


Figure 11: Screening the bioGATA-1 F3 mice by Southern blots and PCRs.
Genomic DNA tail biopsies of 10 F3 mice were digested with HindIII for the 5' probe and BamHI for the 3' probe (see Figure 7) and were also screened by PCR for detection of the Cre recombinase. Top panel: four female mice appear homozygous for the bioGATA-1 knock-in (lanes 3, 5-7), in which the 9.9kb HindIII band has been replaced by a 8.9kb band, due to the neo excision. Three F3 female mice are heterozygous for the bioGATA-1/CMV-Cre transgenes (lanes 4, 8, 9), whereby one of the two X chromosomes carries the neo deleted bioGATA-1 knock-in. Two mice are wild-type C57/BL6 (lanes 1, 2) and one mouse is heterozygous for the bioGATA-1 knock-in, with the neo cassette excised (lane 10). Neo excision is confirmed in the middle panel with a 3' probe: after neo excision, the internal BamHI site disappears, resulting in a single wild type-like 9.9kb band. Bottom panel: Cre recombinase detection by PCR (lanes 3, 5-7, 10).

The next round of matings involved the crossing of the F3 mice with transgenic mice that ubiquitously express 3xHA tagged BirA biotin ligase (Papadopoulos and Strouboulis unpublished) under the promoter of the hnRNPA2 housekeeping gene (Katsantoni et al., 2007) in order to obtain mice that express fully biotinylated GATA-1 (Figure 12A). In addition, in order to initiate the study of GATA-1 functions in the eosinophilic lineage, F3 bioGATA-1 knock-in mice were crossed with interleukin (IL)-5 overexpressing transgenic mice which exhibit lifelong eosinophilia resulting in over 265-fold higher number of eosinophils than normal (Dent et al., 1990) (Figure 12B).

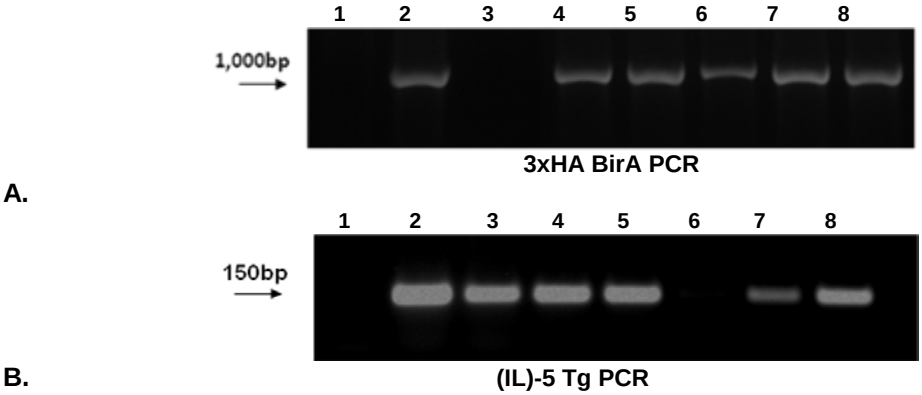


Figure 12: Screening the bioGATA-1 F4 mice by PCRs for the detection of BirA and IL-5 Tg transgenes.

Genomic DNA tail biopsies of 6 F4 mice were screened by PCR for detection of the 3xHA BirA (A) and IL-5 transgenes (B).
A: Five mice carry the BirA transgene (lanes 4-8) and one mouse is wild-type C57/BL6 (lane 3).
B: Five mice carry the IL-5 transgene (lanes 3-5 and 7, 8) and one mouse is wild-type C57/BL6 (lane 6). No DNA template control is loaded in lane 1 and a positive control in lane 2, in both (A) and (B) panels.

All F4 mice developed normally, were fertile and when their blood was analyzed, no signs of anemia or thrombocytopenia were observed (Table 1).

Genotype	Number of animals	Erythrocytes (x10^4/μl)	Hematocrit (%)	Hemoglobin (g/dl)	Platelets (x10^4/μl)
WT	5	842± 70	44.0± 2.6	15.1± 0.6	48.0± 7.6
bioGATA1/Cre/BirA	7	843± 60	45.4± 3.3	14.8± 1.1	49.5± 5.4

Table 1: Peripheral blood indices from C57/BL6 and bioGATA1/Cre/BirA mice.

IV. Physiologically expressed biotin tagged GATA-1 in erythroid cells

We next wished to assess expression of biotin tagged GATA-1 in the erythroid lineage in mice. E12.5 fetal livers (which are ~95% erythropoietic) were extracted from C57/BL6 triple transgenic for bioGATA-1 knock-in, CMV-Cre and BirA pregnant females -intercrossed with males of the same genotype- for nuclear extracts and chromatin preparations. Western blot analysis showed that the endogenously expressed GATA-1 protein in these embryos is tagged, as evidenced by a shift in the molecular weight of the protein compared to the GATA-1 protein isolated from C57/BL6 wild-type and BirA single transgenic littermates (lanes 4 and 5 versus lanes 2 and 3, Figure 13). Moreover, streptavidin-HRP results in the detection of the biotinylated GATA-1 protein only in extracts from the triple transgenic knock-in embryos and from the bioGATA-1 MEL C88 control cells. Nuclear extracts from wild-type and BirA single transgenic embryos do not give a signal when probed with streptavidin-HRP (Figure 13).

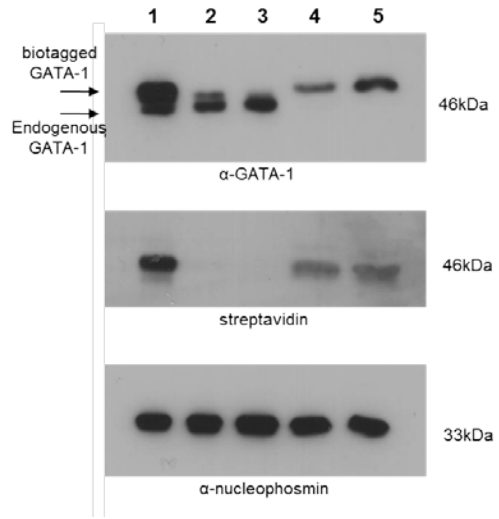


Figure 13: Detecting bioGATA-1 in E12.5 fetal liver cell extracts by Western blots.

Nuclear extracts from E12.5 fetal liver cells isolated from C57/BL6 embryos triple transgenic for bioGATA-1 knock-in, CMV-Cre, BirA (lanes 4, 5), C57/BL6 wild-type (lane 2) and single transgenic BirA embryos (lane 3). Additional upper band seen in C57/BL6 wild-type (lane 2) extracts is most likely due to post-translational modifications previously reported for GATA-1 (Hernandez-Hernandez et al., 2006; Lee et al., 2009; Collavin et al., 2004; de Thonel et al., 2010; Lamonica et al., 2006). Control extracts from MEL C88 bioGATA-1 cells (bottom band corresponds to the endogenous protein and top band to the biotin tagged GATA-1 protein; de Boer et al., 2003) loaded in lane 1.
Protein was detected with the N6 rat monoclonal antibody (Santa Cruz) against the N-terminal region of GATA-1 (top panel) and with streptavidin-HRP (middle panel). Nucleophosmin was used as protein loading control (bottom).

Western blots with four different N- or C-terminal anti-GATA-1 antibodies, also showed that the endogenously expressed GATA-1 protein in embryos triple transgenic for bioGATA-1 knock-in, CMV-Cre and BirA appears to be tagged as evidenced by a shift in the molecular weight compared to the GATA-1 protein isolated from C57/BL6 wild-type littermates (Figure 14).

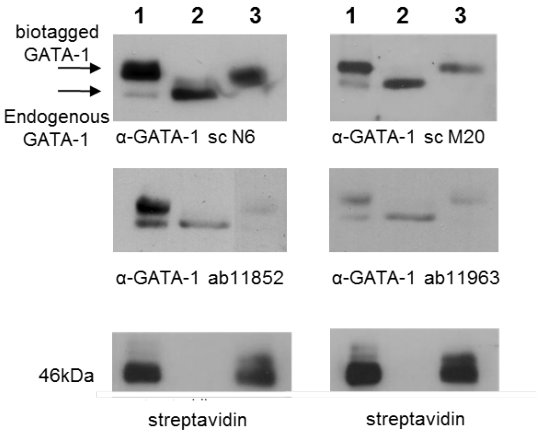


Figure 14: Detecting bioGATA-1 with N- and C-terminal αGATA-1 antibodies by Western blots.
Nuclear extracts from E12.5 fetal liver cells isolated from C57/BL6 embryos triple transgenic for bioGATA-1 knock-in, CMV-Cre, BirA (lane 3) and C57/BL6 wild-type embryos (lane 2). Control extracts from MEL C88 bioGATA-1 cells (bottom band corresponds to the endogenous protein and top band to the biotin tagged GATA-1 protein, de Boer et al., 2003) loaded in lane 1.
Protein was detected with the N6 rat monoclonal (Santa Cruz) and the rabbit polyclonal 11852 antibody (Abcam) against the N-terminal region of GATA-1 (left), the M20 goat polyclonal (Santa Cruz) and the rabbit polyclonal 11963 antibody (Abcam) against the C-terminal region of GATA-1 (right) and with streptavidin-HRP (bottom).

Immunoprecipitations with the N6 GATA-1 antibody against the N-terminal region result in the precipitation of practically all of the GATA-1 protein from soluble nuclear extracts, indicating that the presence of the biotin tag in the N-terminus does not affect the overall structure of the GATA-1 protein. No GATA-1 can be detected in the supernatants after the IPs and no significant cross-reactions are observed in the control IgG conjugates (Figure 15A). Moreover, streptavidin pull-downs precipitated biotin tagged GATA-1 in fetal liver cells from knock-in embryos only (Figure 15B).

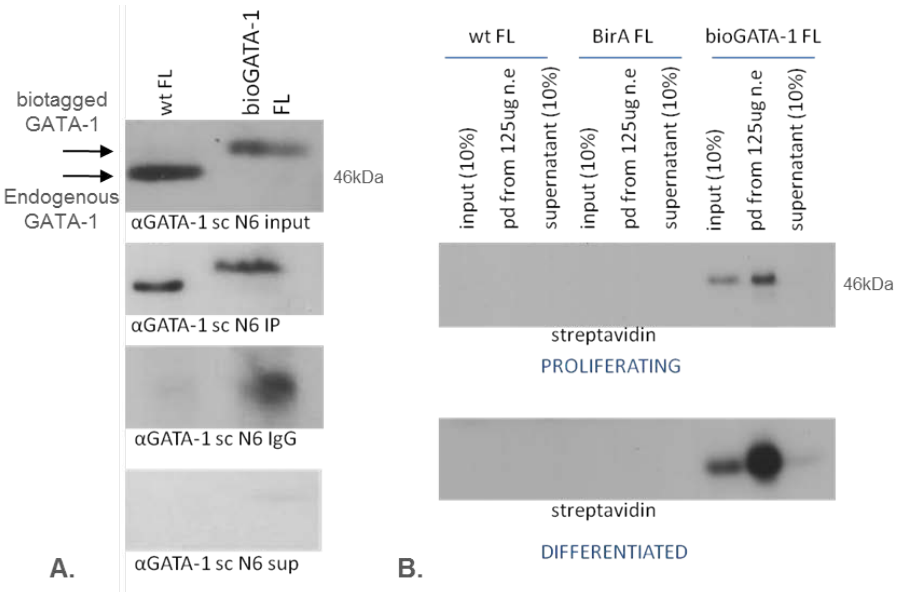


Figure 15: Detecting bioGATA-1 in GATA-1 immunoprecipitations and streptavidin pull-downs by Western blots.

A. Nuclear extracts equivalent to 10% of the material used for the immunoprecipitations (top), immunoprecipitations with the N6 rat monoclonal antibody (Santa Cruz) (IPs), control rat IgG conjugates (IgGs) and supernatants after IPs from E12.5 fetal liver cells isolated from C57/BL6 wild-type and C57/BL6 triple transgenic for bioGATA-1 knock-in, CMV-Cre, BirA embryos. Protein was detected with the N6 αGATA-1 antibody (Santa Cruz).
B. Streptavidin pull-downs from proliferating (top) and differentiated (bottom) E12.5 fetal liver cells isolated from C57/BL6 wild-type (left), C57/BL6 BirA transgenic (middle) and C57/BL6 triple transgenic for bioGATA-1 knock-in, CMV-Cre, BirA embryos (right) blotted with streptavidin-HRP.
Samples were loaded in the following order: nuclear extracts equivalent to 10% of the material used for the pull-downs loaded as control input samples, pull-down samples (pd) and 10% of supernatant after pds.

Following confirmation that the tagged GATA-1 protein is biotinylated in bioGATA-1/BirA mice, we proceeded to isolate *in vivo* GATA-1 protein complexes by streptavidin pull-downs from the nuclear extracts of E12.5 fetal liver cells. To this end, fetal liver proerythroblasts were cultured using two conditions: the first one allows proerythroblasts to proliferate whilst the second one induces them to undergo terminal erythroid differentiation (von Lindern et al. 2001; Schuh et al. 2005). Cells were expanded for three days or more, depending on starting cell number, in proliferating serum-free medium containing stem cell factor (SCF), erythropoietin (Epo) and dexamethasone. Nuclear extracts were isolated from 10^8-10^9 cells, whereas the remaining cells were differentiated using high Epo and transferrin. Large scale streptavidin pull-down experiments on bioGATA-1 knock-in embryos and on single transgenic BirA embryos were set up in duplicate using 10mg of protein per sample. Streptavidin bound material was processed for mass spectrometry and quantitative mass spectrometry (Figure 16).

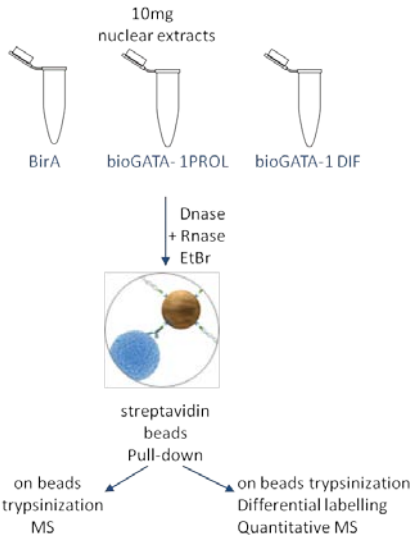


Figure 16: Overview of the streptavidin pull-down protocols applied to prepare samples for mass spectrometry analysis.
10mg of nuclear extracts were isolated from E12.5 fetal liver cells from differentiated BirA transgenic and from proliferating and differentiated bioGATA-1 knock-in embryos. Extracts were treated with DNase I, RNase A and ethidium bromide to remove RNA and DNA and were then used in streptavidin pull-down assays.
Replicate pull-downs were trypsinized on the beads and processed for mass spectrometry and also for on-bead trypsinization, differential labelling and quantitative mass spectrometry. Mass spectrometric analysis was carried out by Dr. Jeroen Demmers at the Erasmus Proteomics Center, Rotterdam, the Netherlands.

Mass spectrometry results are summarized in Table 2. It is noteworthy that FOG-1 (Friend of GATA1), a known partner of GATA-1, is more abundant in differentiated than in proliferating bioGATA-1 knock-in cells. This also accounts for the fact that members of the NURD complex are found exclusively in differentiated cells, as it is known that FOG-1 mediates interactions of NuRD with GATA-1 (Rodriguez et al., 2005). Moreover, this is the first *in vivo* indication of GATA-1 being in complex with a) the RUNX-1 and Lbp1a transcription factors, both of which are involved in hematopoiesis (Okuda et al., 1996; North et al., 1999; Lancrin et al., 2009, reviewed by Lam & Zhang, 2012; Jane et al., 1995; Volker et al., 1997; Zhou et al., 2000; Parekh et al., 2004) and b) members of the SWI/SNF family of chromatin remodeling proteins.

		Unique peptides					
	Protein name	Proerythrobl. bioGATA-1 #1	Proerythrobl. bioGATA-1 #2	Differentiated bioGATA-1 #1	Differentiated bioGATA-1 #2	Differentiated BirA	Comments
Transcription factors	GATA-1	9	9	9	9		
	Zfpm1 (FOG-1)	1	2	7	9		
	Runx1		2	5	3		Essential hematopoietic TF, HSCs and megakaryocytes
	Ubp1/Lbp1a			2	3		Essential for yolk sac angiogenesis
	Ikzf1/Ikaros				1		Essential for lymphoid differentiation, implicated in globin regulation
NuRD members	Chd4/Mi2β			9	7		
	HDAC1			4	4		
	MTA2			2	2		
	RbAp48			1			
	RbAp46				1		
	Chd5?				4		Neuronal specific, NuRD-like complex in brain, tumour suppressor
SWI/SNF members	Smarca4/Brg1	3	3	5	5		9 out of 13 SWI/SNF members identified
	Smarca5/Snf2h	2	3	7	13		
	Smarcc2	3	3	4	4		
	Smarcc1/BAF155		3	6	3		
	Smarcd2	3		2	2		
	Smarce1/BAF57	1	1	2	1		
	Pbrm1		2	2	2		Polybromo 1, PBAF subcomplex
	Smarcb1/SNF5				1		
	Arid1a				1		
	Actin-like 6A		1				

Table 2: Mass spectrometry analysis data summarized.

A preliminary validation of the data from the mass spectrometry analysis was carried out by streptavidin pull-downs and Western blots for the detection of FOG-1. As seen in Figure 17, FOG-1 is barely detected in pull-downs of proliferating bioGATA-1 knock-in cells whereas it is much more abundant in the extracts from differentiated bioGATA-1 knock-in cells, in agreement with the mass spec results.

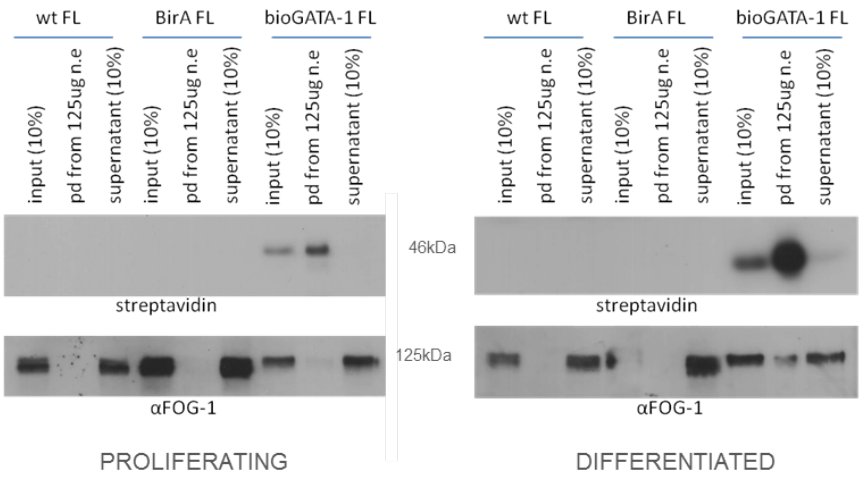


Figure 17: Detecting FOG-1 in streptavidin pull-downs from proliferating and differentiated E12.5 fetal liver cells by Western blots.

Input nuclear extracts equivalent to 10% of the material used for the pull-downs, pull-downs (pd) and 10% of supernatant after pds from proliferating (left) and differentiated (right) E12.5 fetal liver cells isolated from C57/BL6 wild-type, BirA transgenic and triple transgenic for bioGATA-1 knock-in, CMV-Cre, BirA embryos. GATA-1 was detected with streptavidin and FOG-1 with the goat polyclonal αFOG1 antibody M20 (Santa Cruz).

Current work in progress includes further validation and functional characterization of the different protein interaction networks that GATA-1 participates in and the different complexes it binds to in the proliferating versus differentiating erythroid cells. Results will be validated by additional streptavidin pull-downs and our aim is to do an initial characterization of the distinctive GATA-1 partners in the two differentiation stages.

Finally, E12.5 fetal liver cells were isolated from C57/BL6 bioGATA-1/BirA mice and expanded for three days under erythropoietic stress conditions induced by dexamethasone as previously described (von Lindernet al. 2001; Schuh et al. 2005) in order to isolate chromatin for streptavidin deep sequencing experiments. The use of streptavidin allowed us to use unprecedentedly highly stringent washes (2% SDS), which reduces significantly background binding and which will presumably aid in the identification of novel GATA-1 targets (Figure 18). Using the cell surface Ter119 receptor as a marker for mature erythroid cells we separated fetal liver cells into proerythroblasts ($Ter119^-$ fraction) and differentiated erythroid cells ($Ter119^+$ fraction). $5x10^7$ cells per sample were cross-linked with formaldehyde and sonicated chromatin from both $Ter119^-$ and $Ter119^+$ populations was bound to streptavidin beads. The efficiency of streptavidin-ChIP was assessed by qPCR using primers of well-known GATA-1 repressed and activated target genes (Figure 18). The same loci were assayed by GATA-1 antibody ChIP, using antibodies that bind to either end of the protein (antibody 11852 binds to the C-terminal domain of GATA-1 and antibody 11963 to the N-terminal). It should be noted that both GATA-1 antibodies were shown to be the most efficient in ChIP assays out of five commercial antibodies tested in pilot experiments (section 3.2). Overall, streptavidin-ChIP resulted in higher enrichments in practically all genes loci tested compared to antibody-ChIP (streptavidin-ChIP scale is increased by 2 and 1.5 times that of 11963 and 11852 GATA-1 ChIP, respectively). It is of note that streptavidin-ChIP consistently outperforms antibody-ChIP using the 11963 antibody which recognizes the N-terminus of GATA-1 where the boitinylation tag has also been inserted. Streptavidin-ChIP also performs at least as well as the antibody-ChIP using the 11852 antibody which recognizes the C-terminus of GATA-1.

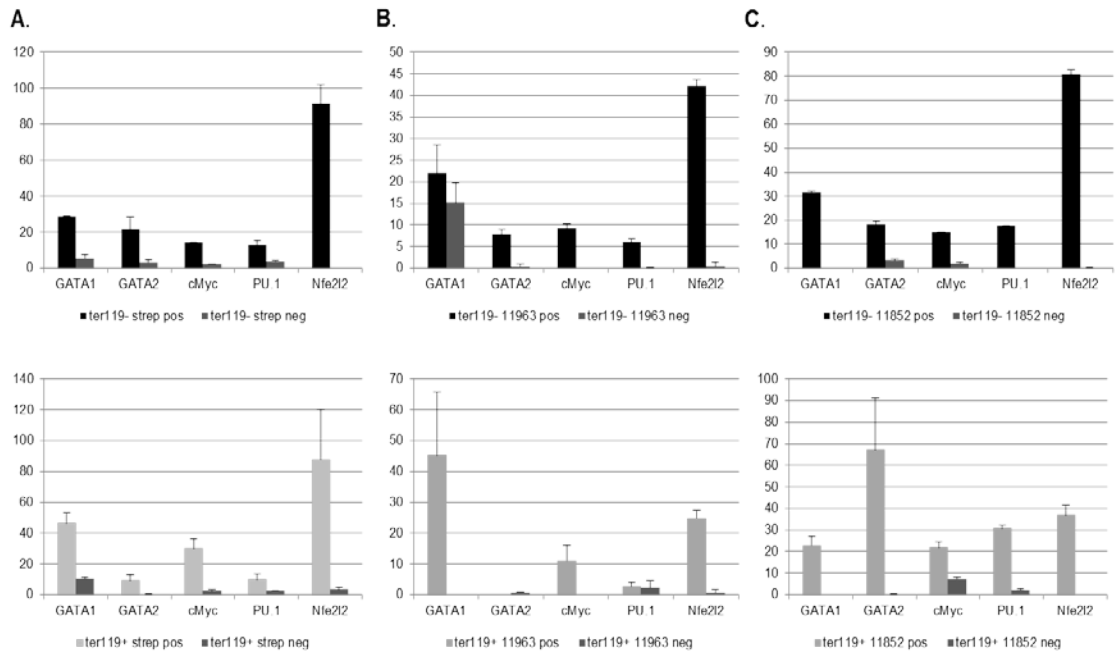


Figure 18: Profile of streptavidin and antibody-ChIP GATA-1 binding on the *gata-1*, *gata-2*, *c-myc*, *pu.1* and *nrf2* genes.
Real-time PCR analysis of $Ter119^-$ (top) and $Ter119^+$ (bottom) streptavidin (A) and GATA-1 (11963, B and 11852, C) ChIP on selected sites of *gata-1*, *gata-2*, *c-myc*, *pu.1* and *nfe2l2* genes (from left to right). The y-axis represents enrichment over input DNA, after normalization with the gapdh gene. The enrichment is the average of three replicate experiments.

V. Summary

In order to gain greater insight into how GATA-1 exerts differential regulatory functions in distinct hematopoietic lineages, we established a system for the expression of physiological levels of biotin-tagged GATA-1 in all hematopoietic lineages where it is active and initially used this system to study GATA-1 functions in the erythroid lineage.

BioGATA-1 knock-in mice were generated following blastocyst injections of successfully targeted S129 ES cells. Nuclear extracts and chromatin preparations were isolated from E12.5 fetal livers (which are ~95% erythropoietic) extracted from triple transgenic for bioGATA-1 knock-in, CMV-Cre and BirA pregnant females. We showed that the endogenously expressed GATA-1 protein in these embryos is tagged, as evidenced by a shift in the molecular weight compared to the GATA-1 protein isolated from C57/BL6 wild-type and BirA single transgenic littermates. Crossing these mice with a ubiquitously expressing 3xHA BirA transgenic line resulted in efficient biotinylation of the physiologically expressed GATA-1 protein in erythroid cells. Several lines of evidence presented here strongly suggest that GATA-1 protein structure and functions are not affected by the inclusion of the small tag. First, all blood parameters assayed for BioGATA-1 knock-in mice were indistinguishable from those for wild type mice. Second, immunoprecipitation experiments using the N6 GATA-1 antibody against the N-terminal region resulted in the purification of practically all of the GATA-1 protein from nuclear extracts of BioGATA-1 knock-in mice, indicating that there are no (gross) alterations in the structure of the protein as a result of the insertion of the tag. Third, streptavidin pull-downs and mass spec identified the well known GATA-1 interacting partner FOG-1 as co-purifying with biotin tagged GATA-1. Fourth, streptavidin ChIP clearly show biotin tagged GATA-1 to be binding to well-known GATA-1 gene targets in erythroid cells.

Mass spectrometry analysis of large scale streptavidin pull-downs from proliferating and differentiated erythroid cells showed that FOG-1 (Friend of GATA1), a known partner of GATA-1, is more abundant in differentiated than in proliferating bioGATA-1 knock-in cells. This also accounts for the fact that members of the NURD complex are found exclusively in differentiated cells, as it is known that FOG-1 mediates interactions of NuRD with GATA-1 (Rodriguez et al., 2005). Western blot analysis revealed that FOG-1 is barely detected in pull-downs of proliferating bioGATA-1 knock-in cells whereas it is much more abundant in the extracts from differentiated bioGATA-1 knock-in cells, in agreement with the mass spec results. Moreover, we found that GATA-1 is in complex with a) the RUNX-1 and Lbp1a transcription factors, both of which are involved in hematopoiesis and b) members of the SWI/SNF family of chromatin remodeling proteins. It is noteworthy that this is the first *in vivo* indication of GATA-1 being in complex with the aforementioned proteins, we now need to validate these results by western blots.

Work from previous reports suggests that GATA-1 undergoes multiple posttranslational modifications, including phosphorylation, sumoylation, ubiquitination and acetylation (Hernandez-Hernandez et al., 2006; Lee et al., 2009; Collavin et al., 2004; de Thonel et al., 2010; Lamonica et al., 2006). However, all evidence is based on work on cell lines or in vitro transfected cells. To these ends, we have initiated the characterization of the *in vivo* GATA-1 post-translational modifications in proliferating versus differentiated erythroid cells. We prepared large scale streptavidin pull-downs of whole cell lysates from E12.5 fetal liver cells, isolated from C57/BL6 bioGATA-1/BirA mice, upon the addition of phosphatase, deacetylase, deubiquitinase and desumoylase inhibitors to ensure that GATA-1 retains the *in vivo* post-translational modifications it undergoes through differentiation. Treatment of the isolated whole cell lysates with phosphatases such as λ-phosphatase or calf intestine phosphatase (CIAP) resulted in a shift in the molecular weight of GATA-1 band as detected by western blots, providing thus evidence for *in vivo* GATA-1 phosphorylation (Figure 19A). Large scale streptavidin pull down samples from both erythroid cell extracts were sent for mass spectrometry (Figure 19B). Our work will afford for the first time important insight into the *in vivo* post-translational modifications of GATA-1 in all tissues where GATA-1 is active.

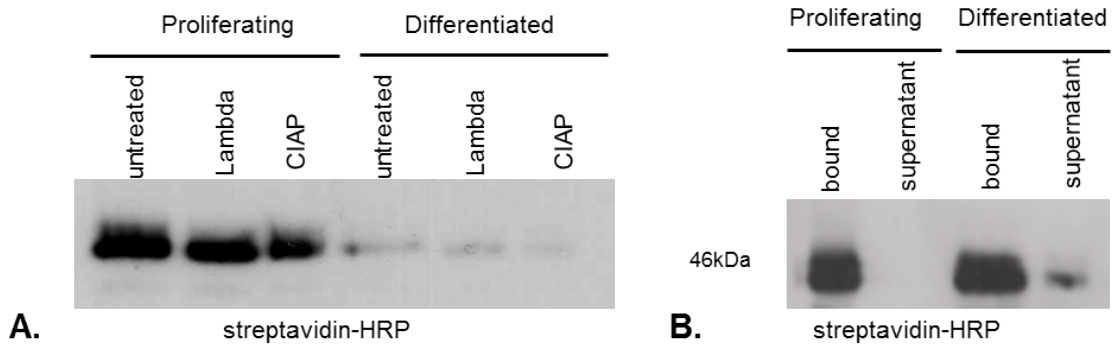


Figure 19: Detecting GATA-1 post-translational modifications (phosphorylation) in whole cell extracts from proliferating and differentiated E12.5 fetal liver cells by Western blots.
A: Total cell extracts from untreated, λ-phosphatase (lambda) and calf intestine phosphatase (CIAP) treated proliferating (left) and differentiated (right) erythroid cells. GATA-1 was detected with streptavidin.
B: Large scale streptavidin pull-downs of whole cell lysates from proliferating (left) and differentiated (right) erythroid cells treated with dephosphorylase, deacetylase, deubiqutinase and desumoylase inhibitors. GATA-1 is detected with streptavidin. Bound and supernatant samples are equivalent to 1/100 of the total bound and supernatant extracts sent for mass spectrometry.

One of the key limitations of ChIP is the need for antibodies that recognize the target protein with high specificity resulting in efficient precipitation of the bound protein from cross-linked chromatin. An ideal control to circumvent these problems and examine the efficiency of the antibody for ChIP experiments would be to perform parallel ChIP experiments with cells that lack the target protein (Carey et al., 2009). This may be possible when working with cell lines however this is not feasible when the experiments include tissues, for example. Previous reports have shown that biotinylation tagging has been used successfully in place of antibodies in ChIP assays, with obvious advantages due το the very high affinities of streptavidin binding allowing high stringencies to be employed (Kim et al., 2008; van Werven & Timmers, 2006; Viens et al., 2004; de Boer et al., 2003). Recent work has led to the further optimization of conditions for streptavidin-ChIP that result in optimal enrichment with lower background and improved ChIP sequencing results (Kolodziej et al., 2009). In order to identify novel GATA-1 gene targets and explore the advantages of biotinylation tagging in ChIP assays we have isolated E12.5 fetal liver cells from bioGATA-1 knock-in females. qPCR experiments of streptavidin-ChIP samples confirmed GATA-1 binding to a set of known GATA-1 gene targets and provided evidence for the suitability of streptavidin-ChIP for massive parallel sequencing applications. Moreover, streptavidin-ChIP resulted in reduced background and higher enrichment than the antibody-ChIP with two different ChIP-applicable GATA-1 antibodies. We now need to proceed with the experimental validation of the novel GATA-1 gene targets identified by deep sequencing analysis of streptavidin-ChIP samples from bioGATA-1 knock-in mice.

Future work will involve the use of the bioGATA-1 knock-in mouse to study the distinct novel GATA-1 gene targets and protein partners in megakaryocytes, eosinophils, mast cells and dendritic cells where GATA-1 also executes essential functions. Overall, this work, when completed, will afford for the first time important insight into the molecular basis of the diverse functions of a single transcription factor in regulating the differentiation of distinct hematopoietic lineages.

3.2 GATA-1 ChIP-Sequencing leads to the identification of GATA-1 novel functions.

GATA-1 is a hematopoietic transcription factor essential for the differentiation of several lineages including the erythroid, megakaryocytic, eosinophilic, mast cells and dendritic cells (Ferreira et al., 2005). Several studies have revealed that GATA-1 interacts with different proteins and binds to different sets of genes in order to orchestrate the distinct lineage specific differentiation programs it regulates (reviewed by Wickrema & Crispino, 2007; Tsiftsoglou et al., 2009; Bresnick et al., 2010; Nakajima, 2011). Chromatin immunoprecipitation (ChIP) coupled to massive parallel sequencing (ChIP-seq) has provided the tools to obtain a genome wide comprehensive view of how transcription factors interact to drive tissue-specific differentiation.

ChIP sequencing studies on GATA-1 using mostly erythroid hematopoietic cell lines have revealed numerous known and novel GATA-1 direct gene targets (reviewed by Kerenyi & Orkin, 2010). Cheng et al. (2009) used a GATA-1 null erythroid cell line (G1E cells) to reveal GATA-1 direct gene targets upon restoration of GATA-1 expression. Based on the pattern of GATA-1 binding, they proposed a model that distinguishes GATA-1 activated versus GATA-1 repressed target genes. In another study, using the same GATA-1 cell line (G1E), it was shown that GATA-1 gene targets, the expression of which is altered upon GATA-1 binding, display a greater number of GATA-binding motifs and more frequent palindromic GATA sites which are bound by GATA-1 closer to their transcription start site compared to GATA-1 gene targets with unaltered expression profile (Yu et al., 2009). GATA-1 ChIP-Sequencing experiments in human K562 erythroleukemia cells, mouse G1E cells and validation experiments in primary cells, defined the occupancy of GATA-1 and GATA-2 in erythroid cells (Fujiwara, 2009). The GATA switch, which involves a chromatin occupancy exchange between GATA-1 and GATA-2, was the subject of research in another study, where the megakaryocytic GATA-1 null cell line (G1ME) was used (Doré, 2012).

Our long-term aim is to unravel the distinct gene targets of GATA-1 in the different hematopoietic lineages where GATA-1 is active. For this purpose, we initially focused on the genome wide analysis of GATA-1 binding in erythroid cells. A number of novel GATA-1 erythroid gene targets have been revealed recently, however all recent findings derive from studies using immortalized erythroid cell lines. We isolated erythroid cells from mouse fetal liver cells in order to identify novel *in vivo* GATA-1 targets in the erythroid lineage and also to use this study as a reference for the physiological targets of GATA-1 in primary hematopoietic cells. Our future plans include the identification of novel targets in other hematopoietic lineages, such as the eosinophils, mast cells, dendritic cells and the comparison of these datasets with the list of genes identified in the erythroid lineage. The aforementioned lineages though, differ in the amount of GATA-1 expressed compared to the erythroid cells, where GATA-1 is abundantly expressed. In order to address this issue of low GATA-1 expression, we are planning to use the biotin GATA-1 knock-in mouse we have generated and take advantage of the high affinity reagents available for the biotin tags which allow high enrichments in ChIP experiments (section 3.1). To these ends, our GATA-1 ChIP-Sequencing data will have established methods for massive parallel sequencing and bioinformatic analysis and will also provide a reference framework when it comes to testing the biotin-streptavidin system for genomics applications in the GATA-1 biotin tag knock-in mice or ES cells.

In this study, we used the cell surface Ter119 receptor which is a marker for mature erythroid cells to separate mouse fetal liver cells into early proerythroblasts ($Ter119^-$ fraction) and differentiated erythroid cells ($Ter119^+$ fraction) in order to identify potential GATA-1 binding sites during differentiation of erythroid cells. The *in silico* analysis of GATA-1 deep sequencing and microarray experiments of the two populations, led to the identification of a novel role for GATA-1 in the Myc/Max/Mad network and in oxidative stress cell response.

I. GATA-1 deep sequencing experiments and *in silico* analysis

E12.5 fetal liver cells were isolated from C57/BL6 wild type mice and expanded for three days under erythropoietic stress conditions induced by dexamethasone as previously described (von Lindern et al. 2001; Schuh et al. 2005) and formaldehyde cross-linked chromatin was prepared from 10^7 cells per sample. Using the cell surface Ter119 receptor as a marker for mature erythroid cells we separated fetal liver cells into proerythroblasts ($Ter119^-$ fraction) and differentiated erythroid cells ($Ter119^+$ fraction). Sonicated chromatin from both $Ter119^-$ and $Ter119^+$ populations was immunoprecipitated with the Abcam 11852 αGATA-1 antibody, which was shown in pilot experiments as being the most efficient in ChIP assays out of five commercial antibodies that were tested (Figure 20A). In order to assess the efficiency of ChIP and also the separation of fetal liver cells into $Ter119^-$ and $Ter119^+$ cells, we carried out qPCR using primers for the major GATA-1 enhancer HS1 and another GATA-1 enhancer located 122kb downstream of the GATA-1 transcription start site (TSS), which are both known to bind GATA-1 *in vivo* (Valverde-Garduno et al., 2004; Im et al., 2005). Hematopoietic exon 1 erythroid (IE exon1) and 5'HS2 were used as negative controls (Figure 20B).

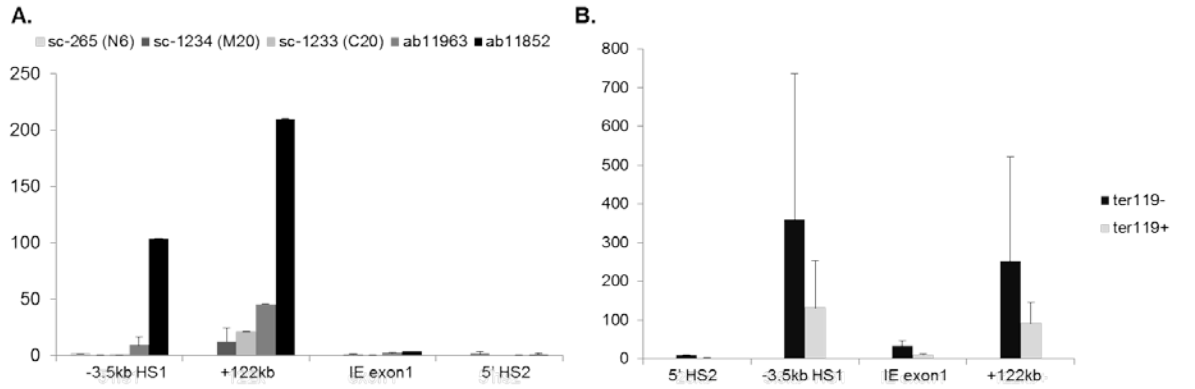


Figure 20: Relative enrichment of GATA-1 in four regions, using five commercially available αGATA-1 antibodies in MEL cells (A) and testing early prerythroblasts and differentiated erythroid cells with the Abcam 11852 αGATA-1 antibody (B).
A: Cross-linked chromatin was prepared from MEL cells and was immunoprecipitated with three C-terminal αGATA-1 antibodies (ab11852, sc C-20, scM-20) and two N-terminal antibodies (ab11963 and sc N6).
B: Cross-linked chromatin was prepared from Ter119 depleted E12.5 fetal liver cells and was immunoprecipitated with αGATA-1 antibody Ab11852 (Abcam).
Two known GATA-1 target sequences (HS1 and m122) and two sequences that are not bound by GATA-1 (IE exon 1 and 5'HS2) were used as positive and negative controls, respectively, in order to assay for relative enrichment for GATA-1 binding. Rat and rabbit IgGs were included as negative controls. Samples were analyzed in duplicate relative to a sequence in the *gapdh* locus, using an ABI Prism 7000 sequence detection system.

Duplicates of the eluted material, as well as input control samples, were sent for massive parallel sequencing using the Illumina platform, in order to identify *in vivo* bound GATA-1 gene targets. A total of 18.2 million $Ter119^-$ and 15.3 million $Ter119^+$ uniquely mapped sequence reads were generated (Figure 21A). All 51-nucleotide sequence reads produced by the Illumina Genome Analyzer II were mapped to the NCBI37/mm9 Mouse Genome Assembly using the Eland software (Illumina). Sequence reads with multiple genome alignments and/or more than 2 nucleotides mismatches were excluded from subsequent analysis. Using the QuEST peak-calling algorithm (Valouev et al., 2008) we assembled the unique, non-redundant sequence reads for each replicate into peaks that identify potential GATA-1 bound regions across the genome. For both samples, we took the union of the peaks of the two replicates, resulting in 9,795 peaks and 14,239 peaks for the $Ter119^-$ and $Ter119^+$ samples, respectively (Figure 21A). This approach is in accordance with previous observations that peaks appearing in only one replicate may reflect low occupancy binding sites (e.g. repressed genes) rather than false positives (Cheng et al., 2009). Visualization in both the $Ter119^-$ and $Ter119^+$ datasets of peaks in known GATA-1 target gene loci, such as the *β-globin*, *gata-1*, *gata-2*, *klf1* or *scl/tal-1* loci (Valverde-Garduno et al., 2004; Martowicz et al., 2005; Rodriguez et al., 2005; Kassouf et al., 2010), or the recently identified GATA-1 gene target *zbtb7* (Yu et al., 2009), provided early validation for our sequencing data (Figure 21B).

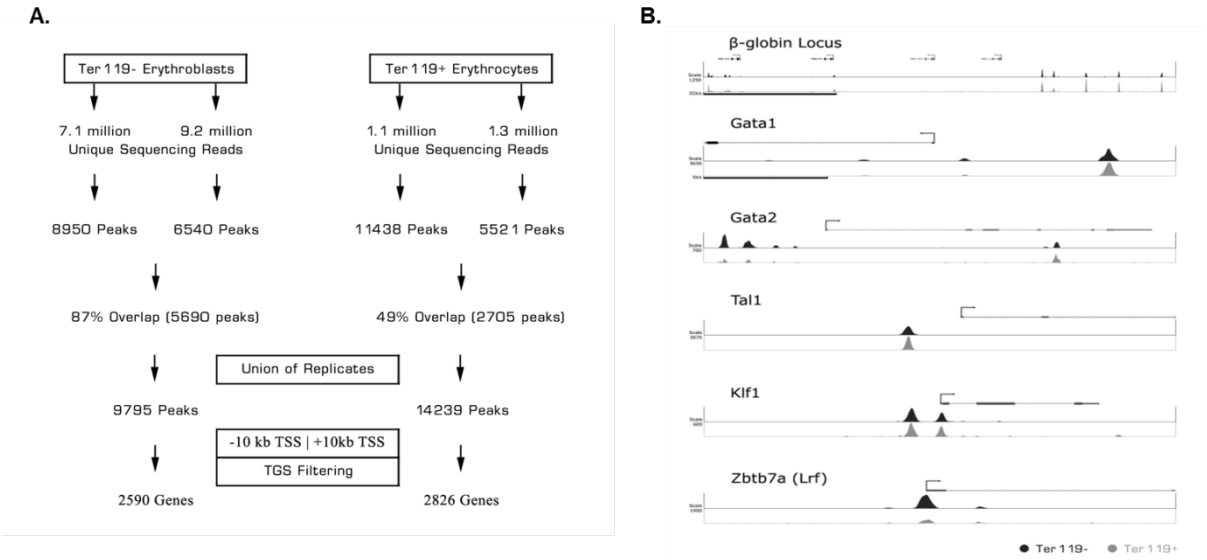


Figure 3: Distribution of GATA-1 chromatin occupancy in Ter119$^-$ and Ter119$^+$ cells.
A: Distribution of the distances of GATA-1 peaks from annotated gene TSSs.
B: Location analysis of GATA-1 peaks. Percentage of intragenic and intergenic peaks (left). Distribution of intragenic GATA-1 peaks in 5'UTR, 3'UTR, introns and exons (middle). Distribution of intergenic GATA-1 peaks in 0-5kb, 5-25kb, 25-50kb and >50kb windows (right).

We next plotted the distances of all identified peaks from annotated gene Transcription Start Sites (TSSs) and found that for both Ter119$^-$ and Ter119$^+$ datasets, the majority of the peaks clustered proximally (within 5kb) to gene TSSs (Figure 22A). Furthermore, both Ter119$^-$ and Ter119$^+$ samples, displayed a similar distribution of approximately 64% and 36% of peaks falling within intergenic and intragenic regions, respectively (Figure 22B). Of the intragenic peaks, 59% fall within introns and 35% within exons, with clear clustering of peaks towards the 5'-most introns and exons (Figure 22B, right panel and data not shown). By contrast only 4% of GATA-1 peaks fall within 5'UTRs and 2% within 3' UTRs *in vivo* (Figure 22B, middle panel). Overall, we do not observe any differences in intergenic or intragenic GATA-1 peak distribution between Ter119$^-$ proerythroblasts and Ter119$^+$ erythroid cells.

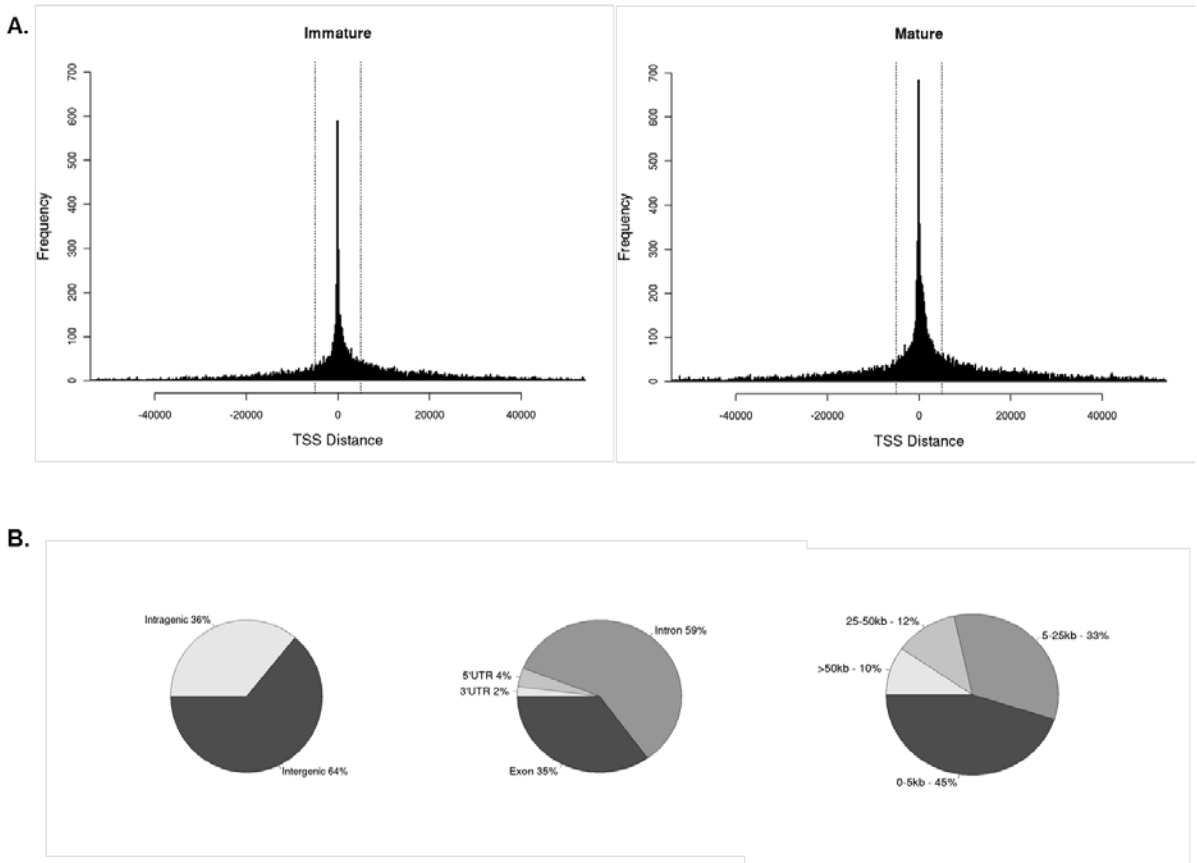


Figure 22: Distribution of GATA-1 chromatin occupancy in Ter119$^-$ and Ter119$^+$ cells.
A: Distribution of the distances of GATA-1 peaks from annotated gene TSSs.
B: Location analysis of GATA-1 peaks. Percentage of intragenic and intergenic peaks (left). Distribution of intragenic GATA-1 peaks in 5'UTR, 3'UTR, introns and exons (middle). Distribution of intergenic GATA-1 peaks in 0-5kb, 5-25kb, 25-50kb and >50kb windows (right).

The mapping of GATA-1 peaks within±10kb of genes' TSSs resulted in the identification of a total of 2,590 and 2,826 potential GATA-1 target genes in the Ter119$^-$ and Ter119$^+$ datasets, respectively. The union of the two datasets yielded a total of 3,651 potential GATA-1 target genes, of which 1,765 genes were common to both Ter119$^-$ and Ter119$^+$ datasets thus giving an intersection of 48.3%. By contrast, 825 (22.6%) and 1,061 (29.1%) genes were unique to the Ter119$^-$ and Ter119$^+$ cells, respectively (Figure 23A). These data reveal a considerable conservation of GATA-1 target genes throughout erythroid differentiation.

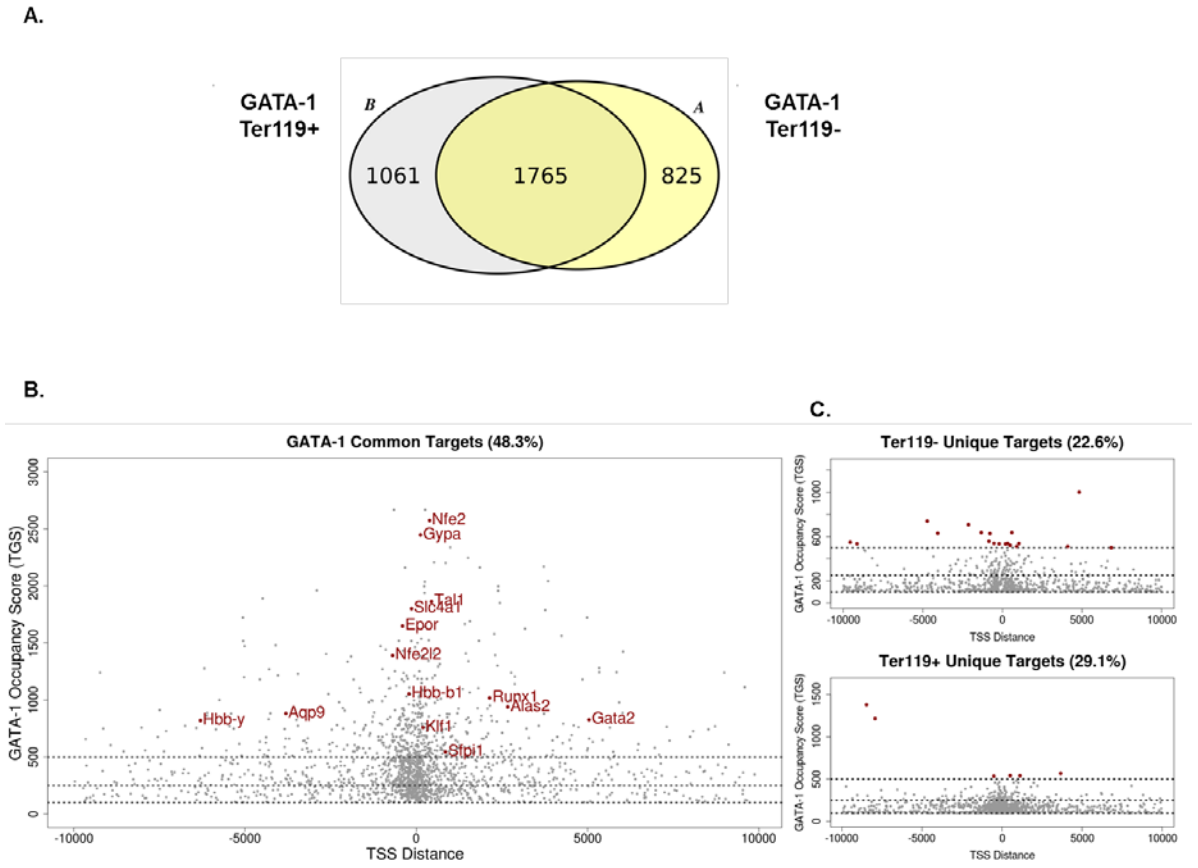


Figure 23: Location Analysis and Score Distribution of GATA-1 occupancy sites in Ter119$^-$ and Ter119$^+$ cells.

A: Venn diagram showing the common and unique GATA-1 target genes identified in Ter119$^-$ and Ter119$^+$ erythroid cells.

B: Distribution of GATA-1 TGS and minimum distance from the target gene TSS. Scatterplot of GATA-1 potential target genes identified in both cell populations (each gene is plotted once selecting for the condition with the highest TGS). Horizontal lines define score thresholds for the 3 classes. Most highly enriched (Class I) target genes are found in the intersection of the two datasets and comprise most of the bibliographically described GATA-1 target genes.

C: GATA-1 potential target genes identified uniquely in Ter119$^-$ or Ter119$^+$ cells are shown in the top and bottom scatterplots, respectively. Despite the fact that target genes unique in either Ter119$^-$ or Ter119$^+$ cells are poor in Class I genes they still show a prominent clustering of GATA-1 occupancy sites near the identified target gene TSS.

In order to further facilitate the differential analysis of potential GATA-1 gene targets in Ter119$^-$ and Ter119$^+$ cells, we classified all genes into three categories on the basis of their total gene score (TGS), which was defined as the sum of the GATA-1 peak scores assigned to each gene. Class I includes genes with a TGS greater than 500, Class II includes genes with a TGS of 500 to 250 and Class III includes genes with a TGS of 250 to 100 (summarized in Table 3).

Not surprisingly, Class I includes most of the well-established GATA-1 target genes implicated in erythroid differentiation and/or function. Genes like the *gata1* locus itself, *gata2*, the *β-globin* locus (especially the Locus Control Region), *epoR*, *nfe2*, *slc4a1*, *gypa*, *tal1*, *lrf*, *klf1*, *nrf2*, *runx1*, *alas2* all have a TGS score greater than 500 (Figure 23B). Furthermore, the high specificity of Class I target genes is further supported by their marked enrichment in erythroid cell related ontologies (data not shown). Thus, Class I genes, corresponding to approximately 15% of all identified genes (Figure 23B, Table 3), include targets that are highly enriched for GATA-1 binding and most likely represent the erythroid transcription program.

Class II and Class III genes include most of the GATA-1 targets that are unique to the Ter119$^-$ or Ter119$^+$ cells (806/825 and 1,055/1,061 genes, respectively; Table 3). The differences in the numbers between these two datasets suggest that low binding GATA-1 targets are being lost from Ter119$^-$ cells and new ones are being acquired in Ter119$^+$ cells with erythroid differentiation (Figure 23C, Table 3).

	Ter119-	Ter119+	Common	Ter119- Unique	Ter119+ Unique
Class I	436	353	199	19	6
Class II	720	715	222	123	143
Class III	1434	1758	511	683	912
Ascending			353		
Descending			480		
Total	**2590**	**2826**	**1765**	**825**	**1061**

Table 3: Number of genes in TGS classes.

Next, we sought to experimentally validate the *in silico* data analysis by checking GATA-1 binding to well-established GATA-1 gene targets, such as *gata2*, by ChIP followed by qPCR (Figure 24). We found more enriched GATA-1 binding to the -3.9kb site, compared to the -2.8kb and the +9.1kb sites identified by QuEST, in both Ter119$^-$ and Ter119$^+$ cells. In addition, we confirmed more enriched GATA-1 binding to all sites in Ter119$^-$ proerythroblasts versus Ter119$^+$ mature erythroid cells, in very good agreement with the peaks identified in the ChIP-Seq profile.

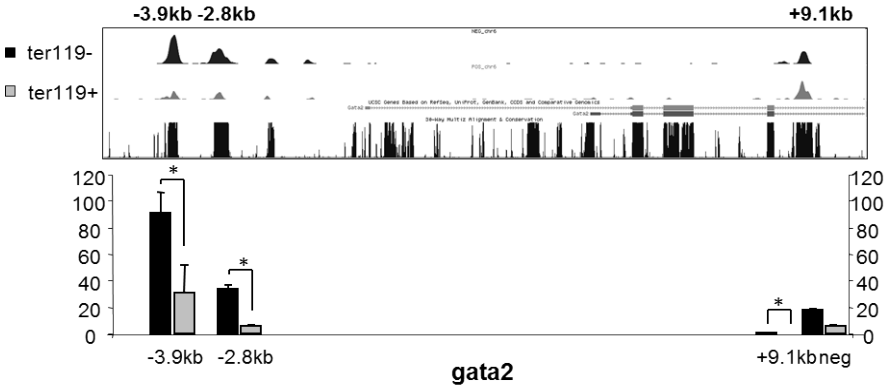


Figure 24: Profile of GATA-1 binding to the *gata2* gene.
Top panel: ChIP-Seq profile in Ter119$^-$ (in black) and Ter119$^+$ (in grey) populations from fetal liver cells, RefSeq annotation of the gene (boxes indicate exons; thin lines introns) and conservation plot.
Bottom panel: Real-time PCR validation of anti-GATA1 (11852, Abcam) ChIP on selected sites of *gata2,* an established GATA-1 bound locus. The y-axis represents enrichment over input DNA, after normalization with the *gapdh* gene. The enrichment is the average of triplicate experiments normalized to the IgG enrichment, compared to negative sites. Asterisk denotes significant difference in the values ($p \leq 0.05$).

II. GATA-1 novel role in the Myc/Mad/Max network

We then incorporated in our meta-analysis, the expression profiles of Ter119$^-$ and Ter119$^+$ cells obtained from microarray experiments, kindly donated by M. Kassouf (Oxford) (unpublished data). This comparison resulted in the identification of subsets of GATA-1 target genes that vary greatly in their expression pattern from one population to the other, in that they are either significantly upregulated or repressed. We focused on two such subsets of GATA-1 genes: the *heme oxygenase 1* (*ho-1*) gene which is upregulated 26-fold with differentiation and the *myc* gene which is repressed by 7-fold. Enrichment of GATA-1 to both loci was validated by Real-Time PCR experiments (Figures 25, 27).

The data analysis indicated GATA-1 binding not only to the *myc* gene (Rylski et al., 2003), but also to the *max*, *mad1* and *mad4* loci in the two cell populations. The Myc/Max/Mad network comprises of the Myc family of proto-oncoproteins (c-, N- and L-Myc), Max and the Mad family of proteins (Mad1, Max, Mad3, Mxd4 and Mnt/Rox) and regulates basic cellular functions such as proliferation, differentiation and apoptosis. Myc and Mad proteins compete for heterodimerization with Max in order to regulate proliferation and differentiation, respectively. The balance between Myc and Mad family proteins is thus critical for cell proliferation (reviewed by Lüscher, 2001; Grinberg et al., 2004; Lüscher & Vervoorts, 2012).

Rylski et al. (2003) used a complementation assay based on the synchronous inducible rescue of GATA-1$^-$ G1E erythroblasts to investigate how GATA-1 expression regulates the cell cycle in differentiating erythroid precursors. This work revealed that GATA-1 not only drives differentiation but also results in rapid synchronous cell cycle arrest and inactivation of cyclin-dependent kinases (Cdks). This G1 cell cycle arrest observed is fully dependent on the GATA-1-mediated repression of *c-myc*, by direct binding of GATA-1 to the promoter of the *myc* proto-oncogene.

We validated enrichment of GATA-1 on *myc*, *mad1*, *max* and *mxd4* by qPCR experiments and noticed that there was significantly different binding of GATA-1 to the selected sites of these loci, in Ter119$^-$ versus Ter119$^+$ erythroid cells. GATA-1 is much more enriched in Ter119$^-$ cells compared to Ter119$^+$ cells. Moreover, it is clear that the pattern of binding predicted from the *in silico* analysis correlates fully with the ChIP experimental data (Figure 25).

GATA-1 binding sites were identified in additional components of the Myc/Max/Mad network, such as the *mad3* and the *mnt/rox* loci (data not shown). Furthermore, we noticed that GATA-1 is more enriched in Ter119$^-$ cells compared to Ter119$^+$ cells in these loci, similarly to what we observed for the rest of the loci involved in the network. Expression analysis indicated that not only *myc* is repressed during differentiation but also, *mxd4*, *max* and *mnt/rox*. By contrast, Mad1, which competes with Myc for heterodimerization with Max in order to promote differentiation versus proliferation, is 4-fold upregulated. Thus, it seems that GATA-1 binds to members of the extended Myc/Max/Mad network at an early differentiation stage, ensuring reduced expression of all but Mad1, which is activated, and then as differentiation progresses either there is reduced requirement for GATA-1 to remain bound to the genes in order to exert its function or there is no great demand for further repression/activation and thus no requirement for high levels of GATA-1 in these loci. Overall, the broad binding of GATA-1 to the Max and Mad family proteins and their subsequent activation/repression during differentiation, make us hypothesize a novel aspect in GATA-1 functions and a broader role for GATA-1 as a direct upstream regulator of the Myc/Max/Mad network in erythroid cell differentiation.

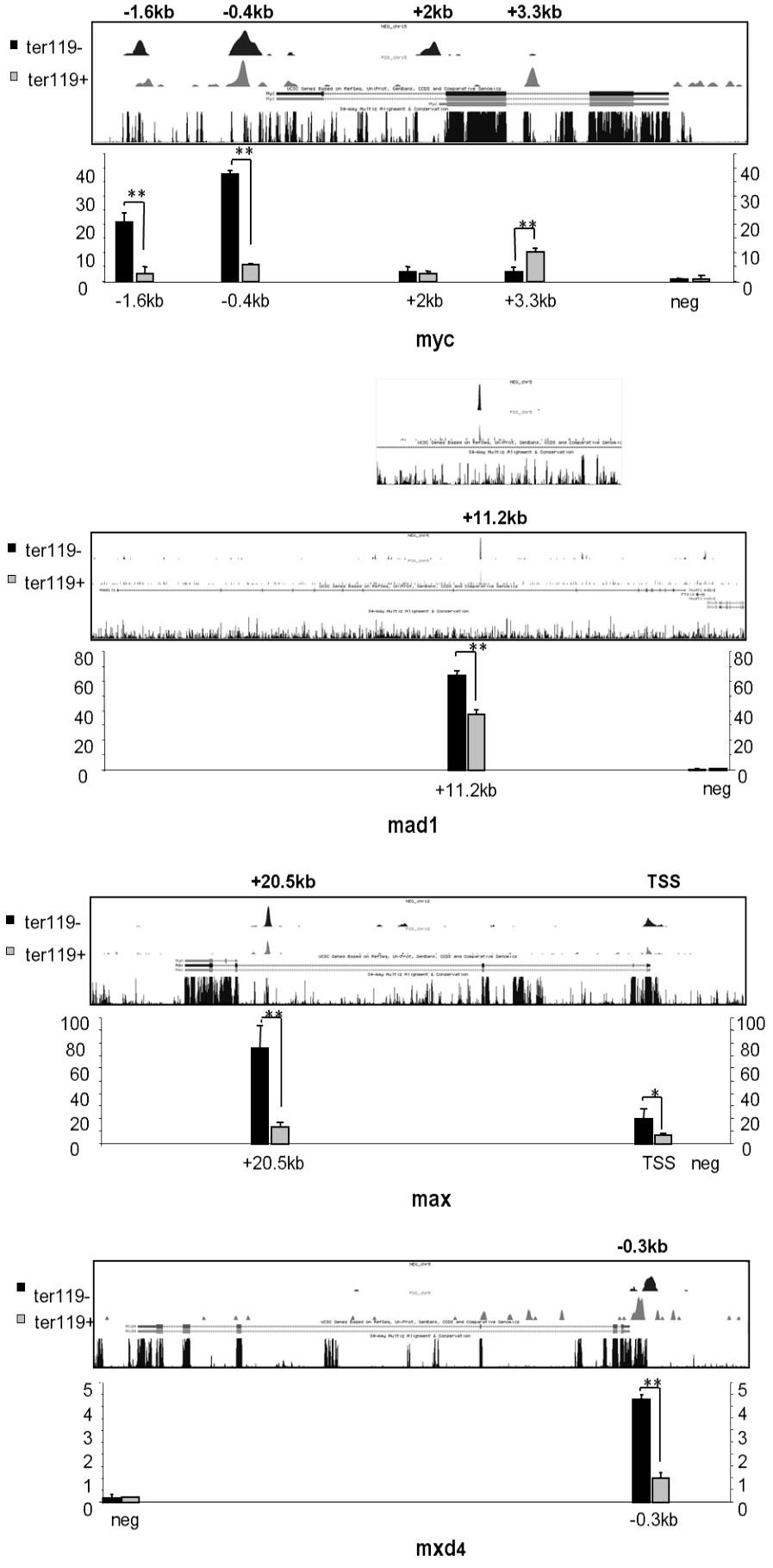


Figure 25: Profile of GATA-1 binding on the *myc/mad/max/mxd4* genes.

Top panels: ChIP-Seq profile in Ter119$^-$ (in black) and Ter119$^+$ (in grey) populations from mouse fetal liver cells, RefSeq annotation of the gene (boxes indicate exons; thin lines introns) and conservation plot.
Bottom panels: Real-time PCR analysis of anti-GATA1 (11852, Abcam) ChIP on selected sites of *myc*, *mad*, *max*, *maxd4* genes (from top to bottom). The y-axis represents enrichment over input DNA, after normalization with the *gapdh* gene. The enrichment is the average of triplicate experiments normalized to the IgG enrichment, compared to negative sites. Asterisks denote significant difference in the values (* p≤0.05, ** p≤0.01).

III. GATA-1 novel role in the anti-oxidant cell response

We also focused on the second subset of GATA-1 genes, which includes *ho-1* and the oxidative stress response pathway. Upregulated 26-fold in Ter119$^+$ cells, *ho-1* is the most highly differentially expressed gene in the microarray list of genes. Furthermore, we also observed a correlation between highly different expression levels and GATA-1 occupancy of the *nrf2* (*nfe2l2*) and *bach1* genes, which are upstream transcriptional regulators of the *ho-1* gene (Figure 26, Sun et al., 2002).

HO-1 is the inducible form of HO, which binds heme at a 1:1 ratio and catalyzes its degradation into biliverdin with the release of free iron and CO (Otterbein et al., 2003). HO-1 is expressed at low basal levels in most cells and tissues and is upregulated by multiple oxidative stress stimuli. Due to its rapid induction, HO-1 is considered to be an adaptive cellular mechanism against the cytotoxic effects of oxidative stress (reviewed by Ryter et al., 2006; Paine et al., 2010). HO-1 promotes cytoprotection by degrading free heme, which would otherwise sensitize cells to undergo programmed cell death (Gozzelino et al., 2010). Heme consists of a tetrapyrrole ring with a central iron ion and is an abundant molecule in almost all living organisms, displaying two functions: when bound to hemoglobin, myoglobin and cytochromes plays a physiological role for the transport of oxygen and mitochondrial electrons and when free can cause toxic effects due to the oxidative stress induction it triggers (reviewed by Mense & Zhang, 2006). Compared to other cells in the body, the erythroid cells exhibit a much more rapid rate of heme synthesis so as to carry adequate levels of heme for active initiation of globin chain synthesis in the ribosomes (Ponka, 1997; reviewed by Cianetti et al., 2010). Cytoprotection of erythroid cells against the toxic effects of heme throughout erythropoiesis is therefore of particular importance and a tight regulation of oxidative stress response during the erythroid differentiation must exist.

HO-1 induction is triggered by a wide variety of stimuli such as sodium arsenite, hydrogen peroxide (H_2O_2), UVA, heme and others. When no stimulus applies, under physiological conditions, ho-1 enhancers E1 and E2 are bound by BACH1 in complex with small MAF proteins, which act as repressors and therefore ho-1 is expressed at low levels. Under the same conditions, NRF2, an activator of the ho-1 gene, is in the cytoplasm, bound and constantly ubiquitinated by the Cul3–KEAP1 ubiquitin E3 ligase complex (Kobayashi et al., 2004). Upon oxidative stresses, KEAP1 becomes modified and NRF2 is stabilized and enters the nucleus, where increased levels of heme displace BACH1 from the enhancers of *ho-1*. It is the turn of NRF2 to now interact with small MAF proteins, access the *ho-1* enhancers and induce expression of the gene. (Sun et al., 2002; Reichard et al., 2007; reviewed by Paine et al., 2010).

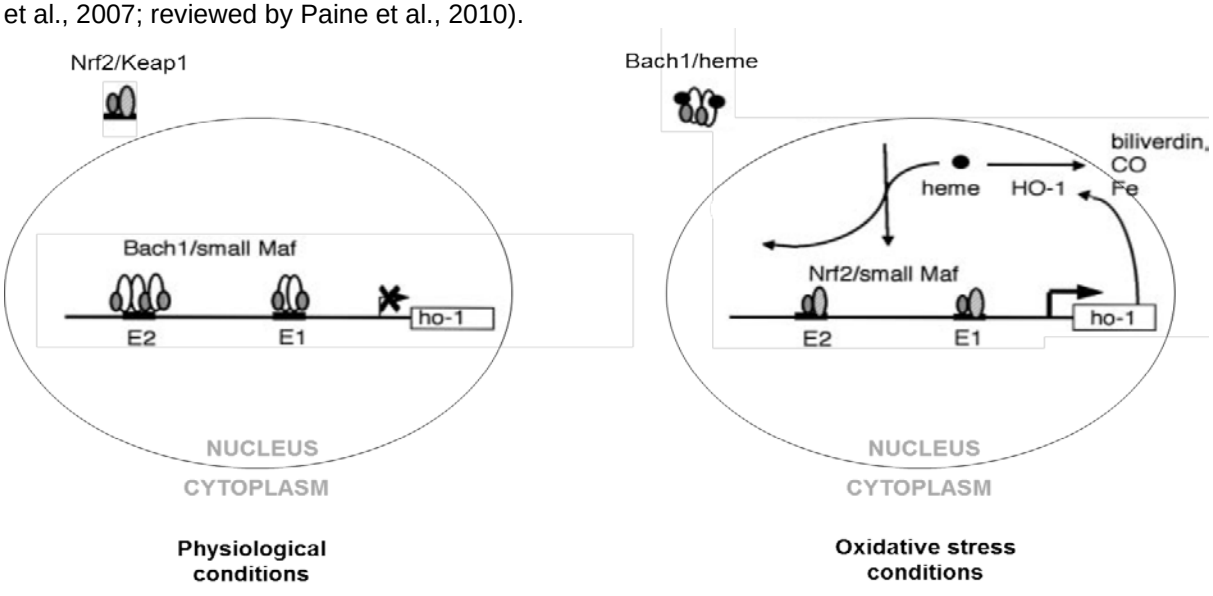


Figure 26: Binding of BACH1, NRF2 and small MAFs on *ho-1* enhancers.

A hypothetical model which describes the regulation of *ho-1* by BACH1, NRF2 and heme. Besides MafK, other Maf-related factors may also serve as partners for BACH1. BACH1 makes enhancers inaccessible to NRF2 and other activators by binding to them. Repression by BACH1 is alleviated upon increase of heme levels, allowing expression of *ho-1* followed by degradation of heme. Figure adapted from Sun et al., 2002.

The data analysis indicated multiple GATA-1 binding sites on the *ho-1* gene, three on *bach1* and a single highly enriched site on *nrf2* (*nfe2l2*) in both cell populations. We validated binding of GATA-1 on all three genes by qPCR experiments and noticed that there was significantly different binding of GATA-1 in many of the selected sites of these loci, in Ter119^{-} versus Ter119^{+} erythroid cells. The pattern of binding on *ho-1* is rather different to the pattern observed on *bach1* and *nrf2*, being more spread across the gene and exhibiting lower enrichment scores. We therefore assume that the enrichment in the binding of GATA-1 to the *ho-1* gene is not high enough to account for the 26-fold induction of *ho-1* expression during erythroid differentiation and GATA-1 may thus exert its influence on *ho-1* induction by regulating the expression of the upstream *ho-1* regulators, bach1 and nrf2 (Figure 27).

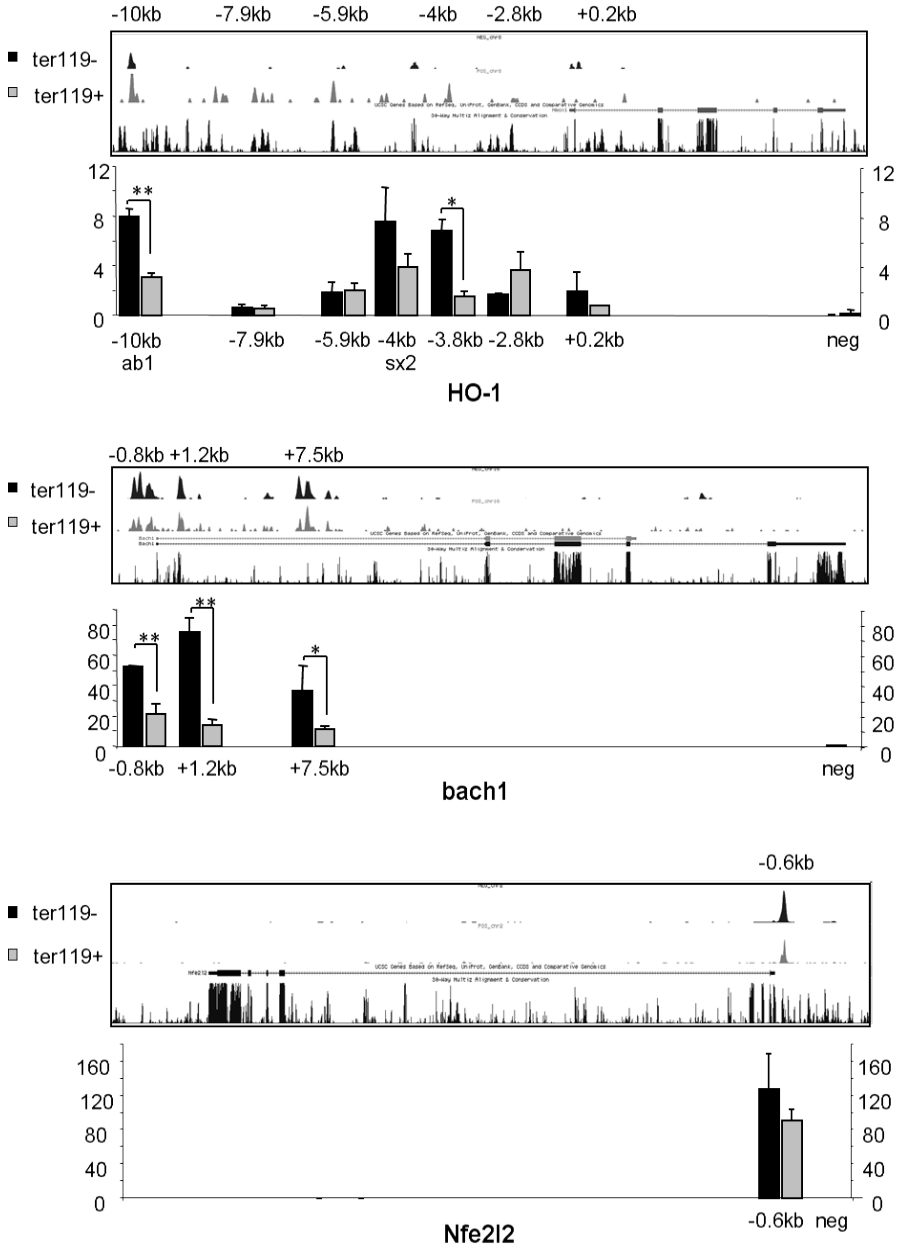


Figure 27: Profile of GATA-1 binding on the *ho-1*, *bach1* and *nfe2l2* genes.

Top panels: ChIP-Seq profile in Ter119$^-$ (in black) and Ter119$^+$ (in grey) populations from fetal liver cells, RefSeq annotation of the gene (boxes indicate exons; thin lines introns) and conservation plot.
Bottom panels: Real-time PCR analysis of anti-GATA1 (11852, Abcam) ChIP on selected sites of *ho-1*, *bach1* and *nfe2l2* genes (from top to bottom). The y-axis represents enrichment over input DNA, after normalization with the *gapdh* gene. The enrichment is the average of triplicate experiments normalized to the IgG enrichment, compared to negative sites. Asterisks denote significant difference in the values (* p≤0.05, ** p≤0.01).

In order to explore the potential role of GATA-1 in the HO-1 pathway during erythropoiesis, we used the erythropoietic cell line G1E-ER. G1E cells (GATA-1$^-$ erythroid) are cells derived from *in vitro* differentiated GATA-1 null murine embryonic stem (ES) cells that proliferate as immature erythroblasts and terminally differentiate upon restoration of GATA-1 expression (Weiss et al., 1997). In G1E-ER cells, GATA-1 is fused to the estrogen-receptor and localized in the cytoplasm; upon addition of β-estradiol or tamoxifen, GATA-1 enters the nucleus and restores terminal differentiation (Rylski et al., 2003). The efficiency of induction in our hands was assessed by the detection of GATA-1 in the nucleus in G1E-ER cells treated with 1μM of estradiol and the subsequent GATA-1 mediated repression of GATA-2. Terminal maturation was validated by cytospins of untreated and estradiol-treated G1E-ER cells, stained with May-Grünwald and benzidine (Figure 28).

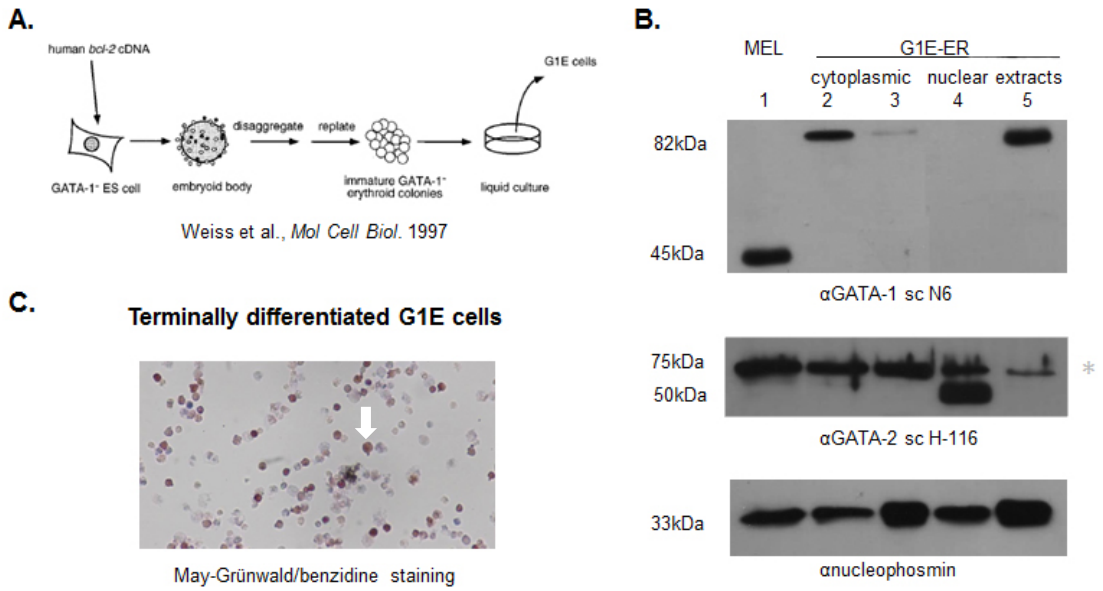


Figure 28: Isolation of G1E cells, a GATA-1$^-$ erythroid cell line (A) and GATA-1 induced differentiation of G1E-ER cells by western blots (B) and cytospins (C).
A: Isolation of G1E cells, a murine GATA-1 null erythroid cell line. Human *bcl-2* cDNA in an erythroid expression vector was introduced into GATA-1 null ES cells so as to protect them from apoptosis upon *in vitro* differentiation towards the erythroid lineage. Stable lines containing the transgene were used to derive pure erythroid cell colonies in methylcellulose cultures by a two-step in vitro differentiation method. Individual erythroid cell colonies were isolated and expanded in liquid medium. Image taken from Weiss et al., 1997.
B: Nuclear (4, 5) and cytoplasmic (2, 3) extracts from untreated G1E-ER cells (2, 4), and β-estradiol treated G1E-ER cells (3, 5). Nuclear extracts from MEL C88 cells were used as control (1). Extracts were immunoblotted with anti GATA-1antibody (top), anti GATA-2 antibody (middle) and anti nucleophosmin antibody as protein loading control (bottom). GATA-1 is detected in the nuclear extracts of G1E-ER treated cells only and GATA-2 in the nuclear extracts of untreated G1E-ER cells. Asterisk denotes a non-specific cross-reacting band (middle panel).
C: May Grünwald/Giemsa and benzidine-stained cytospins of β-estradiol treated G1E-ER cells; arrow indicates a cell that is fully differentiated and therefore positive for benzidine which stains hemoglobin. Cytospins were imaged on a BX60 Olympus microscope with a Q Imaging camera.

Cross-linked chromatin was prepared from 10^7 untreated and β-estradiol treated G1E-ER cells, was immunoprecipitated with the Abcam 11852 αGATA-1 antibody and the eluted DNA was used for qPCR experiments. The same sets of primers for *bach1* and *nfe2l2* -used in Figure 27- were used in

order to assay for GATA-1 binding in G1E-ER cells and compare the enrichment levels and the pattern of binding between the primary cells and the cell line. We found that both the enrichment levels and the pattern of binding, for example greater enrichment on the +1.2kb region on *bach1* compared to the rest of sites, correlated very well between the fetal liver cells and the G1E-ER cells (Figures 27, 29). These data are in agreement with our hypothesis that GATA-1 directly controls the two upstream effectors of HO-1 in the oxidative stress during erythropoiesis.

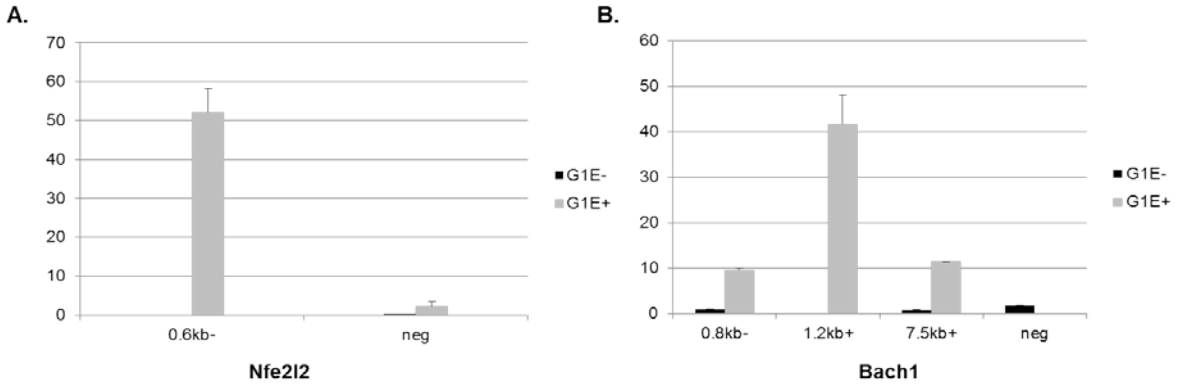


Figure 29: Profile of GATA-1 binding on the *nfe2l2* and *bach1* genes in G1E-ER cells.
qPCR analysis of anti-GATA1 ChIP on selected sites of *nrf2* (A) and *bach1* genes (B). The y-axis represents enrichment over input DNA, after normalization with the *gapdh* gene. The enrichment is the average of three replicate experiments normalized to the IgG enrichment, compared to negative sites.

We then measured the expression levels of bach1 and nrf2 upon GATA-1 expression in order to investigate the role of GATA-1 binding on *bach1* and *nrf2* loci. RNA was isolated from untreated and β-estradiol treated G1E-ER cells followed by cDNA synthesis and qPCR analysis using primers specific for *bach1* and *nrf2* cDNAs. It appears that the nrf2 expression was reduced approximately 2-fold, whereas bach1 expression was reduced 13-fold upon GATA-1 expression in G1E-ER cells (Figure 30). These data indicate that GATA-1 directly controls the expression of *bach1* and to a lesser extent that of *nrf2*; which is in accordance to the current notion that it is alleviation of repression, rather than activation of activators like Nrf2, that is the critical step for ho-1 induction, given that ablation of bach1 results in constitutive expression of ho-1 (Sun et al., 2002).

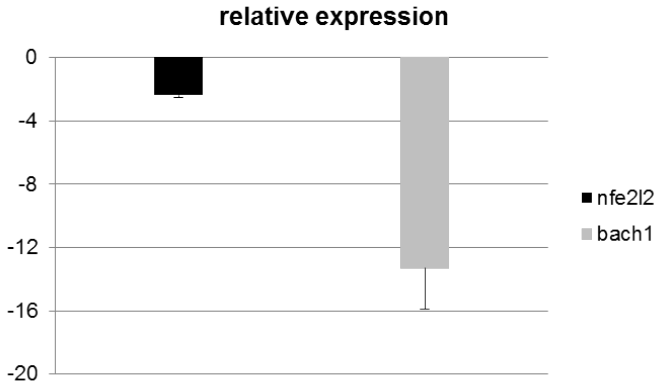


Figure 30: Expression levels of *bach1* and *nfe2l2* genes in G1E-ER cells.
qPCR analysis of cDNA from *nfe2l2* (left) and *bach1* genes (right) in β-estradiol treated G1E-ER cells. The y-axis represents expression levels after normalization with the *gapdh* gene. The expression is the average of two replicate experiments.

In order to test our hypothesis of a novel role for GATA-1 in the regulation of HO-1 expression and the oxidative stress response during erythropoiesis, we measured the levels of reactive oxygen species (ROS) and NRF2 protein in G1E-ER cells exposed to oxidative stress conditions. Oxidative stress was induced by the addition of either H_2O_2 or hemin to G1E-ER cells with and without activation of

GATA-1 by β-estradiol. We assessed induction of oxidative stress and potential cytoprotection by GATA-1 by measuring ROS 24hrs post β-estradiol induction of GATA-1, which is when G1E cells start hemoglobinizing and erythroid genes are highly expressed. The oxidative stress inducers were added 18hrs (H_2O_2) and 20hrs (hemin) post GATA-1 activation so that the collection time point (24hrs post estradiol induction) coincides with the peak of oxidative stress, which is 6hrs for H_2O_2 and 4hrs for hemin (Figure 31).

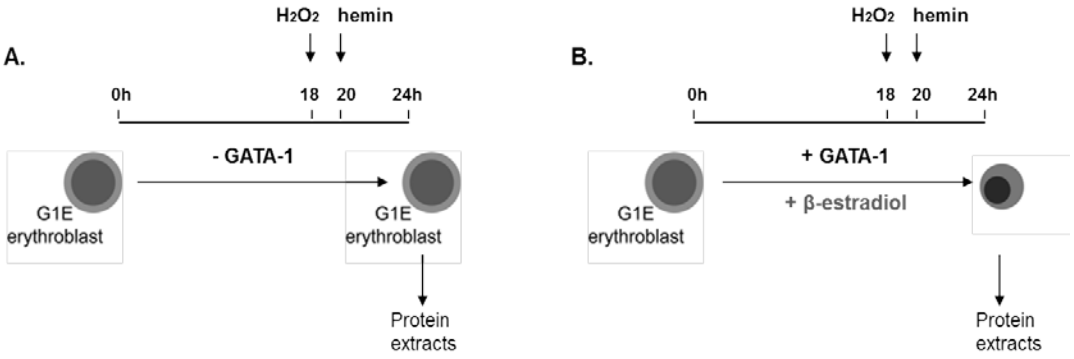


Figure 31: Oxidative stress induction in untreated (A) and β-estradiol treated (B) G1E-ER cells.
GATA-1 null (A) and GATA-1 expressing (B) G1E-ER cells were exposed to oxidative stress conditions by adding H_2O_2 for 6hrs or hemin for 4hrs. Cells were collected 24hrs post β-estradiol addition.

ROS are chemically reactive molecules containing oxygen and are produced as a normal product of cellular metabolism (Ryter & Tyrrell, 2000). Under oxidative stress conditions ROS accumulate and if our hypothesis of a potential GATA-1 role in regulating oxidative stress response stands true, we would expect that G1E cells that express GATA-1 accumulate less ROS than G1E GATA-1 null cells (Figure 32). In addition, we measured the protein levels of NRF2 in the nucleus of both cells, assuming that GATA-1 expression would result in elevated levels of NRF2 in the nucleus of G1E cells that express GATA-1, compared to GATA-1 null cells (Figure 33).

It seems that in both cases where oxidative stress was induced, upon addition of hemin or H_2O_2, GATA-1 expressing G1E cells generated less ROS compared to GATA-1 null G1E cells. The FACS analysis revealed that upon activation of GATA-1, there was a clear shift of the ROS levels towards lower concentrations in the histogram plots in either case of oxidative stress. This shift was obvious even under over-induction of oxidative stress conditions, in which case G1E cells that did not express GATA-1 were fully stained with the ROS detection marker (Figure 32).

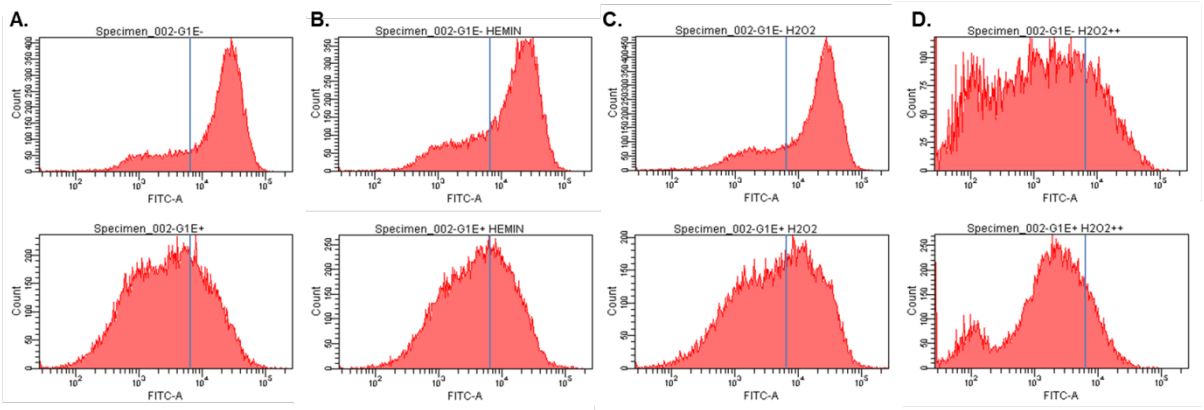


Figure 32: ROS levels in untreated (A) and β-estradiol treated (B) G1E-ER cells, under oxidative stress conditions.
FACS analysis histograms of ROS levels in untreated (top) and β-estradiol treated (bottom) G1E-ER cells, under normal (control) conditions (A) and oxidative stress conditions induced by the addition of hemin (B) or H_2O_2 (C). Over-induction of oxidative stress by the addition of a 10-fold elevated concentration of H_2O_2 was used as control in order to confirm that addition of hemin or H_2O_2 does indeed result in elevated levels of ROS and oxidative stress (D).

We then measured NRF2 in nuclear extracts from GATA-1 null and GATA-1 expressing G1E cells and found that under oxidative stress conditions NRF2 levels are increased upon GATA-1 activation (Figure 33). We can therefore assume that GATA-1 is involved in the cytoprotective mechanism of cells against oxidative stress and presumably does so by regulating the upstream activator of HO-1, NRF2.

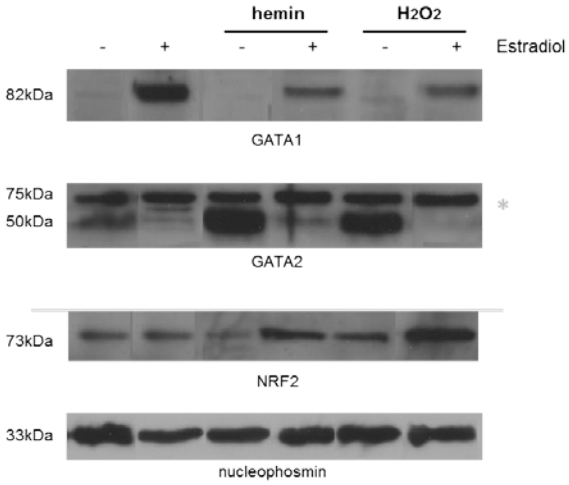


Figure 33: Detection of NRF2 levels in untreated and β-estradiol treated G1E cells, under oxidative stress conditions.
Nuclear extracts from untreated and β-estradiol treated G1E cells upon addition of hemin and H_2O_2. Extracts were immunoblotted with anti GATA-1 antibody (top) and anti GATA-2 antibody (middle) to confirm GATA-1 activation, anti NRF2 (middle) and anti nucleophosmin (lower) as protein loading control. Asterisk denotes a non-specific cross-reacting band (middle panel).

Additional ChIP experiments using antibodies against BACH1 and NRF2 to detect distinct occupancy of these factors in the upstream enhancers of *ho-1* in GATA-1 null and GATA-1 expressing G1E cells will reveal whether GATA-1 is indeed a direct regulator of HO-1, at the transcription level.

IV.Summary

In order to identify the GATA-1 gene targets throughout erythropoiesis we used the cell surface Ter119 receptor to separate mouse fetal liver cells into early proerythroblasts ($Ter119^-$ fraction) and differentiated erythroid cells ($Ter119^+$ fraction). The meta-analysis of GATA-1 deep sequencing and microarray expression profiling of the two populations, led to the identification of genes that differ in their expression levels and in the enrichment for GATA-1.

Overall, we noticed a considerable conservation of GATA-1 target genes throughout erythroid differentiation. Mobility of an appreciable fraction of GATA-1 targets within the three Classes was observed as erythroid differentiation proceeds from $Ter119^-$ to $Ter119^+$ cells. More specifically, of the 1,765 genes that are bound by GATA-1 in both $Ter119^-$ and $Ter119^+$ cells, 480 genes (27%) show reduced GATA-1 binding in mature $Ter119^+$ cells compared to $Ter119^-$ cells, whereas, 353 genes (20%) transitioned to a higher class as a result of higher enrichment for GATA-1 binding with erythroid differentiation. Interestingly, Gene Ontology (GO) analysis using DAVID (Huang et al., 2009) of genes transitioning to lower categories with erythroid differentiation, revealed a relative enrichment for genes involved in immune and early hematopoietic pathways, myeloid differentiation and immune response activation, for example, *pu.1*, *kit*, *nfkb1*, *hhex*, and *zfp36* genes. Though not as clear cut, the GO analysis of genes transitioning to a higher category showed a relative enrichment in oxygen response pathways, chromatin organization and modification and cell cycle regulation, which are all processes associated with mature erythroid physiology. Examples include the *slc4a1*, *cat*, *ho-1*, *urod*, and *trim10/herf1* genes.

Generally, we find that genes representing the erythroid transcription program are highly enriched for GATA-1 binding throughout differentiation. By contrast, reduced GATA-1 binding during differentiation appears to be associated with genes implicated in early hematopoietic and alternative lineage programs. Thus, it seems that GATA-1 binds to its erythroid target genes at an early differentiation stage and orchestrates their expression in a manner dependent on the amount of bound GATA-1 protein. This may reflect a reduction in the amount of GATA-1 protein with differentiation and/or a reduced requirement for GATA-1 to remain bound to genes that it has already repressed, for example.

Alternatively, the fact that there is reduced GATA-1 binding in mature $Ter119^{+}$ cells compared to $Ter119^{-}$ cells could be attributed to the state of chromatin given that terminally differentiated cells, which are smaller and about to exude their nucleus, have more condensed chromatin and may thus be more inaccessible to the antibody. To these ends we have isolated chromatin from fetal livers of GATA-1 biotin tag knock-in embryos and plan to investigate whether this is the case with streptavidin too and also test whether the streptavidin-biotin system is more efficient and results in less background than the antibody, being more suitable thus for massive parallel sequencing.

The identification of the subsets of GATA-1 target genes that vary greatly in their expression pattern from one population to the other, in that they are either significantly upregulated or repressed, led to the identification of the *heme oxygenase 1* (*ho-1*) gene which is upregulated 26-fold, and the *myc* gene which is repressed by 7-fold. Additional qPCR experiments on the *myc* GATA-1-bound sites identified by the deep sequencing data analysis, validated enrichment of GATA-1 in these regions. Our results are in accordance with the microarray data of Rylski et al. (2003), which indicated that myc is repressed by GATA-1 in G1E-ER cells. ChIP experiments in estradiol-stimulated G1E-ER cells revealed GATA-1 inhibition of *myc* transcription by direct binding to the promoter of the myc locus (Rylski et al. 2003). The authors suggest that *myc* probably becomes repressed when GATA-1 protein levels reach a critical threshold and the fact that throughout the course of erythroid differentiation cells divide only two or three times implies that a burst of high-level of GATA-1 activity rather than gradual expression of GATA-1 occurs in primary cells (Rylski et al, 2003). These findings are consistent with our observations that GATA-1 shows a high level of binding in the early steps of erythropoiesis, perhaps to trigger the subsequent events of erythroid differentiation and then at later stages a lower activity of GATA-1 is required (and observed in $Ter119^{+}$ cells).

Microarray data from the same study (Rylski et al., 2003) indicated upregulation of *mad1*, *mad4* and *mxi1*, whose products oppose Myc actions in G1E-ER estradiol-induced cells. Conditional expression of Myc in human leukemia K562 cells revealed inhibition of erythroid differentiation by blocking the upregulation of Mad1 and of a set of genes that trigger erythroid differentiation such as GATA-1 (Acosta et al., 2008). Enforced expression of GATA-1, partially rescues the differentiation inhibition imposed by Myc, indicating that Myc results in the block of differentiation partly by antagonizing GATA-1 function. However, no *myc* repression by GATA-1 was observed in K562 cells (Acosta et al., 2008).

We found that GATA-1 not only binds to the myc locus but also to the other components of the Myc/Max/Mad network, the *max*, *mad1*, *mnt/rox*, *mad3* and *mxd4* genes. Moreover, expression analysis indicated that not only *myc* is repressed during differentiation but also, *mxd4*, *max* and *mnt/rox*. On the contrary, Mad1, which competes with Myc for heterodimerization with Max in order to promote differentiation versus proliferation, is 4-fold upregulated. Our data indicate a key *in vivo* involvement of GATA-1 in the Myc/Max/Mad network and a potential role for GATA-1 as a master regulator of this network in erythroid cells. Additional ChIP experiments in primary cells will provide greater insight into the potential regulation of other members of the network such as Mga proteins by GATA-1. In addition, similar experiments in G1E-ER cells will be of particular use, in revealing the net effects of GATA-1 expression in the Myc/Max/Mad network with the simultaneous restoration of erythroid differentiation.

Another aspect of this study was the oxidative stress pathway as indicated by the great upregulation of *ho1* in differentiated erythroid cells. qPCR experiments demonstrated multiple GATA-1 binding sites across the *ho-1* locus with rather low enrichment scores. In addition, we found three GATA-1 peaks in the *bach1* gene and a single highly enriched GATA-1 binding site in the *nrf2* gene, in both cell populations, which were validated by qPCR. These data indicate that the binding of GATA-1 on *ho-1* is not highly enriched so as to account for the great upregulation in *ho-1* expression and that

GATA-1 may also exert this 26-fold *ho-1* upregulation in expression, by controlling the expression of the upstream *ho-1* regulators bach1 and nrf2. To these ends, we explored the role of GATA-1 in regulating bach1 and nrf2 directly, by using the erythropoietic cell line G1E-ER. This cell line allowed us to directly assess potential GATA-1 functions throughout erythropoiesis and under oxidative stress conditions.

We found great conservation in both the enrichment levels and the pattern of GATA-1 binding to the *nrf2* and *bach1* loci between the primary cells and the G1E-ER cell line. The measurement of the levels of nrf2 and bach1 transcripts by qPCR revealed a 2-fold and a 13-fold reduction respectively, in G1E-ER cells upon GATA-1 activation. These data indicate that GATA-1 directly controls the expression of bach1 and to a lesser extent of nrf2; which is in accordance to previous findings that it is alleviation of repression, rather than upregulation of activators like Nrf2, that is the critical step for ho-1 induction, given that ablation of bach1 results in constitutive expression of ho-1 (Sun et al., 2002).

Miyazaki et al. (2010) treated the human leukemic cell line U937 and control monocytes from healthy donors with lipopolysaccharide (LPS) –which is known to affect *ho-1* expression- to investigate how *ho-1* expression is regulated in a cell line and in a primary cell context. They found that LPS upregulated HO-1 in U937 and suppressed it in monocytes without altering the expression levels of bach1 or nrf2 in whole cell lysates. HO-1 expression levels were altered as a consequence of an increase in the levels of NRF2 (U937) or BACH1 (monocytes) in the nucleus exclusively. The intracellular localization of NRF2 or BACH1 by immunoblotting was detected in primary cells only, as the levels of expression in the cell line were below the detection level (Miyazaki et al., 2010).

In order to examine the phenotype of GATA-1 in the regulation of HO-1 expression we exposed the G1E-ER cells to oxidative stress conditions. Oxidative stress was induced by the addition of either H_2O_2 or hemin to G1E-ER cells with and without activation of GATA-1 by β-estradiol. Following the Miyazaki et al. observations (2010) we separately analyzed the HO-1, BACH1 and NRF2 levels in the cytoplasmic and nuclear fraction of G1E-ER cells by immunoblots. We were unable to detect HO-1 or BACH1 proteins in either of the nuclear or cytoplasmic fractions, even under oxidative stress conditions. Lack of detection of HO-1 protein has been reported in previous studies by Alves et al. (2011), who used the human K562 erythroleukemic cell line to analyze the heme degradation pathway under heme-induced erythroid differentiation. They found that HO-1 mRNA and protein expression levels were below limits of detection in K562 cells and that heme was unable to induce HO-1, at the protein and mRNA profiles. Moreover, they observed lack of significant expression of HO-1 in erythroid precursors during normal erythropoiesis in human bone marrow (Alves et al., 2011). Previous studies by Lavrovsky et al. (1994) however, have shown significant induction of HO-1 in heme-treated K562 cells, by Nothern blots.

Nevertheless, we were able to detect NRF2 levels in nuclear extracts from G1E-ER cells and observed a clear induction of NRF2 protein in the nucleus of GATA-1 expressing G1E cells versus GATA-1 null G1E cells under oxidative stress. Moreover, the FACS analysis of G1E cells under oxidative stress revealed that upon activation of GATA-1, there was less accumulation of ROS compared to GATA-1 null G1E cells. These data together, are in agreement with previous findings from studies in macrophages according to which, upon treatment of cells with electrophiles, NRF2 is liberated from KEAP1 in the cytoplasm, translocates to the nucleus and activates transcription of cytoprotective genes, including *ho-1* (Ishii et al., 2000; Itoh et al., 2003). Besides, studies on renal proximal tubular epithelial cells revealed that heme resulted in the up-regulation of NRF2 by stabilizing the protein; no alteration in the levels of mRNA or of Maf-dimerization was observed (Alam et al., 2003). We therefore propose that GATA-1 is involved in the cytoprotective mechanism of cells against oxidative stress and presumably does so by regulating the upstream activator of HO-1, NRF2.

Controversial data have been published concerning the interplay between Bach1 and Nrf2 (reviewed by Paine et al., 2010). Collectively, in certain Bach1 knock-down studies ho-1 induction was triggered without induction of Nrf2-dependent genes (MacLeod et al., 2009) whereas in other studies *bach1* repression induced *ho-1* expression in coordination with Nrf2-dependent genes (Dhakshinamoorthy et al., 2005). Recent work on human Huh-7 cells indicated that heme treatment resulted in the up-regulation of *ho-1*, by regulating bach1 (Shan et al., 2004). Additional studies on the same cells, treated with a potent inducer of ho-1, CoPP, indicated *ho-1* gene activation by post-transcriptional destabilization of BACH1 and stabilization of NRF2 (Shan et al., 2006).

We now wish to investigate whether GATA-1 confers cytoprotection against oxidative stress not only by post-transcriptionally regulating NRF2 but also BACH1.The stability of the two proteins could be examined by monitoring the decay of NRF2 and BACH1 in the presence or absence of oxidative stress inducers –heme or H_2O_2– and upon GATA-1 activation after inhibition of protein synthesis by cycloheximide. Moreover, we are planning to examine whether GATA-1 regulates expression of other important antioxidant enzymes, Nrf2-dependent and independent, the action of which protects red blood cells from oxidative stress. For this purpose the occupancy of GATA-1 on superoxide dismutase (SOD), glutathione peroxidase (GPx), glutathione reductase (GR) and others will be investigated by qPCR in fetal liver cells and G1E-ER cells. Additional ChIP experiments using antibodies against BACH1 and NRF2 to detect distinct occupancy of these factors in the upstream enhancers of *ho-1* and of the antioxidant genes, in GATA-1 null and GATA-1 expressing G1E cells under oxidative stress will provide greater insight in the role of GATA-1 in the cell anti-oxidant mechanism.

In conclusion, this is the first genome wide analysis of *in vivo* GATA-1 gene targets in primary murine adult stage erythroid cells by massive parallel sequencing. While searching for potential GATA-1 binding sites that differ throughout the differentiation of erythroid cells, we came across a novel role for GATA-1 in the Myc/Max/Mad network and in oxidative stress cell response.

The bioinformatics analysis of the deep sequencing results took place in the BSRC ‘Alexander Fleming’ with the help of G. Papadopoulos.

3.3 A knockdown approach reveals a role for Dnmt1 in cell cycle regulation during erythroid cell differentiation?

Gene expression is regulated by the synergistic action of transcription factors and epigenetic regulators such as DNA and histone modifications. One such epigenetic modification is DNA methylation which was in fact the first modification identified and has been studied for more than half a century. DNA methylation occurs predominantly at CpG dinucleotides and has generally been inversely correlated with transcriptional activity (Spada et al., 2006; Kouzarides et al., 2007; Qin et al., 2011). Genome wide DNA methylation patterns change during development and differentiation and have been implicated in various biological processes such as gene regulation, chromatin structure, DNA replication and in developmental processes such as parental imprinting, X-chromosome inactivation in females and in carcinogenesis. DNA methylation in mammals is conferred by the DNA methyl transferase family of proteins which includes the DNMT3a and DNMT3b *de novo* methyltransferases and the DNMT1 maintenance methyltransferase which adds methyl groups on newly synthesized hemimethylated strands of DNA following replication (Robertson et al., 2000; Fatemi et al., 2001; Qin et al., 2011; Arand et al., 2012).

The methylation status of regulatory DNA sequences correlates with the transcriptional activity of genes. In higher eukaryotes, transcriptionally active chromatin (euchromatin) is associated with DNA demethylation (unmethylated first exons and promoters), acetylated histones and the methylated K4 residue of histone H3. On the contrary, transcriptionally inactive chromatin (heterochromatin) is associated with methylated DNA, deacetylated and methylated K9 residues of histone H3 (reviewed by Cheung & Lau, 2005; Turek-Plewa & Jagodziński, 2005). DNMT family activity is crucial for the initiation and maintenance of chromatin remodeling and gene expression regulation. DNMT1 in particular, recruits chromatin-modifying enzymes including deacetylases -HDAC1 and HDAC2- and histone methyltransferases (Suv39h1) in order to maintain the methylation status and induce the appropriate histone H3 modification and subsequent silencing of genes (Fuks et al., 2000, 2001; Espada et al., 2004). Moreover, recent studies showed that for a set of genes, DNMT1 acts as a repressor independently of its methyltrasferase activity, thus providing evidence that DNMT1 may also act as a scaffolding protein for the recruitment of additional repressive partners (Milutinovic et al., 2004; Clements et al., 2012).

Recent work from the Strouboulis lab provided evidence for the interaction of DNMT1 with transcription factors involved in erythropoiesis: a) DNMT1 was co-purified with biotin-tagged TR2/TR4 nuclear receptors and b) DNMT1 was independently co-purified with biotin-tagged GATA-1.

Nuclear receptors TR2 and TR4 (TR2/TR4) bind to direct repeats in mouse and human embryonic and fetal β-type globin promoters and have been proposed to act as adult-stage specific repressors for these genes. Mass spectrometry analysis of nuclear extracts from mouse erythroleukemic (MEL) cells expressing biotin-tagged TR2 and TR4 revealed interactions of TR2/TR4 with DNMT1 (Cui et al., 2011). Moreover, mass spectrometry analysis of nuclear extracts from MEL cells expressing biotin-tagged GATA-1 revealed an interaction of GATA-1 with DNMT1 independently of the TR2/TR4 complex (Papageorgiou, Amaral-Psarris and Strouboulis, unpublished). Importantly, the mass spec data indicating an interaction between GATA-1 and DNMT1 have been confirmed by immunoprecipitations (Papageorgiou, Amaral-Psarris and Strouboulis, unpublished).

The fact that DNMT1 interacts in two separate complexes with transcription factors implicated in erythropoiesis and globin gene regulation raises the prospect of DNMT1, and DNA methylation in general, may play an important role in erythroid differentiation.

To address a potential implication of DNMT1 in erythroid cell differentiation, we performed DNMT1 knock-down experiments in mouse fetal liver derived proerythroblasts and in mouse erythroleukemic (MEL) cells. Preliminary expression profiling, cytospin preparations and FACS analysis in a set of at least three independent experiments showed a clear arrest in erythroid differentiation with defects in cell cycle regulation as a result of reduced levels of DNMT1.

I. DNMT1 knock-down screening

An initial screening was set up to assess the efficiency of five different lentiviral vectors expressing shRNAs against DNMT1, namely clones 3545-3549, for knocking down DNMT1 in MEL cells. Three out of the 5 shRNA clones tested, clones 3546-3548, resulted in the reduction of DNMT1 protein, whereas clones 3545 and 3549 did not show an appreciable reduction in DNMT1 protein levels. Following quantitation, (Figure 34C) we concluded that shRNA clone 3548 was the most efficient in the DNMT1 knock down in MEL cells and in mouse fetal liver cells (Figure 34).

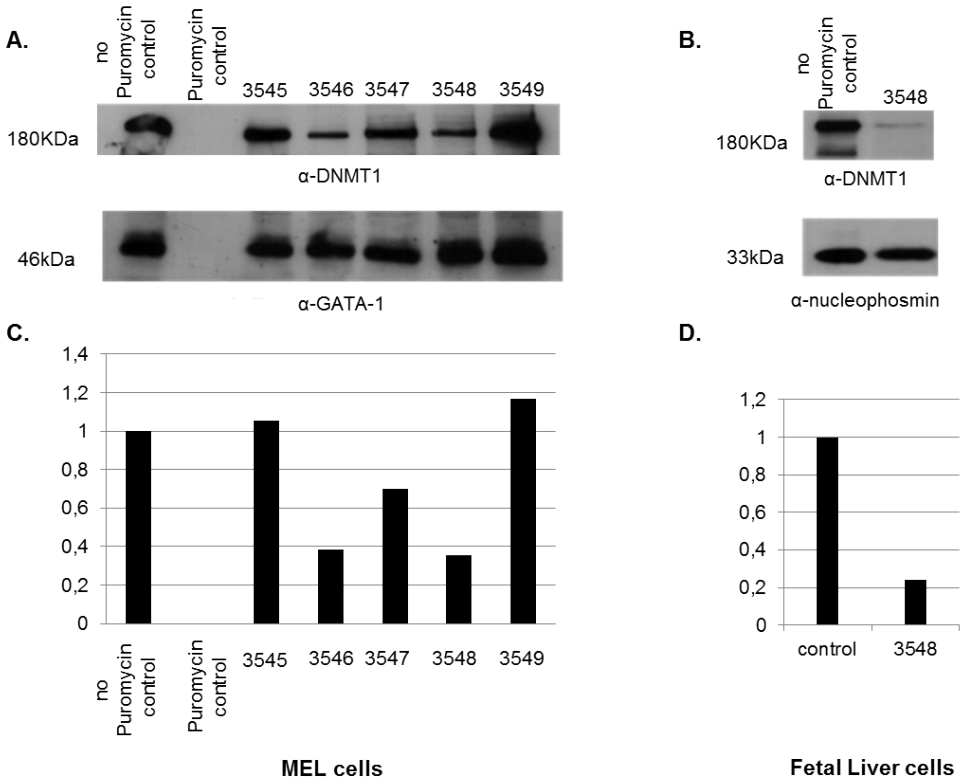


Figure 34: Screening of DNMT1 shRNA clones for their efficiency to knock down DNMT1 by Western blots.

A: Nuclear extracts from MEL cells transduced with viral supernatant media from HEK cells transfected with five different DNMT1 shRNA plasmids. Extracts were immunoblotted with anti DNMT1 antibody (top) and anti GATA-1 (lower) antibody as protein loading control. Extracts from non-transduced cells with and without puromycin were used as controls.

B: Nuclear extracts from mouse fetal liver cells transduced with viral supernatant media from HEK cells transfected with DNMT1 shRNA clone 3548. Extracts were immunoblotted with anti DNMT1 (top) and anti nucleophosmin (lower) as protein loading control.

C, D: Relative levels of DNMT1 protein in MEL cells (C) and fetal liver cells (D) following lentiviral mediated shRNA DNMT1 knock-down. Samples were normalized to GATA-1 (MEL cells) or nucleophosmin (fetal liver cells) and quantitated using the ImageJ image processing package.

In order to exclude the possibility that the reduced levels of the DNMT1 protein observed are due to lentivirus induced cell death, we measured apoptosis and cell death by FACS analysis. Control non-transduced MEL cells and cells transduced with either a scrambled shRNA virus (plko.1) or the 3548 shRNA virus were stained with annexin to determine the percentage of cells undergoing apoptosis and with Propidium Iodide (PI) for the measurement of nonviable cells (Figure 35).

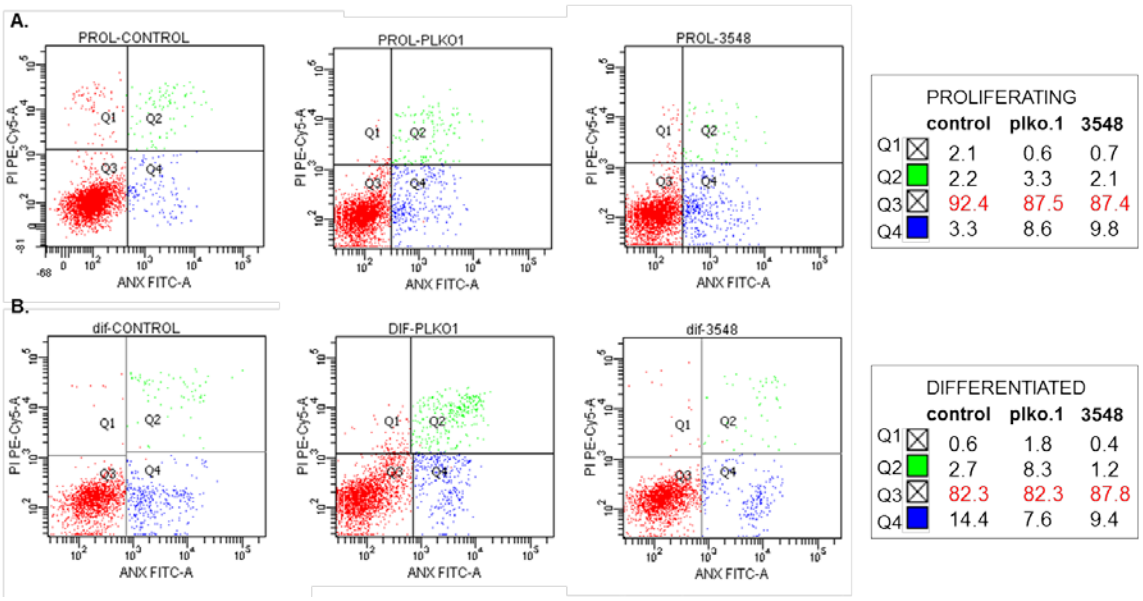


Figure 35: Cell viability as measured by FACS analysis.
MEL cells non-transduced (left) and transduced with scrambled shRNA virus (plko.1) (middle) and 3548 shRNA virus against DNMT1 (right) stained with FITC Annexin V (X-axis) and propidium iodide (Y-axis). Viable proliferating (A) or differentiated (B) cells fall in quadrant Q3; actual percentages shown in red in top and bottom tables to the right. Data shown obtained from one experiment of a series of at least three separate experiments with similar results.

The majority of cells stained negative for both annexin and PI in proliferating and differentiated cells transduced with scrambled control and 3548 viruses. In all cases, the percentage of viable cells ranged between 82-92%, which led us to the conclusion that there is no overt viral toxicity effect in our knock down assays. Similar percentages of viable cells were observed when mouse fetal liver cells were virally transduced with the same vectors (data not shown).

As a further test for the efficiency and specificity of the 3548 shRNA construct, we measured DNA methlytransferase activity in crude nuclear extracts of 3548 and plko.1 transduced fetal liver cells or in DNMT1 immunoprecipitates from the same extracts. We observed a 55% reduction in DNA methyltransferase activity in nuclear extracts and an 85% reduction in activity in DNMT1 immunoprecipitates (Figure 36), thus providing concrete evidence as to the efficiency and specificity of the DNMT1 knockdown. Moreover, the fact that DNMT3a protein is not detected in the DNMT1 immunoprecipitates (Figure 36A, middle panel) clearly suggests that the reduction in DNA methyltransferase in the knock down extracts is specifically due to the reduction in DNMT1 protein levels.

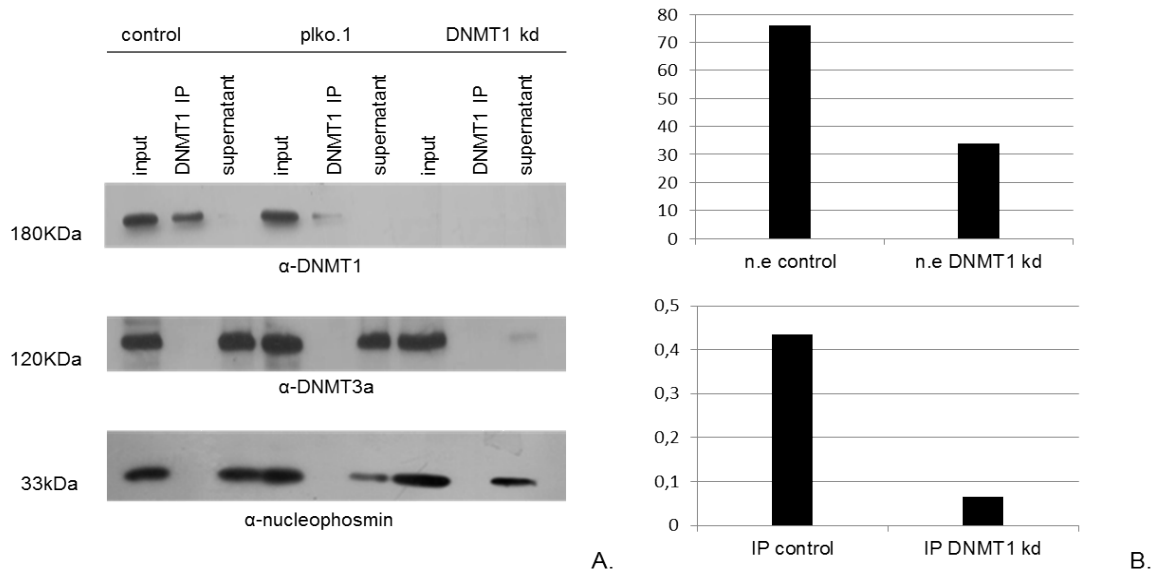


Figure 36: DNA methyltransferase activity assays of DNMT1 nuclear extracts and immunoprecipitates from E12.5 fetal liver cells.

A: anti-DNMT1 immunoprecipitations of nuclear extracts from non-transduced (control), scrambled shRNA (plko.1) and 3548 shRNA (DNMT1 kd) transduced mouse fetal liver derived proerythroblasts immunoblotted with anti-DNMT1 antibody. Probing with anti-DNMT3a antibody (middle) confirms that the reduction in DNA methyltransferase activity observed in the DNMT1 immunoprecipitate is specifically due to the knockdown of DNMT1. Nucleophosmin was used as a loading control (bottom). Input, Supernatant: 10% of the nuclear extract used for the IPs.

B: DNA methyltransferase activity as assessed with an ELISA-based method using nuclear extracts (top) or DNMT1 immunoprecipitates (bottom) from non-transduced and 3548 shRNA transduced mouse fetal liver cells.

II. Analysis of the DNMT1 knock-down phenotype in MEL cells

We next proceeded to assess the effects of knocking down DNMT1 on erythroid differentiation. The short lifespan and the limited number of viable fetal liver cells by the end of the viral transduction protocol did not allow us to carry these assays in fetal liver erythroid differentiation. Thus, we focused on MEL cells which are a proerythroblastic cell line that can be chemically induced, e.g. by dimethyl sulfoxide (DMSO) treatment, to undergo terminal erythroid differentiation (Antoniou, 1991). Following shRNA lentiviral transduction, MEL cells were induced to differentiate using DMSO and were examined for a phenotype upon DNMT1 knock-down by cytospins and microarray expression analyses. We again confirmed the efficiency of transduction in both proliferating and differentiated MEL cells by Western blots (Figure 37).

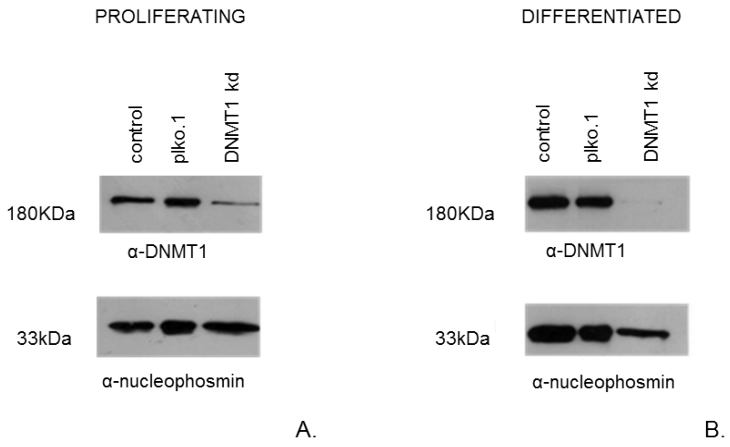


Figure 37: DNMT1 shRNA knock-down in MEL cell differentiation assayed by Western blots.

Nuclear extracts from proliferating (A) and differentiated (B) non-transduced (control), scrambled shRNA (plko.1) and 3548 shRNA (DNMT1 kd) transduced MEL cells blotted with anti-DNMT1 antibody. Nucleophosmin was used as protein loading control (lower panel).

Cytospins of proliferating and differentiated MEL cells transduced with either the scrambled shRNA or the DNMT1 kd shRNA revealed that DNMT1 is required for the cells to terminally differentiate. DNMT1 knock-down proliferating cells appear more disordered than control cells (Figure 38C) whereas a more severe phenotype is observed in the differentiated cells where the majority of cells are arrested in an earlier stage of differentiation (Figure 38F). In the scrambled shRNA differentiated slides we observe reticulocytes and red blood cells as expected after 3 days in differentiation conditions whereas in the DNMT1 shRNA slides, the cells resemble polychromatic and orthochromatic erythroid cells (Figure 38).

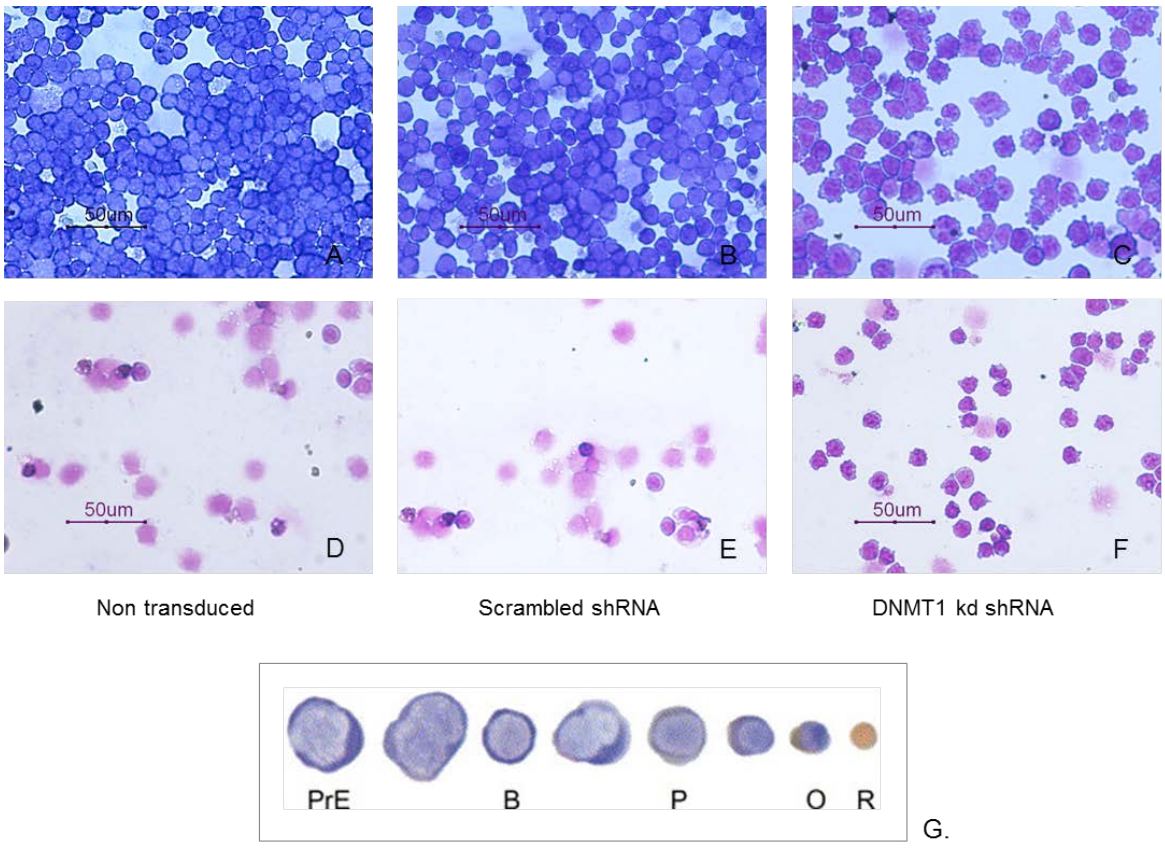


Figure 38: Cytospins of proliferating and differentiated non-transduced, scrambled shRNA control and DNMT1 kd MEL cells.
May Grünwald/Giemsa-stained cytospins of proliferating (A, B, C) and differentiated (D, E, F) non transduced (left), scrambled shRNA (plko.1) (middle) and 3548 shRNA (DNMT1 kd) (right) MEL cells.
G: Stages of erythroid cell differentiation from the proerythroblast stage (PrE, left) to reticulocyte (R, right) passing through dividing proerythroblast, basophilic (B, plus dividing basophilic cell), polychromatic (P), and orthochromatic (O) stages. Image taken from Gutiérrez et al., 2005.

The cytospin preparations suggest that DNMT1 kd cells are arrested at an earlier stage of differentiation compared to the non-transduced or the scrambled RNA transduced cells. This observation was further confirmed by staining cells with Propidium Iodide (PI) to measure DNA content by flow cytometry, so as to estimate the percentage of proliferating versus G1 arrested terminally differentiated cells. It appears that upon differentiation, only 60% of the DNMT1 kd cells have stopped dividing and have entered the resting phase G0/G1 of the cell cycle, as opposed to 90% of the control cells (non-transduced or scrambled shRNA control cells). Moreover, 20% of the

DNMT1 kd cells are in the G2 phase, i.e synthesizing proteins and continue to grow to enter the mitotic phase (M), as opposed to 10% of the control cells (Figure 39B). Thus, it seems that although proliferating non-transduced and transduced MEL cells have the same cell cycle phenotype (Figure 39A), upon differentiation, more than one third of the knock down cells do not exit the cell cycle and continue to proliferate as opposed to the control cells which are mostly arrested in the G0 resting phase of the cell cycle (Figure 39B). We can therefore assume that the reduced levels of DNMT1 do not have an impact on proliferating erythroid cells, however, they result in impaired differentiation of MEL cells.

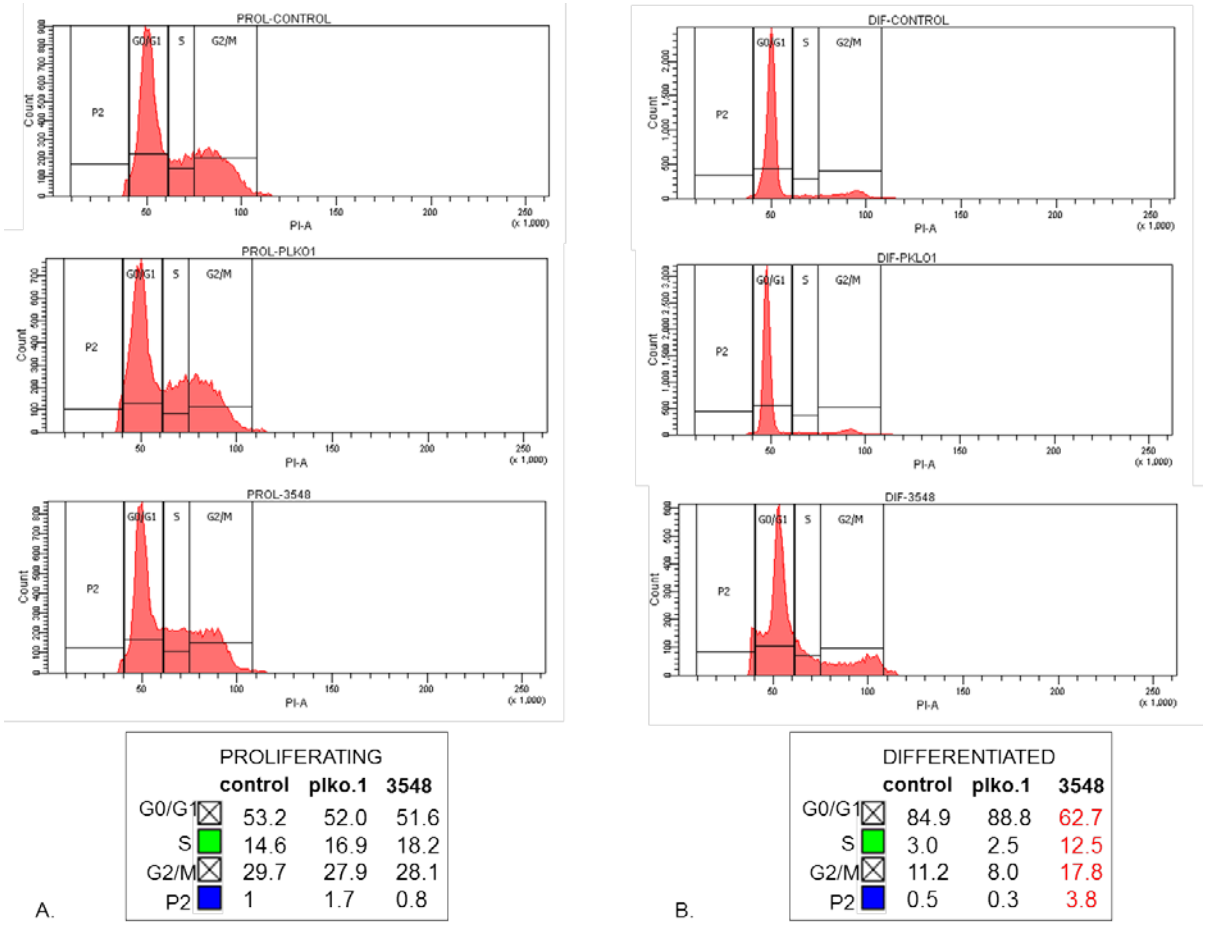


Figure 39: Cell cycle analysis by DNA content.
MEL cells non-transduced (top) or transduced with scrambled shRNA virus (plko.1) (middle) and 3548 shRNA virus against DNMT1 (bottom) were stained with PI. Cell cycle distribution of proliferating (A) and differentiated (B) cells as measured by FACS analysis; actual percentages shown in bottom left and right tables.
Data shown were obtained from one experiment of a series of at least three separate experiments with similar results.

In addition, preliminary expression profiling analysis of proliferating and differentiated non-transduced cells and MEL cells transduced with either the scrambled shRNA or the DNMT1 shRNA provided molecular evidence for the defects in MEL cell differentiation observed by the morphology of the cells and the cell cycle analysis (Figures 38, 39). Total RNA was extracted from non-transduced cells and MEL cells transduced with either the scrambled shRNA or the 3548 DNMT1 shRNA, in proliferating and differentiating conditions. Microarray based expression profiling showed that 162 genes were upregulated in non-transduced control MEL cells upon differentiation, versus 326 and 261 genes in the scrambled shRNA and DNMT1 shRNA transduced MEL cells, respectively. Moreover, 629 genes

were repressed in non-transduced MEL cells upon differentiation, versus 486 and 233 genes in scrambled shRNA and DNMT1 shRNA transduced MEL cells, respectively. As seen in Table 4, the ratio between repressed versus induced cells is 3.88 in non-transduced MEL cells, 1.49 in scrambled shRNA transduced MEL cells and 0.89 in DNMT1 shRNA transduced MEL cells.

	Induced	Repressed	Ratio (R/I)
MEL wt cells	**162**	**629**	**3.88**
MEL scrambled shRNA	**326**	**486**	**1.49**
MEL DNMT1 shRNA	**261**	**233**	**0.89**

Table 4: Gene expression analysis data summarized.
Total numbers of induced and repressed genes and repressed versus induced ratio in non-transduced and scrambled or DNMT1 shRNA transduced MEL cells.

A total of 299 repressed genes, commonly shared between non-transduced and scrambled shRNA transduced MEL cells were not repressed in DNMT1 shRNA transduced MEL cells. Gene Ontology analysis revealed that a significant proportion of the 299 non-repressed genes in the DNMT1 knock-down cells (22%) are related to cell cycle progression. In contrast to the repressed genes, which seem to differ significantly between the DNMT1 knock-down cells and the control cells, all genes related to erythroid differentiation were activated in both cell populations (Figure 40). It appears thus that the DNMT1 knock-down does not block the activation of the erythroid genes, such as globin genes, during differentiation but, rather, fails to inhibit genes responsible for cell proliferation arrest, thus leading to cell cycle defects and impaired erythropoiesis.

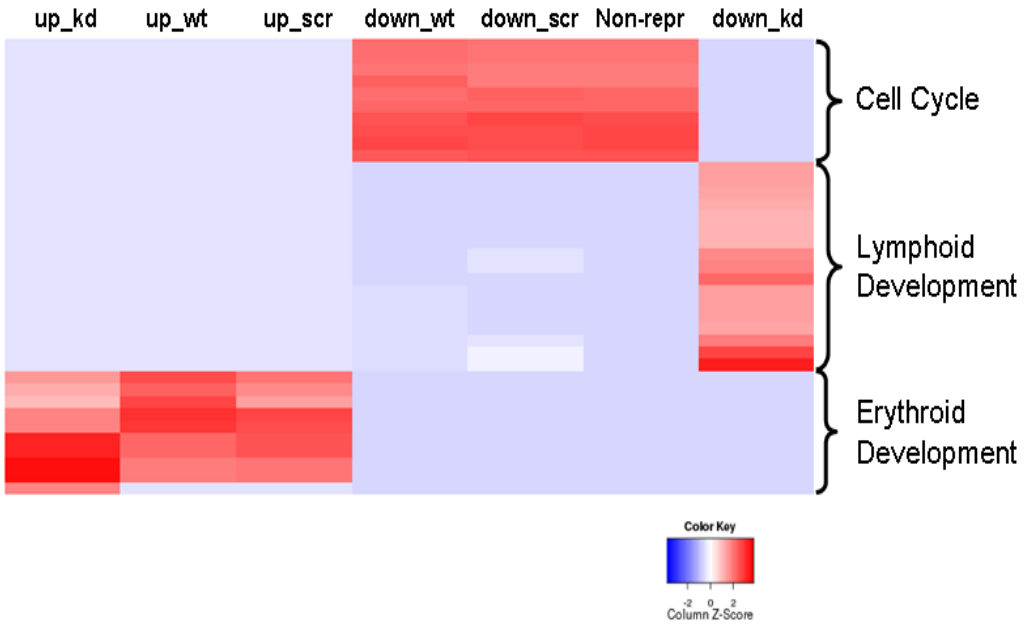


Figure 40: Heatmap of the expression profiles of non-transduced and scrambled or DNMT1 shRNA transduced MEL cells.
Clustering of up- and down-regulated genes in non-transduced (wt), scrambled (scr) or DNMT1 shRNA(kd) transduced MEL cells. Broad classification of gene categories based on their GO function is shown to the right.

III. Summary

Recent evidence from the Strouboulis lab showing that DNMT1 interacts in two separate complexes with transcription factors involved in erythropoiesis and globin gene regulation, raised the prospect of DNMT1 being implicated in erythroid differentiation. To these ends, we knocked down DNMT1 in MEL and mouse fetal liver cells and induced them to differentiate. Reduced levels of DNMT1 were observed by Western blots and by measuring DNA methyltransferase activity in nuclear extracts and DNMT1 immunoprecipitates. MEL cells were induced to differentiate and impaired erythropoiesis was observed by three different assays. First, cytospin preparations revealed that the majority of DNMT1 knock down cells were arrested at an earlier stage in differentiation compared to control cells. Only a limited number of DNMT1 knock down cells resembled red blood cells or reticulocytes, as would be expected for differentiated cells. Most cells seemed to be polychromatic or orthochromatic erythroid cells. Second, staining cells with PI showed that more than one third of the knock down cells continued proliferating upon differentiation conditions, whereas the vast majority of control cells (90%) had exited the cell cycle and had differentiated. Third, preliminary microarray experiments revealed that 299 repressed genes, commonly shared between non-transduced and scrambled shRNA transduced MEL cells were not repressed in DNMT1 shRNA transduced MEL cells. Further analysis showed that 22% of these genes are related to the cell cycle, whereas all genes related to the erythroid genes were activated, as expected. We therefore propose that the DNMT1 knock-down results in impaired erythropoiesis due to cell cycle defects deriving from failure to inhibit the genes responsible for cell proliferation arrest. It remains presently unclear as to how DNMT1 is recruited to the cell cycle related genes. However, since GATA-1 has been previously reported to regulate cell cycle related genes (Rylski et al., 2003), it is possible that at least part of DNMT1's function in cell cycle regulation may be taking place through its interactions with GATA-1.

Several studies on cell lines and mice homozygous for Dnmt1 null mutations have also underscored the importance of DNMT1 in maintaining DNA methylation during differentiation. Dnmt1 null embryonic stem (ES) cells are restricted in their differentiation potential, whereas dnmt1 homozygous null mice die prior to the 8-somite stage (Lei et al., 1996). Conditional homozygous dnmt1 null somatic cells, exhibiting reduced methylation levels, die due to increased apoptosis at day 5 (Jackson-Grusby et al., 2001). These data imply that despite the fact that *de novo* methylation is sufficient to sustain undifferentiated cells, lineage commitment and survival of differentiated cells requires the function of DNMT1. Therefore for differentiation to occur, the maintenance of the methylation pattern by DNMT1 –previously established by the *de novo* methylases- is required (Spada et al., 2006).

The same impaired differentiation phenotype was observed in dnmt1 ablation studies in hematopoietic stem cells (HSCs). Experiments on mice with gradually diminished dnmt1 expression revealed that distinct methylation threshold levels are required for the differentiation potential of HSCs. HSCs from dnmt1 null homozygous mice resulted in no myeloid or lymphoid colony formation in *in vitro* differentiation assays, due to increased apoptosis. Hypomorphic expression of dnmt1 rescued differentiation and revealed that alternative stem cell programs require distinct methylation levels. Stem cell commitment towards the lymphoid lineage for example, requires high methylation status, which however is no longer needed for the survival of the committed lymphoid cells. HSCs self-renewal program also requires sustained high levels of methylation. In accordance to these data, gene expression profiling studies on the hypomethylated dnmt1 HSCs resulted in increased expression of myeloerythroid genes versus decreased expression of 'stemness' and lymphoid genes (Bröske et al., 2009).

Moreover, data from Trowbridge et al. (2009), using dnmt1 null homozygous HSCs, showed that DNMT1 is required for presumptive self-renewal divisions of HSCs as well as for differentiation, with severe defects observed in the ability of DNMT1 null HSCs to generate mature myeloid cells. In another study from the same group, where they used leukemic granulocyte-macrophage progenitors, they showed that DNA methylation-mediated silencing of bivalent chromatin domains is required for apoptosis, tumor suppression and the hematopoietic lineage restriction (Trowbridge et al, 2012). In another study, from Shearstone et al. (2011), where they used reduced representation bisulfite sequencing (RRBS) to measure DNA methylation in differentiating mouse erythroblasts, it appears that global demethylation is continuous throughout differentiation and that demethylation of the erythroid specific β-globin locus coincides with the global demethylation observed. However, the increased levels of demethylation are not accompanied by increased levels of transcription. It was then shown that only the levels of the *de novo* methylatransferases Dnmt3a and Dnmt3b decrease during differentiation, whereas Dnmt1 mRNA and protein levels remain unaltered. In addition,

erythroid cells showed an overall less severe degree of demethylation compared to primordial cells, probably because genes upregulated during differentiation, such as β-globin, exhibited lower levels of methylation in the early steps of differentiation. In the same study, it was shown that impaired demethylation blocks the rapid induction of erythroid genes such as β-globin, a phenotype which is reversed upon DNMT1 inhibition by 5-aza 2'-deoxycytidine (5-aza). Moreover, Dnmt1 knock-down experiments in fetal liver cells showed that decreased levels of DNMT1 resulted in the acceleration of the erythroid specific gene induction. Thus, it appears that global demethylation may lead to the rapid removal of methylation marks at sites of massively induced erythroid genes (Shearstone et al., 2011).

Our data collectively support the notion that DNMT1 is implicated in erythropoiesis as evidenced by mass spectrometry analysis in MEL cells expressing biotin-tagged GATA-1 or TR2/TR4. We were able to efficiently knock-down DNMT1 in fetal liver and MEL cells, nevertheless we chose to study the role of DNMT1 in MEL cells as they could be maintained in culture for longer periods. DNMT1 knock-down studies in fetal liver cells were also conducted by Shearstone et al. (2011), however the cells were cultured in differentiation medium for a single day prior to collecting them for expression profiling analysis. We followed a different experimental scheme, whereby cells were isolated at day 0, expanded in proliferating medium for 5 days, transduced with lentivirus for 24hrs, cultured in selection medium and induced to differentiate for 3 days in order to obtain fully differentiated erythrocytes, as suggested by bibliography (Zhang et al., 2003).

Three lines of evidence confirm the distorted differentiation phenotype we observed: a) the morphology of the cells as seen by cytospin preparations, b) the cell cycle arrest as measured by DNA content and c) the microarray data analysis. DNMT1 knock-down MEL cells appear to proliferate at the same rate with control cells, however their morphology differs from the control phenotype. DNMT1 RNAi experiments in near-diploid human tumor cells (HCT116) and primary human fibroblasts (IMR90) resulted in two different phenotypes: proliferation was slowed down in HCT116 cells, whereas IMR90 cells were arrested in G1 phase. Proliferation was decelerated in HCT116 compared to control cells at 72hrs however resumed after 144hrs (Barra et al, 2012). Thus, it seems that different cell lines respond differently to DNMT1 knock-down assays and display distinct phenotypes at discrete collection time points. Therefore discrepancies observed between different experiments could be attributed to the distinct nature of each cell line used as well as to the degree of the reduction of the dnmt1 expression.

In accordance to previous studies (Shearstone et al., 2011), our data indicate that DNMT1 is dispensable for the induction of the genes of the erythroid program, however it is required for normal erythropoiesis to occur, by repressing the genes responsible for proliferation. It remains to investigate the molecular mechanism underlying the role of DNMT1 in erythropoiesis and identify the protein partners it interacts with to exert its function.

Chapter 4

Discussion

4.1 Overview of thesis

Hematopoiesis is the process for the generation of all mature blood cells from the hematopoietic stem cells and it involves regulation and interplay by a number of transcription factors (Zon, 2001). GATA-1 is the master regulator of erythropoiesis and a key determinant of the megakaryocytic, eosinophilic, dendritic and mast cell lineages (reviewed by Morceau et al., 2004). The long term scope of this work is to study in a systematic way the *in vivo* GATA-1 functions in regulating the differentiation of the aforementioned distinct hematopoietic lineages. GATA-1 knockout mice are embryonic lethal, thus precluding their use for such studies (Fujiwara et al., 1996). To these ends, we proceeded with biotinylation tagging of the endogenously expressed GATA-1 protein as a potentially uniform platform to characterize the different protein interactions that GATA-1 undergoes and the distinct genetic programs that GATA-1 regulates in different hematopoietic lineages.

This thesis summarizes the work that has been done to obtain the biotin tagged GATA-1 knock-in mice and recent findings concerning the functions of GATA-1 in regulating the erythroid lineage, in particular. Moreover, the role of a novel partner of GATA-1, DNMT1, is explored in the red blood cells and data from the knock-down approach we followed are presented.

Chapter 1 provided an overview of hematopoiesis, erythropoiesis and of known GATA-1 gene targets and protein partners. GATA-1 post-translational modifications and mutations associated with syndromes and diseases were briefly discussed.

Chapter 2 synopsizes the materials and methods used in the experiments described in this thesis.

In Chapter 3, section 1, the targeting strategy we followed to knock-in the biotin tag in the coding region of GATA-1 in ES cells and mice, was described. Evidence for the physiological expression of the biotin tagged GATA-1 in the red blood cells was provided and the mass spectrometry and ChIP-sequencing assays we performed, were introduced. The identification of novel *in vivo* GATA-1 partners and the elimination of the background binding seen in relevant ChIP experiments were documented.

In section 3.2, the results from the *in silico* analysis of GATA-1 deep sequencing in early proerythroblasts and differentiated erythroid cells were described. The identification of a novel role for GATA-1 in the Myc/Max/Mad network and in the oxidative stress cell response was presented.

Section 3.3 included the analysis of the data obtained from the knock-down assays on DNMT1 in mouse erythroleukemic cells. From the preliminary expression profiling, the cytospin preparations and the FACS analysis, a clear arrest in erythroid differentiation with defects in cell cycle regulation was observed.

Finally, in chapter 4 a synopsis of the current findings and the future prospects will be discussed.

4.2 Discussion

Biotinylation tagging has been widely applied not only for characterizing protein complexes but also to identify gene targets in ChIP assays, taking advantage of the very high biotin-streptavidin affinity, which allows high stringencies to be employed in both assays. Current bibliography includes publications with biotin tagged GATA-1, which has been used for nuclear extracts (de Boer et al., 2003; Rodriguez et al., 2005) and chromatin preparations (Yu et al., 2009) in erythroid cells. However, all previous findings derive from *in vitro* experiments, where cultured cell lines were used and GATA-1 levels fluctuated depending on the transfection efficiency.

To these ends, we biotin tagged GATA-1 in ES cells and obtained mice by blastocyst injections aiming to work with physiological levels of GATA-1 expression, in all hematopoietic lineages where it is active. The erythroid lineage, where GATA-1 is abundantly expressed, was the first lineage examined and provided evidence that GATA-1 protein structure and functions are not affected by the inclusion of the small tag. All blood parameters assayed for BioGATA-1 knock-in mice were indistinguishable from those for wild type mice. Immunoprecipitation experiments revealed no (gross) alterations in the structure of the protein, whereas streptavidin pull-downs and mass spec identified well-known GATA-1 interacting partners and streptavidin ChIP clearly showed biotin tagged GATA-1 to be binding to well-known GATA-1 gene targets in erythroid cells.
From our mass spec analysis, we identified GATA-1 in complex with a) the RUNX-1 and Lbp1a transcription factors, both of which are involved in hematopoiesis and b) members of the SWI/SNF family of chromatin remodeling proteins.

The RUNX-1 gene was first cloned from an acute myeloid leukemia (AML) patient with a chromosomal translocation in chromosome 21 (Miyoshi et al., 1991). It is a master regulator of hematopoiesis, specifying the definitive hematopoietic stem cell (HSC). RUNX-1 is fundamental in the derivation of HSCs from the hemogenic endothelium, and as such is one of the earliest markers of hematopoiesis (Lancrin et al., 2009). Various studies have provided evidence for the physical interaction of GATA-1 with RUNX-1. Elagib et al. (2003) used K562 leukemic cells to show that both N- and C-terminal regions of the GATA-1 protein are critical for RUNX-1 binding. Ectopic expression studies in *Drosophila* suggested that the interaction is mediated through the GATA-1 zinc-finger (Waltzer et al., 2003). Xu et al. (2006) provided further evidence that GATA-1 binds to RUNX-1 through its zinc-finger and that the C-finger is indispensable for this association. Transcriptional cross-talk between RUNX-1 and GATA-1 is of particular importance in megakaryopoiesis, where both factors are involved. It was shown that GATA-1 targets RUNX-1 for phosphorylation by cyclin-dependent kinases and converts RUNX-1 from a repressor to an activator (reviewed by Elagib & Goldfarb, 2007; Goldfarb, 2009). It is noteworthy that all previous reports have been established by studies in *in vitro* systems, using cultured cell lines. We report here that GATA-1 is *in vivo* in complex with the RUNX-1 protein and now need to validate this indication by western blot.

LBP1 and its highly related orthologue CP2 are ubiquitously expressed members of the grainyhead transcription family, share high sequence homology, common DNA binding motif and regulate a set of common genes. Studies in K562 cells have associated both factors with the regulation of the murine and the human α- and β-globin genes (Jane et al., 1995; Zhou et al., 2000). Loss of either LBP1 or CP2 had no effect in erythropoiesis, presumably due to their redundant functions (Parekh et al., 2004). However, overexpression of an antisense CP2 mRNA in MEL cells undergoing *in vitro* erythroid differentiation resulted in impaired α- and β-globin expression and hemoglobinization, suggesting an indispensable role for CP2 in globin transcription in erythroid cells (Chae et al., 1999; Chae & Kim, 2003). Multiple erythroid genes contain binding sites for CP2 adjacent to GATA-1 binding motifs in their regulatory regions. Bosé et al. (2006) showed that CP2 binds such regions in the presence of GATA-1 in K562 erythroid cells and by coimmunoprecipitations and GST-pull-down experiments demonstrated that the two factors physically interact, suggesting a novel module for erythroid genes regulation. Our mass spec data indicated that GATA-1 is *in vivo* in complex with the LBP1 protein and we now need to validate this observation by western blot.

SWI/SNF complexes are involved in both activation and repression of transcription by selectively regulating chromatin remodeling through direct interaction with specific DNA-binding proteins. Kadam et al. (2000) proposed that one mechanism by which mammalian SWI/SNF regulates specific subsets of genes is by interacting with distinct zinc finger proteins, including GATA-1, through the BRG1 and BAF155 or BAF170 subunits, and targeting the native complex to nucleosomal sites. In erythroid

cells, SWI/SNF facilitates remodeling and transcription of chromatin-assembled β-globin genes, by direct interaction with EKLF and GATA-1. Kadam and Emerson (2003) used recombinant GATA-1 protein for GST-pull-down experiments in MEL cells and identified that SWI/SNF exclusively interacts with the GATA-1 C-finger. Hypomorhic brg1 mutant mouse embryos develop normally until midgestation and then exhibit impaired erythroid lineage development and die due to lethal anemia, stressing out the essential role of SWI/SNF proteins and of BRG1 in particular, during erythropoiesis (Bultman et al., 2005). In addition, BRG1 interacts with GATA-1 in another complex that includes TAL-1/SCL, E47, LMO2 and Ldb1 and which stimulates P4.2 (protein 4.2) transcription in erythroid progenitors via two E box–GATA elements in the gene's proximal promoter. When BRG1 associates with this complex, it becomes recruited to these sites and represses P4.2 expression (Xu et al., 2006). All previous findings derive from *in vitro* experiments, whereas from our mass spec analysis we have identified GATA-1 in complex *in vivo* with multiple members of the SWI/SNF family, including brg1, snf2h, smarcc2, smarcc1, smarcd2, smarce1, pbrm1, snf5 and arid1a proteins. Further validation by western blots is in our immediate plans.
Work from previous reports suggests that GATA-1 undergoes multiple posttranslational modifications, including phosphorylation, sumoylation, ubiquitination and acetylation (Hernandez-Hernandez et al., 2006; Lee et al., 2009; Collavin et al., 2004; de Thonel et al., 2010; Lamonica et al., 2006). However, all evidence is based on work in cell lines or *in vitro* transfected cells. To these ends, we have initiated the characterization of the *in vivo* GATA-1 post-translational modifications in proliferating versus differentiated erythroid cells. Treatment of the isolated whole cell lysates with phosphatases such as λ-phosphatase or calf intestine phosphatase (CIAP) resulted in a shift in the molecular weight of GATA-1 band, thus providing evidence for *in vivo* GATA-1 phosphorylation. Large scale streptavidin pull down samples from differentiated erythroid cell extracts have been sent for mass spectrometry. The *in vivo* acetylated form of GATA-1 failed to be detected by the antibody raised by Boyes et al. (1998), used in *in vitro* studies in chicken models, whereas no antibodies are available for the other post-translational modifications. Once we identify the *in vivo* PTMs of GATA-1, we plan to raise antibodies or perform mutagenesis experiments in order to validate the peptides isolated from the mass spec. The development of modification-specific antibodies will also allow the functional dissection of specific GATA-1 PTMs in erythroid cell differentiation. This approach can also be extended in the other hematopoietic lineages where GATA-1 plays a role. Overall, our work will afford for the first time important insight into the *in vivo* post-translational modifications of GATA-1 in all tissues where GATA-1 is active.

Chromatin immunoprecipitation (ChIP) coupled to massive parallel sequencing (ChIP-seq) has provided the tools to obtain a genome wide comprehensive view of how transcription factors interact to drive tissue-specific differentiation. ChIP sequencing studies on GATA-1 using various, mostly erythroid cells, hematopoietic cell lines have revealed numerous known and novel GATA-1 direct gene targets (reviewed by Kerenyi & Orkin, 2010). All erythroid genes described include GATA binding sites in their promoters, including gata-1 itself, gata-2, klf1 and scl/tal1 loci (reviewed by Tsiftsoglou et al., 2009). Moreover, GATA-1 binds to and directly regulates antiapoptotic genes, such as Bcl-2 and Bcl-xL, globins, cell cycle genes, such as c-myb and c-myc and microRNAs. However, most of the newly identified GATA-1 gene targets derive from studies using immortalized erythroid cell lines. Our long-term aim is to unravel the distinct *in vivo* gene targets of GATA-1 in the different hematopoietic lineages where GATA-1 is active. To these ends, we isolated erythroid cells from mouse fetal liver cells in order to identify novel *in vivo* GATA-1 targets in the erythroid lineage and also to use this study as a reference for the physiological targets of GATA-1 in primary hematopoietic cells.

In this study, we used the cell surface Ter119 receptor, which is a marker for mature erythroid cells, to separate mouse fetal liver cells into early proerythroblasts ($Ter119^-$ fraction) and differentiated erythroid cells ($Ter119^+$ fraction) in order to identify potential GATA-1 binding sites during differentiation of erythroid cells. The *in silico* analysis of GATA-1 deep sequencing and microarray experiments of the two populations, led to the identification of a novel role for GATA-1 in the Myc/Max/Mad network and in oxidative stress cell response.

We verified GATA-1 binding not only to the *myc* gene (Rylski et al., 2003), but also to the *max*, *mad1, mad3, mnt/rox* and *mad4* loci in the two cell populations. Expression analysis indicated that not only *myc* is repressed during differentiation but also, *mxd4*, *max* and *mnt/rox*. By contrast, Mad1, which competes with Myc for heterodimerization with Max in order to promote differentiation versus proliferation, is 4-fold upregulated. Notably, binding of GATA-1 to the aforementioned loci, appeared

to be more enriched in the proerythroblast stage versus the more differentiated cells. Thus, it seems that GATA-1 binds to members of the extended Myc/Max/Mad network at an early differentiation stage, ensuring reduced expression of all but Mad1, which is activated, and then as differentiation progresses either there is reduced requirement for GATA-1 to remain bound to the genes in order to exert its function, or there is no great demand for further repression/activation and thus no requirement for high levels of GATA-1 in these loci. This idea is consistent with the general role of GATA-1 in repressing cyclins and c-myc at the G1-to-S transition phase of the cell cycle, forcing cells to stop proliferating, exit the cell cycle and induce them to differentiate (Rylski et al., 2003). This rapid arrest of cells at the G1 phase most probably requires an outburst of GATA-1 for the cells to exit the cell cycle and then presumably the threshold of GATA-1 is no longer essential, so GATA-1 levels decrease, allowing the cells to divide for another 2-3 times while differentiating. Overall, the broad binding of GATA-1 to the Max and Mad family proteins and their subsequent activation/repression during differentiation, make us hypothesize a novel aspect in GATA-1 functions and a broader role for GATA-1 as a direct upstream regulator of the Myc/Max/Mad network in erythroid cell differentiation. Additional ChIP experiments in primary cells will provide greater insight into the potential regulation of other members of the network such as Mga proteins by GATA-1. In addition, similar experiments in G1E-ER cells will be of particular use, in revealing the net effects of GATA-1 expression in the Myc/Max/Mad network with the simultaneous restoration of erythroid differentiation.

Another aspect of this study was the oxidative stress pathway and in particular *ho-1*, which is a key player against the cytotoxic effects of oxidative stress (reviewed by Ryter et al., 2006; Paine et al., 2010) and was found to be greatly upregulated in the differentiated erythroid cells.

HO-1 is the inducible form of HO, which binds heme at a 1:1 ratio and catalyzes its degradation into biliverdin with the release of free iron and CO (Otterbein et al., 2003). HO-1 is expressed at low basal levels in most cells and tissues and is upregulated by multiple oxidative stress stimuli. HO-1 promotes cytoprotection by degrading free heme, which would otherwise sensitize cells to undergo programmed cell death (Gozzelino et al., 2010). Heme consists of a tetrapyrrole ring with a central iron ion and is an abundant molecule in almost all living organisms, displaying two functions: when bound to hemoglobin, myoglobin and cytochromes plays a physiological role for the transport of oxygen and mitochondrial electrons and when free can cause toxic effects due to the oxidative stress induction it triggers (reviewed by Mense & Zhang, 2006). Compared to other cells in the body, the erythroid cells exhibit a much more rapid rate of heme synthesis so as to carry adequate levels of heme for active initiation of globin chain synthesis in the ribosomes (Ponka, 1997; reviewed by Cianetti et al., 2010). Cytoprotection of erythroid cells against the toxic effects of heme throughout erythropoiesis is therefore of particular importance and a tight regulation of oxidative stress response during the erythroid differentiation must exist.

qPCR experiments revealed a number of GATA-1 binding sites across the *ho-1* locus, albeit with low enrichment scores compared to the fewer but highly enriched GATA-1 binding sites observed in the gene loci of the transcriptional regulators of *ho-1*, bach1 and nrf2. Thus, we hypothesized that most probably the great upregulation in *ho-1* expression we observed is exerted by GATA-1 regulating the expression of bach1 and nrf2. To directly assess potential GATA-1 functions throughout erythropoiesis and under oxidative stress conditions we used the GATA-1 null proerythroblastic cell line G1E-ER. We detected a dramatic decrease in the levels of bach1 mRNA upon restoration of GATA-1 levels, whereas nrf2 transcripts did not significantly change, in accordance with previous findings showing that it is alleviation of repression, rather than upregulation of activators like Nrf2, that is the critical step for ho-1 induction, given that ablation of bach1 results in constitutive expression of ho-1 (Sun et al., 2002).

We then wished to examine what happens to the levels of HO-1, BACH1 and NRF2 in the presence or absence of GATA-1, under normal and oxidative stress conditions. We exposed G1E-ER cells to oxidative stress induced by heme or H_2O_2 and quantitated the cytoplasmic and nuclear levels of the proteins by western blots. HO-1 and BACH1 were below the levels of detection, as seen before (Lavrovsky et al., 1994; Alves et al., 2012) however we were able to measure a clear accumulation of NRF2 in the nucleus of G1E-ER cells, under oxidative stress and upon restoration of GATA-1. The cytoprotective role of GATA-1 against oxidative stress was also revealed by the decreased levels of ROS in the same cells, measured by FACS analysis. Our findings are in agreement with previous reports from studies in macrophages in which, upon treatment of cells with electrophiles, NRF2 is

liberated from KEAP1 in the cytoplasm, translocates to the nucleus and activates transcription of cytoprotective genes, including *ho-1* (Ishii et al., 2000; Itoh et al., 2003).

Another important parameter, currently under investigation, is the stability of BACH1 and NRF2, which also contributes to the interplay between the two proteins. We are planning to assess the stability of the two proteins by monitoring the decay of NRF2 and BACH1 in the presence or absence of oxidative stress inducers –heme or H_2O_2– and upon GATA-1 activation after inhibition of protein synthesis by cycloheximide.

Overall, we believe that GATA-1 is clearly involved in the cytoprotective mechanism of erythroid cells against oxidative stress and probably confers this role by regulating transcription and/or stabilization of various anti-oxidant players. Our *in silico* analysis has indicated GATA-1 binding to a number of Nrf2-dependent and independent anti-oxidant genes, such as peroxiredoxins, including Prdx1, superoxide dismutase (SOD), glutathione peroxidase (GPx), glutathione reductase (GR) and we suspect that GATA-1 might be a key regulator of the cellular antioxidant response by controlling their expression.

Moreover, ChIP experiments using antibodies against BACH1 and NRF2 to detect distinct occupancy of these factors in the upstream enhancers of *ho-1* and of the antioxidant genes in erythroid cells are in our immediate plans.

In conclusion, this is the first genome wide analysis of *in vivo* GATA-1 gene targets in primary murine adult stage erythroid cells by massive parallel sequencing. We were able to identify and initially characterize the novel roles of GATA-1 in the Myc/Max/Mad network and in oxidative stress cell response. Our aim is to identify novel targets in other hematopoietic lineages where GATA-1 is active, such as the eosinophils, mast cells, dendritic cells and compare them to the list of genes identified in the erythroid lineage. The aforementioned lineages though, differ in the amount of GATA-1 expressed compared to the erythroid cells, where GATA-1 is more abundantly expressed. In order to address this issue of low GATA-1 expression, we are planning to use the biotin GATA-1 knock-in mouse we have generated and take advantage of the high affinity reagents available for the biotin tags which allow high enrichments in ChIP experiments. qPCR experiments of streptavidin-ChIP samples from erythroid cells confirmed GATA-1 binding to a set of known GATA-1 gene targets and provided evidence for the suitability of streptavidin-ChIP for massive parallel sequencing applications. Moreover, streptavidin-ChIP resulted in reduced background and higher enrichment than the antibody-ChIP with two different ChIP-competent GATA-1 antibodies. To these ends, our GATA-1 ChIP-Sequencing data has established methods for massive parallel sequencing and bioinformatic analysis and will also provide a reference framework for genomics applications in the GATA-1 biotin tag knock-in mice.

Biotinylation tagging of GATA-1 in MEL cells provided evidence for a novel interaction of GATA-1 with the DNA maintenance methyltransferase DNMT1 (Papageorgiou, Amaral-Psarris and Strouboulis, unpublished). Moreover, DNMT1 was co-purified in another complex with biotin-tagged TR2/TR4 nuclear receptors, which bind to direct repeats in mouse and human embryonic and fetal β-type globin promoters and have been proposed to act as adult-stage specific repressors for these genes (Cui et al., 2011). The fact that DNMT1 was found to interact in two separate complexes with transcription factors implicated in erythropoiesis and globin gene regulation raised the prospect of DNMT1 being implicated in erythroid differentiation. In order to investigate a potential role for DNMT1 in erythropoiesis, we knocked-down DNMT1 in mouse fetal liver derived proerythroblasts and in mouse erythroleukemic (MEL) cells. Western blots and activity assays provided evidence for the efficiency of the DNMT1 knock-down, the normal expression of the DNMT3 methyltransferases, whereas FACS analysis confirmed no viral secondary effects. Three different assays were employed to quantitate cell differentiation in control and DNMT1 knock-down MEL cells. Preliminary expression profiling, cytospin preparations and FACS analysis in a set of at least three independent experiments showed a clear arrest in erythroid differentiation with defects in cell cycle regulation as a result of reduced levels of DNMT1.

Cytospins showed that DMSO induced DNMT1 knock-down cells did not fully differentiate and rather resembled polychromatic/ orthocromatic red cells. FACS analysis showed increased rates of proliferation and microarray analysis showed impaired repression of the cell cycle-related genes. Notably, all genes responsible for the erythroid transcription program were normally expressed. We

therefore assume that the DNMT1 knock-down impaired erythropoiesis phenotype results from cell cycle defects due to failed repression of the genes responsible for cell proliferation arrest. It remains presently unclear as to how DNMT1 is recruited to the cell cycle related genes. However, GATA-1 has been reported to regulate cell cycle related genes (Rylski et al., 2003) and thus it is possible that at least part of DNMT1's function in the cell cycle regulation may be mediated by its interaction with GATA-1.

Recruitment of DNMT1 to cell cycle genes will be explored by ChIP assays, using control and DNMT1 knock-down MEL cells and by examining targets identified by our microarray experiments. Moreover, we are planning to refine the list of the genes perturbed by performing RNA-Seq analysis so as to obtain a more quantitative perspective of the DNMT1 knock-down phenotype.

The indispensable role of DNMT1 during differentiation has been underscored by several studies. Dnmt1 null ES cells are restricted in their differentiation potential, whereas dnmt1 null mice are lethal (Lei et al., 1996). HSCs also require dnmt1 for their self-renewal program and in order to differentiate into myeloid or lymphoid cells. Hypomethylated dnmt1 partially rescued the HSC phenotype, resulting in increased expression of myeloerythroid genes versus decreased expression of 'stemness' and lymphoid genes (Bröske et al., 2009). Similar observations derive from the analysis by Trowbridge et al. (2009), which showed that DNMT1 null HSCs fail to generate mature myeloid cells. Moreover, when they used leukemic granulocyte-macrophage progenitors, they showed that DNA methylation-mediated silencing of bivalent chromatin domains is required for apoptosis, tumor suppression and the hematopoietic lineage restriction (Trowbridge et al, 2012). According to Spada et al. (2006) despite the fact that *de novo* methylation is sufficient to sustain undifferentiated cells, lineage commitment and survival of differentiated cells requires the function of DNMT1. Therefore for differentiation to occur, the maintenance of the methylation pattern by DNMT1 –previously established by the *de novo* methylases- is required. Somatic cells also require dnmt1 for their maintenance, since conditional homozygous dnmt1 null somatic cells die due to increased apoptosis at day 5 (Jackson-Grusby et al., 2001).

All previous findings support the hypothesis that Dnmt1 is required for normal differentiation to occur. Nevertheless, dnmt1 knock-down studies by Shearstone et al. (2011) in fetal liver cells showed that decreased levels of DNMT1 resulted in the acceleration of the erythroid specific gene induction suggesting that global demethylation may lead to the rapid removal of methylation marks at sites of massively induced erythroid genes.

We also observed normal activation of the erythroid genes, as reported by Shearstone et al. (2011) however we detected impaired erythropoiesis due to a block in the inhibition of the cell cycle genes that promote proliferation versus differentiation. We believe that the different protocols applied in the two studies account for the observed discrepancies. In the first study, fetal liver cells were cultured in differentiation medium for a single day prior to collecting them for expression profiling analysis. We followed a different experimental scheme whereby cells were isolated at day 0, expanded in proliferating medium for 5 days, transduced with lentivirus for 24hrs, cultured in selection medium and induced to differentiate for 3 days in order to obtain fully differentiated erythrocytes, as suggested by bibliography (Zhang et al., 2003). To these ends, we are planning to collect the cells at discrete time points to see whether the phenotype observed by Shearstone et al., appears to also be the case with MEL cells in the first couple of days of the assay.

Collectively, our data indicate that DNMT1 is dispensable for the induction of the genes of the erythroid program; however it is required for normal erythropoiesis to occur, by repressing the genes responsible for proliferation. It remains to investigate the molecular mechanism underlying the role of DNMT1 in erythropoiesis and to identify the protein partners it interacts with to exert its function.

4.3 Conclusions – Perspectives

The key interest of this thesis is the molecular basis of transcription factor functions in cellular differentiation. To address this question, this study focused on the function of a key hematopoietic transcription factor in hematopoiesis, namely, GATA-1.

GATA-1 is a master transcriptional regulator of erythroid and megakaryocytic cell differentiation and an essential transcription factor for the differentiation of several other hematopoietic lineages, including eosinophils, mast cells dendritic cells. The GATA-1 gene knockout results in embryonic lethality in mice due to lethal anemia as a result of failure of erythropoiesis. Due to the severe knockout phenotype, GATA-1 functions have been extensively studied in erythroid cells, albeit largely in *in vitro* cellular models (MEL cells, G1E cells).

Outstanding fundamental questions related to GATA-1 (and all other transcription factor) functions include a clear understanding of the molecular basis of its transcriptional regulatory functions and addressing the question of how a single transcription factor regulates different transcription programs in distinct hematopoietic lineages.

In this thesis, we addressed these questions by taking *in vivo* approaches. We have initially focused on mouse fetal liver erythropoiesis in order to describe the GATA-1 gene targets during erythroid differentiation *in vivo*. This resulted in expanding the role of GATA-1 in regulating the myc network during erythroid cell differentiation. We found that GATA-1 binds not only to the *myc* gene (Rylski et al., 2003), but also to the *max*, *mad1, mad3, mnt/rox* and *mad4* loci. Previous studies have shown the dual role of GATA-1 in driving differentiation and also regulating rapid synchronous cell cycle arrest and inactivation of cyclin-dependent kinases (Cdks) (Rylski et al., 2003). Expression analysis indicated that all *myc*, *mxd4*, *max* and *mnt/rox* are repressed during differentiation. By contrast, Mad1, which competes with Myc for heterodimerization with Max in order to promote differentiation versus proliferation, is 4-fold upregulated. Notably, binding of GATA-1 to the aforementioned loci, appeared to be more enriched in the proerythroblast stage versus the more differentiated cells. These observations reinforce the general role of GATA-1 in repressing cyclins and c-myc at the G1-to-S transition phase of the cell cycle, forcing cells to stop proliferating, exit the cell cycle and induce them to terminally differentiate (Rylski et al., 2003). Considering that the GATA-1s (short) isoform which lacks the first N-terminal 83 amino acids, has been implicated in myeloproliferative disorders and myeloid leukemias in trisomy 21 patients, it is tempting to speculate that the regulatory GATA-1 functions related to the myc network are somehow perturbed in these cases, thus leading to deregulated cell cycle control. It would be of interest to identify how protein interactions and/or gene targets differ between the full length and the short isoforms of GATA-1 *in vivo* (see below). In addition, since cell cycle arrest is a feature of terminal differentiation in general, a question arises as to whether GATA-1 plays similar roles in repressing the myc network in the other hematopoietic lineages where it functions. The use of the biotin tag knock-in mouse we have generated in this study will be useful in addressing this question and in uncovering the protein partners that GATA-1 interacts with in cell cycle control.

In addition, our work uncovered a previously unknown role for GATA-1 in regulating the oxidative stress response in erythroid cells. We provided evidence for GATA-1 binding not only to the ho-1 gene, the key player against the cytotoxic effects of oxidative stress, but also to its upstream regulators, bach1 and nrf2 and to a number of Nrf2-dependent and independent anti-oxidant genes, such as peroxiredoxins, including Prdx1, superoxide dismutase (SOD), glutathione peroxidase (GPx), glutathione reductase (GR). We have also provided evidence that in the absence of GATA-1, proerythroblastic cells are highly susceptible to reactive oxygen damage. This work suggests a novel role for GATA-1 as a positive regulator of the oxidative stress response genes in erythroid cells. In contrast to the regulation of the myc network, we believe that this cytoprotective role of GATA-1 is most likely erythroid-specific due to the increased cytoprotective needs of the erythroid cell due to hemoglobin synthesis, its exposure to oxygen and its role as scavenger of reactive species when circulating in tissues.

In order to further address GATA-1 *in vivo* functions in all hematopoietic lineages where GATA-1 has been genetically shown to fulfill essential roles, we generated a knock-in transgenic mouse that physiologically expresses biotin-tagged GATA-1, thus serving as a very good tool for the affinity purification of GATA-1 protein complexes and gene targets. We have used fetal liver erythropoiesis,

where GATA-1 is most abundantly expressed, in order to provide evidence that *in vivo* biotinylation tagging of physiologically expressed GATA-1 is functional. We have proceeded to purify GATA-1 protein complexes from immature and mature mouse fetal liver derived erythroid cells and to analyse them by mass spectrometry to identify potential novel GATA-1 interactors such as RUNX1, LBP1a or the SWI/SNF chromatin remodeling complex. The functional implications of these novel GATA-1 interactions in erythropoiesis are currently under investigation. In addition, we have preliminary evidence for quantitative changes in GATA-1 protein interactions with erythroid differentiation and we are in the process of identifying GATA-1 post-translational modifications *in vivo*. Further analysis of this mouse in the erythroid lineage will help identify the molecular basis of the novel GATA-1 functions described above, i.e. in the repression of the myc network or the upregulation oxidative stress response genes, for example, by identifying the GATA-1 protein partners in regulating GATA-1 gene target subsets and/or in identifying specific GATA-1 post-translationally modified *in vivo* isoforms that may distinguish between protein interaction partners of gene targets. It is of note that as a result of a collaboration with the lab of Dr. Paresh Vyas (Weatherall Institute for Molecular Medicine, Oxford), we now have knock-in mice physiologically expressing biotin tagged GATA-1s which can be used to address differences between the two GATA-1 isoforms in erythroid cells (and not only).

The generation of the biotin tag GATA-1 knock-in mouse by this work, also means that we can address the same line of questions regarding the molecular basis of GATA-1 functions in differentiation in all other hematopoietic lineages where it is expressed, i.e. megakaryocytes, eosinophils, mast cells and dendritic cells. We can look for overlapping or distinct protein interactions, post-translational modifications and gene target networks. For example, it would be of interest to see whether GATA-1 is involved in transcriptional repression in lineages where FOG-1 (which mediates the interaction of GATA-1 with the repressive NuRD complex) is not expressed, such as eosinophils or mast cells. It would also be of interest to investigate the GATA-1 interaction with RUNX1 which we describe for the first time in fetal liver erythropoiesis. RUNX1 is a master regulator of hematopoiesis, specifying the definitive hematopoietic stem cell (HSC) and, as such, is one of the earliest markers of hematopoiesis (Lancrin et al., 2009). Since RUNX1 is expressed in the early stages of the hematopoietic tree and prior to GATA-1 expression, we suspect that GATA-1 will be found in complex with RUNX1 in all other lineages where GATA-1 is expressed. Our hypothesis stands true at least for the megakaryocytic lineage, where it has been shown that GATA-1 *in vitro* targets RUNX-1 for phosphorylation by cyclin-dependent kinases and converts RUNX-1 from a repressor to an activator (reviewed by Elagib & Goldfarb, 2007; Goldfarb, 2009). By contrast, Lbp1a has been shown to play an indispensable role in globin transcription in erythroid cells only (Chae et al., 1999; Chae & Kim, 2003). Moreover, multiple genes of the erythroid lineage specifically contain binding sites for CP2 adjacent to GATA-1 binding motifs in their regulatory regions. Therefore, we suspect that there will be no association of GATA-1 with Lbp1a in non-erythroid lineages. In addition, SWI/SNF complexes which are involved in both activation and repression of transcription by selectively regulating chromatin remodeling through direct interaction with specific DNA-binding proteins, are abundantly expressed in all cells. We thus expect to find GATA-1 in complex with SWI/SNF members in the different lineages where GATA-1 is expressed and it would be of interest to elucidate how the same GATA-1/SWI/SNF protein complex may execute differential functions in distinct hematopoietic lineages.

A second line of evidence that GATA-1 epigenetically regulates cellular differentiation derives from experiments in mouse erythroleukemic (MEL) cells, where it was shown that GATA-1 interacts with the DNA maintenance methyltransferase DNMT1. In order to further investigate the role of DNMT1 in erythropoiesis, we knocked-down DNMT1 in mouse fetal liver derived proerythroblasts and in MEL cells. Preliminary expression profiling, cytospin preparations and FACS analysis in a set of at least three independent experiments showed a clear arrest in erythroid differentiation with defects in cell cycle regulation as a result of reduced levels of DNMT1.

Cytospins showed that DMSO induced DNMT1 knock-down cells did not fully differentiate and rather resembled polychromatic/orthocromatic red cells. FACS analysis showed increased rates of proliferation and microarray analysis showed impaired repression of cell cycle-related genes. Notably, all genes responsible for the erythroid transcription program were unaffected. We therefore assume that the DNMT1 knock-down impaired erythropoiesis phenotype results from cell cycle defects due to failed repression of the genes responsible for cell proliferation arrest. It remains presently unclear as to how DNMT1 is recruited to the cell cycle related genes. However, the genome wide analysis of the *in vivo* GATA-1 gene targets we conducted, in primary murine adult stage erythroid cells, by massive

parallel sequencing indicated GATA-1 binding in numerous loci of the Myc/Max/Mad pathway, which regulates cellular proliferation versus differentiation. It is therefore conceivable that the role of GATA-1 in cell cycle regulation through the repression of the Myc/Max/Mad pathway might account for the impaired erythropoiesis we observed in the DNMT1 knock-down studies we performed. Moreover, this regulation of the cell cycle genes might be mediated by the synergistic action of GATA-1 and DNMT1 and perhaps involves the physical association of GATA-1 with DNMT1 for binding to certain genes.

In summary, our work has provided novel insight as to GATA-1 *in vivo* functions in erythroid differentiation, for example in cell cycle regulation and oxidative stress response, has identified potential novel proteins partners and has provided a unique transgenic mouse model that will serve as an extremely useful tool in elucidating the basis for the differential GATA-1 functions in different hematopoietic lineages. Lastly, this work has also provided novel evidence as to the implication of the DNMT1 DNA methyltransferase in erythroid differentiation. These functions may in part be mediated through interactions with GATA-1. Thus, this work has gone some way in addressing the original questions of molecular basis of GATA-1 functions and in setting the stage for addressing in the long term the question of how does a single transcription factor, GATA-1, execute diverse functions in distinct hematopoietic lineages.

Acknowledgements

First of all, I'd like to thank my supervisor, Dr. John Strouboulis, for recruiting me, helping me out, advising me, teaching me and collaborating with me. The research I was involved in, during my PhD, has been a fascinating and stimulating journey, which I hope will carry on for the years to come. Working in the JS310 lab has been hard at times, enjoyable at other times however always felt like home to me. Thank you.

My acknowledgements to my committee, Prof. Achilleas Gravanis and Prof. George Mavrothalassitis, for their supervision, their useful suggestions and comments. I'd also like to thank the rest of the members of the committee Lecturer Ioannis Charalambopoulos, Prof. Dimitris Kardassis, Prof. Eleni Papadaki and Assosiate Prof. Christos Tsatsanis.

Many thanks to my best friend in the bench, Dimitris Papageorgiou, who has made this journey a great fun and with whom we shared so many. Also thanks to the rest of the JS310 members, for creating a great environment, Grigoris, Giorgos, Dimitra, Carla, Marina, Effie and Dimitra.

I'd like to say thank you to the secretarial office in the Medical School of the University of Crete, who have always been very accessible and helpful.

I wish to thank Prof. Paresh Vyas and Dr. Catherine Porcher and the people in their labs for hosting me and teaching me the basics behind molecular hematology during my 6-month stay in the WIMM, Oxford. Special thanks to Lin Zhang, Julia Draper, Mira Kassouf, Hedia Chagraoui and Gaëtan Juban.

I am honoured to have received the Fleming scholarship for the first 4 years of my PhD and I'd like to thank all the people at the Fleming Institute for being there: people at the reception desk, the Fleming services, the animal caretakers, the cleaning ladies, the I.T support and the technical and procurement office.

I'm also grateful and happy to have met some of my best friends in the Skoulakis lab -Maria, Kleio, Rania, Eirini, Alexandros, Maribel and all the rest- and in the Panayotou lab, Marina and Katerina.

There are so many others to thank inside and outside the Fleming Institute but I'll stop here and will finish by thanking my partner, Christos, for his patience, tolerance and emotional and moral support throughout the past couple of years and finally, my beloved parents, Vasilis and Verena and my brother Nikos for their love and support.

Thank you all!

APPENDIX

Table 5 (Real-Time PCR primer sequences):

Forward primer of ab1 of *Ho-1* -10kb	CGGAGCTGTGCCTTTTCTG
Reverse primer of ab1 of *Ho-1* -10kb	GAGCAACGGGAACCATGACT
Forward primer of *Ho-1* -7.9kb	GTTTTCGTTTTCTTGCTGCTTGT
Reverse primer of *Ho-1* -7.9kn	CTCTTATACTTCTATCCCACACAGAAAAGT
Forward primer of *Ho-1* -5.9kb	TTTTGACACAGGGTCTCTCTTTGTA
Reverse primer of *Ho-1* -5.9kb	AAGACAAAAAGATCCCAGCACTCT
Forward primer of sx2 of *Ho-1* -4kb	GGGTGACTCAGCAAAA TCTGTCTT
Reverse primer of sx2 of *Ho-1* -4kb	AGCTGCTTTTATGCTGTGTCATG
Forward primer of *Ho-1* -3.8kb	CACAGGAGCTGAACTTTGTTTTTC
Reverse primer of *Ho-1* -3.8kb	CAGCATAAAAGCAGCTGAGGAA
Forward primer of *Ho- 1* -2.8kb	TCTGCGGCTCCTAGAGAACAA
Reverse primer of *Ho-1* -2.8kb	CACCTCTCAGAGACACGCTTACTG
Forward primer of *Ho-1* +0.2kb	GAGTTACCGCCCAGTCTACCTGTA
Reverse primer of *Ho-1* +0.2kb	CCAGCGTGTCAGAGGAGTGA
Forward primer of *Ho-1* negative site	AATTTATTTTTGGTAGTGAGACGTGAAC
Reverse primer of *Ho-1* negative site	CCCTCAGCCCCACAGTAGAG
Forward primer of *Bach1* -0.8kb	TTCCGGATATGAAGCCCAGTA
Reverse primer of *Bach1* -0.8kb	GCCACGACGGACTGCATT
Forward primer of *Bach1* +1.2kb	GCGGGCTGTGTGTTTTGAT
Reverse primer of *Bach1* +1.2kb	CCTTGGCCACGTCTTAAGCA
Forward primer of *Bach1* +7.5kb	CCTCCCAGGGAGGCAAGA
Reverse primer for *Bach1* +7.5kb	GGGTATCTCTCCTTCCAGAAAACTT
Forward primer of *Bach1* negative site	TCCCCAAACGCACACAAAG
Reverse primer of *Bach1* negative site	CCAGCTCCAGCAGCTGTGTA
Forward primer of *Nfe2l2* -0.6kb	AGCTGCCACTCCCTGATTTG
Reverse primer of *Nfe2l2* -0.6kb	GCCGGTGGAGCGAGATAA
Forward primer of *Nfe2l2* negative site	GGCTCCCAACCCCAAATCT
Reverse primer of *Nfe2l2* negative site	TGAGGGTGTGTTGTACAGTGATCA
Forward primer of *Myc* -1.6kb	TGTGCAATGAGCTCGATGAAG
Reverse primer of *Myc* -1.6kb	TTGTTATTAGCCACTGAGTTTGCAA
Forward primer of *Myc* -0.4kb	TTCCCCGCGAGATGGA
Reverse primer of *Myc* -0.4kb	CAGCAACTCACTGCCACGTATAC
Forward primer of *Myc* +2kb	TCCTTCAGGTGGCGCAAA
Reverse primer of *Myc* +2kb	GGAGGCGACTGTAGGGAATACA
Forward primer of *Myc* +3.3kb	CCTTTACCCCTCCTAAGCATTTTAA
Reverse primer of *Myc* +3.3kb	CCCCTCCAGATCAGTTCCTTT
Forward primer of *Myc* negative site	ACACACACATACGAAGGCA
Reverse primer of *Myc* negative site	ACCGTTAACCCCTTCCTCCC
Forward primer of *Max* TSS	GGCCCTCT GTTATCTGGGTTT
Reverse primer of *Max* TSS	CGGCGAAATCGGTCCTT
Forward primer of *Max* +20.5kb	CTACCGTCTTTGGGAATATTACACTTT
Reverse primer of *Max* +20.5kb	TGATGTCTTGAGCAGATCACTTTG
Forward primer of *Max* negative site	GGTGATTCAGAACCCGAACAG
Reverse primer of *Max* negative site	TCCCCCTTTCTGGGTTTCTT
Forward primer of *Mad* +11.2kb	CCTGCCAGGCCCATCTG
Reverse primer of *Mad* +11.2kb	TTTCTGGCTCCTGCTCTCTGA
Forward primer of *Mad* negative site	TGAGGGCACGGAACACAGA
Reverse primer of *Mad* negative site	TCCAAAGGCAGCCTTGCT
Forward primer of *Mxd4* -0.3kb	CGGCCATCGTGCATGTTA
Reverse primer of *Mxd4* -0.3kb	GGACTCAATTTACATAGACCGTTCTG
Forward primer of *Mxd4* negative site	CCAGGCTACATGGGATCCAT
Reverse primer of *Mxd4* negative site	CCCCTTGGCTGAATTTTGTG
Forward primer of *Nrf2* cDNA	CTCAGCATGATGGACTTGGA
Reverse primer of *Nrf2* cDNA	CATGTCCTGCTGGGACTGTA
Forward primer of *Bach1* cDNA	TCCGGTCGATGACAGTGAG
Reverse primer of *Bach1* cDNA	CAACTGGGACCAGGGAGAG
Forward primer of 2.95kb GATA-1 fragment	TGAAAGAAACTAGTAGTAAAACATG
Reverse primer of 2.95kb GATA-1 fragment	TATGTATTTCTTCTAGACTATGTAT

Forward primer of 2.1kb GATA-1 fragment	ATACATAGTCTAGAAGAAATACATA
Reverse primer of 2.1kb GATA-1 fragment	AGTGCTGATGCTGCGGCCGTGGCTG
Forward primer of 2.3kb GATA-1 fragment	CAGCCACGGCCGCAGCATCAGCACT
Reverse primer of 2.3kb GATA-1 fragment	ATCACCGCGGGACCACCACAGGGGTTACCAGCTC
Forward primer of *PU.1* promoter	CCAGGGCTGCCCTTTGA
Reverse primer of *PU.1* promoter	TTGCATAAATCTCTTGCGCTACA
Forward primer of *PU.1* negative site	GCATCTGGTGGGTGGACAAG
Reverse primer of *PU.1* negative site	GCGCGCCATCTTCTGGTA
Forward primer of *gapdh*	CCAATGTGTCCCGTCGTGGATCT
Reverse primer of *gapdh*	GTTGAAGTCGCAGGAGACAAC
Forward primer of *gata-2* -39kb	GAGATGAGCTAATCCCGCTGTA
Reverse primer of *gata-2* -3.9kb	AAGGCTTGATTTTTCCAGGCC
Forward primer of *gata-2* negative site	TCCATCCAGCAGCTTTAGGAA
Reverse primer of *gata-2* negative site	GGGTTCGAAGCCACTCCAA
Forward primer of *gata-1* IE	ACTCTGGGTGTCACCTCAGTTTC
Reverse primer of *gata-1* IE	ACAGGGACAGTGCCTCAACTTC
Forward primer of *gata-1* negative site	CACTAGCAGCTGGGTGGGTTA
Reverse primer of *gata-1* negative site	TGCCGCTTGCCTTTGTAAG
Forward primer of *klf* promoter	TATCGCACACACCCCTCCTT
Reverse primer of *klf* promoter	CCCACATCTGATTGGCTGTCT
Forward primer of *klf* negative site	TGCTCCCCACTATGATAATGGA
Reverse primer of *klf* negative site	GCCACAACCAAAGAAGACATTTT

REFERENCES

Acosta, J. C., N. Ferrándiz, et al. (2008). Myc inhibits p27-induced erythroid differentiation of leukemia cells by repressing erythroid master genes without reversing p27-mediated cell cycle arrest. *Mol Cell Biol* 28(24): 7286-95.

Alam, J., E. Killeen, et al. (2003). Heme activates the heme oxygenase-1 gene in renal epithelial cells by stabilizing Nrf2. *Am J Physiol Renal Physiol* 284(4): F743-52.

Allen, T. D. and T. M. Dexter. (1982). Ultrastructural aspects of erythropoietic differentiation in long-term bone marrow culture. *Differentiation* 21(2): 86-94.

Alves, L. R., E. S. Costa, et al. (2011). Heme-oxygenases during erythropoiesis in K562 and human bone marrow cells. *PLoS One* 6(7): e21358.

Anderson, K. P., S. C. Crable, et al. (1998). Multiple proteins binding to a GATA-E box-GATA motif regulate the erythroid Kruppel-like factor (EKLF) gene. *J Biol Chem* 273(23): 14347-14354.

Antoniou, M. (1991). Induction of Erythroid-Specific Expression in Murine Erythroleukemia (MEL) Cell Lines. *Methods Mol Biol* 7: 421-34.

Arand, J., D. Spieler, et al. (2012). In Vivo Control of CpG and Non-CpG DNA Methylation by DNA Methyltransferases. *PLoS Genet* 8(6): e1002750.

Armstrong, J. A., J. J. Bieker, et al. (1998). A SWI/SNF-related chromatin remodeling complex, E-RC1, is required for tissue-specific transcriptional regulation by EKLF in vitro. *Cell* 95(1): 93–104.

Ashkenazi, A. and V. M. Dixit. (1998). Death receptors: Signaling and modulation. *Science* 281(5381): 1305-1308.

Baron, M. H. and S. M. Farrington. (1994). Positive regulators of the lineage-specific transcription factor GATA-1 in differentiating erythroid cells. *Mol Cell Biol* 14(5): 3108-14.

Baron, M. H., J. Isern, et al. (2012). The embryonic origins of erythropoiesis in mammals. *Blood* 119(21): 4828-37.

Barra, V., T. Schillaci, et al. (2012). Bypass of cell cycle arrest induced by transient DNMT1 post-transcriptional silencing triggers aneuploidy in human cells. *Cell Div* 7(1): 2.

Bessis, M. (1958). Erythroblastic island, functional unity of bone marrow. *Rev Hematol* 13(1): 8-11.

Bessis, M., C. Mize, et al. (1978). Erythropoiesis: comparison of in vivo and in vitro amplification. *Blood Cells* 4(1-2): 155-74.

Bieker, J. J. (2005). Probing the onset and regulation of erythroid cell-specific gene expression. *Mt Sinai J Med* 72(5): 333-8.

Bose, F., C. Fugazza, et al. (2006). Functional interaction of CP2 with GATA-1 in the regulation of erythroid promoters. *Mol Cell Biol* 26(10): 3942-54.

Boyes, J., P. Byfield, et al. (1998). Regulation of activity of the transcription factor GATA-1 by acetylation. *Nature* 396(6711): 594-8.

Bresnick, E. H., H. Y. Lee, et al. (2010). GATA switches as developmental drivers. *J Biol Chem* 285(41): 31087-93.

Bröske, A. M., L. Vockentanz, et al. (2009). DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction. *Nat Genet* 41(11): 1207-15.

Bruchova, H., D. Yoon, et al. (2007). Regulated expression of microRNAs in normal and polycythemia vera erythropoiesis. *Exp Hematol* 35(11): 1657-67.

Bultman, S. J., T. C. Gebuhr, et al. (2005). A Brg1 mutation that uncouples ATPase activity from chromatin remodeling reveals an essential role for SWI/SNF-related complexes in beta-globin expression and erythroid development. *Genes Dev* 19(23): 2849–2861.

Calligaris, R., S. Bottardi, et al. (1995). Alternative translation initiation site usage results in two functionally distinct forms of the GATA-1 transcription factor. *Proc Natl Acad Sci U S A* 92(25): 11598-11602.

Cantor, A. B. and S. H. Orkin. (2002). Transcriptional regulation of erythropoiesis: an affair involving multiple partners. *Oncogene* 21(21): 3368-3376.

Cantor, A. B., H. Iwasaki, et al. (2008). Antagonism of FOG-1 and GATA factors in fate choice for the mast cell lineage. *J Exp Med* 205(3): 611-24.

Carey, M. F., C. L. Peterson, et al. (2009). Chromatin immunoprecipitation (ChIP). *Cold Spring Harb Protoc* (9):pdb.prot5279.

Carotta, S., S. Pilat, et al. (2004). Directed differentiation and mass cultivation of pure erythroid progenitors from mouse embryonic stem cells. *Blood* 104(6): 1873-80.

Chae, J. H., Y. H. Lee, et al. (1999). Transcription factor CP2 is crucial in hemoglobin synthesis during erythroid terminal differentiation in vitro. *Biochem Biophys Res Commun* 263(2): 580-3.

Chae, J. H. and C. G. Kim (2003). CP2 binding to the promoter is essential for the enhanced transcription of globin genes in erythroid cells. *Mol Cells* 15(1): 40-7.

Chang, A. N., A. B. Cantor, et al. (2002). GATA-factor dependence of the multitype zinc-finger protein FOG-1 for its essential role in megakaryopoiesis. *Proc Natl Acad Sci U S A* 99(14): 9237-42.

Chang, T. C. and J. T. Mendell. (2007). microRNAs in vertebrate physiology and human disease. *Annu Rev Genomics Hum Genet* 8: 215-39.

Chateauvieux, S., C. Grigorakaki, et al. (2011). Erythropoietin, erythropoiesis and beyond. *Biochem Pharmacol* 82(10): 1291-303.

Cheng, Y., W. Wu, et al. (2009). Erythroid GATA1 function revealed by genome-wide analysis of transcription factor occupancy, histone modifications, and mRNA expression. *Genome Res* 19(12): 2172-84.

Cheung, P. and P. Lau (2005). Epigenetic regulation by histone methylation and histone variants. *Mol Endocrinol* 19(3): 563-73.

Chiba, T., Y. Nagata, et al. (1993). Induction of erythroid-specific gene expression in lymphoid cells. *Proc Natl Acad Sci U S A* 90(24): 11593-11597.

Cianetti, L., M. Gabbianelli, et al. (2010). Ferroportin and Erythroid Cells: An Update. *Adv Hematol* pii: 404173.

Ciovacco, W. A., W. H. Raskind, et al. (2008). Human phenotypes associated with GATA-1 mutations. *Gene* 427(1-2): 1-6.

Clarke, A., E. R. Maandag, et al. (1992). Requirement for a functional Rb-1 gene in murine development. *Nature* 359(6393): 328-30.

Clark, A., K. Doyle, et al. (2004). Cell-intrinsic requirement for pRb in erythropoiesis. *Blood* 104(5): 1324-6.

Clements, E. G., H. P. Mohammad, et al. (2012). DNMT1 modulates gene expression without its catalytic activity partially through its interactions with histone-modifying enzymes. *Nucleic Acids Res* 40(10): 4334-46.

Cohen-Kaminsky, S., L. Maouche-Chretien, et al. (1998). Chromatin immunoselection defines a TAL-1 target gene. *Embo J* 17(17): 5151–5160.

Collavin, L., M. Gostissa, et al. (2004). Modification of the erythroid transcription factor GATA-1 by SUMO-1. *Proc Natl Acad Sci U S A* 101(24): 8870-5.

Costa, G., V. Kouskoff et al. (2012). Origin of blood cells and HSC production in the embryo. *Trends Immunol* 33(5): 215-23.

Crispino, J. D., M. B. Lodish, et al. (1999). Use of altered specificity mutants to probe a specific protein-protein interaction in differentiation: the GATA-1:FOG complex. *Mol Cell* 3(2): 219-228.

Crossley, M. and S. H. Orkin. (1994). Phosphorylation of the erythroid transcription factor GATA-1. *J Biol Chem* 269(24): 16589-96.

Crossley, M., M. Merika, et al. (1995). Self-association of the erythroid transcription factor GATA-1 mediated by its zinc finger domains. *Mol Cell Biol* 15(5): 2448-2456.

Cui, S., K. E. Kolodziej, et al. (2011). Nuclear Receptors TR2 and TR4 Recruit Multiple Epigenetic Transcriptional Corepressors That Associate Specifically with the Embryonic β-Type Globin Promoters in Differentiated Adult Erythroid Cells. *Mol Cell Biol* 31(16): 3298-3311.

Cull, M. G. and P. J. Schatz (2000). Biotinylation of proteins in vivo and in vitro using small peptide tags. *Methods Enzymol* 326: 430-40.

de Boer, E., P. Rodriguez, et al. (2003). Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. *Proc Natl Acad Sci U S A* 100(13): 7480-5.

Dent, L. A., M. Strath, et al. (1990). Eosinophilia in transgenic mice expressing interleukin 5. *J Exp Med* 172(5): 1425-31.de Thonel, A., J. Vandekerckhove, et al. (2010). HSP27 controls GATA-1 protein level during erythroid cell differentiation. *Blood* 116(1):85-96.

de Thonel, A., J. Vandekerckhove, et al. (2010). HSP27 controls GATA-1 protein level during erythroid cell differentiation. *Blood* 116(1): 85-96.

Dhakshinamoorthy, S., A. K. Jain, et al. (2005). Bach1 competes with Nrf2 leading to negative regulation of the antioxidant response element (ARE)-mediated NAD(P)H:quinoneoxidoreductase 1 gene expression and induction in response to antioxidants. *J Biol Chem* 280(17): 16891-900.

Doan, L. L., S. D. Porter, et al. (2004). Targeted transcriptional repression of Gfi1 by GFI1 and GFI1B in lymphoid cells. *Nucleic Acids Res* 32(8): 2508-19.

Dore, L. C., J. D. Amigo, et al. (2008). A GATA-1-regulated microRNA locus essential for erythropoiesis. *Proc Natl Acad Sci U S A* 105(9): 3333-8.

Doré, L. C., T. M. Chlon, et al. (2012). Chromatin occupancy analysis reveals genome-wide GATA factor switching during hematopoiesis. *Blood* 119(16): 3724-33.

Driegen, S., R. Ferreira, et al. (2005). A generic tool for biotinylation of tagged proteins in transgenic mice. *Transgenic Res* 14(4): 477-82.

Drissen, R., M. von Lindern, et al. (2005). The erythroid phenotype of EKLF-null mice: defects in hemoglobin metabolism and membrane stability. *Mol Cell Biol* 25(12): 5205-14.

Du, J., M. J. Stankiewicz, et al. (2002). Novel combinatorial interactions of GATA-1, PU.1, and C/EBPepsilon isoforms regulate transcription of the gene encoding eosinophil granule major basic protein. *J Biol Chem* 277(45): 43481-94.

Du, W. and J. S. Searle. (2009). The rb pathway and cancer therapeutics. *Curr Drug Targets* 10(7): 581-9.

Duan, Z. and M. Horwitz. (2003). Gfi-1 oncoproteins in hematopoiesis. *Hematology* 8(5): 339-44.

Eilken, H. M., S. Nishikawa, et al. (2009). Continuous single-cell imaging of blood generation from haemogenic endothelium. *Nature* 457(7231): 896-900.

Elagib, K. E., F. K. Racke, et al. (2003). RUNX1 and GATA-1 coexpression and cooperation in megakaryocytic differentiation. *Blood* 101(11): 4333-41.

Elagib, K. E., A. N. Goldfarb (2007). Regulation of RUNX1 transcriptional function by GATA-1. *Crit Rev Eukaryot Gene Expr* 17(4): 271-80.

Espada, J., E. Ballestar, et al. (2004). Human DNA methyltransferase 1 is required for maintenance of the histone H3 modification pattern. *J Biol Chem* 279(35): 37175-84.

Evans, T, M. Reitman, et al. (1988). An erythrocyte-specific DNA-binding factor recognizes a regulatory sequence common to all chicken globin genes. *Proc Natl Acad Sci U S A* 85(16): 5976-5980.

Fatemi, M., A. Hermann, et al. (2001). The activity of the murine DNA methyltransferase Dnmt1 is controlled by interaction of the catalytic domain with the N-terminal part of the enzyme leading to an allosteric activation of the enzyme after binding to methylated DNA. *J Mol Biol* 309(5): 1189-99.

Feng, W., C. Southwood, et al. (1994). Analyses of beta-thalassemia mutant DNA interactions with erythroid Kruppel-like factor (EKLF), an erythroid cell-specific transcription factor. *J Biol Chem* 269(2): 1493-500.

Ferkowicz, M. J. and M. C. Yoder. (2005). Blood island formation: longstanding observations and modern interpretations. *Exp Hematol* 33(9): 1041-7.

Ferreira, R., K. Ohneda, et al. (2005). GATA1 function, a paradigm for transcription factors in hematopoiesis. *Mol Cell Biol* 25(4): 1215-27.

Fox, A. H., K. Kowalski, et al. (1998). Key residues characteristic of GATA N-fingers are recognized by FOG. *J Biol Chem* 273(50): 33595-33603.

Fox, A. H., C. Liew, et al. (1999). Transcriptional cofactors of the FOG family interact with GATA proteins by means of multiple zinc fingers. *Embo J* 18(10): 2812-2822.

Fraser, S. T., J. Isern, et al. (2007). Maturation and enucleation of primitive erythroblasts during mouse embryogenesis is accompanied by changes in cell-surface antigen expression. *Blood* 109(1): 343-52.

Fujiwara, Y., C. P. Browne, et al. (1996). Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. *Proc Natl Acad Sci U S A* 93(22): 12355-12358.

Fujiwara, T., H. O'Geen, et al. (2009). Discovering hematopoietic mechanisms through genome-wide analysis of GATA factor chromatin occupancy. *Mol Cell* 36(4): 667-81.

Fuks, F., W. A. Burgers, et al. (2000). DNA methyltransferase Dnmt1 associates with histone deacetylase activity. *Nat Genet* 24(1): 88-91.

Fuks, F., W. A. Burgers, et al. (2001). Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. *Embo J* 20(10): 2536-44.

Goardon, N., J. A. Lambert, et al. (2006). ETO2 coordinates cellular proliferation and differentiation during erythropoiesis. *Embo J* 25(2): 357–366.

Goldfarb A. N. (2009). Megakaryocytic programming by a transcriptional regulatory loop: A circle connecting RUNX1, GATA-1, and P-TEFb. *J Cell Biochem* 107(3): 377-82.

Gordon, G. M. and W. Du. (2011). Conserved RB functions in development and tumor suppression. *Protein Cell* 2(11): 864-78.

Gozzelino, R. V. Jeney, et al. (2010). Mechanisms of Cell Protection by Heme Oxygenase-1. *Annu Rev Pharmacol Toxicol* 50: 323-354.

Graf, T. (2002). Differentiation plasticity of hematopoietic cells. *Blood* 99(9): 3089-3101.

Gregory, R. C., D. J. Taxman, et al. (1996). Functional interaction of GATA1 with erythroid Kruppel-like factor and Sp1 at defined erythroid promoters. *Blood* 87(5): 1793-1801.

Gregory, T., C. Yu, et al. (1999). GATA-1 and erythropoietin cooperate to promote erythroid cell survival by regulating bcl-xL expression. *Blood* 94(1): 87-96.

Grinberg, A. V., C. D. Hu, et al. (2004). Visualization of Myc/Max/Mad family dimers and the competition for dimerization in living cells. *Mol Cell Biol* 24(10): 4294-308.

Gutierrez, L, F. Lindeboom, et al. (2004). Homotypic signalling regulates Gata1 activity in the erythroblastic island. *Development* 131(13): 3183-3193.

Gutiérrez, L., F. Lindeboom, et al. (2005). A hanging drop culture method to study terminal erythroid differentiation. *Erythropoiesis* 33(10): 1083-1091.

Hamlett, I., J. Draper, et al. (2008). Characterisation of megakaryocyte GATA1-interacting proteins: the co-repressor ETO2 and GATA1 interact to regulate terminal megakaryocyte maturation. *Blood* 112(7): 2738-49.

Hernandez-Hernandez, A., P. Ray, et al. (2006). Acetylation and MAPK phosphorylation cooperate to regulate the degradation of active GATA-1. *Embo J* 25(14):3264-74.

Heyworth, C., S. Pearson, et al. (2002). Transcription factor-mediated lineage switching reveals plasticity in primary committed progenitor cells. *Embo J* 21(14): 3770-81.

Hirasawa, R., R. Shimizu, et al. (2002). Essential and instructive roles of GATA factors in eosinophil development. *J Exp Med* 195(11): 1379-86.

Huang, D. Y., Y. Y. Kuo, et al. (2005). GATA-1 mediates auto-regulation of Gfi-1B transcription in K562 cells. *Nucleic Acids Res* 33(16): 5331-42.

Huang, da W., B. T. Sherman, et al. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4(1): 44-57.

Hung, H. L., J. Lau, et al. (1999). CREB-Binding protein acetylates hematopoietic transcription factor GATA-1 at functionally important sites. *Mol Cell Biol* 19(5): 3496-505.

Im, H., J. A. Grass, et al. (2005). Chromatin domain activation via GATA-1 utilization of a small subset of dispersed GATA motifs within a broad chromosomal region. *ProcNatlAcadSci U S A* 102(47): 17065-70.

Isern, J., S. T. Fraser, et al. (2008). The fetal liver is a niche for maturation of primitive erythroid cells. *Proc Natl Acad Sci U S A* 105(18): 6662-7.

Ishii, T., K. Itoh, et al. (2000). Transcription factor Nrf2 coordinately regulates a group of oxidative stress-inducible genes in macrophages. *J Biol Chem* 275(21): 16023-9.

Itoh, K., N. Wakabayashi, et al. (2003). Keap1 regulates both cytoplasmic-nuclear shuttling and degradation of Nrf2 in response to electrophiles. *Genes Cells* 8(4): 379-91.

Iwasaki, H., S. Mizuno, et al. (2003). GATA-1 converts lymphoid and myelomonocytic progenitors into the megakaryocyte/erythrocyte lineages. *Immunity* 19(3): 451-62.

Jackson-Grusby, L., C. Beard, et al. (2001). Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. *Nat Genet* 27(1): 31–9.

Jane, S. M., A. W. Nienhuis, et al. (1995). Hemoglobin switching in man and chicken is mediated by a heteromeric complex between the ubiquitous transcription factor CP2 and a developmentally specific protein. *Embo J* 14(1): 97-105.

Kadam, S., G. S. McAlpine, et al. (2000). Functional selectivity of recombinant mammalian SWI/SNF subunits. *Genes Dev* 14(19): 2441–2451.

Kadam, S. and B. M. Emerson (2003). Transcriptional specificity of human SWI/SNF BRG1 and BRM chromatin remodeling complexes. *Mol Cell* 11(2): 377-89.

Kadri, Z., R. Shimizu, et al. (2009). Direct binding of pRb/E2F-2 to GATA-1 regulates maturation and terminal cell division during erythropoiesis. *PLoS Biol* 7(6): e1000123.

Kassouf, M. T., J. R. Hughes, et al. (2010). Genome-wide identification of TAL1's functional targets: insights into its mechanisms of action in primary erythroid cells. *Genome Res* 20(8):1064-83.

Katsantoni, E. Z., N. E. Anghelescu, et al. (2007). Ubiquitous expression of the rtTA2S-M2 inducible system in transgenic mice driven by the human hnRNPA2B1/CBX3 CpG island. *BMC Dev Biol* 27(7): 108.

Keller, G., M. Kennedy, et al. (1993). Hematopoietic commitment during embryonic stem cell differentiation in culture. *Mol Cell Biol* 13(1): 473-86.

Kerenyi, M. A. and S. H. Orkin (2010). Networking erythropoiesis. *JEM* 207(12): 2537-2541.

Kim, S. I. and E. H. Bresnick. (2007). Transcriptional control of erythropoiesis: emerging mechanisms and principles. *Oncogene* 26(47): 6777-94.

Kim, J., J. Chu, et al. (2008). An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* 132(6): 1049-61.

Kingsley, P. D., J. Malik, et al. (2006). "Maturational" globin switching in primary primitive erythroid cells. *Blood* 107(4): 1665-72.

Kitajima, K., M. Tanaka, et al. (2003). In vitro differentiation of mouse embryonic stem cells to hematopoietic cells on an OP9 stromal cell monolayer. *Methods Enzymol* 365: 72-83.

Kobayashi, A., M. I. Kang, et al. (2004). Oxidative stress sensor Keap1 functions as an adaptor for Cul3-based E3 ligase to regulate proteasomal degradation of Nrf2. *Mol Cell Biol* 24(16): 7130-9.

Kolodziej, K. E., F. Pourfarzad, et al. (2009). Optimal use of tandem biotin and V5 tags in ChIP assays. *BMC Mol Biol* 10(6).

Komatsu, N., K. Kirito, et al. (1997). Cell-cycle-dependent regulation of erythropoietin receptor gene. *Blood* 89(4): 1182-8.

Kouzarides, T. (2007). Chromatin Modifications and Their Function. *Cell* 128(4): 693–705.

Kulessa, H., J. Frampton, et al. (1995). GATA-1 reprograms avian myelomonocytic cell lines into eosinophils, thromboblasts, and erythroblasts. *Genes Dev* 9(10): 1250-62.

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

Lam, K. and D. E. Zhang. (2012). RUNX1 and RUNX1-ETO: roles in hematopoiesis and leukemogenesis. *Front Biosci* 17: 1120-39.

Lamonica, J. M., C. R. Vakoc, et al. (2006). Acetylation of GATA-1 is required for chromatin occupancy. *Blood* 108(12): 3736-8.

Lancrin, C., P. Sroczynska, et al. (2009). The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage. *Nature* 457(7231): 892-5.

Lavery, D. J. and U. Schibler (1993). Circadian transcription of the cholesterol 7 α hydroxylase gene may involve the liver-enriched bZIP protein DBP. *Genes & Dev* 7: 1871–1884.

Lavrovsky, Y., M. L. Schwartzman, et al. (1994). Identification of binding sites for transcription factors NF-kappa B and AP-2 in the promoter region of the human hemeoxygenase 1 gene. *Proc Natl Acad Sci U S A* 91(13): 5987-91.

Lecointe, N., O. Bernard, et al. (1994). GATA-and SP1-binding sites are required for the full activity of the tissue-specific promoter of the tal-1 gene. *Oncogene* 9(9): 2623-2632.

Lee, H. Y., K. D. Johnson, et al. (2009). Controlling hematopoiesis through sumoylation-dependent regulation of a GATA factor. *Mol Cell* 36(6): 984-95.

Lei, H., S. Oh, et al. (1996). De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* 122(10): 3195–205.

Lemischka, I. R., D. H. Raulet, et al. (1986). Developmental potential and dynamic behavior of hematopoietic stem cells. *Cell* 45(6): 917-27.

Letting, D. L., Y. Y. Chen, et al. (2004). Context-dependent regulation of GATA-1 by friend of GATA-1. *Proc Natl Acad Sci U S A* 101(2): 476-81.

Liew, C. W., K. D. Rand, et al. (2006). Molecular analysis of the interaction between the hematopoietic master transcription factors GATA-1 and PU.1. *J Biol Chem* 281(38): 28296-306.

Lin, S. C., S. X. Skapek, et al. (1996). Genes in the RB pathway and their knockout in mice. Semin *Cancer Biol* 7(5): 279-89.

Litt, M. D., M. Simpson, et al. (2001). Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. *Embo J* 20(9): 2224–2235.

Lowry, J. A. and J. P. Mackay. (2006). GATA-1: one protein, many partners. *Int J Biochem Cell Biol* 38(1): 6-11.

Lu, S. J., Q. Feng, et al. (2008). Biologic properties and enucleation of red blood cells from human embryonic stem cells. *Blood* 112(12): 4475-84.

Lüscher, B. (2001). Function and regulation of the transcription factors of the Myc/Max/Mad network. *Gene* 277(1-2): 1-14.

Lüscher, B. and J. Vervoorts (2012). Regulation of gene transcription by the oncoprotein MYC. *Gene* 494(2): 145-60.

Mackay, J. P., K. Kowalski, et al. (1998). Involvement of the N-finger in the self-association of GATA-1. *J Biol Chem* 273(46): 30560-30567.

MacLeod, A. K., M. McMahon, et al. (2009). Characterization of the cancer chemopreventive NRF2-dependent gene battery in human keratinocytes: demonstration that the KEAP1-NRF2 pathway, and not the BACH1-NRF2 pathway, controls cytoprotection against electrophiles as well as redox-cycling compounds. *Carcinogenesis* 30(9): 1571-80.

Mantalaris, A., P. Keng, et al. (1998). Engineering a human bone marrow model: a case study on ex vivo erythropoiesis. *Biotechnol Prog* 14(1): 126-33.

Manwani, D. and J. J. Bieker. (2008). The erythroblastic island. *Curr Top Dev Biol* 82: 23-53.

Margot, J. B., A. M. Aguirre-Arteta, et al. (2000). Structure and function of the mouse DNA methyltransferase gene: Dnmt1 shows a tripartite structure 1. *J Mol Biol* 297(2): 293-300.

Martin, D. I. and S. H. Orkin. (1990). Transcriptional activation and DNA binding by the erythroid factor GF-1/NFE1/Eryf 1. *Genes Dev* 4(11): 1886-1898.

Martowicz, M. L. , J. A. Grass, et al. (2005). Dynamic GATA factor interplay at a multicomponent regulatory region of the GATA-2 locus. *J BiolChem* 280(3): 1724-32.

Masaki, S., R. Ohtsuka, et al. (2007). Expression patterns of microRNAs 155 and 451 during normal human erythropoiesis. *Biochem Biophys Res Commun* 364(3): 509-14.

McDevitt, M. A., R. A. Shivdasani, et al. (1997). A “knockdown” mutation created by cis-element gene targeting reveals the dependence of erythroid cell maturation on the level of transcription factor GATA-1. *Proc Natl Acad Sci U S A* 94(13): 6781-6785.

McGrath K. E., J. M. Frame, et al. (2011). A transient definitive erythroid lineage with unique regulation of the β-globin locus in the mammalian embryo. *Blood* 117(17): 4600-8.

McKercher, S., B. E. Torbett, et al. (1996). Targeted disruption of the PU.1 gene results in multiple hematopoietic abnormalities. *Embo J* 15(20): 5647-58.

McNagny, K. and T. Graf (2002). Making eosinophils through subtle shifts in transcription factor expression. *J Exp Med* 195(11): F43-7.

Mead, P. E., A. E. Deconinck, et al. (2001). Primitive erythropoiesis in the Xenopus embryo: the synergistic role of LMO-2, SCL and GATA-binding proteins. *Development* 128(12): 2301–2308.

Meier, N., S. Krpic, et al. (2006). Novel binding partners of Ldb1 are required for haematopoietic development. *Development* 133(24): 4913–4923.

Mense, S. M. and L. Zhang (2006). Heme: a versatile signaling molecule controlling the activities of diverse regulators ranging from transcription factors to MAP kinases. *Cell Res* 16(8): 681-92.

Merika, M. and S. Orkin. (1995). Functional synergy and physical interactions of the erythroid transcription factor GATA-1 with the Kruppel family proteins Sp1 and EKLF. *Mol Cell Biol* 15(5): 2437-47.

Milutinovic, S., S. E. Brown, et al. (2004). DNA methyltransferase 1 knock down induces gene expression by a mechanism independent of DNA methylation and histone deacetylation. *J Biol Chem* 279(27): 27915-27.

Miyazaki, T., Y. Kirino, et al. (2010). Expression of heme oxygenase-1 in human leukemic cells and its regulation by transcriptional repressor Bach1. *Cancer Sci* 101(6): 1409-16.

Miyoshi, H., K. Shimizu, et al. (1991). t(8;21) breakpoints on chromosome 21 in acute myeloid leukemia are clustered within a limited region of a single gene, AML1. *Proc Natl Acad Sci U S A* 88(23): 10431-4.

Moffat, J., D. A. Grueneberg, et al. (2006). A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* 124(6): 1283-98.

Morceau, F., M. Schnekenburger, et al. (2004). GATA-1: friends, brothers, and coworkers. *Ann N Y Acad Sci* 1030: 537-54.

Motoyama, N., T. Kimura, et al., (1999). bcl-x prevents apoptotic cell death of both primitive and definitive erythrocytes at the end of maturation. *J Exp Med* 189(11): 1691-8.

Mundschau, G. and J. Crispino. (2006). GATA1s goes germline. *Nat Genet* 38(7): 741-2.

Nagata, S. (1997). Apoptosis by death factor. *Cell* 88(3): 355-65.

Nakajima, H. (2011). Role of transcription factors in differentiation and reprogramming of hematopoietic cells. *Kleio J Med* 60(2): 47-55.

Nakano, T. (1996). In vitro development of hematopoietic system from mouse embryonic stem cells: a new approach for embryonic hematopoiesis. *Int J Hematol* 65(1): 1-8.

Nerlov, C., E. Querfurth, et al. (2000). GATA-1 interacts with the myeloid PU.1 transcription factor and represses PU.1-dependent transcription. *Blood* 98(8): 2543-51.

Nichols, K. E., J. D. Crispino, et al. (2000). Familial dyserythropoietic anaemia and thrombocytopenia due to an inherited mutation in GATA1. *Nat Genet* 24(3): 266-70.

Nishikawa, K., M. Kobayashi, et al. (2003). Selfassociation of Gata1 enhances transcriptional activity in vivo in zebra fish embryos. *Mol Cell Biol* 23(22): 8295-8305.

North, T., T. L. Gu, et al. (1999). Cbfa2 is required for the formation of intra-aortic hematopoietic clusters. *Development* 126(11): 2563-75.

Nuez, B., D. Michalovich, et al. (1995). Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. *Nature* 375(6529): 316-8.

O'Carroll, D., I. Mecklenbrauker, et al. (2007). A Slicer-independent role for Argonaute 2 in hematopoiesis and the microRNA pathway. *Genes Dev* 21(16): 1999-2004.

Okuda, T., J. van Deursen, et al. (1996). AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. *Cell* 84(2): 321-30.

Olsen, A. L., D. L. Stachura, et al. (2006). Designer blood: creating hematopoietic lineages from embryonic stem cells. *Blood* 107(4): 1265-75.

Orelio, C., K. N. Harvey, et al. (2004). The role of apoptosis in the development of AGM hematopoietic stem cells revealed by Bcl-2 overexpression. *Blood* 103(11): 4084-92.

Otterbein, L. E., M. P. Soares, et al. (2003). Heme oxygenase-1: unleashing theprotective properties of heme. *Trends Immunol* 24(8): 449-55.

Paine, A., B. Eiz-Vesper, et al. (2010). Signaling to heme oxygenase-1 and its anti-inflammatory therapeutic potential. *Biochem Pharmacol* 80(12): 1895-903.

Pal, S., A. B. Cantor, et al. (2004). Coregulator-dependent facilitation of chromatin occupancy by GATA-1. *Proc Natl Acad Sci U S A* 101(4): 980-5.

Palis, J. and G. B. Segel. (1998). Developmental biology of erythropoiesis. *Blood Rev*, 12(2): 106-14.

Palis, J and M. C. Yoder. (2001). Yolk-sac hematopoiesis: the first blood cells of mouse and man. *Exp Hematol* 29(8): 927-936.

Palis, J. (2004). Developmental biology: no red cell is an island. *Nature* 432(7020): 964-5.

Parekh, V., A. McEwen, et al. (2004). Defective extraembryonic angiogenesis in mice lacking LBP-1a, a member of the grainyhead family of transcription factors. *Mol Cell Biol* 24(16): 7113-29.

Pase, L., J. E. Layton, et al. (2009). miR-451 regulates zebrafish erythroid maturation in vivo via its target gata2. *Blood* 113(8): 1794-804.

Perkins, A., A. Sharpe, et al. (1995). Lethal beta-thalassaemia in mice lacking the erythroid CACCCtranscription factor EKLF. *Nature* 375(6529): 318-22.

Pevny, L., M. C. Simon, et al. (1991). Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature* 349(6306): 257-260.

Pevny, L., C. S. Lin, et al. (1995). Development of hematopoietic cells lacking transcription factor GATA-1. *Development* 121(1): 163-172.

Philipsen, S., S. Pruzina, et al. (1993). The minimal requirements for activity in transgenic mice of hypersensitive site 3 of the beta globin locus control region. *Embo J* 12(3): 1077-1085.

Pilon, A. M., M. O. Arcasoy, et al. (2008). Failure of terminal erythroid differentiation in EKLF-deficient mice is associated with cell cycle perturbation and reduced expression of E2F2. *Mol Cell Biol* 28(24): 7394-401.

Ponka, P. (1997). Tissue-Specific Regulation of Iron Metabolism and Heme Synthesis: Distinct Control Mechanisms in Erythroid Cells. *Blood* 89(1): 1-25.

Porcher, C., W. Swat, et al. (1996). The T cell leukemia oncoprotein SCL/tal-1 is essential for development of all hematopoietic lineages. *Cell* 86(1): 47-57.

Poznic, M. (2009). Retinoblastoma protein: a central processing unit. *J Biosci* 34(2): 305-12.

Qin, W., H. Leonhardt, et al. (2011). Regulation of DNA methyltransferase 1 by interactions and modifications. *Nucleus* 2(5): 392-402.

Rathjen, T., C. Nicol, et al. (2006). Analysis of short RNAs in the malaria parasite and its red blood cell host. *FEBS Lett* 580(22): 5185-8.

Reichard J. F., T. Gregory, et al. (2007). Heme oxygenase-1 induction by NRF2 requires inactivation of the transcriptional repressor BACH1. *Nucleic Acids Res* 35(21):7074–7086.

Rekhtman, N., F. Radparvar, et al. (1999). Direct interaction of hematopoietic transcription factors PU.1 and GATA-1: functional antagonism in erythroid cells. *Genes Dev* 13(11): 1398-411.

Richmond, T. D., M. Chohan, et al. (2005). Turning cells red: signal transduction mediated by erythropoietin. *Trends Cell Biol* 15(3): 146-55.

Robb, L., I. Lyons, et al. (1995). Absence of yolk sac hematopoiesis from mice with a targeted disruption of the scl gene. *Proc Natl Acad Sci U S A* 92(15): 7075-7079.

Robertson, K. D. and A. P. Wolffe (2000). DNA methylation in health and disease. *Nature Rev Genet* 1: 11-19.

Rodriguez, P., E. Bonte, et al. (2005). GATA-1 forms distinct activating and repressive complexes in erythroid cells. *Embo J* 24(13): 2354-66.

Rodriguez P., H. Braun, et al. (2006). Isolation of transcription factor complexes by in vivo biotinylation tagging and direct binding to streptavidin beads. *Methods Mol Biol* 338: 305-23.

Rylski, M., J. J. Welch, et al. (2003). GATA-1-mediated proliferation arrest during erythroid maturation. *Mol Cell Biol* 23(14): 5031-42.

Ryter, S. W. and R. M. Tyrrell (2000). The heme synthesis and degradation pathways: role in oxidant sensitivity. Hemeoxygenase has both pro- and antioxidant properties. *Free Radic Biol Med* 28(2): 289-309.

Ryter, S. W., J. Alam, et al. (2006). Heme oxygenase-1/carbon monoxide: from basic science to therapeutic applications. *Physiol Rev* 86(2): 583-650.

Saleque, S., S. Cameron, et al. (2002). The zinc-finger proto-oncogene Gfi-1b is essential for development of the erythroid and megakaryocytic lineages. *Genes Dev* 16(3): 301-6.

Sankaran, V. G., J. Xu, et al. (2010). Advances in the understanding of haemoglobin switching. *Br J Haematol* 149(2): 181-94.

Schatz, P. J. (1993). Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: a 13 residue consensus peptide specifies biotinylation in Escherichia coli. *Biotechnology (N Y)* 11(10):1138-43.

Schroeder, T. (2010). Hematopoietic stem cell heterogeneity: subtypes, not unpredictable behavior. *Cell Stem Cell* 6(3): 203-7.

Schuh, A. H., A. J. Tipping, et al. (2005). ETO-2 associates with SCL in erythroid cells and megakaryocytes and provides repressor functions in erythropoiesis. *Mol Cell Biol* 25(23): 10235–10250.

Schwenk, F., U. Baron, et al. (1995). A cre-transgenic mouse strain for the ubiquitous deletion of loxP-flanked gene segments including deletion in germ cells. *Nucleic Acids Res* 23(24): 5080-1.

Scott, E., M. Simon, et al. (1994). Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science* 265(5178): 1573-7.

Shan, Y., R. W. Lambrecht, et al. (2004). Role of Bach-1 in regulation of heme oxygenase-1 in human liver cells: insights from studies with small interfering RNAS. *J Biol Chem* 279(50): 51769-74.

Shan, Y., R. W. Lambrecht, et al. (2006). Role of Bach1 and Nrf2 in up-regulation of the heme oxygenase-1 gene by cobalt protoporphyrin. *FASEB J* 20(14): 2651-3.

Shearstone, J. R., R. Pop et al. (2011). Global DNA demethylation during mouse erythropoiesis in vivo. *Science* 334(6057): 799-802.

Shimizu, R., S. Takahashi, et al. (2001). In vivo requirements for GATA-1 functional domains during primitive and definitive erythropoiesis. *Embo J* 20(18): 5250-5260.

Shimizu, R., T. Kuroha, et al. (2004). Leukemogenesis caused by incapacitated GATA-1 function. *Mol Cell Biol* 24(24): 10814-25.

Shimizu, R., C. D. Trainor, et al. (2007). GATA-1 self-association controls erythroid development in vivo. *J Biol Chem* 282(21): 15862-71.

Shivdasani, R., E. Mayer, et al. (1995). Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. *Nature* 373(6513): 432-4.

Shivdasani, R. A., Y. Fujiwara, et al. (1997). A lineage-selective knockout establishes the critical role of transcription factor GATA-1 in megakaryocyte growth and platelet development. *Embo J* 16(13): 3965-3973.

Siti-Ismail, N., A. Samadikuchaksaraei, et al. (2012). Development of a novel three-dimensional, automatable and integrated bioprocess for the differentiation of embryonic stem cells into pulmonary alveolar cells in a rotating vessel bioreactor system. *Tissue Eng Part C Methods* 18(4): 263-72.

Soler, E., C. Andrieu-Soler, et al. (2011). A systems approach to analyze transcription factors in mammalian cells. *Methods* 53(2): 151-62.

Sood, R. and P. Liu. (2012). Novel insights into the genetic controls of primitive and definitive hematopoiesis from zebrafish models. *Adv Hematol* 2012: 830703.

Spada, F., U. Rothbauer, et al. (2006). Regulation of DNA methyltransferase 1. *Adv Enzyme Regul* 46: 224-34.

Stamatoyannopoulos, G., P. Constantoulakis, et al. (1987). Coexpression of embryonic, fetal, and adult globins in erythroid cells of human embryos: relevance to the cell-lineage models of globin switching. *Dev Biol* 123(1):191-7.

Sun, J., H. Hoshino, et al. (2002). Hemoprotein Bach1 regulates enhancer availability of hemeoxygenase-1 gene. *Embo J* 21(19): 5216-24.

Suwabe, N., T. Takahashi, et al. (1998). GATA-1 regulates growth and differentiation of definitive erythroid lineage cells during in vitro ES cell differentiation. *Blood* 92(11): 4108-18.

Takahashi, S., K. Onodera, et al. (1997). Arrest in primitive erythroid cell development caused by promoter-specific disruption of the GATA-1 gene. *J Biol Chem* 272(19): 12611-12615.

Takahashi, S., T. Komeno, et al. (1998). Role of GATA-1 in proliferation and differentiation of definitive erythroid and megakaryocytic cells in vivo. *Blood* 92(2): 434-42.

Tanaka, H., I. Matsumura, et al. (2000). GATA-1 blocks IL-6-induced macrophage differentiation and apoptosis through the sustained expression of cyclin D1 and bcl-2 in a murine myeloid cell line M1. *Blood* 95(4): 1264-73.

Testa, U. (2004). Apoptotic mechanisms in the control of erythropoiesis. *Leukemia* 18(7): 1176-99.

Trowbridge, J. J., J. W. Snow, et al. (2009). DNA methyltransferase 1 is essential for and uniquely regulates hematopoietic stem and progenitor cells. *Cell Stem Cell* 5(4): 442-9.

Trowbridge, J. J., A. U. Sinha, et al. (2012). Haploinsufficiency of Dnmt1 impairs leukemia stem cell function through derepression of bivalent chromatin domains. *Genes Dev* 26(4): 344-9.

Tsang, A. P., J. E. Visvader, et al. (1997). FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. *Cell* 90(1): 109-119.

Tsang, A. P., Y. Fujiwara, et al. (1998). Failure of megakaryopoiesis and arrested erythropoiesis in mice lacking the GATA-1 transcriptional cofactor FOG. *Genes Dev* 12(8): 1176-1188.

Tsiftsoglou, A. S., I. S. Vizirianakis, et al. (2009). Erythropoiesis: model systems, molecular regulators, and developmental programs. *IUBMB Life* 61(8): 800-30.

Turek-Plewa, J. and P. P. Jagodziński (2005). The role of mammalian DNA methyltransferases in the regulation of gene expression. *Cell Mol Biol Lett* 10(4): 631-47.

Vakoc, C. R., D. L. Letting, et al. (2005). Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. *Mol Cell* 17(3): 453-62.

Valouev, A., D. S. Johnson, et al. (2008). Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data. *Nat Methods* 5(9): 829-34.

Valverde-Garduno, V., B. Guyot, et al. (2004). Differences in the chromatin structure and cis-element organization of the human and mouse GATA1 loci: implications for cis-element identification. *Blood* 104(10): 3106-16.

van Werven, F. J. and H. T. Timmers (2006). The use of biotin tagging in Saccharomyces cerevisiae improves the sensitivity of chromatin immunoprecipitation. *Nucleic Acids Res* 34(4): e33.

Viens, A., U. Mechold, et al. (2004). Use of protein biotinylation in vivo for chromatin immunoprecipitation. *Anal Biochem* 325(1): 68-76.

Volker, J. L., L. E. Rameh, et al. (1997). Mitogenic stimulation of resting T cells causes rapid phosphorylation of the transcription factor LSF and increased DNA-binding activity. *Genes Dev* 11(11): 1435-46.

von Lindern, M., E. M. Deiner, et al. (2001). Leukemic transformation of normal murine erythroid progenitors: v- and c-ErbB act through signaling pathways activated by the EpoR and c-Kit in stress erythropoiesis. *Oncogene* 20(28): 3651-64.

Vyas, P., K. Ault, et al. (1999). Consequences of GATA-1 deficiency in megakaryocytes and platelets. *Blood* 93(9): 2867-2875.

Vyas, P., M. A. McDevitt, et al. (1999). Different sequence requirements for expression in erythroid and megakaryocytic cells within a regulatory element upstream of the GATA-1 gene. *Development* 126(12): 2799-2811.

Wadman, I., J. Li, et al. (1994). Specific in vivo association between the bHLH and LIM proteins implicated in human T cell leukemia. *Embo J* 13(20): 4831–4839.

Wadman, I. A., H. Osada, et al. (1997). The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. *Embo J* 16(11): 3145-3157.

Wall, L., E. deBoer, et al. (1988). The human beta-globin gene 3' enhancer contains multiple binding sites for an erythroid-specific protein. *Genes Dev* 2(9): 1089-1100.

Waltzer, L., G. Ferjoux, et al. (2003). Cooperation between the GATA and RUNX factors Serpent and Lozenge during Drosophila hematopoiesis. *Embo J* 22(24): 6516-25.

Weiss, M. J., G. Keller, et al. (1994). Novel insights into erythroid development revealed through in vitro differentiation of GATA-1 embryonic stem cells. *Genes Dev* 8(10): 1184-1197.

Weiss, M. J., C. Yu, et al. (1997). Erythroid-cell-specific properties of transcription factor GATA-1 revealed by phenotypic rescue of a gene-targeted cell line. *Mol Cell Biol* 17(3): 1642-51.

Whyatt, D. J., A. Karis, et al. (1997). The level of the tissue-specific factor GATA-1 affects the cell-cycle machinery. *Genes Funct* 1(1): 11-24.

Whyatt, D., F. Lindeboom, et al. (2000). An intrinsic but cell-nonautonomous defect in GATA-1-overexpressing mouse erythroid cells. *Nature* 406(6795): 519-524.

Wickrema, A. and J. D. Crispino (2007). Erythroid and megakaryocytic transformation. *Oncogene* 26(47): 6803-6815.

Wijgerde, M., J. Gribnau, et al. (1996). The role of EKLF in human beta-globin gene competition. *Genes Dev* 10(22): 2894-902.

Wiles, M. V. and G. Keller (1991). Multiple hematopoietic lineages develop from embryonic stem (ES) cells in culture. *Development* 111(2): 259-67.

Xu, G., R. Kanezaki, et al. (2006). Physical association of the patient-specific GATA1 mutants with RUNX1 in acute megakaryoblastic leukemia accompanying Down syndrome. *Leukemia* 20(6): 1002-8.

Xu, Z., S. Huang, et al. (2003). Identification of a TAL1 target gene reveals a positive role for the LIM domain binding protein Ldb1 in erythroid gene expression and differentiation. *Mol Cell Biol* 23(21): 7585-99.

Xu, Z., X. Meng, et al. (2006). Recruitment of the SWI/SNF protein Brg1 by a multiprotein complex effects transcriptional repression in murine erythroid progenitors. *Biochem J* 399(2): 297-304.

Yu, C., A. B. Cantor, et al. (2002). Targeted deletion of a high-affinity GATA-binding site in the GATA-1 promoter leads to selective loss of the eosinophil lineage in vivo. *J Exp Med* 195(11): 1387-1395.

Yu, Y. L., Y. J. Chiang, et al. (2005). MAPK-mediated phosphorylation of GATA-1 promotes Bcl-XL expression and cell survival. *J Biol Chem* 280(33): 29533-42.

Yu, M., L. Riva, et al. (2009). Insights into GATA-1-mediated gene activation versus repression via genome-wide chromatin occupancy analysis. *Mol Cell* 36(4): 682-95.

Zhan, M., C. P. Miller, et al. (2007). MicroRNA expression dynamics during murine and human erythroid differentiation. *Exp Hematol* 35(7): 1015-25.

Zhang, P., G. Behre, et al. (1999). Negative cross-talk between hematopoietic regulators: GATA proteins repress PU.1. *Proc Natl Acad Sci U S A* 96(15): 8705-10.

Zhang, P., X. Zhang, et al. (2000). PU.1 inhibits GATA-1 function and erythroid differentiation by blocking GATA-1 DNA binding. *Blood* 86(8): 2641-8.

Zhang, J., M. Socolovsky, et al. (2003). Role of Ras signaling in erythroid differentiation of mouse fetal liver cells: functional analysis by a flow cytometry-based novel culture system. *Blood* 102(12): 3938-46.

Zhou, W., D. R. Clouston, et al. (2000). Induction of human fetal globin gene expression by a novel erythroid factor, NF-E4. *Mol Cell Biol* 20(20): 7662-72.

Zon, L. I., H. Youssoufian, et al. (1991). Activation of the erythropoietin receptor promoter by transcription factor GATA-1. *Proc Natl Acad Sci U S A* 88(23): 10638-10641.

Zon, L. I. (2001). Hematopoiesis. New York, Oxford University Press, Inc.