Characterization of Transcription Factor Complexes involved in Globin Gene Regulation

Katarzyna Ewa Kołodziej

26th March 2008

Cover: **Rodota**

The work presented in this thesis was performed at the Department of Cell Biology at the Erasmus University Medical Center Rotterdam. This department is member of the Medisch Genetisch Centrum Zuid-West Nederland. The research was partially supported by the Netherlandse Organizatie voor Wetenschappelijk Onderzoek (NWO) and by EU.

Characterization of Transcription Factor Complexes involved in Globin Gene Regulation

Karakterizering van transcriptie factor complexen betrokken bij de regulatie van globine genen

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de Rector Magnificus
Prof.dr. S.W.J. Lamberts
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
woensdag 26 maart 2008 om 15.45 uur.

door

Katarzyna Ewa Kołodziej

geboren te Wrocław, Polen

Promotiecommissie

Promotor:

Prof.dr. F.G. Grosveld

Overige laden:

Prof.dr. J.N.J. Philipsen

Prof.dr. J.D. Engel

Dr.ir. D.N. Meijer

Copromotor:

Dr. J. Strouboulis

Moim rodzicom

&

for Luc

TABLE OF CONTENTS

Chapter 1 : Introduction 9

Chapter 2 : Isolation of transcription factor complexes by in vivo biotinylation tagging and direct binding to streptavidin beads (Patrick Rodriguez, Harald Braun, Katarzyna E. Kolodziej, Ernie de Boer, Jennifer Campbel, Edgar Bonte, Sjaak Philipsen and John Strouboulis Methods Mol Bio. 2006; 338: 305-23) 33

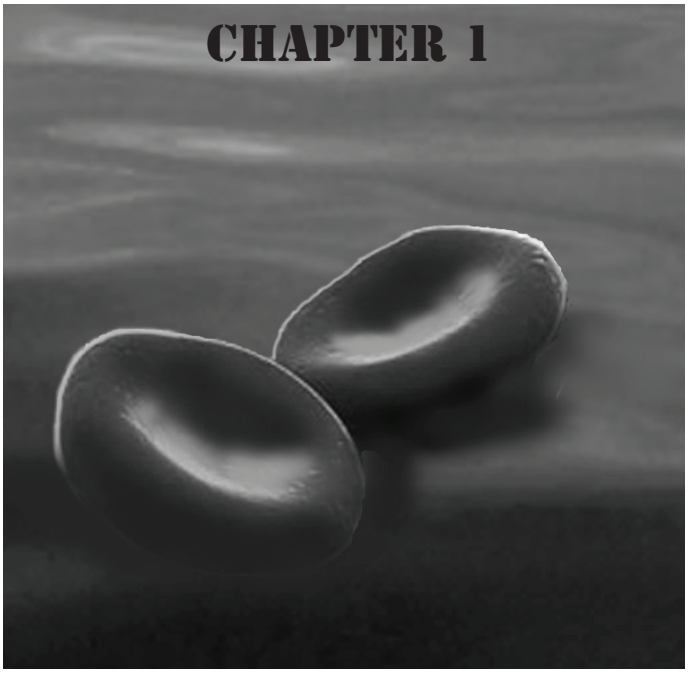
Chapter 3 : GATA-1 forms distinct activating and repressive complexes in erythroid cells (Patrick Rodriguez, Edgar Bonte, Jeroen Krijgsveld, Katarzyna E. Kolodziej, Boris Guyot, Albert Heck, Paresh Vyas, Ernie de Boer, Frank Grosveld and John Strouboulis, EMBO J. 2005; 24:2354-66) 53

Chapter 4 : Characterization of TR2/TR4 complexes in mouse erythroleukemic (MEL) cells by in vivo biotinylation tagging and mass spectrometry 81

Chapter 5 : Optimization of ChIP assays using in vivo biotin tagged transcription 113

Chapter 6 : Discussion 127

INTRODUCTION



INTRODUCTION

- **function of hemoglobin**

Blood cells can be classified in two groups: red blood cells or erythrocytes, that transport O_2 and CO_2 bound to hemoglobin and white blood cells or leukocytes which combat infections and in some cases phagocytose and digest debris (1). In addition blood contains a large number of platelets that participate in the process of blood clotting. Remarkably all these cells are ultimately generated from a common stem cell in the bone marrow (1, 99, 165-167). A schematic representation of the main lineage commitment steps in hematopoesis is depicted in Figure 1 (after (18)).

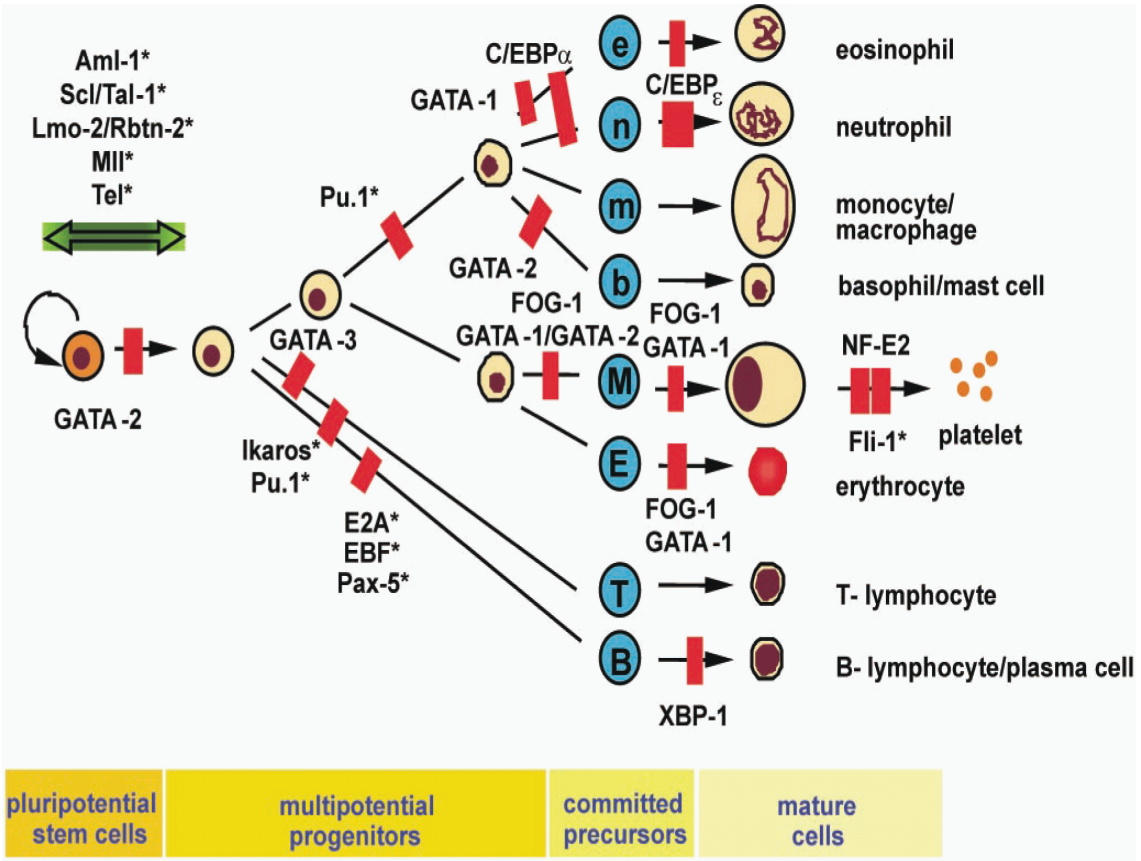


Figure 1.Schematic representation of hematopoiesis. A number of transcription factors involved in hematopoiesis are indicated (after (18)).

Hemoglobin is a tetramer composed of four globin peptide chains, two α-like and two β-like, with each globin chain binding one heme molecule (1). Oxygen binds to the iron ion complexed with the heme molecule. The globin chains prevent this process from becoming irreversible by folding and creating a protein pocket which prevents two heme molecules from coming together and the iron ion from being oxidized (11).The α-hemoglobin chain consists of 141 amino acids and is found in fetal hemoglobin and in normal adult hemoglobin A. The β-globin chain consist of 146 amino acids and is found in normal adult hemoglobin A (1).

- **globin gene structure**

All animals that use hemoglobin for oxygen transport have different hemoglobin species expressed in early and later developmental stages (144, 145) . The major hemoglobin in human fetal life is HbF ($\alpha_2\gamma_2$), and the major hemoglobin in adult life is HbA ($\alpha_2\beta_2$). Adult hemoglobin A2 ($\alpha_2\delta_2$) is a minor hemoglobin expressed at less than 2% of the total hemoglobin (6, 145).

In man, two gene clusters direct the synthesis of hemoglobin chains: (1) the α-globin locus containing the embryonic ζ globin gene and the two adult α1 and α2 globin genes and (2) the β-globin locus, which consists of the ε-Gγ-Aγ-δ-β globin genes arranged in the order in which they are expressed during development (Figure 2A) (144, 145). The β-like globin genes extend over 50 kb on

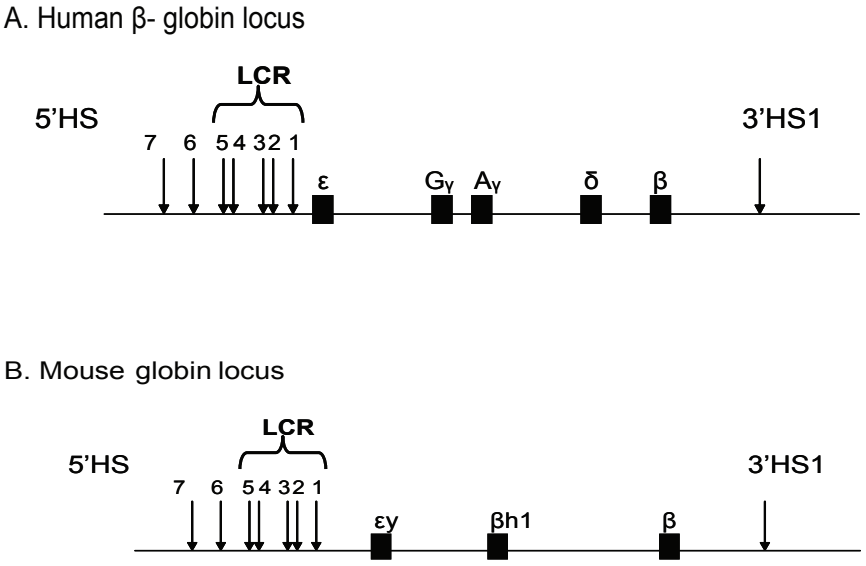


Figure 2. A) Scheme of human β-globin locus B) scheme of mouse β-globin locus

chromosome 11 and evolutionarily arose from a single ancestral globin gene as a result of successive duplication events. The evolution of γ-globin genes is most likely related to the increased oxygen affinity of HbF over HbA, which facilitate oxygen delivery to the fetus in the placental circulation (153).

The mouse model has been extensively used to study the regulation of the human globin gene expression primarily through the use of transgenic animals (144, 145, 149). The regulation of human globin genes in transgenic mouse models can be assessed using the endogenous mouse globin genes as reference. This species has no distinct fetal stage of erythropoiesis and hence no γ-globin genes. The mouse β-globin locus consists of the εy-βh1-β_{min}-β_{maj} globin genes (Figure 2B) (reviewed in (6, 144, 145)), and erythroid development involves a single switch from embryonic to adult globin chains. In mice the embryonic to adult switch takes place at around 11.5 days of gestation, with the emryonic εy gene being orthologous to the human ε-globin gene and the βh1 gene orthologous to the human γ gene (Figure 3B). In contrast to man the fetal liver stage of erythropoiesis in mouse produces α globin and β major and β minor, but no embryonic globin chains (reviewed in (67, 145)). It is thus considered an adult stage for globin chain synthesis in erythropoiesis.

Erythropoiesis in man (Fig 3A) begins in the yolk sac but at about five weeks of gestation the site of hematopoiesis changes from the yolk sac to the fetal liver, which remains the main site of erythopoiesis until about the twentieth week of gestation. Subsequently, the site of hematopoiesis shifts to the spleen and the bone marrow (around the thirtieth week of gestation) and by the time of birth, the bone marrow is the main hematopoietic organ ((145) and references therein). The shifting sites of erythropoiesis coincide with changes in the hemoglobin composition of the red cells and also in other morphological and biochemical characteristics. Expression of ε-globin is restricted to the yolk sac, whereas the γ and β globins are restricted to erythroblasts of liver origin (145). During the second (γ to β globin) switch, which occurs gradually around birth, the adult stage δ and β globin genes are activated, whereas γ globin is suppressed to very low levels (1%-2%) (78, 145). All normal adults produce a small quantity (~1%) of fetal hemoglobin, which appears to be confined to a small population of red cells which are called F cells (125, 172).

When the 70kb human β locus was analyzed in transgenic mice, the expression pattern was shown to be similar to the pattern observed in humans (149) (Figure 3C). The ε-globin gene showed a similar expression pattern to the mouse εy gene, but the level of expression was much lower than that of εy or of γ-globin genes. Importantly, it was observed that the γ-globin genes were expressed in high levels throughout the embryonic yolk sac stage, with significant levels of γ-globin expression remaining in the early fetal liver stage. Expression of γ-globin is switched off at day 16 of development in the fetal liver. These observations suggest that the human γ-globin genes detect a distinct fetal stage in the mouse fetal liver, which is not

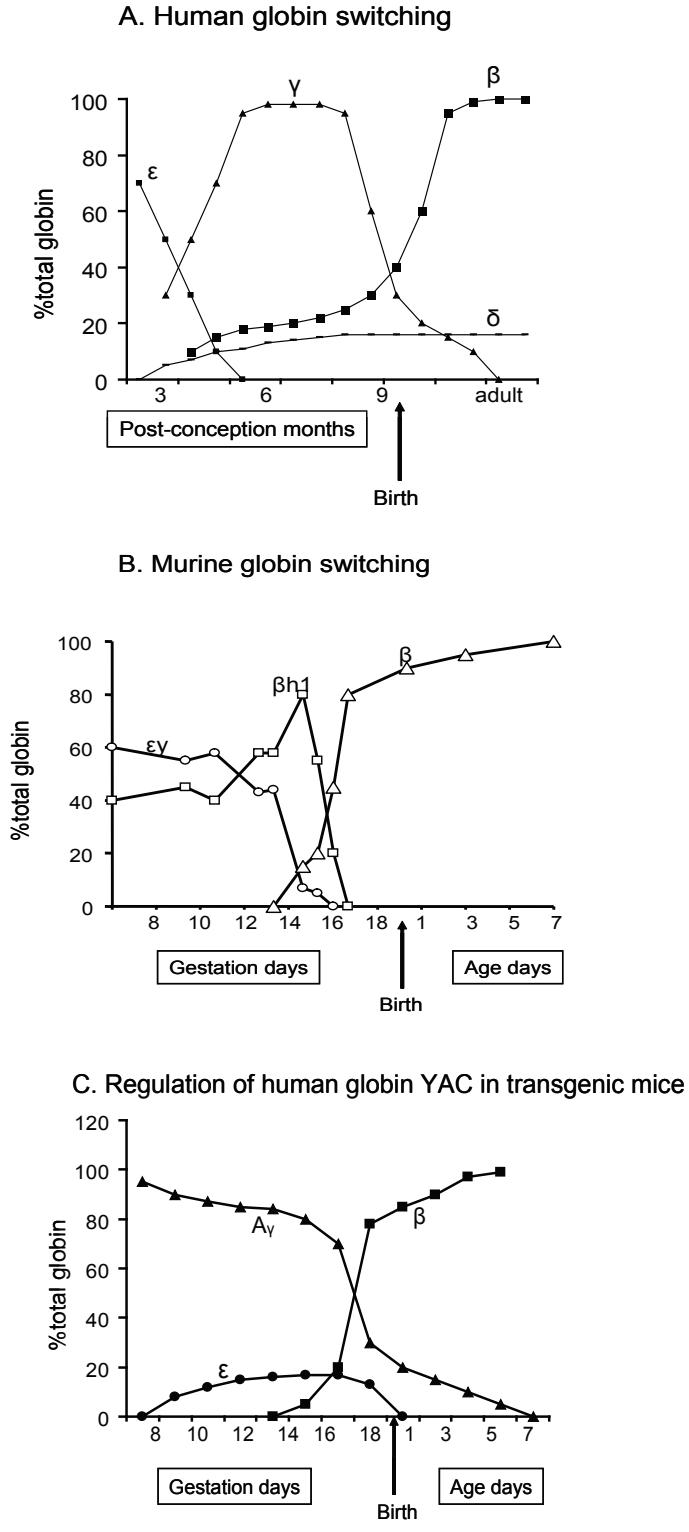


Figure3. A) Human globin expression levels measured throughout development. Squares represent β globin expression levels, triangles γ globin expression levels and dots ε globin expression levels. B) mouse globin gene expression during development: circles- εy globin expression levels, squares βh1 expression levels, triangles β major globin expression levels. C) Expression levels of human globin genes in the blood of transgenic mice. Circles, human ε- globin; triangles, human Aγ globin; squares, human β globin (after (133)).

detected by any of the endogenous mouse globin genes. This difference between the human and mouse fetal globin genes could be accounted for by changes in the transcription factor environment that can be detected by the human γ-globin regu;atory elements but not by those of the mouse embryonic globin genes. The same set of experiments showed that expression of human β globin gene follows that of β_{maj} almost completely (149). The developmental pattern of expression of the human β-globin locus introduced in transgenic mice as a 130kb YAC was similar to the one observed with the 70kb transgene and is shown in Figure 3C (15, 57, 149, 154).

- **Hemoglobinopathies**

Thalassaemias, which are the commonest monogenic diseases in man, result from over 200 different mutations in the human α- and β- globin genes. Thalassaemia is an autosomal recessively inherited group of disorders (31), associated with a wide spectrum of symptoms, ranging from intrauterine death to very mild, symptom-free anemia (163, 164). Thalassaemias are classified according to the particular globin chain which is ineffectively produced, as α, β (these two are by far the most important), δβ, and γδβ thalassaemias. The most important and probably the most intensively studied hemoglobinopathies are sickle cell anemia and hemoglobin E thalassaemia (disease usually found in people of Southeast Asian ancestry, especially in Cambodia, Laos and Thailand, it is believed to be the most common hemoglobinopathy in the world) (163). Importantly, treatment with regular blood transfusions and adequate chelation therapy has improved prognosis for the severe forms of thalassaemia. Table 1 shows the thalassaemias typology (after (145))

Many insights into the regulation of hemoglobin switching have been obtained by the study of mutations that increase fetal hemoglobin production during adult life. These mutations are clinically relevant, because increased synthesis of HbF in individuals that have co-inherited sickle cell anemia or β thalassemia reduces disease severity (145).

All normal adults produce a small quantity (~1%) of fetal hemoglobin, which appears to be confined to a small population of red cells which are called F cells (125, 172). Hereditary persistence of fetal hemoglobin (**HPFH**) characterizes a condition where production of variable levels at (>1%) of HbF persists into childhood and adult life (125). In molecular terms, there are deletion HPFHs, characterized by deletions in the 3' of the human β globin locus resulting in the total absence of the δ and β globin genes, and non-deletion HPFHs caused by mutations affecting the promoters of the $^{G}\gamma$ or $^{A}\gamma$ genes (145). Some of these mutations are listed in Table 1 (based on (145) and references therein). Importantly, mutations that increase HbF levels have been found to ameliorate the effects of β-thalassemia and sickle cell disease, when co-inherited (145). Thus, studying the HPFH-associated mutations can be potentially useful for the development of therapeutic approaches to treat thalassemias.

A number of different agents to treat thalassemic patients have been already tested in clinical trials for their efficiency in reactivating or augmenting fetal globin production in the adult stage. Hydroxyurea, an S-phase-specific agent that inhibits ribonucleotide reductase and activated guanylyl cyclase (raising the levels of cGMP) has been shown to increase HbF even in pediatric patients. However, hydroxyurea has many other side effects, including increased red cell size, presumably by altering cellular hydration and reducing the white cell count which may also play a role in reducing the number of painful crises (123, 125). There have also been several studies to assess the effect of recombinant erythropoietin (Epo) (a cytokine produced by the kidney that regulates red blood cell production), either alone or in combination with hydroxuurea in treating thalassaemias, as there is indirect evidence that rapid erythroid expansion may favor HbF production (109, 110, 123). The administration of butyrate fatty acid compounds has also been associated with varying responses in increasing HbF levels. Continuous butyrate infusions administered to patients induced γ-globin expression, F cells, F reticulocytes and HbF but the effects declined with the prolonged administration. As butyrate is also known to cause growth arrest in G1 phase of the cell cycle, a pulse regimen was successfully tested to avoid its anti-proliferative effects (3, 4). Butyrate is known to inhibit histone deacetylases (HDACs) and thus may work by interfering with repressive chromatin associated complexes binding to the γ-globin promoter and thus de-repressing the γ-globin expression (24, 35, 132). Finally, 5-azacytidine and its derivatives are agents that result in DNA hypomethylation and have been shown to induce HbF production in patients with sickle cell disease and thalassaemia (38-40, 49, 68). However, concern about its potential carcinogenic effects (the same concern as for hydroxyurea) halted its development as a therapeutic agent in patients with hemoglobin disorders (49). Nevertheless, the use of agents that affect histone acetylation (butyrate) and DNA methylation in inducing HbF

suggest an epigenetic basis in the silencing of γ globin expression in the adult which can be reversed.

- **Regulation of globin genes**

The expression of the human and mouse β-like globin genes, is controlled by the complex interaction between:

A. ***cis*-acting** sequences : the locus control region (LCR) and downstream regulatory sequences

B. **trans-acting** factors: transcription factors and chromatin remodeling and modifying cofactors and complexes

- **Cis-regulation of globin genes**

The **LCR** consists of five erythroid specific DNaseI hypersensitive sites (HS1-5) located from 8 to 22 kbp 5'to the ε- globin gene (see also Figure 2A) (61, 158, 159). The functional core of each HS is approximately 250 nucleotides long and consists of several binding sites for various transcription factors (such as GATA-1, EKLF, NF-E2) (reviewed in (17, 67)). There is evidence that distinct HS sites contribute specific functions, e.g. HS5 acts as an insulator (developmental-specific chromatin border) (161) and HS3 as a critical activator of the entire gene locus (117, 119, 130). The order of the genes relative to the LCR is conserved in mammals, so the genes which are active early in development are closer to the LCR than genes which are activated later (41, 63, 66, 131, 134).

Deletion of the LCR in humans causes β-thalassaemia (144) with no globin expression and no evidence of an open chromatin domain organization in the locus (83, 144). In transgenic mice bearing the globin gene locus without an LCR, expression levels are severely affected but temporal globin expression or switching does not appear to be affected (147). The position of globin genes with respect to the LCR is crucial for their activation as shown for the ε-globin gene, which when relocated to the 3'end of the locus becomes transcriptionally silent at all developmental stages (136, 139, 140, 154). Moreover when the β-globin gene was placed in a proximal position to the LCR it becomes transcribed at all developmental stages (15, 28, 41, 64).

Various models explaining how the LCR activates downstream globin genes have been previously proposed: a looping (60), a tracking (13), a facilitated tracking (13), and a linking (14) model. Presently most of the experimental evidence available supports a looping model (41, 59, 128, 130, 156, 171).

According to this model, the LCR and the distal 5' and 3' hypersensitive regions are thought to interact together to form a developmentally stable three dimensional core called the Active Chromatin Hub (**ACH**) (36, 42, 128, 156). This represents the spatial clustering of regulatory elements that appear to concentrate transcription factors and chromatin remodelers at one site. The globin genes undergo developmental switches in their interaction with the ACH, correlating with the switch in their transcriptional activity. The intervening chromatin that domains that contain the inactive genes loops out of the ACH at any given developmental stage (36).

Several experimental approaches provided evidence for a competitive mechanism as the basis of the γ to β globin gene switching. It was shown that the LCR interacts with only one globin gene promoter at a time "flip-flopping" between two or more promoters depending on the developmental stage (171). The competition is explained by the preferential interaction between the LCR and the γ globin genes (presumably mediated by stage specific transcription factors) which outcompete the β globin gene during the embryonic/fetal stage of development. A gene is likely to achieve competitive dominance when it interacts more frequently and more strongly with the LCR than other competing genes (41, 60). Thus, it seems necessary to evolve promoter-mediated silencing for the proximal genes, which is what is observed for the ε and γ globin genes. When the γ globin gene is silenced at the adult stage, the LCR preferentially interacts with the β gene (10, 41, 45, 59).

ε- globin gene silencing. Experiments in transgenic mice showed the LCR-dependent yolk sac erythropoiesis-restricted expression of the ε-globin gene. This led to the

concept of autonomous silencing, whereby all the elements responsible for turning off ε gene expression in the fetal/adult stage are contained within the canonical gene or adjacent sequences (136). Further experimental investigations identified sequences that participate in silencing, in both the proximal and distal ε-globin gene promoter (53, 103, 135, 155). Thus, mutation of the GATA-1, YY1 or Sp1 sites upstream of the ε promoter abolished ε-globin silencing (135), raising the possibility that globin gene silencing is the result of the combinatorial action of several transcriptional factors participating in the formation of a silencing complex.

γ-globin gene silencing. Figure 4 shows the transcription factor binding sites that have been mapped in the γ globin promoter (after (145)). The identification of DR1 (direct repeat 1) sites in the γ globin but also in the ε globin gene promoters led to the suggestion that nuclear receptors such as COUP-TFII or TR2 and TR4 could be implicated in globin gene silencing (53, 150). Further studies in transgenic mice have localized a silencing element in the -378 to -730 region upstream of the γ promoter (146). It has been suggested that this sequence contains elements that contribute to the autonomous silencing of the γ gene (144). This hypothesis was strengthened by the further identification of a GATA-1 site mutation in -378 to -730 sequences that resulted in HPFH (75, 144). Mutational studies using human γ gene hybrid promoters suggest that the CACCC box is the main proximal motif that participates in γ globin silencing (102).

Several transcription factors have been implicated in globin gene silencing. For

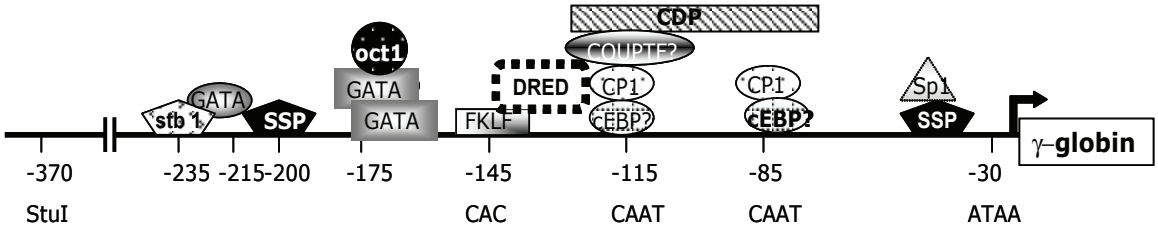


Figure 4. Scheme of transcription factor binding sites on the γ globin gene promoter (145)

example, Mbd2 was recently reported to mediate γ-globin silencing, as Mbd2 knock-out mice continue expression of γ-globin in adulthood. Mbd2 does not however bind directly to the promoter (137). Other transcription factors that are part of complexes and are implicated in γ globin silencing are discussed further in this chapter. A number of studies have demonstrated that nuclear receptors also have a profound impact on the development of hematopoietic cells, e.g., RXR and TR/c-erb-A effectively regulate erythroid cell growth and differentiation (7, 8). Also the v-erbA oncoprotein an oncogenic version of TR/c-erbA, arrests differentiation of erythroblasts and cooperates with v-erbB in producing fatal erythroleukemia in chickens (183, 184). Additionally, activation of both GR and ER is required for the efficient induction and self-renewal of erythroid progenitor cells (9, 148, 170). Interestingly, introduction of dominant-negative RAR in bone marrow cells immortalized a lymphohematopoietic progenitor, which can be induced to differentiate into lymphoid, myeloid, and erythroid lineages (85, 157).

Chromatin modification marks have been shown to characterize the chromatin of active and silenced globin genes. Histone hyperacetylation has been associated with the opening of the chromatin structure and activation of genes including genes in the β-globin locus (54, 81, 130). Histone hyperacetylation at the LCR and β-globin promoter has been associated with the formation of the Active Chromatin Hub (βACH) in the adult stage (54, 156). Histone acetyltrasferases (HATs) together with the chromatin remodeling SWI/SNF complexes have been shown to be active at the murine and human β-globin locus in facilitating transcriptional activation (2, 121, 122). In addition, the specific inhibition of HDACs was shown to selectively increase acetylation at a hypoacetylated promoter in fetal liver cells, suggesting that active deacetylation contributes to the silencing of the β- locus promoters, particularly γ globin (16, 25, 174). These observations also suggest that γ-globin is epigenetically regulated and that its suppression is reversible. ChIP studies

also show that methylation of histone 3 lysine K79 (H3-K79me), at the active murine β globin gene in adult erythroid cells is dependent on the presence of p45/NF-E2 (74). Moreover, recent experiments have shown that methylation of histone H3 at lysine 4 (H3-K4me4) is mediated by the MLL complex, and that H3 acetylation at lysine 9 (H3-K9ac) marks the active β major globin gene. These modifications are in dynamic equilibrium with repressive the H3-K9me3 mark (37).

Several protein complexes have been reported to bind to the promoters of the globin genes in the β-locus, here we briefly discuss some of them and we concentrate on one, namely, the DRED complex.

p22NF-E4 has been recently shown to be part of the stage selector protein (**SSP**) complex and to have a role in globin switching (188, 189). Enforced expression of human p22NF-E4 increases γ-globin expression in K562 cells, whereas over-expression of p22NF-E4 in transgenic mice carrying the human β- globin locus YAC results in a delayed human γ – to β- globin switch (189). In addition, NF-E4 interacts with and is a target of the co-activator PCAF acetyltrasferase. Acetylation prolongs the half-life of NF-E4 and reduces interaction with HDAC1 maximizing the activating ability of NF-E4 at the γ globin promoter (188).

The **PYR** chromatin remodeling complex was isolated from MEL cells and shown to be present only in adult hematopoietic cells. It binds to a 250 bp polypyrymidine (PYR)-rich DNA sequence 1kb upstream of the human δ-globin gene (120). Deletion of the δ-globin gene promoter fragment, which includes the PYR binding sequence results in a delay of γ- to β- switch in transgenic mice (45, 121). A transcription factor –Ikaros- has been recently identified as the DNA binding subunit of PYR complex (122). Interestingly, Ikaros knock out mice that are also transgenic for the human β-globin locus YAC show delayed human γ- to β- globin switching (108). The PYR complex can have an activating or repressive function, as it contains subunits of the SWI/SNF complex and subunits of the NuRD complex (including HDACs) (122). A simplified scheme of human globin gene switching is depicted in Figure 5.

The CAC box and the CCAAT box are conserved motifs found in the globin promoters. The distal CCAAT box in γ-globin promoter contains an approximate direct repeat (DR1) TGACCAATAGCC. Importantly, the CAAT box regions of the ε and ζ genes as well as murine εy and βh1 also contain a DR-1 type sequence (53). Hereditary persistence of fetal hemoglobin (HPFH) has been associated with deletions/mutations in distinct plates in the γ- globin promoter (like e.g. the -117 Greek HPFH (32)). It was noted that of the sixteen different mutations identified in the Gγ and Aγ promoters that are associated with HPFH six are located within DR1 elements. These observations suggested an essential role for the DR1 elements in both ε- and γ-globin gene silencing in definitive erythroid cells (53, 153, 155).

EKLF was the first candidate protein thought to be responsible for the stage- specific repression of ε-globin, which binds to CACCC element in ε-globin promoter (initially it was identified to bind CACCC element of and activate the β-globin gene promoter (116)). However analysis of transgenic EKLF knock-out mice containing a mutation introducing an EKLF binding site in the ε-globin gene promoter did not reveal any changes in ε- globin expression (155). It was subsequently suggested that there must be a definitive erythroid stage specific repressor which binds to Direct Repeats in the ε-globin gene promoter, and this entity was named **DRED** (DIRECT REPEAT ERYTHROID DEFINITIVE). The γ-globin gene promoter also contains a single DR1 sequence, which was also speculated to have a repressive function. So far no DR1 elements have been identified in the adult β-globin promoter (150, 154, 155).

An alternative model for globin switching presented in Lee et al (93), suggested that CCTTG elements, located very close to the DR elements, would play a key role in the γ-globin silencing. Comparison of transgenic mice bearing a mutDR1 YAC, a mutCCTTG YAC, and the wild-type human β-globin locus YAC revealed a 10 fold increase in Aγ- globin mRNA levels in adult spleen when DR1 was disrupted. Disruption of the CCTTG motif caused only

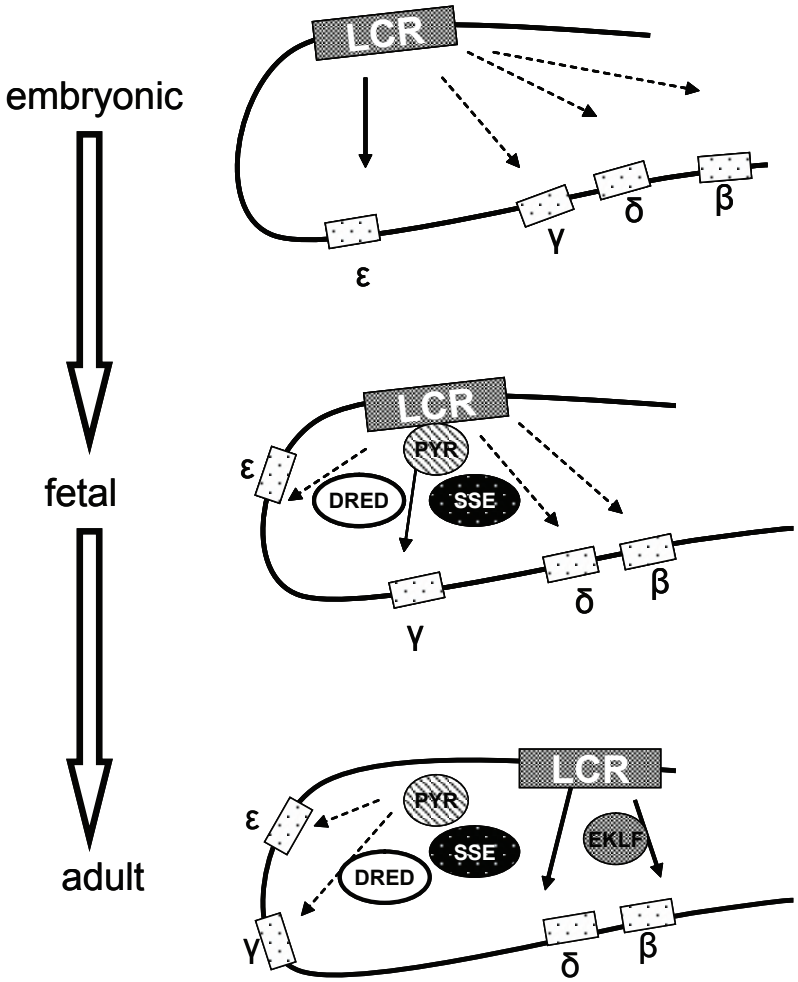


Figure 5. Simplified model of human globin gene switching from embryonic via fetal to adult stage, as indicated by the vertical arrows. The LCR has a preference for acting on the closest ε globin promoter at the embryonic stage. In the fetal stage DRED and other transcription factor complexes silence ε globin which no longer competes for the LCR with the γ globin gene. In this way γ globin expression can be activated. In the adult stage, the DRED, PYR, SSE and possibly other transcription factor complexes silence γ globin and block interactions with the LCR. The β globin promoter can be now activated by the LCR. In addition, EKLF has been indicated as it has been described to play role in β globin activation (after (6))

slight elevation of Aγ-globin mRNA and only in fetal liver cells. Together those observations supported the hypothesis that DR motifs are important regulators of globin gene switching and that the elements in the Aγ promoter contribute significantly to the mechanism controlling globin switching (126).

DRED (Direct Repeat Erythroid Definitive), a 540 kDa protein complex, has been isolated from MEL cells and was shown to bind to direct repeats present in the ε- and γ-globin promoters. Importantly, DRED binding is reduced 3-fold when a point mutation that gives rise to HPFH is present in the γ-globin promoter at position -117. DNA-affinity purification of the DRED complex revealed the presence of five proteins: CDP, USF-1, LBP-1a, and the orphan nuclear receptors TR2 and TR4. DRED DNA binding activity was confirmed by EMSA only for TR2 and TR4 (150). This complex is further described in Chapter 4.

TR2 and TR4- nuclear orphan receptors

- **Cloning the TR2/TR4 genes and their regulation and expression**

The human testicular orphan nuclear receptors TR2 and TR4 were isolated from human prostate and testis cDNA libraries using an oligonucleotide probe homologous to the DNA binding domain (DBD) of the steroid hormone receptor (SR) (21). There are three TR2 isoforms that arise as the result of alternative splicing. The TR2 gene was mapped to human chromosome 12 at band q22, whereas the TR4 gene was mapped to human chromosome 3 at band q24,3 but their cDNA is highly homologous (105, 180).

TR2/TR4 genes have been identified in many species: mouse, rat, human, Drosophila, sea urchin and amphibian (76). SmTR2/TR4 a *Schistosoma mansoni* worm homologue of the TR2/TR4 group of nuclear receptors is a 1,943 amino acid protein, the largest nuclear receptor reported

to date and it binds to the DR3 consensus hormone response element (70).

TR2 has a higher abundance in the mouse and rat reproductive organs, while TR4 is highly expressed in brain and male reproductive organs. Both TR2 and TR4 are highly expressed in the proliferation-active population of most developing organs, whereas expression is dramatically reduced in the more differentiated cell types (150, 160). In contrast to TR2, which is widely expressed throughout animal development (and can be detected as early as day 9 *post coitum*), TR4 has been detected with distinct abundance in specific brain regions (21). At embryonic days 14.5 and 19.5 high expression of TR4 was found in the central nervous system with lower expression detected throughout the embryo (160).

Several factors have been found to regulate TR2 and TR4 expression. For example the Ciliary Neurotrophic Factor α (CNTFRα) can increase TR2 and TR4 expression and TR2 then induces CNTFRα transcription through binding to a DR element within the CNTFR gene (85). In addition, all-trans retinoic acid (RA) stimulates cellular proliferation in 3T3-L1 preadipocytes by activating TR2 through an IR0-type RA response element in the TR2 promoter (62, 181). In contrast, irradiation and over-expression of p53 can repress transcription and translation of TR2 gene in human MCF-7 cells (106).

- **TR2/TR4 protein structure**

The cDNA sequence analysis shows that the human TR4 protein consists of 615 aa with a calculated MW of 67.3kDa. The TR2 cDNA encodes for a 603 aa polypeptide, and a calculated MW of 67kDa (76). The TR2 and TR4 receptors share high homology in their N-terminal domains and in the proximal box (P-box) of the DNA binding domain but differ in the length of their ligand binding domains (LBD) (see Figure 6). These similarities suggest that TR2 and TR4 may act on the same hormone response element (HRE) (72, 106, 182).

Both TR2 and TR4 have been reported to be posttranslationally phosphorylated, and SUMO-ylated. TR4 activity can be modulated by MAPK-mediated phosphorylation. Hyperphosphorylation of the N-terminal activation domain (AF-1) renders TR4 a repressor, whereas hypophosphorylation of this domain makes it an activator. The TR4 opposing transcriptional activities as an activator or repressor are mediated through the specific recruitment of the P/CAF coactivator or the RIP140 corepressor by hypophosphorylated and hyperphosphorylated forms of the protein, respectively (71).

In vivo metabolic labeling showed that the LBD of TR2 could be phosphorylated by protein kinase C (PKC). Phosphorylation increased the stability and nuclear accumulation of TR2. (79). Further experiments have shown multiple serines in the DBD of TR2 to be phosphorylated. PKC mediated phosphorylation of the DBD facilitates DNA binding by TR2 and the recruitment of the P/CAF coactivator. Thus, TR2 serves as an activator in the phosphorylated state (in contrast to TR4 which when phosphorylated has repressive function) (80). A more recent report has shown

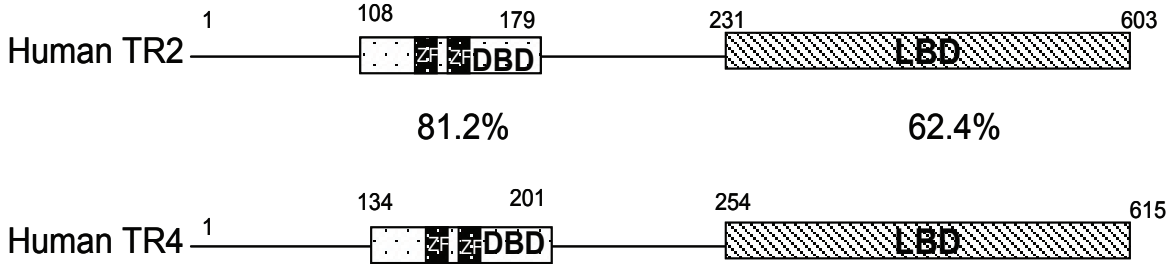


Figure 6. TR2 and TR4 orphan nuclear receptors. The homology between their DBD and LBD domains in % is indicated (97)

that Lys238 located in the LBD of TR2 can be SUMO-ylated resulting in the replacement of coregulators recruited to the Oct4 promoter (129). More specifically non- SUMO-ylated TR2 is localized to the promyelocytic leukemia (PML) bodies and activates Oct4. In contrast, SUMO-ylated TR2 seems to be released from PML bodies to act as an Oct4 repressor. SUMOylation of

TR2 induces an exchange of coregulators: corepressor RIP140 replaces coactivator P/CAF which switches TR2 from and activator to a repressor (129).

- **TR2/TR4 as transcriptional regulators**

TR2 and TR4 bind to the AGGTCA repeat motifs spaced by 0-6 nucleotides (direct repeats: DR0- DR6) in hormone response elements (HREs) with different affinity: DR1>DR2: DR5, DR4, DR6> DR3 (107). TR2 and TR4 may also compete for binding to target sites with other nuclear receptors such as RAR, RXR, vitD receptor, and TRα1 (72).

There have been several examples in literature showing that TR2 and TR4 can act as a **repressors** or **activators** (97) depending on the HRE that they bind to and also on their post-translational modification (discussed above). TR4 was shown to increase expression from the proximal promoter of the human oxitocin gene (by binding to DR4 element in the promoter) (162) and luteinising hormone (by binding to DR0 element) (186) activity. On the other hand, TR4 can repress the 21-OHase gene (by binding to monomeric AGGTCA element) (91). Rabbit TR4 was shown to act as a repressor in the estrogen- receptor mediated signaling pathway (65, 76). Finally, TR4 can bind bipartite repeats of many target genes including SV40 (88), the ciliary neurothrophic factor receptor gene (179), a DR4 of the thyroid hormone target genes (94) and a DR3 of the vitamin D3 target genes (96), the Bcl2 promoter (82), the promoter of the peroxisome proliferators- activated receptor α (PPARα) target genes (69, 175, 187).

TR2 has a modulatory effect on several signaling pathways involving retinoic acid (107), thyroid hormone (23) and ciliary neurothrophic factor (178), erythropoietin gene (92) aldolase gene (22), histamine receptor gene (90, 107) and simian virus 40 (89), in the human RARβ gene (27).TR2 was also shown to induce the expression of HPV-16 gene (by binding DR4 element)(34).

Both TR2 and TR4 can suppress retinoic acid induced transactivation of human HaCaT keratinocytes. They can suppress RA- mediated but not vitamin D mediated transcriptional activity in these cells (95).

- **TR2 and TR4 knock-out phenotypes**

TR2 knock out mice revealed no apparent developmental or pathological abnormalities (141). By contrast, TR4 null mice were born at lower than Mendelian ratios, with a significantly lower proportion of female knockout mice compared to male knockouts. These mice display reduced postnatal growth with TR4 null dams exhibiting defects in maternal behavior in that they do not build nests or collect pups to a single location, or crouch over pups, or nurse their offspring. An additional phenotype observed among TR4 null mice is the accumulation of fluid surrounding the eye (33). Closer investigation of TR4 null mice revealed reduced myelination which was particularly obvious in the forebrains and in early developmental stages (185). The cerebellum was also significantly reduced in size (26). Mouse embryonic fibroblast (MEF) cells prepared from TR4 null mice are more susceptible to UV-irradiation mediated apoptosis compared to the TR4- wild-type littermates (82).

Double TR2 and TR4 knock out mice die before 7.5 day *post coitum* (C. Chang, unpublished and (152)).

- **Known protein interactions**

It was previously demonstrated that, TR2 and TR4 can either act as homodimers or heterodimers, however they preferentially heterodimerize as coexpression of these two receptors exerts a much stronger repressive activity on a DR5 containing reporter compared to expressing either receptor alone. In developing testis, TR2 and TR4 are coexpressed in the same testicular cell populations and exhibit a parallel pattern of expression during development (107).

TR2 and TR4 have been shown to heterodimerise with other nuclear receptors (98). The androgen receptor can interact with TR4 and function as repressor to down-regulate the TR4 target genes by preventing TR4 binding to its target genes. The heterodimerization of

AR and TR4 also allows TR4 to repress AR target gene expression demonstrating a bidirectional suppression of their target genes (98).

Receptor Interacting Protein 140 (RIP140) was shown to interact with TR2 using yeast two-hybrid system using the LBD of TR2 as bait. This interaction was further confirmed by immunolocalization experiments in COS cells (87). The RIP-interacting domain of TR2 maps to the C-terminal 10- to 20-amino-acid sequence of TR2, but the C terminus of TR2 alone is not sufficient for this interaction (87). RIP140 was also shown to interact with several other nuclear receptors (including ERα and AR) (5, 19, 20, 73). In addition Luciferase assays in COS cells showed that RIP140 potentiates the trans repression via a TR2 IR7 (inverted repeat) regulatory element (87). It has also been previously shown that TR2 co-immunoprecipitates with HDACs 3 and 4. Dissection of TR2 domains showed that the DNA binding domain (DBD) was responsible for the interaction with HDACs 3 and 4 (27, 55, 56).

TR4 associated protein TRA16 is the first reported specific co-repressor of TR4, isolated by screening a human testis library by yeast two-hybrid using full length TR4 as the bait. This interaction was also further confirmed by immunofluorescence (177). TRA16 is a novel protein of 139 aa, with a total molecular weight of 16 kDa. TRA16 strongly interacts with TR4 (176). Luciferase reporter assays in COS cells showed that TRA16 can repress TR4-mediated activation in a dose dependent manner (177). It was further suggested that TRA16 may confer its co-repressor function by interrupting the binding of TR4 to its target DNA sequence (177). Another yeast two-hybrid system, using the hinge-LBD domain of TR4 as a bait, identified TIP27 as a co-repressor (118). This interaction was confirmed by in vitro and in vivo pull-down assays (118). Luciferase reporter assays showed that TIP27 repressed TR4-mediated transactivation on a DR1-Luc reporter construct but that this repression was not mediated byTR4 binding to the DR1 element. The exact mechanism of TIP27 function is still unknown (118).

- **TR2 and TR4 in hematopoiesis**

TR4 is highly expressed in hematopoietic cells. In fact, TR4 mRNA is very abundant in primary cultures of chicken bone marrow cells. Ectopic expression of TR4 in bone marrow cells resulted in an effective outgrowth in the myeloid compartment. Interestingly, although erythroid cells showed prominent expression of TR4 in all species analyzed, an outgrowth of erythroid progenitor cells was not observed. In addition, if introduced into established erythroid progenitors, TR4 had only a marginal growth promoting activity (84).

TR2 and TR4 have been shown to be the core DNA binding activity of the DRED complex. EMSA experiments, measuring the Ki for the DR1 elements in β-like globin gene promoters, showed that TR2 and TR4 bind different DRs with different affinities in the following order: human ε distal > ε proximal> γ; mouse εγ distal> βh1> εγ proximal. Data from these experiments also indicate that the affinity of the ε globin distal DR1 element for the TR2 homodimer is lower than that for the TR4 homodimer or TR2/TR4 heterodimer, whose affinities are very similar (151).

In transgenic mice bearing a dominant negative dnTR4 mutant that does not bind DNA and a human β-globin locus YAC, the human ε-globin gene was activated in both primitive and definitive erythroid cells (126, 151). In contrast the human fetal γ-globin gene was only slightly affected during primitive erythropoiesis in the yolk sac, but was induced in the fetal liver (definitive erythropoiesis) stage. These data suggest that TR2/TR4 comprise a stage-specific, gene-selective repressor of the human embryonic and fetal globin genes.

In addition, recent publications have shown that forced expression of TR2 and TR4 under the control of the GATA-1–HRD in mice caused a reduction of murine εy transcription in a dose dependent manner (150).

Aims of the thesis

Despite the many attempts, so far there has been no general agreement in the field regarding the nature of the mechanisms governing β-globin switching (51, 52, 111, 138). Many transcriptional regulators have been suggested to be key factors in this process (12, 46, 53, 113, 114, 116, 143, 155, 173, 188, 189) and have become targets for pharmaceutical development (24, 29, 30, 35, 38, 50, 58, 77, 86, 100, 101, 104, 112, 123, 124, 132, 142, 169, 190). In addition, the experimental and clinical observations support the idea that the chromatin structure of the globin gene promoters, particularly of the embryonic ε-globin and the fetal γ-globin genes, is an important parameter in developmental globin gene switching (36, 43, 44, 47, 48, 67, 108, 115, 127, 128, 130, 146, 156, 168).

DRED is a recently identified complex (150) that has been implicated in embryonic ε-and fetal γ-globin gene silencing in adult (126, 150, 151). The core DRED component is the orphan nuclear receptor heterodimer TR2/TR4 which has been shown to bind to DR1 elements in the ε and γ globin gene promoters (126, 150, 151), however the mechanism of their function in erythroid cells is not well understood.

The aim of this work was to purify TR2/TR4 interacting partners in erythroid cells, to gain more insight into how TR2/TR4 can silence (and also possibly to activate) the different globin genes in a temporal and spatial regulatory manner. To this end we used an *in vivo* biotinylation tagging technique which, as we further demonstrate, proves to be a very powerful tool for efficient protein complexes purification. It can also be very efficiently used for the mapping of transcription factor binding sites *in vivo* by chromatin immunoprecipitation (ChIP) experiments.

Table 1. Types of mutations in non- deletion HPFH, after (145) and the references therein.

Type and ethnic group	Mutation
$^{G}\gamma$ HPFH	
Japanese	Gγ- 114 C to T
Australian	Gγ -114 C to G
Black/Sardinian	Gγ-175 T to C
Tunisian	Gγ-200 +C
Black	Gγ-202 C to G
$^{A}\gamma$ HPFH	
Georgian	Aγ-114 C to T
Black	Aγ-114 to -120 deleted
Greek	Aγ-117 G to A
Cretan	Aγ-158 C to T
Black	Aγ-175 T to C
Brazilian	Aγ-195 C to G
Chinese/Italian	Aγ-196 C to T
British	Aγ-198 T to C
Georgian	Aγ-202 C to T

Table 2. Thallassaemias (based on (145))

(δβ)^{o} thalassemia	**β^{o} thalassemia (deletions that remove the β globin promoter)**
Japanese	*Turkish/Bulgarian*
Spanish	*African American*
Indian	*Croatian*
Turkish	*Thai*
Macedonian/Turkish	*Czech*
Black	*Turkish*
Eastern European	*Asian Indian*
Laotian	*Australian*
Thai	*Dutch*
Mediterranean/Sicilian	*Southeast Asian*
Hb Lepore	*Filipino*
(Aγδβ)^{o} thalassemia	*Italian*
Cantonese	**Deletions removing sequences of the LCR**
Malaysian (1) and (2)	*Dutch*
Chinese	*Anglo-Saxon*
Yuannanese	*English*
Thai	*Hispanic*
German	**(εγδβ)^{o} thalassemia**
Belgian	*Mexican*
Italian	*Canadian*
Turkish	*Irish*
American black	*Dutch*
Indian	

References

1. **Alberts B, Johnson A, J. L, Raff M, Roberts K, and Walter P** editors. *Molecular Biology of the Cell*. Garland Science, 2002.
2. **Armstrong JA, Bieker JJ, and Emerson BM**. A SWI/SNF-related chromatin remodeling complex, E-RC1, is required for tissue-specific transcriptional regulation by EKLF in vitro. *Cell* 95: 93-104, 1998.
3. **Atweh GF, DeSimone J, Sauntharajah Y, Fathallah H, Weinberg RS, Nagel RL, Fabry ME, and Adams RJ**. Hemoglobinopathies. *Hematology / the Education Program of the American Society of Hematology American Society of Hematology* 14-39, 2003.
4. **Atweh GF, and Loukopoulos D**. Pharmacological induction of fetal hemoglobin in sickle cell disease and beta-thalassemia. *Seminars in hematology* 38: 367-373, 2001.
5. **Augereau P, Badia E, Carascossa S, Castet A, Fritsch S, Harmand PO, Jalaguier S, and Cavailles V**. The nuclear receptor transcriptional coregulator RIP140. *Nuclear receptor signaling* 4: e024, 2006.
6. **Bank A**. Regulation of human fetal hemoglobin: new players, new complexities. *Blood* 107: 435-443, 2006.
7. **Bartunek P, and Zenke M**. Retinoid X receptor and c-cerbA/thyroid hormone receptor regulate erythroid cell growth and differentiation. *Molecular endocrinology (Baltimore, Md* 12: 1269-1279, 1998.
8. **Bauer A, Mikulits W, Lagger G, Stengl G, Brosch G, and Beug H**. The thyroid hormone receptor functions as a ligand-operated developmental switch between proliferation and differentiation of erythroid progenitors. *The EMBO journal* 17: 4291-4303, 1998.
9. **Bauer A, Tronche F, Wessely O, Kellendonk C, Reichardt HM, Steinlein P, Schutz G, and Beug H**. The glucocorticoid receptor is required for stress erythropoiesis. *Genes & development* 13: 2996-3002, 1999.
10. **Behringer RR, Ryan TM, Palmiter RD, Brinster RL, and Townes TM**. Human gamma- to beta-globin gene switching in transgenic mice. *Genes & development* 4: 380-389, 1990.
11. **Berg J, Tymoczko JL, and Stryer L** editors. *Biochemistry*. whfreeman.
12. **Bieker JJ, Ouyang L, and Chen X**. Transcriptional factors for specific globin genes. *Annals of the New York Academy of Sciences* 850: 64-69, 1998.
13. **Blackwood EM, and Kadonaga JT**. Going the distance: a current view of enhancer action. *Science (New York, NY* 281: 60-63, 1998.
14. **Bulger M, and Groudine M**. Looping versus linking: toward a model for long-distance gene activation. *Genes & development* 13: 2465-2477, 1999.
15. **Bungert J, Dave U, Lim KC, Lieuw KH, Shavit JA, Liu Q, and Engel JD**. Synergistic regulation of human beta-globin gene switching by locus control region elements HS3 and HS4. *Genes & development* 9: 3083-3096, 1995.
16. **Burgess-Beusse B, Farrell C, Gaszner M, Litt M, Mutskov V, Recillas-Targa F, Simpson M, West A, and Felsenfeld G**. The insulation of genes from external enhancers and silencing chromatin. *Proceedings of the National Academy of Sciences of the United States of America* 99 Suppl 4: 16433-16437, 2002.
17. **Cantor AB, and Orkin SH**. Hematopoietic development: a balancing act. *Current opinion in genetics & development* 11: 513-519, 2001.
18. **Cantor AB, and Orkin SH**. Transcriptional regulation of erythropoiesis: an affair involving multiple partners. *Oncogene* 21: 3368-3376, 2002.
19. **Castet A, Boulahtouf A, Versini G, Bonnet S, Augereau P, Vignon F, Khochbin S, Jalaguier S, and Cavailles V**. Multiple domains of the Receptor-Interacting Protein 140 contribute to transcription inhibition. *Nucleic acids research* 32: 1957-1966, 2004.
20. **Cavailles V, Dauvois S, L'Horset F, Lopez G, Hoare S, Kushner PJ, and Parker MG**. Nuclear factor RIP140 modulates transcriptional activation by the estrogen receptor. *The EMBO journal* 14: 3741-3751, 1995.
21. **Chang C, Kokontis J, Acakpo-Satchivi L, Liao S, Takeda H, and Chang Y**. Molecular cloning of new human TR2 receptors: a class of steroid receptor with multiple ligand-binding domains. *Biochem Biophys Res Commun* 165: 735-741, 1989.
22. **Chang C, Lee HJ, and Lee YF**. Identification of the human aldolase A gene as the first induced target for the TR2 orphan receptor, a member of the steroid hormone receptor superfamily. *Biochem Biophys Res Commun* 235: 205-211, 1997.
23. **Chang C, and Pan HJ**. Thyroid hormone direct repeat 4 response element is a positive regulatory element for the human TR2 orphan receptor, a member of steroid receptor superfamily. *Molecular and cellular biochemistry* 189: 195-200, 1998.
24. **Chen JS, Faller DV, and Spanjaard RA**. Short-chain fatty acid inhibitors of histone deacetylases: promising anticancer therapeutics? *Current cancer drug targets* 3: 219-236, 2003.
25. **Chen X, and Bieker JJ**. Stage-specific repression by the EKLF transcriptional activator. *Molecular and cellular biology* 24: 10416-10424, 2004.

26. **Chen YT, Collins LL, Uno H, Chou SM, Meshul CK, Chang SS, and Chang C**. Abnormal cerebellar cytoarchitecture and impaired inhibitory signaling in adult mice lacking TR4 orphan nuclear receptor. *Brain Res* 1168: 72-82, 2007.
27. **Chinpaisal C, Lee C-H, and Wei L-N**. Mechanisms of the Mouse Orphan Nuclear Receptor TR2-11-mediated Gene Suppression. 1998, p. 18077-18085.
28. **Choi OR, and Engel JD**. Developmental regulation of beta-globin gene switching. *Cell* 55: 17-26, 1988.
29. **Christman JK**. 5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. *Oncogene* 21: 5483-5495, 2002.
30. **Cihak A**. Biological effects of 5-azacytidine in eukaryotes. *Oncology* 30: 405-422, 1974.
31. **Colah RB, Surve R, Sawant P, D'Souza E, Italia K, Phanasgaonkar S, Nadkarni AH, and Gorakshakar AC**. HPLC studies in hemoglobinopathies. *Indian journal of pediatrics* 74: 657-662, 2007.
32. **Collins FS, Metherall JE, Yamakawa M, Pan J, Weissman SM, and Forget BG**. A point mutation in the A gamma-globin gene promoter in Greek hereditary persistence of fetal haemoglobin. *Nature* 313: 325-326, 1985.
33. **Collins LL, Lee Y-F, Heinlein CA, Liu N-C, Chen Y-T, Shyr C-R, Meshul CK, Uno H, Platt KA, and Chang C**. Growth retardation and abnormal maternal behavior in mice lacking testicular orphan nuclear receptor 4. 2004, p. 15058-15063.
34. **Collins LL, Lin DL, Mu XM, and Chang C**. Feedback regulation between orphan nuclear receptor TR2 and human papilloma virus type 16. *The Journal of biological chemistry* 276: 27316-27321, 2001.
35. **Davie JR**. Inhibition of histone deacetylase activity by butyrate. *The Journal of nutrition* 133: 2485S-2493S, 2003.
36. **de Laat W, and Grosveld F**. Spatial organization of gene expression: the active chromatin hub. *Chromosome Res* 11: 447-459, 2003.
37. **Demers C, Chaturvedi CP, Ranish JA, Juban G, Lai P, Morle F, Aebersold R, Dilworth FJ, Groudine M, and Brand M**. Activator-mediated recruitment of the MLL2 methyltransferase complex to the beta-globin locus. *Mol Cell* 27: 573-584, 2007.
38. **DeSimone J, Heller P, Hall L, and Zwiers D**. 5-Azacytidine stimulates fetal hemoglobin synthesis in anemic baboons. *Proceedings of the National Academy of Sciences of the United States of America* 79: 4428-4431, 1982.
39. **DeSimone J, Heller P, Schimenti JC, and Duncan CH**. Fetal hemoglobin production in adult baboons by 5-azacytidine or by phenylhydrazine-induced hemolysis is associated with hypomethylation of globin gene DNA. *Progress in clinical and biological research* 134: 489-500, 1983.
40. **DeSimone J, Schimenti JC, Duncan CH, Heller P, and Lavelle D**. DNA methylation and expression of baboon globin genes. *Progress in clinical and biological research* 165: 59-72, 1984.
41. **Dillon N, Trimborn T, Strouboulis J, Fraser P, and Grosveld F**. The effect of distance on long-range chromatin interactions. *Mol Cell* 1: 131-139, 1997.
42. **Drissen R, Palstra RJ, Gillemans N, Splinter E, Grosveld F, Philipsen S, and de Laat W**. The active spatial organization of the beta-globin locus requires the transcription factor EKLF. *Genes & development* 18: 2485-2490, 2004.
43. **Duan Z, Stamatoyannopoulos G, and Li Q**. Role of NF-Y in in vivo regulation of the gamma-globin gene. *Molecular and cellular biology* 21: 3083-3095, 2001.
44. **Duan ZJ, Fang X, Rohde A, Han H, Stamatoyannopoulos G, and Li Q**. Developmental specificity of recruitment of TBP to the TATA box of the human gamma-globin gene. *Proceedings of the National Academy of Sciences of the United States of America* 99: 5509-5514, 2002.
45. **Enver T, Raich N, Ebens AJ, Papayannopoulou T, Costantini F, and Stamatoyannopoulos G**. Developmental regulation of human fetal-to-adult globin gene switching in transgenic mice. *Nature* 344: 309-313, 1990.
46. **Enver T, Zhang JW, Papayannopoulou T, and Stamatoyannopoulos G**. DNA methylation: a secondary event in globin gene switching? *Genes & development* 2: 698-706, 1988.
47. **Fang X, Sun J, Xiang P, Yu M, Navas PA, Peterson KR, Stamatoyannopoulos G, and Li Q**. Synergistic and additive properties of the beta-globin locus control region (LCR) revealed by 5'HS3 deletion mutations: implication for LCR chromatin architecture. *Molecular and cellular biology* 25: 7033-7041, 2005.
48. **Fang X, Xiang P, Yin W, Stamatoyannopoulos G, and Li Q**. Cooperativeness of the higher chromatin structure of the beta-globin locus revealed by the deletion mutations of DNase I hypersensitive site 3 of the LCR. *Journal of molecular biology* 365: 31-37, 2007.
49. **Fathallah H, and Atweh GF**. DNA hypomethylation therapy for hemoglobin disorders: molecular mechanisms and clinical applications. *Blood reviews* 20: 227-234, 2006.
50. **Fathallah H, and Atweh GF**. Induction of fetal hemoglobin in the treatment of sickle cell disease. *Hematology / the Education Program of the American Society of Hematology American Society of Hematology* 58-62, 2006.
51. **Fathallah H, Portnoy G, and Atweh GF**. Epigenetic analysis of the human alpha- and beta-globin gene clusters. *Blood cells, molecules & diseases* 2007.

52. **Fathallah H, Sutton M, and Atweh GF**. Pharmacological induction of fetal hemoglobin: Why haven't we been more successful in thalassemia? *Annals of the New York Academy of Sciences* 1054: 228-237, 2005.
53. **Filipe A, Li Q, Deveaux S, Godin I, Romeo PH, Stamatoyannopoulos G, and Mignotte V**. Regulation of embryonic/fetal globin genes by nuclear hormone receptors: a novel perspective on hemoglobin switching. *The EMBO journal* 18: 687-697, 1999.
54. **Forsberg EC, Downs KM, Christensen HM, Im H, Nuzzi PA, and Bresnick EH**. Developmentally dynamic histone acetylation pattern of a tissue-specific chromatin domain. *Proceedings of the National Academy of Sciences of the United States of America* 97: 14494-14499, 2000.
55. **Franco PJ, Farooqui M, Seto E, and Wei LN**. The orphan nuclear receptor TR2 interacts directly with both class I and class II histone deacetylases. *Molecular endocrinology (Baltimore, Md* 15: 1318-1328, 2001.
56. **Franco PJ, Li G, and Wei LN**. Interaction of nuclear receptor zinc finger DNA binding domains with histone deacetylase. *Molecular and cellular endocrinology* 206: 1-12, 2003.
57. **Gaensler KM, Kitamura M, and Kan YW**. Germ-line transmission and developmental regulation of a 150-kb yeast artificial chromosome containing the human beta-globin locus in transgenic mice. *Proceedings of the National Academy of Sciences of the United States of America* 90: 11381-11385, 1993.
58. **Goren A, Simchen G, Fibach E, Szabo PE, Tanimoto K, Chakalova L, Pfeifer GP, Fraser PJ, Engel JD, and Cedar H**. Fine tuning of globin gene expression by DNA methylation. *PLoS ONE* 1: e46, 2006.
59. **Gribnau J, de Boer E, Trimborn T, Wijgerde M, Milot E, Grosveld F, and Fraser P**. Chromatin interaction mechanism of transcriptional control in vivo. *The EMBO journal* 17: 6020-6027, 1998.
60. **Grosveld F**. Activation by locus control regions? *Current opinion in genetics & development* 9: 152-157, 1999.
61. **Grosveld F, van Assendelft GB, Greaves DR, and Kollias G**. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell* 51: 975-985, 1987.
62. **Gupta P, Park SW, Farooqui M, and Wei LN**. Orphan nuclear receptor TR2, a mediator of preadipocyte proliferation, is differentially regulated by RA through exchange of coactivator PCAF with corepressor RIP140 on a platform molecule GRIP1. 2007, p. 2269-2282.
63. **Hanscombe O, Vidal M, Kaeda J, Luzzatto L, Greaves DR, and Grosveld F**. High-level, erythroid-specific expression of the human alpha-globin gene in transgenic mice and the production of human hemoglobin in murine erythrocytes. *Genes & development* 3: 1572-1581, 1989.
64. **Hanscombe O, Whyatt D, Fraser P, Yannoutsos N, Greaves D, Dillon N, and Grosveld F**. Importance of globin gene order for correct developmental expression. *Genes & development* 5: 1387-1394, 1991.
65. **Harada H, Kuboi Y, Miki R, Honda C, Masushige S, Nakatsuka M, Koga Y, and Kato S**. Cloning of Rabbit TR4 and Its Bone Cell-Specific Activity to Suppress Estrogen Receptor-Mediated Transactivation. 1998, p. 204-212.
66. **Hardison R, Slightom JL, Gumucio DL, Goodman M, Stojanovic N, and Miller W**. Locus control regions of mammalian beta-globin gene clusters: combining phylogenetic analyses and experimental results to gain functional insights. *Gene* 205: 73-94, 1997.
67. **Harju S, McQueen KJ, and Peterson KR**. Chromatin structure and control of beta-like globin gene switching. *Experimental biology and medicine (Maywood, NJ* 227: 683-700, 2002.
68. **Heller P, and DeSimone J**. 5-Azacytidine and fetal hemoglobin. *American journal of hematology* 17: 439-447, 1984.
69. **Hirose T, Apfel R, Pfahl M, and Jetten AM**. The orphan receptor TAK1 acts as a repressor of RAR-, RXR- and T3R-mediated signaling pathways. *Biochem Biophys Res Commun* 211: 83-91, 1995.
70. **Hu R, Wu W, Niles EG, and LoVerde PT**. SmTR2/4, a Schistosoma mansoni homologue of TR2/TR4 orphan nuclear receptor. *International journal for parasitology* 36: 1113-1122, 2006.
71. **Huq MDM, Gupta P, Tsai N-P, and Wei L-N**. Modulation of Testicular Receptor 4 Activity by Mitogen-activated Protein Kinase-mediated Phosphorylation. 2006, p. 2072-2082.
72. **Hwang SB, Burbach JP, and Chang C**. TR4 orphan receptor crosstalks to chicken ovalbumin upstream protein-transcription factor and thyroid hormone receptor to induce the transcriptional activity of the human immunodeficiency virus type 1 long-terminal repeat. *Endocrine* 8: 169-175, 1998.
73. **Ikonen T, Palvimo JJ, and Janne OA**. Interaction between the amino- and carboxyl-terminal regions of the rat androgen receptor modulates transcriptional activity and is influenced by nuclear receptor coactivators. *The Journal of biological chemistry* 272: 29821-29828, 1997.
74. **Im H, Park C, Feng Q, Johnson KD, Kiekhaefer CM, Choi K, Zhang Y, and Bresnick EH**. Dynamic regulation of histone H3 methylated at lysine 79 within a tissue-specific chromatin domain. *The Journal of biological chemistry* 278: 18346-18352, 2003.
75. **Ingle J, Adewoye A, Dewan R, Okoli M, Rollins L, Eung SH, Luo HY, Chui DH, and Steinberg MH**. Hb Hope [beta136(H14)Gly-->Asp (GGT-->GAT)]: interactions with Hb S [beta6(A3)Glu-->Val (GAG-->GTG)], other variant hemoglobins and thalassemia. *Hemoglobin* 28: 277-285, 2004.

76. **Inui S, Lee Y-F, Haake AR, Goldsmith LA, and Chang C**. Induction of TR4 Orphan Receptor by Retinoic Acid in Human HaCaT Keratinocytes. 112: 426-431, 1999.
77. **Juttermann R, Li E, and Jaenisch R**. Toxicity of 5-aza-2'-deoxycytidine to mammalian cells is mediated primarily by covalent trapping of DNA methyltransferase rather than DNA demethylation. *Proceedings of the National Academy of Sciences of the United States of America* 91: 11797-11801, 1994.
78. **Katsantoni EZ, Langeveld A, Wai AW, Drabek D, Grosveld F, Anagnou NP, and Strouboulis J**. Persistent gamma-globin expression in adult transgenic mice is mediated by HPFH-2, HPFH-3, and HPFH-6 breakpoint sequences. *Blood* 102: 3412-3419, 2003.
79. **Khan SA, Park SW, Huq M, and Wei LN**. Protein kinase C-mediated phosphorylation of orphan nuclear receptor TR2: effects on receptor stability and activity. *Proteomics* 5: 3885-3894, 2005.
80. **Khan SA, Park SW, Huq MD, and Wei LN**. Ligand-independent orphan receptor TR2 activation by phosphorylation at the DNA-binding domain. *Proteomics* 6: 123-130, 2006.
81. **Kim A, and Dean A**. Developmental stage differences in chromatin subdomains of the beta-globin locus. *Proceedings of the National Academy of Sciences of the United States of America* 101: 7028-7033, 2004.
82. **Kim E, Ma W-L, Lin D-L, Inui S, Chen Y-L, and Chang C**. TR4 orphan nuclear receptor functions as an apoptosis modulator via regulation of Bcl-2 gene expression. *Biochemical and Biophysical Research Communications* 361: 323-328, 2007.
83. **Kioussis D, Vanin E, deLange T, Flavell RA, and Grosveld FG**. Beta-globin gene inactivation by DNA translocation in gamma beta-thalassaemia. *Nature* 306: 662-666, 1983.
84. **Koritschoner NP, Madruga J, Knespel S, Blendinger G, Anzinger B, Otto A, Zenke M, and Bartunek P**. The nuclear orphan receptor TR4 promotes proliferation of myeloid progenitor cells. *Cell Growth Differ* 12: 563-572, 2001.
85. **Koritschoner NP, Madruga J, Knespel S, Blendinger G, Anzinger B, Otto A, Zenke M, and Bartunek P**. The Nuclear Orphan Receptor TR4 Promotes Proliferation of Myeloid Progenitor Cells. 2001, p. 563-572.
86. **Lavelle D, DeSimone J, and Heller P**. Fetal hemoglobin reactivation in baboon and man: a short perspective. *American journal of hematology* 42: 91-95, 1993.
87. **Lee C-H, Chinpaisal C, and Wei L-N**. Cloning and Characterization of Mouse RIP140, a Corepressor for Nuclear Orphan Receptor TR2. 1998, p. 6745-6755.
88. **Lee H-J, Lee Y, Burbach JPH, and Chang C**. Suppression of Gene Expression on the Simian Virus 40 Major Late Promoter by Human TR4 Orphan Receptor. 1995, p. 30129-30133.
89. **Lee HJ, and Chang C**. Identification of human TR2 orphan receptor response element in the transcriptional initiation site of the simian virus 40 major late promoter. *The Journal of biological chemistry* 270: 5434-5440, 1995.
90. **Lee HJ, Lee YF, and Chang C**. Identification of the histamine H1 receptor gene as a differentially repressed target of the human TR2 orphan receptor. *Molecular and cellular biochemistry* 194: 199-207, 1999.
91. **Lee HJ, Lee YF, and Chang C**. TR4 orphan receptor represses the human steroid 21-hydroxylase gene expression through the monomeric AGGTCA motif. *Biochem Biophys Res Commun* 285: 1361-1368, 2001.
92. **Lee HJ, Young WJ, Shih CY, and Chang C**. Suppression of the human erythropoietin gene expression by the TR2 orphan receptor, a member of the steroid receptor superfamily. *The Journal of biological chemistry* 271: 10405-10412, 1996.
93. **Lee JS, Ngo H, Kim D, and Chung JH**. Erythroid Kruppel-like factor is recruited to the CACCC box in the beta-globin promoter but not to the CACCC box in the gamma-globin promoter: the role of the neighboring promoter elements. *Proceedings of the National Academy of Sciences of the United States of America* 97: 2468-2473, 2000.
94. **Lee Y-F, Pan H-J, Burbach JPH, Morkin E, and Chang C**. Identification of Direct Repeat 4 as a Positive Regulatory Element for the Human TR4 Orphan Receptor. A MODULATOR FOR THE THYROID HORMONE TARGET GENES. 1997, p. 12215-12220.
95. **Lee Y-F, Young W-J, Burbach JPH, and Chang C**. Negative Feedback Control of the Retinoid-Retinoic Acid/Retinoid X Receptor Pathway by the Human TR4 Orphan Receptor, a Member of the Steroid Receptor Superfamily. 1998, p. 13437-13443.
96. **Lee Y-F, Young W-J, Lin W-J, Shyr C-R, and Chang C**. Differential Regulation of Direct Repeat 3 Vitamin D3 and Direct Repeat 4 Thyroid Hormone Signaling Pathways by the Human TR4 Orphan Receptor. 1999, p. 16198-16205.
97. **Lee YF, Lee HJ, and Chang C**. Recent advances in the TR2 and TR4 orphan receptors of the nuclear receptor superfamily. *The Journal of steroid biochemistry and molecular biology* 81: 291-308, 2002.
98. **Lee YF, Shyr CR, Thin TH, Lin WJ, and Chang C**. Convergence of two repressors through heterodimer formation of androgen receptor and testicular orphan receptor-4: a unique signaling pathway in the steroid receptor superfamily. *Proceedings of the National Academy of Sciences of the United States of America* 96: 14724-14729, 1999.
99. **Lemischka IR, Raulet DH, and Mulligan RC**. Developmental potential and dynamic behavior of hematopoietic stem cells. *Cell* 45: 917-927, 1986.

100. **Ley TJ, Anagnou NP, Young NS, Nienhuis AW, DeSimone J, and Heller P**. 5-azacytidine for beta thalassaemia? *Lancet* 1: 467, 1983.
101. **Ley TJ, DeSimone J, Noguchi CT, Turner PH, Schechter AN, Heller P, and Nienhuis AW**. 5-Azacytidine increases gamma-globin synthesis and reduces the proportion of dense cells in patients with sickle cell anemia. *Blood* 62: 370-380, 1983.
102. **Li Q, Fang X, Han H, and Stamatoyannopoulos G**. The minimal promoter plays a major role in silencing of the galago gamma-globin gene in adult erythropoiesis. *Proceedings of the National Academy of Sciences of the United States of America* 101: 8096-8101, 2004.
103. **Li Q, Peterson KR, and Stamatoyannopoulos G**. Developmental control of epsilon- and gamma-globin genes. *Annals of the New York Academy of Sciences* 850: 10-17, 1998.
104. **Liakopoulou E, Blau CA, Li Q, Josephson B, Wolf JA, Fournarakis B, Raisys V, Dover G, Papayannopoulou T, and Stamatoyannopoulos G**. Stimulation of fetal hemoglobin production by short chain fatty acids. *Blood* 86: 3227-3235, 1995.
105. **Lin DL, and Chang C**. Identification of an essential cis-acting element (TR2-PACE) in the 5' promoter of human TR2 orphan receptor gene. *Endocrine* 12: 89-97, 2000.
106. **Lin DL, and Chang C**. p53 is a mediator for radiation-repressed human TR2 orphan receptor expression in MCF-7 cells, a new pathway from tumor suppressor to member of the steroid receptor superfamily. *The Journal of biological chemistry* 271: 14649-14652, 1996.
107. **Lin TM, Young WJ, and Chang C**. Multiple functions of the TR2-11 orphan receptor in modulating activation of two key cis-acting elements involved in the retinoic acid signal transduction system. *The Journal of biological chemistry* 270: 30121-30128, 1995.
108. **Lopez RA, Schoetz S, DeAngelis K, O'Neill D, and Bank A**. Multiple hematopoietic defects and delayed globin switching in Ikaros null mice. *Proceedings of the National Academy of Sciences of the United States of America* 99: 602-607, 2002.
109. **Loukopoulos D, Voskaridou E, Kalotychou V, Schina M, Loutradi A, and Theodoropoulos I**. Reduction of the clinical severity of sickle cell/beta-thalassemia with hydroxyurea: the experience of a single center in Greece. *Blood cells, molecules & diseases* 26: 453-466, 2000.
110. **Loukopoulos D, Voskaridou E, Stamoulakatou A, Papassotiriou Y, Kalotychou V, Loutradi A, Cozma G, Tsiarta H, and Pavlides N**. Hydroxyurea therapy in thalassemia. *Annals of the New York Academy of Sciences* 850: 120-128, 1998.
111. **Mabaera R, Greene MR, Richardson CA, Conine SJ, Kozul CD, and Lowrey CH**. Neither DNA hypomethylation nor changes in the kinetics of erythroid differentiation explain 5-azacytidine's ability to induce human fetal hemoglobin. 2007, p. blood-2007-2006-093948.
112. **Mankidy R, Faller DV, Mabaera R, Lowrey CH, Boosalis MS, White GL, Castaneda SA, and Perrine SP**. Short-chain fatty acids induce gamma-globin gene expression by displacement of a HDAC3-NCoR repressor complex. *Blood* 108: 3179-3186, 2006.
113. **Martin DI, Tsai SF, and Orkin SH**. Increased gamma-globin expression in a nondeletion HPFH mediated by an erythroid-specific DNA-binding factor. *Nature* 338: 435-438, 1989.
114. **Merika M, and Orkin SH**. DNA-binding specificity of GATA family transcription factors. *Molecular and cellular biology* 13: 3999-4010, 1993.
115. **Miles J, Mitchell JA, Chakalova L, Goyenechea B, Osborne CS, O'Neill L, Tanimoto K, Engel JD, and Fraser P**. Intergenic transcription, cell-cycle and the developmentally regulated epigenetic profile of the human beta-globin locus. *PLoS ONE* 2: e630, 2007.
116. **Miller IJ, and Bieker JJ**. A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Kruppel family of nuclear proteins. *Molecular and cellular biology* 13: 2776-2786, 1993.
117. **Milot E, Strouboulis J, Trimborn T, Wijgerde M, de Boer E, Langeveld A, Tan-Un K, Vergeer W, Yannoutsos N, Grosveld F, and Fraser P**. Heterochromatin effects on the frequency and duration of LCR-mediated gene transcription. *Cell* 87: 105-114, 1996.
118. **Nakajima T, Fujino S, Nakanishi G, Kim Y-S, and Jetten AM**. TIP27: a novel repressor of the nuclear orphan receptor TAK1/TR4. 2004, p. 4194-4204.
119. **Navas PA, Peterson KR, Li Q, Skarpidi E, Rohde A, Shaw SE, Clegg CH, Asano H, and Stamatoyannopoulos G**. Developmental specificity of the interaction between the locus control region and embryonic or fetal globin genes in transgenic mice with an HS3 core deletion. *Molecular and cellular biology* 18: 4188-4196, 1998.
120. **O'Neill D, Bornschlegel K, Flamm M, Castle M, and Bank A**. A DNA-binding factor in adult hematopoietic cells interacts with a pyrimidine-rich domain upstream from the human delta-globin gene. *Proceedings of the National Academy of Sciences of the United States of America* 88: 8953-8957, 1991.
121. **O'Neill D, Yang J, Erdjument-Bromage H, Bornschlegel K, Tempst P, and Bank A**. Tissue-specific and

developmental stage-specific DNA binding by a mammalian SWI/SNF complex associated with human fetal-to-adult globin gene switching. *Proceedings of the National Academy of Sciences of the United States of America* 96: 349-354, 1999.

122. **O'Neill DW, Schoetz SS, Lopez RA, Castle M, Rabinowitz L, Shor E, Krawchuk D, Goll MG, Renz M, Seelig HP, Han S, Seong RH, Park SD, Agalioti T, Munshi N, Thanos D, Erdjument-Bromage H, Tempst P, and Bank A**. An ikaros-containing chromatin-remodeling complex in adult-type erythroid cells. *Molecular and cellular biology* 20: 7572-7582, 2000.

123. **Olivieri NF, Rees DC, Ginder GD, Thein SL, Brittenham GM, Waye JS, and Weatherall DJ**. Treatment of thalassaemia major with phenylbutyrate and hydroxyurea. *Lancet* 350: 491-492, 1997.

124. **Olivieri NF, Rees DC, Ginder GD, Thein SL, Waye JS, Chang L, Brittenham GM, and Weatherall DJ**. Elimination of transfusions through induction of fetal hemoglobin synthesis in Cooley's anemia. *Annals of the New York Academy of Sciences* 850: 100-109, 1998.

125. **Olivieri NF, and Weatherall DJ**. The therapeutic reactivation of fetal haemoglobin. *Human molecular genetics* 7: 1655-1658, 1998.

126. **Omori A, Tanabe O, Engel JD, Fukamizu A, and Tanimoto K**. Adult stage gamma-globin silencing is mediated by a promoter direct repeat element. *Molecular and cellular biology* 25: 3443-3451, 2005.

127. **Orkin SH**. Globin gene regulation and switching: circa 1990. *Cell* 63: 665-672, 1990.

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

129. **Park SW, Hu X, Gupta P, Lin YP, Ha SG, and Wei LN**. SUMOylation of Tr2 orphan receptor involves Pml and fine-tunes Oct4 expression in stem cells. *Nature structural & molecular biology* 14: 68-75, 2007.

130. **Patrinos GP, de Krom M, de Boer E, Langeveld A, Imam AM, Strouboulis J, de Laat W, and Grosveld FG**. Multiple interactions between regulatory regions are required to stabilize an active chromatin hub. *Genes & development* 18: 1495-1509, 2004.

131. **Perkins AC, Gaensler KM, and Orkin SH**. Silencing of human fetal globin expression is impaired in the absence of the adult beta-globin gene activator protein EKLF. *Proceedings of the National Academy of Sciences of the United States of America* 93: 12267-12271, 1996.

132. **Perrine SP, Rudolph A, Faller DV, Roman C, Cohen RA, Chen SJ, and Kan YW**. Butyrate infusions in the ovine fetus delay the biologic clock for globin gene switching. *Proceedings of the National Academy of Sciences of the United States of America* 85: 8540-8542, 1988.

133. **Peterson KR, Li QL, Clegg CH, Furukawa T, Navas PA, Norton EJ, Kimbrough TG, and Stamatoyannopoulos G**. Use of yeast artificial chromosomes (YACs) in studies of mammalian development: production of beta-globin locus YAC mice carrying human globin developmental mutants. *Proceedings of the National Academy of Sciences of the United States of America* 92: 5655-5659, 1995.

134. **Porcu S, Kitamura M, Witkowska E, Zhang Z, Mutero A, Lin C, Chang J, and Gaensler KM**. The human beta globin locus introduced by YAC transfer exhibits a specific and reproducible pattern of developmental regulation in transgenic mice. *Blood* 90: 4602-4609, 1997.

135. **Raich N, Clegg CH, Grofti J, Romeo PH, and Stamatoyannopoulos G**. GATA1 and YY1 are developmental repressors of the human epsilon-globin gene. *The EMBO journal* 14: 801-809, 1995.

136. **Raich N, Enver T, Nakamoto B, Josephson B, Papayannopoulou T, and Stamatoyannopoulos G**. Autonomous developmental control of human embryonic globin gene switching in transgenic mice. *Science (New York, NY* 250: 1147-1149, 1990.

137. **Rupon JW, Wang SZ, Gaensler K, Lloyd J, and Ginder GD**. Methyl binding domain protein 2 mediates gamma-globin gene silencing in adult human betaYAC transgenic mice. *Proceedings of the National Academy of Sciences of the United States of America* 103: 6617-6622, 2006.

138. **Russell JE**. A post-transcriptional process contributes to efficient gamma-globin gene silencing in definitive erythroid cells. *European journal of haematology* 79: 516-525, 2007.

139. **Shih DM, Wall RJ, and Shapiro SG**. Developmentally regulated and erythroid-specific expression of the human embryonic beta-globin gene in transgenic mice. *Nucleic acids research* 18: 5465-5472, 1990.

140. **Shimotsuma M, Matsuzaki H, Tanabe O, Campbell AD, Engel JD, Fukamizu A, and Tanimoto K**. Linear distance from the locus control region determines epsilon-globin transcriptional activity. *Molecular and cellular biology* 27: 5664-5672, 2007.

141. **Shyr C-R, Collins LL, Mu X-M, Platt KA, and Chang C**. Spermatogenesis and Testis Development Are Normal in Mice Lacking Testicular Orphan Nuclear Receptor 2. 2002, p. 4661-4666.

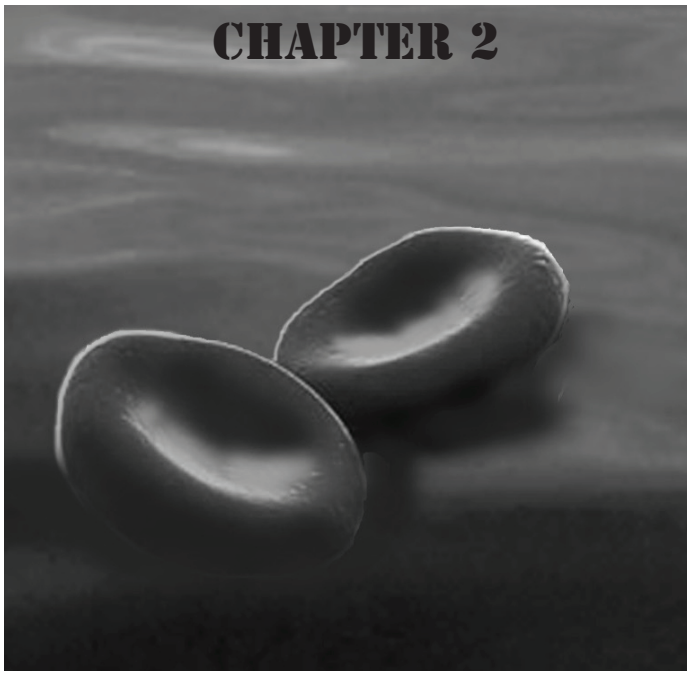
142. **Singer ST, Kuypers FA, Olivieri NF, Weatherall DJ, Mignacca R, Coates TD, Davies S, Sweeters N, and Vichinsky EP**. Fetal haemoglobin augmentation in E/beta(0) thalassaemia: clinical and haematological outcome. *British journal of haematology* 131: 378-388, 2005.

143. **Song CZ, Keller K, Murata K, Asano H, and Stamatoyannopoulos G**. Functional interaction between coactivators CBP/p300, PCAF, and transcription factor FKLF2. *The Journal of biological chemistry* 277: 7029-7036,

2002.
144. **Stamatoyannopoulos G**. Control of globin gene expression during development and erythroid differentiation. *Experimental hematology* 33: 259-271, 2005.
145. **Stamatoyannopoulos G, and Grosveld F**. *Hemoglobin Switching*. . Philadelphia: W.B. Saunders Publishing Co, 2001, p. 135-182.
146. **Stamatoyannopoulos G, Josephson B, Zhang JW, and Li Q**. Developmental regulation of human gamma-globin genes in transgenic mice. *Molecular and cellular biology* 13: 7636-7644, 1993.
147. **Starck J, Sarkar R, Romana M, Bhargava A, Scarpa AL, Tanaka M, Chamberlain JW, Weissman SM, and Forget BG**. Developmental regulation of human gamma- and beta-globin genes in the absence of the locus control region. *Blood* 84: 1656-1665, 1994.
148. **Steinlein P, Wessely O, Meyer S, Deiner EM, Hayman MJ, and Beug H**. Primary, self-renewing erythroid progenitors develop through activation of both tyrosine kinase and steroid hormone receptors. *Curr Biol* 5: 191-204, 1995.
149. **Strouboulis J, Dillon N, and Grosveld F**. Developmental regulation of a complete 70-kb human beta-globin locus in transgenic mice. *Genes & development* 6: 1857-1864, 1992.
150. **Tanabe O, Katsuoka F, Campbell AD, Song W, Yamamoto M, Tanimoto K, and Engel JD**. An embryonic/fetal beta-type globin gene repressor contains a nuclear receptor TR2/TR4 heterodimer. *The EMBO journal* 21: 3434-3442, 2002.
151. **Tanabe O, McPhee D, Kobayashi S, Shen Y, Brandt W, Jiang X, Campbell AD, Chen YT, Chang C, Yamamoto M, Tanimoto K, and Engel JD**. Embryonic and fetal beta-globin gene repression by the orphan nuclear receptors, TR2 and TR4. *The EMBO journal* 26: 2295-2306, 2007.
152. **Tanabe O, Shen Y, Liu Q, Campbell AD, Kuroha T, Yamamoto M, and Engel JD**. The TR2 and TR4 orphan nuclear receptors repress Gata1 transcription. *Genes & development* 21: 2832-2844, 2007.
153. **Tanimoto K, and Engel JD**. In vivo modulation of human beta-globin gene switching. *Trends in cardiovascular medicine* 10: 15-19, 2000.
154. **Tanimoto K, Liu Q, Bungert J, and Engel JD**. Effects of altered gene order or orientation of the locus control region on human beta-globin gene expression in mice. *Nature* 398: 344-348, 1999.
155. **Tanimoto K, Liu Q, Grosveld F, Bungert J, and Engel JD**. Context-dependent EKLF responsiveness defines the developmental specificity of the human epsilon-globin gene in erythroid cells of YAC transgenic mice. *Genes & development* 14: 2778-2794, 2000.
156. **Tolhuis B, Palstra RJ, Splinter E, Grosveld F, and de Laat W**. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell* 10: 1453-1465, 2002.
157. **Tsai S, Bartelmez S, Sitnicka E, and Collins S**. Lymphohematopoietic progenitors immortalized by a retroviral vector harboring a dominant-negative retinoic acid receptor can recapitulate lymphoid, myeloid, and erythroid development. *Genes & development* 8: 2831-2841, 1994.
158. **Tuan D, and London IM**. Mapping of DNase I-hypersensitive sites in the upstream DNA of human embryonic epsilon-globin gene in K562 leukemia cells. *Proceedings of the National Academy of Sciences of the United States of America* 81: 2718-2722, 1984.
159. **Tuan D, Solomon W, Li Q, and London IM**. The "beta-like-globin" gene domain in human erythroid cells. *Proceedings of the National Academy of Sciences of the United States of America* 82: 6384-6388, 1985.
160. **van Schaick HS, Rosmalen JG, Lopes da Silva S, Chang C, and Burbach JP**. Expression of the orphan receptor TR4 during brain development of the rat. *Brain research* 77: 104-110, 2000.
161. **Wai AW, Gillemans N, Raguz-Bolognesi S, Pruzina S, Zafarana G, Meijer D, Philipsen S, and Grosveld F**. HS5 of the human beta-globin locus control region: a developmental stage-specific border in erythroid cells. *The EMBO journal* 22: 4489-4500, 2003.
162. **Wang C-P, Lee Y-F, Chang C, and Lee H-J**. Transactivation of the proximal promoter of human oxytocin gene by TR4 orphan receptor. *Biochemical and Biophysical Research Communications* 351: 204-208, 2006.
163. **Weatherall DJ**. The thalassaemias. *BMJ (Clinical research ed* 314: 1675-1678, 1997.
164. **Weatherall DJ, and Clegg JB**. Genetic disorders of hemoglobin. *Seminars in hematology* 36: 24-37, 1999.
165. **Weatherall DJ, and Clegg JB**. Molecular genetics of human hemoglobin. *Annual review of genetics* 10: 157-178, 1976.
166. **Weatherall DJ, and Clegg JB**. Recent developments in the molecular genetics of human hemoglobin. *Cell* 16: 467-479, 1979.
167. **Weatherall DJ, Wood WG, Jones RW, and Clegg JB**. The developmental genetics of human hemoglobin. *Progress in clinical and biological research* 191: 3-25, 1985.
168. **Wei G-H, Zhao G-W, Song W, Hao D-L, Lv X, Liu D-P, and Liang C-C**. Mechanisms of human [gamma]-globin transcriptional induction by apicidin involves p38 signaling to chromatin. *Biochemical and Biophysical Research Communications* 363: 889-894, 2007.
169. **Weinberg RS, Ji X, Sutton M, Perrine S, Galperin Y, Li Q, Liebhaber SA, Stamatoyannopoulos G, and**

Atweh GF. Butyrate increases the efficiency of translation of gamma-globin mRNA. *Blood* 105: 1807-1809, 2005.
170. **Wessely O, Deiner EM, Beug H, and von Lindern M**. The glucocorticoid receptor is a key regulator of the decision between self-renewal and differentiation in erythroid progenitors. *The EMBO journal* 16: 267-280, 1997.
171. **Wijgerde M, Grosveld F, and Fraser P**. Transcription complex stability and chromatin dynamics in vivo. *Nature* 377: 209-213, 1995.
172. **Wood WG**. Increased HbF in adult life. *Bailliere's clinical haematology* 6: 177-213, 1993.
173. **Xue L, Chen X, Chang Y, and Bieker JJ**. Regulatory elements of the EKLF gene that direct erythroid cell-specific expression during mammalian development. *Blood* 103: 4078-4083, 2004.
174. **Xue Y, Wong J, Moreno GT, Young MK, Cote J, and Wang W**. NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. *Mol Cell* 2: 851-861, 1998.
175. **Yan ZH, Karam WG, Staudinger JL, Medvedev A, Ghanayem BI, and Jetten AM**. Regulation of peroxisome proliferator-activated receptor alpha-induced transactivation by the nuclear orphan receptor TAK1/TR4. *The Journal of biological chemistry* 273: 10948-10957, 1998.
176. **Yang Y, Wang X, Dong T, Kim E, Lin W-J, and Chang C**. Identification of a Novel Testicular Orphan Receptor-4 (TR4)-associated Protein as Repressor for the Selective Suppression of TR4-mediated Transactivation. 2003, p. 7709-7717.
177. **Yang Y, Wang X, Dong T, Kim E, Lin W-J, and Chang C**. Identification of a novel testicular orphan receptor-4 (TR4) associated protein (TRA16) as repressor for the selective suppression of TR4-mediated transactivation. 2002, p. M207116200.
178. **Young W-J, Lee Y-F, Smith SM, and Chang C**. A Bidirectional Regulation between the TR2/TR4 Orphan Receptors (TR2/TR4) and the Ciliary Neurotrophic Factor (CNTF) Signaling Pathway. 1998, p. 20877-20885.
179. **Young W-J, Smith SM, and Chang C**. Induction of the Intronic Enhancer of the Human Ciliary Neurotrophic Factor Receptor (CNTFRalpha) Gene by the TR4 Orphan Receptor. A MEMBER OF STEROID RECEPTOR SUPERFAMILY. 1997, p. 3109-3116.
180. **Young WJ, Smith SM, and Chang C**. Induction of the intronic enhancer of the human ciliary neurotrophic factor receptor (CNTFRalpha) gene by the TR4 orphan receptor. A member of steroid receptor superfamily. *The Journal of biological chemistry* 272: 3109-3116, 1997.
181. **Yu K, Bayona W, Kallen CB, Harding HP, Ravera CP, McMahon G, Brown M, and Lazar MA**. Differential activation of peroxisome proliferator-activated receptors by eicosanoids. *The Journal of biological chemistry* 270: 23975-23983, 1995.
182. **Yu Z, Lee CH, Chinpaisal C, and Wei LN**. A constitutive nuclear localization signal from the second zinc-finger of orphan nuclear receptor TR2. 1998, p. 53-60.
183. **Zenke M, Kahn P, Disela C, Vennstrom B, Leutz A, Keegan K, Hayman MJ, Choi HR, Yew N, Engel JD, and et al.** v-erbA specifically suppresses transcription of the avian erythrocyte anion transporter (band 3) gene. *Cell* 52: 107-119, 1988.
184. **Zenke M, Munoz A, Sap J, Vennstrom B, and Beug H**. v-erbA oncogene activation entails the loss of hormone-dependent regulator activity of c-erbA. *Cell* 61: 1035-1049, 1990.
185. **Zhang Y, Chen Y-T, Xie S, Wang L, Lee Y-F, Chang S-S, and Chang C**. Loss of Testicular Orphan Receptor 4 Impairs Normal Myelination in Mouse Forebrain. 2007, p. 908-920.
186. **Zhang Y, and Dufau ML**. Dual mechanisms of regulation of transcription of luteinizing hormone receptor gene by nuclear orphan receptors and histone deacetylase complexes. *The Journal of steroid biochemistry and molecular biology* 85: 401-414, 2003.
187. **Zhang Y, and Dufau ML**. Nuclear Orphan Receptors Regulate Transcription of the Gene for the Human Luteinizing Hormone Receptor. 2000, p. 2763-2770.
188. **Zhao Q, Cumming H, Cerruti L, Cunningham JM, and Jane SM**. Site-specific acetylation of the fetal globin activator NF-E4 prevents its ubiquitination and regulates its interaction with the histone deacetylase, HDAC1. *The Journal of biological chemistry* 279: 41477-41486, 2004.
189. **Zhou W, Zhao Q, Sutton R, Cumming H, Wang X, Cerruti L, Hall M, Wu R, Cunningham JM, and Jane SM**. The role of p22 NF-E4 in human globin gene switching. *The Journal of biological chemistry* 279: 26227-26232, 2004.
190. **Zitnik G, Peterson K, Stamatoyannopoulos G, and Papayannopoulou T**. Effects of butyrate and glucocorticoids on gamma- to beta-globin gene switching in somatic cell hybrids. *Molecular and cellular biology* 15: 790-795, 1995.

ISOLATION OF TRANSCRIPTION FACTOR COMPLEXES BY in vivo BIOTINYLATION TAGGING AND DIRECT BINDING TO STREPTAVIDIN BEADS



Methods Mol Biol 2006

Isolation of transcription factor complexes by in vivo biotinylation tagging and direct binding to streptavidin beads

Patrick Rodriguez, Harald Braun, Katarzyna E. Kolodziej, Ernie de Boer, Jennifer Campbell, Edgar Bonte, Frank Grosveld, Sjaak Philipsen and John Strouboulis*

Department of Cell Biology, Erasmus Medical Center, P.O. Box 1738, 3000DR Rotterdam, The Netherlands

* Corresponding author e-mail: i.strouboulis@erasmusmc.nl

Abstract

Efficient tagging methodologies are an integral aspect of protein complex characterization by proteomic approaches. Due to biotin's very high affinity for avidin and streptavidin, biotinylation tagging offers an attractive approach for the efficient purification of protein complexes. The very high affinity of the biotin/(strept)avidin system also offers the potential for the single-step capture of lower abundance protein complexes, such as transcription factor complexes. The identification of short peptide tags that are efficiently biotinylated by the bacterial BirA biotin ligase, led to an approach for the single-step purification of transcription factor complexes by specific in vivo biotinylation tagging. A short sequence tag fused N-terminally to the transcription factor of interest is very efficiently biotinylated by BirA co-expressed in the same cells, as was demonstrated by the tagging of the essential hematopoietic transcription factor GATA-1. The direct binding to streptavidin of biotinylated GATA-1 in nuclear extracts resulted in the single-step capture of the tagged factor and associated proteins, which were eluted and identified by mass spectrometry. This led to the characterization of several distinct GATA-1 complexes with other transcription factors and chromatin remodeling co-factors, which are involved in activation and repression of gene targets. Thus, BirA-mediated tagging is an efficient approach for the direct capture and characterization of transcription factor complexes.

Key words: biotinylation tagging; BirA; transcription factors; chromatin; mass spectrometry; size fractionation; GATA-1

1. Introduction

Completion of the sequencing of an ever-increasing number of genomes has led to a shift of focus towards the characterization of the protein complement of cells, i.e. the proteome. A key aspect of proteomic analysis is the development of simple methodologies for the efficient isolation of protein complexes for peptide analysis and identification by powerful mass spectrometric approaches. This is particularly challenging for the analysis of nuclear proteins involved in transcriptional regulation such as transcription factors and their chromatin associated co-factors due to their relatively lower abundance, the different parallel functions that they execute (e.g. activation and repression involving different partners) and the often transient nature of their interactions. Transcription factor purification approaches involving several pre-purification steps are laborious and costly and most likely result in the isolation of only the most abundant of the protein complexes formed by the factor. We describe here the application of in vivo biotinylation tagging as a simple approach for the efficient direct purification of transcription factor complexes from crude nuclear extracts (1).

Biotin is a naturally occurring cofactor essential for certain metabolic enzymes such as carboxylases. Specific protein-biotin ligases are responsible for covalently attaching biotin to these enzymes. The key to using biotinylation lies in the fact that biotinylated substrates can be bound very tightly by the naturally occurring proteins avidin and streptavidin ($K_d = 10^{-15}$), a fact that has been widely exploited in many affinity-based biochemical applications. In addition, in vivo biotinylation tagging offers a number of advantages for protein purification purposes. Firstly, there are few naturally biotinylated proteins (mostly cytoplasmic and mitochondrial) ensuring that non-specific background binding remains low. Secondly, the very high affinity of (strept)avidin for biotin allows high stringencies to be employed during purification without fear of losing binding of the tagged protein.

The biotinylation tagging approach described here is based on previous work on the screening of a combinatorial synthetic peptide library for efficient biotinylation by the bacterial BirA biotin ligase (2). This led to the identification of a number of short sequence tags that can be very efficiently biotinylated in vitro (2, 3). Such tags were subsequently utilized for the efficient in vivo biotinylation of tagged proteins in bacterial cells through the co-expression of BirA (4, 5). We have applied this approach in mammalian cells and demonstrated its efficiency in specifically biotinylating nuclear proteins in cultured cells and transgenic mice through the co-expression of the BirA biotin ligase together with the tagged protein (Figure 1A) (1). Our work is focused primarily on the biotinylation tagging of hematopoietic transcription factors in erythroid cells. Most of our work to date has been carried out with GATA-1, a critical transcription factor for erythroid cell differentiation. We have been able to very efficiently biotinylate GATA-1 in cultured mouse erythroleukemic cells (Figure 1B and C; (1)) leading to the isolation and characterization of GATA-1 protein complexes by direct binding of nuclear extracts to streptavidin beads. Using this approach we identified a number of GATA-1 complexes, containing other essential hematopoietic transcription factors (FOG-1, Gfi-1b and TAL-1) and chromatin remodeling and modification complexes (MeCP1 and ACF/WCRF). These complexes were implicated in the transcriptional activation and repression of different subsets of target genes (6). Thus, biotinylation tagging has proven to be a very efficient method for the single-step purification and characterization of transcription factor complexes. It should also be noted that we have no evidence so far that biotinylation tagging adversely affects the physiological properties of the tagged protein (1).

In this chapter we describe protocols for the binding of nuclear extracts expressing a specific biotin-tagged protein to streptavidin beads and the preparation of the eluted material for analysis by mass spectrometry. We do not provide protocols for the stable transfection of cultured cells as these will vary depending on the cell line/type used in each case. We routinely prepare large scale nuclear extracts from a few liters of cultured cells, we test for the presence of the biotin tagged

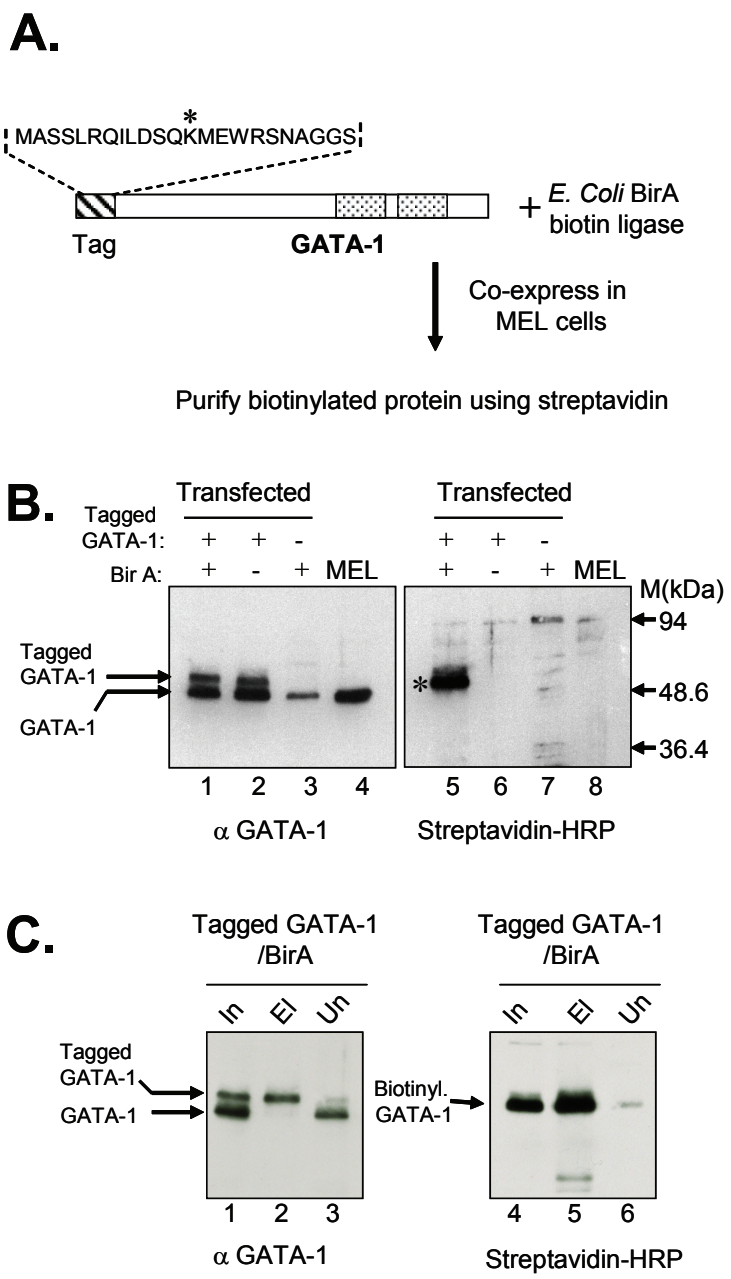


Figure 1: (A) Scheme for the specific biotinylation of tagged GATA-1 by BirA biotin ligase in mouse erythroleukemic (MEL) cells. The sequence of the 23aa peptide tag fused to the N-terminus of GATA-1 is shown. The asterisk indicates the lysine residue that becomes specifically biotinylated by BirA. Speckled boxes indicate the positions of the two GATA-1 Zinc-fingers. (B) Biotinylation of tagged GATA-1 in MEL cells. Left panel: Western blot with an anti-GATA-1antibody to detect endogenous and tagged GATA-1 proteins. Right panel: Western blot of the same extracts with streptavidin-HRP conjugate to detect biotinylated GATA-1. Biotinylated GATA-1 (asterisk) is clearly visible in the right panel only in the lane of the double transfected cells. (C) Efficiency of GATA-1 biotinylation and binding to streptavidin beads. Left panel: Western blot using anti-GATA-1 antibody to detect binding of tagged GATA-1 to streptavidin beads Input and unbound material are shown in lanes 1 and 3. Right panel: the same filter stripped and re-probed with streptavidin-HRP to detect the binding of biotinylated GATA-1 to streptavidin beads (lane 5). Lane 6 shows that very little tagged GATA-1 remains unbound by streptavidin. In: Input (nuclear extract); El: Eluted material; Un: Unbound material. Reproduced with permission from (1). Copyright (2003) National Academy of Sciences, USA.

protein in high molecular weight complex(es) by gel filtration using a Superose 6 column and then carry out the binding of the tagged factor to streptavidin paramagnetic beads.

We normally check the efficiency of the biotin tagging and the binding to streptavidin beads

by testing the nuclear extract (input), the bound material (eluate) and the flowthrough (unbound) by Western blotting using first an antibody against the tagged protein followed, by streptavidin-HRP

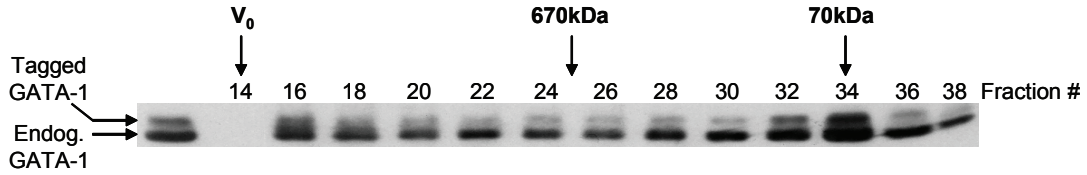


Figure 2: Superose 6 fractionation profiles of nuclear extracts from MEL cells expressing biotin-tagged GATA-1 detected by Western blotting. The tagged GATA-1 protein is migrating with a slower mobility to that of the endogenous GATA-1 due to the extra tag sequences fused to it. The elution of the molecular weight markers is indicated at the top. V_0: void volume.

A.

```
GGCCGCCATC ATG GCT GGT GGC CTG AAT GAC ATC TTT GAG GCC CAG AAG ATC GAG TGG CAT GAG AAC CTG TAC TTC CAG GGA GCC
           MET Ala Gly Gly Leu Asn Asp Ile Phe Glu Ala Gln Lys Ile Glu Trp His Glu Asn Leu Tyr Phe Gln Gly Ala
<----------------->
```

Kozak sequence **Biotinylation tag** **TEV protease cleavage site**

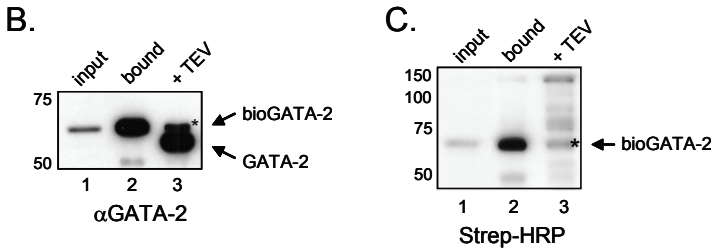


Figure 3: TEV protease cleavage of bioGATA-2 bound to streptavidin beads. Panel (A): sequence and translation of the shorter biotinylation tag and the TEV protease cleavage site. Panels (B) and (C): Lane 1: nuclear extract from MEL cells expressing tagged GATA-2 (there is no endogenous GATA-2 expressed in MEL cells). Lane 2: biotinylated GATA-2 (bioGATA-2) bound to the beads. Lane 3: TEV protease cleavage of bound GATA-2. (A) Detection with anti-GATA-2 antibody. (B) Detection of the blot shown in panel A with streptavidin-HRP. TEV-cleavage of bioGATA-2 results in a downshift in size of the protein and loss of the biotin-tag (compare lanes 2 and 3 in panels A and B, respectively). Remaining uncleaved bioGATA-2 is indicated with an asterisk in panel B.

conjugate on the same blot. In this way the fraction of the tagged protein that becomes biotinylated in vivo and subsequently captured by the streptavidin beads can be determined. As shown in Figure 1, for GATA-1 the biotinylation efficiency and capture by the beads is nearly 100%. The proteins eluted from the beads are fractionated by SDS-PAGE, the gel is stained and photographed (Figure 4). The gel lane with the fractionated proteins is excised and cut into small pieces (or gel plugs) along its entire length. The gel plugs are then processed for protein identification by mass spectrometry. The following sections describe in detail all of these techniques. We also provide an overview of the background binding in experiments using nuclear extracts and the specific enrichment observed when purifying biotin tagged transcription factors (Figure 4). Lastly, we provide protocols for the size fractionation of nuclear extracts using a preparative Superose 6 column and for the cleavage of proteins bound to streptavidin beads using TEV protease (Figure 3). Both approaches are presented with the aim of reducing the background in protein purification by biotinylation tagging. Given the increase in the potential applications of biotinylation tagging for protein purification, the description of such protocols may prove a useful resource.

2. Materials

2.1 Cell culture

1. Dulbecco's Modified Eagle's Medium (DMEM) (Cambrex Bio Science, Belgium), supplemented with 10 % Fetal Bovine Serum (FBS, Hyclone, Belgium).
2. Penicillin (used at 100u/ml final concentration) and streptomycin (used at 100µg/ml final concentration), stored at –20°C (100x stock from Cambrex Bio Science, Belgium).
3. 100% Dimethylsulfoxide (DMSO, Merck, Germany), used at 2% final concentration.
4. Neomycin (Gibco-BRL, UK), stock prepared as 100mg/ml in phosphate buffered saline (PBS, see below), filter-sterilized, aliquoted, stored at –20°C and used at 400µg/ml final concentration.
5. Puromycin (Sigma, St. Louis, MO), 1000x stock aliquoted and stored as above and added to 1µg/ml final concentration.

2.2 Nuclear extract preparation

1. Dulbecco's Phosphate Buffered Saline (PBS, Cambrex Bio Science, Belgium)
2. Protease inhibitors: Complete (Roche, Germany). Use 1 tablet for 50ml of solution
3. Cell resuspension buffer: 2.2 M sucrose in 10 mM HEPES-KOH, pH 7.9, 25 mM KCl, 0.15 mM Spermine, 0.5 mM Spermidine, 1 mM EDTA (with protease inhibitors added as above).
4. Standard household blender with rotating blades for homogenizing cells.
5. Nuclear lysis buffer: 10 mM HEPES, pH 7.9, 100 mM KCl, 3 mM $MgCl_2$, 0.1 mM EDTA, 20% glycerol (with protease inhibitors added as above).
6. Coomassie Plus Assay reagent (Pierce, IL).
7. Protein standards: dilutions of 0, 100, 200, 300, 400 and 500 µg/ml of chicken egg albumin (Sigma-Aldrich, MO) in ddH_2O prepared from a 20 mg/ml stock solution. 10 µl of each standard diluted in 1 ml (final volume) of ddH_2O is used for obtaining a standard curve.
8. Spectrophotometer: Ultraspec II (LKB Biochrom), cuvettes (Sterna, Germany).
9. Conductivity meter: Philips PW 9526.
10. Standards for determining salt concentration: 100, 200 300, 400 mM KCl diluted in ddH_2O from a 1M KCl stock solution. 10 µl of each standard are diluted in 1 ml (final volume) of ddH_2O for obtaining a standard curve.

2.3 Size fractionation by Superose 6 gel filtration

1. Superose 6 analytical grade column: HR 10/30 with a total bed volume of 24 ml. Preparative grade column: XK 50/600 with a bed height of 30.3 cm, total bed volume of 589 mL, both purchased form Amersham Biosciences (UK). Range of separation of 5,000 to 5,000,000 Da. 20% ethanol is used as preservative. The column is connected to an AKTA FLPC system (Amersham Biosciences).
2. Gel filtration calibration kit: dextran blue and high molecular weight standards (Amersham Biosciences, UK).
3. Column running buffer: 20mM HEPES, 0.5mM EGTA, 1mM $MgCl_2$, 200mM KCl, 10% glycerol. All buffers used for gel filtration column should be filtered prior to use.
4. 100% Trichloroacetic acid (TCA, Sigma-Aldrich, MO).

2.4 SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

1. Pre-cast NuPAGE 4-12% Bis-Tris gel (Invitrogen, UK).
2. Gel electrophoresis buffer: 1 X MOPS buffer diluted from 20 X stock solution and NuPAGE antioxidant (both from Invitrogen, UK).
3. Sample loading buffer, final concentration: 62.5 mM Tris-HCl pH 6.8, 25% glycerol (v/v from 100% stock, Sigma-Aldrich, St. Louis, MO), 2% SDS (v/v from 20% stock in ddH_2O), 0.01% bromophenol bl (w/v, Sigma-Aldrich, St. Louis, MO), 5 % β-mercaptoethanol (v/v, 100% stock, Merck, Germany). Can be prepared as a 4X stock solution.
4. Broad range pre-stained SDS-PAGE molecular weight standards (BioRad, Hercules, CA).
5. SimplyBlue Safestain gel staining solution (Invitrogen, UK).

2.5 Western blotting

1. ProTran nitrocellulose membrane (Schleicher & Schuell, Germany).
2. Gel-blotting paper (Schleicher & Schuell, Germany).
3. Blotting buffer: 25 mM Tris (made directly from the solid), 192 mM Glycine (made directly from the solid), 20% methanol.
4. Tris-Buffered Saline (TBS): 10 mM Tris-HCl, pH 7.4, 150 mM NaCl. Can be prepared as a 10X stock and stored at room temperature.
5. Blocking buffer: 5% Bovine Serum Albumin (BSA, Roche, Germany) in 1X TBS, prepared fresh.
6. Washing buffer 1x TBS adjusted to 0.5M NaCl (using a 5M NaCl stock solution), 0.3% Triton X-100 (Sigma-Aldrich, MO).
7. Primary and secondary antibody dilution: in blocking buffer with 0.2% NP-40.
8. Secondary antibodies: anti-rabbit (1/50000 dilution), anti-mouse (1/15000 dilution) from Amersham Biosciences (UK), anti-rat (1/3000 dilution) and anti goat (1/4000 dilution) from DakoCytomation (Denmark). All antibodies are purchased as horseradish peroxidase (HRP) conjugates.
9. Enhanced Chemiluminescence (ECL) kit (Amersham Biosciences, UK).
10. Bio-Max MR film (Kodak, Rochester, NY).

2.6 Streptavidin binding

1. Streptavidin paramagnetic beads (Dynabeads M280, Dynal, Sweden).
2. Chicken Egg Albumin (Sigma-Aldrich, MO).
3. Binding buffer: 1X TBS, 0.3% 0.3% NP-40 (Nonidet-40, Sigma-Aldrich, MO).
4. HENG buffer: 10 mM HEPES-KOH, pH 9, 1.5 mM $MgCl_2$, 0.25 mM EDTA, 20 % glycerol, 1mM PMSF (phenyl methyl sulfonyl fluoride, prepared as a 100x stock in ethanol and stored at –20°C).
5. Wash buffer: HENG buffer with 250 mM KCl and 0.3% NP-40.
6. Elution buffer: 1X sample loading buffer.
7. Magnets: Dynal MPC-1 and MPC-S magnets, for large and small volumes respectively (Dynal, Sweden).

2.7 TEV protease cleavage

1. Nuclear extract dilution buffer: 20mM Tris-HCl pH7.5, 0.45% NP-40.
2. Tobacco Etch Virus (TEV) protease (AcTEV Protease, Invitrogen, Scotland).

2.8 Sample preparation for mass spectrometry

1. Trypsin, sequencing grade (Roche, Germany). 10x stock made by dissolving lyophilized powder in 1mM HCl to 100ng/µl final concentration.
2. 50mM Ammonium bicarbonate (Sigma-Aldrich, MO). Dissolved in ddH_2O and filter sterilized.
3. 100% Acetonitrile (Sigma-Aldrich, MO).
4. 100% Formic acid (Sigma-Aldrich, MO).
5. Gel slice destaining solution: 25mM ammonium bicarbonate in 50% acetonitrile prepared by mixing equal volumes of the stock solutions given above.
6. 50mM iodoacetamide (Sigma-Aldrich, MO) prepared in 50mM ammonium bicarbonate.
7. 6.5mM DTT (dithiothreitol, (Sigma-Aldrich, MO) prepared in 50mM ammonium bicarbonate.

3. Methods

3.1 Cell culture

1. Mouse Erythroleukemia (MEL) cells are Friend virus transformed erythroid progenitors arrested at the proerythroblast stage of differentiation (7). MEL cells are cultured in DMEM supplemented with 10% FBS at 37°C and 5% CO_2. Cells are grown to a maximum of $2x10^6$ cells/ml and routinely split to a density of $5x10^4$ cells/ml. Cells appear to be semi-adherent and rounded with a smooth surface (*see* **Note 1**).
2. For induction, cells are diluted to $5x10^5$ cells/ml and cultured in DMEM with 2% DMSO (v/v) for at least 3 days. Cells become smaller but remain round and when pelleted appear pink/red due to hemoglobinization.

3.2 Nuclear extract preparation

1. Cells are harvested in 1 liter centrifuge bottles by centrifugation at 640 x *g* for 40 min at 4°C in a Beckman J4 centrifuge and washed once with 100 ml of ice-cold PBS. Resuspended cells are transferred into 50ml Falcon tubes and re-pelleted in an Eppendorf 5810R benchtop centrifuge at 2540 x *g* for 10 min at 4°C. The supernatant is discarded.
2. The cell pellet is gently resuspended by pipetting up and down in 200 ml of cell resuspension buffer with protease inhibitors added. Cells are equilibrated to the new osmotic conditions for 20 min on ice.
3. Cells are lysed in a blender using a single 30 second pulse at setting 3 (*see* **Note 2**). Excessive foaming should be avoided.
4. Lysis efficiency is checked under the microscope by staining a 10 µl aliquot with an equal volume of Unna stain (Methylgreen-Pyronin). Nuclei appear blue whereas intact cells appear with a blue nucleus surrounded by a non-stained cytoplasm. Optimal lysis should result in more than 90% of the nuclei appearing free of cytoplasm.
5. Nuclei are pelleted by ultracentrifugation at 141,000 x *g* using the SW28 rotor for 2 h at 4°C. A clean white pellet corresponding to the nuclei should be visible at the bottom of the tube. The top layer (cellular debris/cytoplasm) is discarded.
6. Nuclei are resuspended in 15 ml of nuclear lysis buffer and proteins are extracted by the drop wise addition of a 3.3M KCl solution with gentle agitation on ice, until the final concentration is ~350-400mM (*see* **Note 3**). Nuclear lysis and protein extraction are allowed to proceed by incubating on ice for 20 min. Two phases should be visible: one is clear and represents the soluble nuclear extract fraction whereas the other phase appears viscous and represents the insoluble fraction of mostly chromatin fragments.
7. Insoluble material is removed by ultracentrifugation at 300,000 x *g* using the SW50.1 rotor for 1 h at 4°C. The supernatant (soluble nuclear extract, approximately 17-18ml) is aliquoted in 1-5ml aliquots.
8. A small aliquot is used to measure protein concentration using the Bradford method. First,

a standard curve is obtained by adding 200µl of Bradford reagent to 800 µl of each of the chicken egg albumin standards, incubating for 5 min at room temperature and measuring the absorbance at 595nm in a spectrophotometer. The absorbance of dilutions of the aliquot of the nuclear extract are measured in the same way. The concentration of the nuclear extract is determined using the standard curve, taking the dilution of the sample into account.

9. The salt concentration of the solution is determined by measuring the conductance of the sample with a conductivity meter. First, a calibration curve of conductance is obtained using the KCl dilutions as standards. The conductivity of a diluted aliquot of the nuclear extract is then measured and its approximate salt concentration is estimated using the calibration curve.
10. Nuclear extracts are snap frozen in liquid nitrogen and stored at –70°C (*see* **Note 4**).

3.3 Analytical Superose 6 gel filtration (*see* **Note 5**)

1. An aliquot of nuclear extract is thawed on ice and centrifuged for 5 min at full speed using a microcentrifuge at 4°C. The volume of extract loaded on the column should not exceed 1% of the column volume.
2. The pump and the column are equilibrated with column running buffer. The running program is set up as follows: 100 µl/min flow rate; sample volume loop 200µl; fraction volume 500 µl; elution length 1 column volume of running buffer; alarm pressure set at 0.5 MPa (*see* **Notes 6 and 7**). After each run the column is washed with 2 column volumes of running buffer.
3. The calibration standards are run through the column to establish the elution volume of protein complexes and of free protein monomers (*see* **Notes 8-10**).
4. Collected fractions are concentrated by trichloroacetic acid precipitation (TCA), as follows: 125 µl of cold 100 % TCA are added to each 500 µl fraction, mixed well and incubated on ice for 30 min.
5. Proteins are pelleted by centrifugation at full speed in a microcentrifuge at 4^0C for 20 min.
6. The supernatants are discarded and the pellets washed with at least 500µl 1% ice cold TCA (in ddH_2O) and re-centrifuged as above.
7. The supernatants are discarded again and the pellets are washed with ice-cold acetone and re-centrifuged as above.
8. The pellets are air dried (on ice) and re-suspended in 50 µl of sample loading buffer (*see* **Note 11**). The fractionation patterns of specific proteins are determined by SDS-PAGE and Western blotting, as described below. An example of a GATA-1 Superose 6 fractionation profile of nuclear extracts from MEL cells expressing biotin-tagged GATA-1, is shown in Figure 2. From this, it is clear that both endogenous and tagged GATA-1 elute in high molecular weight (>670kDa) fractions. The fractionation profile of tagged GATA-1 follows that of the endogenous GATA-1 protein.

3.4 Preparative gel filtration by Superose 6

1. Steps 1 and 3 are as above.
2. Running program: 4ml/min flow rate; sample volume loop 5ml; fraction volume 10ml; elution length 1.5 column volumes; alarm pressure set at 0.65Mpa (*see* **Note 12**).
3. Fractions are collected in 15 ml Falcon tubes and used for precipitation (see below) or for binding to streptavidin beads (section 3.6).
4. Proteins are precipitated with 20% TCA (2.5 ml of 100 % TCA are added to every 10ml fraction) and kept on ice for 1h. The tubes are centrifuged in an Eppendorf 5810R benchtop centrifuge at maximum speed (2540 x *g*) for 20 min at 4°C. The pellet is subsequently washed with ice-cold 1% TCA (in ddH_2O). At this step the pellet can be carefully resuspended and

transferred into microfuge tubes. The samples are re-pelleted by spinning as above, or in a microfuge for 20 min, full speed at 4ºC. The pellets are washed with ice-cold acetone, centrifuged again as above and air dried on ice. The protein pellets are finally resuspended in 50µl of sample loading buffer and denatured by boiling before loading on an SDS-PAGE gel (section 3.8).

3.5 Binding to streptavidin beads

1. 5-10mg of nuclear extract is thawed on ice and diluted to 150mM KCl final concentration by the dropwise addition of ice-cold HENG with gentle shaking (*see* **Note 13**). NP-40 is adjusted to a 0.3% final concentration.
2. We routinely use 200 µl of resuspended streptavidin beads per 5mg of nuclear extract. The beads are blocked with 200ng/ml chicken egg albumin (CEA) in a 1ml final volume (made up with HENG buffer), for 1 hour at room temperature on a rotating platform.
3. The beads are immobilized using a magnetic rack and the blocking solution is removed. The beads are then resuspended in the diluted nuclear extract and incubated on a rotating platform 4°C for 2 hours to overnight.
4. The beads are immobilized on ice using the magnetic rack. The supernatant, corresponding to the unbound fraction or flowthrough, is collected and saved.
5. The beads are washed in 1 ml of washing buffer as follows: 2 quick rinses followed by 3 washes, 10 min each at room temperature on a rotating platform.
6. After the last wash, the beads are resuspended in 50 µl of sample buffer. Bound proteins are eluted by boiling the beads for 5 min.
7. The eluted material is fractionated by SDS-PAGE and processed for analysis by mass spectrometry as described below.

3.6 Binding of preparative Superose 6 fractions to streptavidin beads

1. After determining the fractionation profile of the protein(s) of interest by SDS-PAGE and Western blotting (as below), the peak fractions are pooled in a suitable sterile container (we conveniently use a sterilized glass measuring cylinder). All work is carried out on ice or in the cold room.
2. KCl and NP-40 concentrations are adjusted to 150mM and 0.3%, respectively (as above) with gentle mixing.
3. The diluted fractions are divided equally into separate 50ml Falcon tubes, so that tubes are not more than ¾ full, and resuspended streptavidin M280 beads (equilibrated and blocked as above) are added followed by overnight incubation at 4ºC on a rotating platform. We use approximately 10µl of streptavidin M280 beads for every 10ml fraction.
4. The beads are immobilized using the Dynal MPC-1 magnet and the supernatant corresponding to the unbound fraction is removed and saved. The immobilized beads from each tube are resuspended in washing buffer and pooled by transferring to a microfuge tube.
5. The beads are washed 4-5 times at room temperature in 1ml washing buffer for 5-10 min each.
6. The bound proteins are eluted by boiling the beads in sample buffer (50µl of sample buffer per 20µl of immobilized beads). The eluted material is fractionated by SDS-PAGE and processed for analysis by mass spectrometry as described below.

3.7 TEV protease cleavage

With the aim of reducing the non-specific background observed in streptavidin binding experiments,

we developed a modified version of the biotin tag for the N-terminal tagging of fusion proteins. This tag consists of a shorter (14aa) amino acid sequence than the one previously used but which is also very efficiently biotinylated by the BirA biotin ligase (3). The biotin tag sequence is followed by a 7aa cleavage site for the highly specific TEV-protease (Tobacco Etch Virus protease) (Figure 3A). In this way, the biotinylated protein and associated complexes can be specifically released from the streptavidin beads by cleaving off with the TEV protease (Figure 3B and C).

1. Immediately after the washing steps (step 5, section 3.5) the streptavidin paramagnetic beads are resuspended in 1xTBS/0.3% NP-40 buffer (95µl of buffer for every 50µl of resuspended beads used in the outset of the experiment).
2. 5-10% (v/v) of TEV-protease is added to the resuspended beads followed by incubation for 1-3h at 16°C with shaking (*see* **Note 14**).
3. The supernatant can be collected (it should contain the cleaved protein) and, if necessary, can be concentrated by TCA precipitation, as above.
4. The efficiency of protease cleavage can be monitored by testing an aliquot of the supernatant by SDS-PAGE and Western blotting (Figure 3B and C). The material that remains bound to the beads after TEV protease cleavage can be eluted by boiling in sample buffer and tested by SDS-PAGE. Successful cleavage results in the loss of the biotin-tag (as visualized by Streptavidin-HRP, Figure 3C) accompanied by a downshift in the size of the tagged protein thus resulting in faster migration by SDS-PAGE (as visualized by the tagged protein-specific antibody, Figure 3B).

3.8 SDS-PAGE

We preferably use NuPAGE pre-cast gels since they give clear and reproducible results in terms of resolution and sharpness of the protein bands. This if of particular importance if the gel is to be processed for mass spectrometry. There is also the added advantage of using gradient precast gels (e.g. 4-12%) for resolving proteins in a wide range of molecular weights. We use the Invitrogen electrophoresis system for running the NuPAGE gels.

1. Pre-cast gels are removed from the plastic envelope and rinsed in ddH_2O. The sticker near the bottom of the gel and the comb are removed carefully.
2. 800 ml of 1X MOPS electrophoresis buffer are prepared.
3. The gel is placed in the electrophoresis system (Invitrogen). The outside chamber is filled with 600 ml of 1X MOPS.
4. 500 µl of antioxidant are added to the remaining 200 ml of the 1X MOPS buffer, mixed and used to fill the inner chamber.
5. The wells of the gel are rinsed twice with the running buffer.
6. The samples in sample loading buffer are boiled for 5 min to denature the proteins and loaded directly onto the gel. 5-10 µg of nuclear extract are loaded per lane.
7. The gel is electrophoresed at 200V constant voltage for 60-75 min.

3.9 Western blotting

Samples that have been separated by SDS-PAGE are electrophoretically transferred onto nitrocellulose membrane by "wet blotting". For protein transfer we use the Trans-Blot electrophoretic transfer cell (Bio-Rad, CA).

1. Four gel-blotting papers are cut to the size of the gel. A "sandwich" is set up consisting of a sponge, 2 pieces of blotting paper, the gel, the membrane, another 2 pieces of blotting paper and a sponge (*see* **Note 15**). The membrane is pre-wetted in water followed by

transfer buffer.

2. The "sandwich" is placed in the transfer tank containing transfer buffer pre-chilled at 4^{0}C, such that the membrane is between the gel and the anode.
3. Blotting is carried out under constant amperage at 390 mA for 70 min in the cold room (*see* **Note 16**).
4. At the end of the transfer, the "sandwich" is disassembled and the membrane is rinsed in 1x TBS / 0.05% NP-40 (*see* **Note 17**).
5. The membrane is blocked at room temperature in freshly prepared blocking buffer for 1 hour on a rocking platform.
6. The blocking buffer is discarded, replaced by the primary antibody (in this case anti-GATA-1 N6 antibody diluted 1:5000) in blocking buffer / 0.2 % NP-40 and incubated overnight at 4°C on a rotating wheel.
7. The primary antibody is removed and the membrane washed three times with 50 ml of washing buffer, 15 min each wash at room temperature (*see* **Note 18**).
8. A freshly prepared secondary antibody (in this case anti-rat diluted 1:3000 in blocking buffer) is added to the membrane and incubated for one hour at room temperature on a rocking platform.
9. The secondary antibody is discarded and the membrane is washed as in step 8, followed by one wash for 5 min in 1xTBS/0.05% NP-40.
10. The membrane is lifted out of the wash buffer and excess liquid is removed by touching it on a clean tissue and placed in a clean tray such as a plastic weigh boat.
11. 3 ml of ECL solution (per filter) is prepared according to the manufacturer's instructions and immediately added to the membrane and shaken gently for 1 min to ensure an even coverage of the membrane by the liquid.
12. Excess liquid is again removed by touching the membrane on a clean tissue. The membrane is wrapped in cling wrap and exposed to film in an autoradiography cassette in a dark room, *as soon as possible*.

3.10 Preparation of samples for mass spectrometry (*see* **Note 19**)

The procedures described below are for the analysis by liquid chromatography-tandem mass spectrometry (LC-MS/MS) using a Q-Tof Ultima API mass spectrometer. The treatment of samples may vary depending on the type of analysis and the instrument used. It is best to consult with the mass spectrometry facility where the analysis is to be carried out for the processing of samples.

1. Following electrophoresis by SDS-PAGE the gel is stained overnight with Colloidal Blue, according to the manufacturer's instructions.
2. The gel is destained in several changes of ddH_2O until the background (i.e. the non-protein containing part of the gel) is completely destained. This usually takes several hours (i.e. more than 12 hours).
3. The destained gel is photographed to provide a record of the purification experiment.
4. 20-25 microfuge tubes are rinsed in 60% acetonitrile.
5. The entire lane is cut out lengthwise and divided into at least 20 gel slices. Each gel slice is placed in a separate tube.
6. Each gel slice is destained in 100 µl of destaining solution (25mM ammonium bicarbonate in 50% acetonitrile) for 20-30min. This step is repeated until the gel slice becomes completely destained (usually 3-4 times). Alternatively, gel slices can be destained overnight at 4^{o}C.
7. Each gel slice is dehydrated in 100 µl of 100% acetonitrile for 5-10min at room temperature. The plug become hard and white at this step.
8. The gel slices are reduced with freshly prepared 6.5mM DTT solution for 45-60min at

37°C.

9. The solution is discarded and proteins in the gel slices are alkylated by adding 100µl of 54 mM iodoacetamide solution and incubating for 60min at room temperature in the dark.
10. The solution is discarded and the gel slices are washed in 100µl of gel slice destaining solution for 15 min at room temperature. This step is repeated once more.
11. The washing solution is discarded and the gel slices are dried in 100 µl of 100% acetonitrile for 10min. The solution is again discarded and the gel slices are dried at room temperature.
12. Proteins are in-gel digested in 15µl of 10 ng/µl modified trypsin at (diluted from the 100x stock in 50 mM ammonium bicarbonate) for 30 min on ice (*see* **Note 20**). 15 µl of 50 mM ammonium bicarbonate are added to the samples followed by overnight incubation at 37°C.
13. Samples are equilibrated to room temperature. 30µl of 2% acetonitrile in 0.1% formic acid are added to the samples and incubated at room temperature for 15 min. The samples are then vortexed briefly and sonicated for 1 minute.
14. The supernatants are collected in separate tubes and the remaining gel slices are treated with 30µl of 50% acetonitrile in 0.1% formic acid and incubated as above. Samples are again vortexed and sonicated as above and the supernatants are collected and pooled with the corresponding supernatants from step 14.
15. The samples are vacuum dried in a vacuum centrifuge for 45-60 minutes until they are dry.
16. The eluted peptides are now ready for analysis by mass spectrometry.

Figure 4A shows an example of a preparative binding of nuclear extracts from MEL cells expressing biotinylated GATA-1 (lane 3) and control extracts from cells expressing the BirA biotin ligase only (lane 5). The control binding experiment shows that background consists of a few strongly stained bands against a backdrop of more faintly staining bands. Analysis by mass spectrometry identified the strongly staining bands as corresponding to naturally biotinylated proteins such as carboxylases (1). The bulk of the remaining background proteins corresponded to abundant nuclear proteins such as splicing factors, proteins involved in ribosome biogenesis etc. (Figure 4B). The low background binding (<1% of the total) of transcription factors and chromatin remodeling and modification proteins is of note (Figure 4B). By contrast, analysis of the GATA-1 purification gel slice shows a large increase in the binding of transcription factors and chromatin remodeling and modification proteins, thus indicating specific co-purification with biotin-tagged GATA-1 (Figure 4C) (6).

A number of these interactions have been validated by independent immunoprecipitation experiments using nuclear extracts from non-transfected MEL cells and shown to include essential hematopoietic transcription factors such as FOG-1, TAL-1 and Gfi-1b in addition to chromatin remodeling complexes such as MeCP1 and WCRF/ACF (6). In addition, the analysis of the GATA-1 binding experiment identified abundant chromatin associated proteins, such as topoisomerases, as background due to their indirect co-purification with GATA-1 by virtue of their association with chromatin. Thus, we have defined background in these experiments as consisting primarily of naturally biotinylated proteins, abundant nuclear proteins associated with RNA metabolism and ribosome biogenesis and abundant chromatin associated proteins that are indirectly co-purified with chromatin-bound transcription factors.

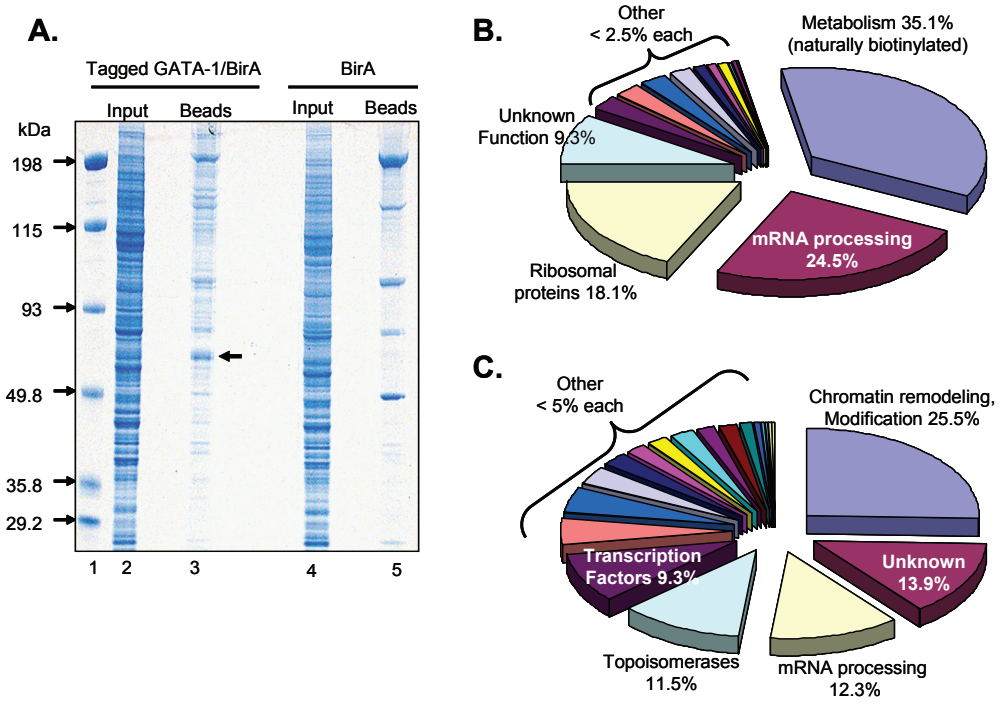


Figure 4: (A) Colloidal Blue-stained gel of a binding experiment of crude nuclear extracts to streptavidin beads. Lane 1: Marker (M). Lane 2: input nuclear extract from tagged GATA-1/BirA double transfected cells. Lane 3: proteins eluted after binding to streptavidin beads. Lane 4: input nuclear extract from BirA transfected cells. Lane 5: proteins eluted after binding to streptavidin beads. Arrow in lane 3 indicates protein band containing biotinylated GATA-1, as determined by mass spectrometry. (B) Classification according to Gene Ontology criteria of proteins identified by mass spectrometry from the control experiment using extracts from cells expressing BirA (around 500 peptide sequences were identified). This represents the background binding. (C) Classification according to Gene Ontology criteria of proteins identified by mass spectrometry using extracts from cells expressing biotin tagged GATA-1 (more than 1000 peptide sequences were identified). A significant increase in the identification of chromatin remodelling proteins and transcription factors is clearly notable compared to the control experiment shown in (B). Fig. 4A reproduced with permission from (1). Copyright (2003) National Academy of Sciences, USA.

Future prospects

We have shown that biotinylation tagging is highly efficient in cultured cells (Figure 1) and transgenic mice (1) and we have used this approach to identify a number of different complexes formed by the essential hematopoietic transcription factor GATA-1 (6). Due to its efficiency and ease of application, biotinylation tagging offers the prospect of rapidly expanding the characterization of transcription factor complexes. For example, the biotinylation tagging of the hematopoietic transcription factor partners of GATA-1 and the characterization of their protein complexes will lead to the rapid elucidation of the distinct and overlapping transcriptional networks these factors regulate in hematopoiesis. Similarly, the biotinylation tagging of chromatin co-factors will lead to a better understanding of their interactions with tissue-specific transcription factors and the molecular basis of their functions (i.e. chromatin remodeling and modification in activation and repression). Furthermore, efforts in reducing the background along the lines described here (i.e. a pre-purification steps such as gel filtration or the use of protease cleavage) will help in further expanding the utility of biotinylation tagging, for example in preserving the native properties of complexes or in determining stoichiometries. The utility of biotinylation tagging will be further increased through the development of additional tools such as the recent derivation of a transgenic mouse strain that expresses BirA ubiquitously in all tissues (8), or the construction of a codon-optimized version of BirA for the efficient expression in mammalian cells (9). The recent description of the biotinylation of cell surface proteins (10) should also serve to expand the utility of this approach. Lastly, it should be noted that in vivo biotinylation tagging can also be employed (e.g. instead of antibodies) in all other applications involving an affinity purification or detection step, such as immunofluorescence (1), immunoprecipitations (1, 11) and chromatin immnoprecipitation (ChIP) assays (1, 12).

5. Notes

1. We routinely screen 12-20 stable transfected MEL cell clones by SDS-PAGE in order to select a clone that expresses the tagged protein at no more than 50% of the expression level of the endogenous protein. This is in order to ensure that the physiological interactions and functions of the protein of interest are not disturbed as a result of the overexpression of the tagged protein.
2. The specific lysis conditions will depend on the make of blender employed. It is recommended that conditions are optimized for cell density, length of lysis time and speed setting of the blender.
3. The final salt concentration is critical for the extraction of nuclear proteins.
4. We routinely obtain around 100 mg of nuclear extract from 4 liters of MEL cell culture at a density of $2x10^6$ cells/ml.
5. There are a large variety of column matrices commercially available for gel filtration, with each matrix having different optimal separation ranges and physicochemical properties (e.g. ability to withstand high pressure in the column). Thus, the choice of matrix will depend on the desired range of fractionation and the liquid chromatography operating system available to the user (e.g. FPLC or HPLC).
6. Users must also refer to the manufacturer's instructions and training for use of the column and the FPLC apparatus.
7. The resolution efficiency of new columns, expressed as the number of theoretical plates per meter of column under normal running conditions, should be tested first. This can be done by injecting a sample of acetone (5mg/mL) in ddH_2O water. Indicative efficiency for the analytical grade column is 11100 theoretical plates/m.
8. While loading the extract, care must be taken that no air bubbles enter the loop. Air bubbles as well as cell debris can damage the column bed.
9. Once a new column is installed, the void (V_0) volume is determined by the peak of elution of dextran blue. In order to further calibrate the column, a mixture of at least two proteins of known molecular weight should also be injected. Recommended standards: bovine serum albumin (67kDa), thyroglobulin (669kDa), aldolase (158kDa).
10. If there is any suspicion that the column bed has been damaged, it is best to run the calibration standards again.
11. If the blue color of the sample loading buffer turns yellow, it is due to the protein sample being acidic which will also affect migration of the sample during SDS-PAGE. A few microliters of Tris-HCl pH9.0 are usually sufficient to neutralize the sample.
12. To avoid pressure build up the run can be started at a flow rate of 1ml/min. It is also better to inject the sample with the lower flow rate.
13. The concentration of 150mM KCl is critical for the efficient binding of biotinylated proteins to streptavidin beads. We have found that even modest increases in salt concentration severely affect binding efficiency.
14. Protease cleavage also works well with shorter incubation times (5-30min) and a broader temperature range (4-37^{o}C).
15. Avoid handling membrane directly, use gloves and forceps.
16. Under these transfer conditions, the temperature of the buffer can rise significantly and frothing may occur. This does not affect the transfer.
17. The gel can be stained after blotting in order to visualize residual proteins as a test for the efficiency of transfer as well as an indication of the amount of protein loaded per lane.
18. The primary antibody can be stored and re-used. Sodium azide is added to the antibody solution to a 0.02% final concentration and stored at 4^0C (sodium azide stock: 10% w/v in ddH_2O. **Caution: sodium azide is highly toxic**.

19. In order to reduce the risk of contaminating the samples for mass spectrometry, particularly with keratins, work is carried out in a hood wearing double gloves, a lab coat and always using sterile plasticware.
20. The volume of trypsin solution added will depend on the size of the gel slice. The volumes given above are for appr. 4x2mm gel slices. At this stage, gel slices should swell and little solution should remain visible.

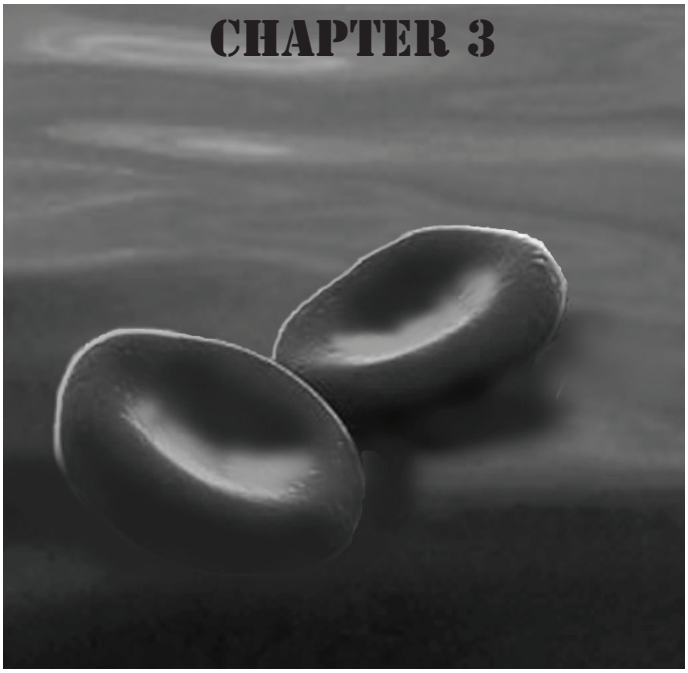
Acknowledgements

We are indebted to Dr. Jeroen Krijgsveld (Utrecht University) and Dr. Jeroen Demmers (Erasmus Medical Center) for expert mass spectrometry analysis. Work in our laboratory has been supported by grants from the Dutch Research Organization (NWO), the European Union (grant HRPN-CT-2000-00078), the NIH (grant RO1 HL 073445-01) and the Netherlands Proteomic Center.

References

1. de Boer, E., Rodriguez, P., Bonte, E., Krijgsveld, J., Katsantoni, E., Heck, A., Grosveld, F. and Strouboulis, J. (2003) Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. *Proc Natl Acad Sci U S A,* **100,** 7480-5.
2. 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,** 1138-43.
3. Beckett, D., Kovaleva, E. and Schatz, P.J. (1999) A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Sci,* **8,** 921-9.
4. Smith, P.A., Tripp, B.C., DiBlasio-Smith, E.A., Lu, Z., LaVallie, E.R. and McCoy, J.M. (1998) A plasmid expression system for quantitative in vivo biotinylation of thioredoxin fusion proteins in Escherichia coli. *Nucleic Acids Res,* **26,** 1414-20.
5. Cull, M.G. and Schatz, P.J. (2000) Biotinylation of proteins in vivo and in vitro using small peptide tags. *Methods Enzymol,* **326,** 430-40.
6. Rodriguez, P., Bonte, E., Krijgsveld, J., Kolodziej, K.E., Guyot, B., Heck, A.J.R., Vyas, P., de Boer, E., Grosveld, F., Strouboulis, J. (2005) GATA-1 forms distinct activating and repressive complexes in erythroid cells. *EMBO J,* **in press**.
7. Singer, D., Cooper, M., Maniatis, G.M., Marks, P.A. and Rifkind, R.A. (1974) Erythropoietic differentiation in colonies of cells transformed by Friend virus. *Proc Natl Acad Sci U S A,* **71,** 2668-70.
8. Driegen, S., Ferreira, R., van Zon, A., Strouboulis, J., Jaegle, M., Grosveld, F., Philipsen, S., Meijer, D. (2005) A generic tool for biotinylation of tagged proteins in transgenic mice. *Transgenic Research,* **in press**.
9. Mechold, U., Gilbert, C. and Ogryzko, V. (2005) Codon optimization of the BirA enzyme gene leads to higher expression and an improved efficiency of biotinylation of target proteins in mammalian cells. *J Biotechnol,* **116,** 245-9.
10. Chen, I., Howarth, M., Lin, W. and Ting, A.Y. (2005) Site-specific labeling of cell surface proteins with biophysical probes using biotin ligase. *Nat Methods,* **2,** 99-104.
11. Koutsodontis, G. and Kardassis, D. (2004) Inhibition of p53-mediated transcriptional responses by mithramycin A. *Oncogene,* **23,** 9190-200.
12. Viens, A., Mechold, U., Lehrmann, H., Harel-Bellan, A. and Ogryzko, V. (2004) Use of protein biotinylation in vivo for chromatin immunoprecipitation. *Anal Biochem,* **325,** 68-76.

GATA-1 FORMS DISTINC ACTIVATING AND REPRESSIVE COMPLEXES IN ERYTHROID CELLS



EMBO Journal 24: 2354-2366, 2005

GATA-1 forms distinct activating and repressive complexes in erythroid cells

Patrick Rodriguez [1], Edgar Bonte [1], Jeroen Krijgsveld [2], Katarzyna Kolodziej[1], Boris Guyot[3], Albert J.R. Heck [2], Paresh Vyas[3], Ernie de Boer [1], Frank Grosveld [1] and John Strouboulis [1,4]

[1] Department of Cell Biology, Erasmus University Medical Center, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

[2] Department of Biomolecular Mass Spectrometry, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Sorbonnelaan 16, 3584 CA Utrecht, The Netherlands

[3] Department of Haematology, The Weatherall Institute of Molecular Medicine, University of Oxford, Oxford OX3 9DU, United Kingdom

[4] Corresponding author: i.strouboulis@erasmusmc.nl

Abstract

GATA-1 is essential for the generation of the erythroid, megakaryocytic, eosinophilic and mast cell lineages. It acts as an activator and repressor of different target genes, e.g. in erythroid cells it represses cell proliferation and early hematopoietic genes while activating erythroid genes, yet it is not clear how both of these functions are mediated. Using a biotinylation tagging/proteomics approach in erythroid cells we describe distinct GATA-1 interactions with the essential hematopoietic factor Gfi-1b, the repressive MeCP1 complex and the chromatin remodeling ACF/WCRF complex, in addition to the known GATA-1/FOG-1 and GATA-1/TAL-1 complexes. Importantly, we show that FOG-1 mediates GATA-1 interactions with the MeCP1 complex, thus providing an explanation for the overlapping functions of these two factors in erythropoiesis. We also show that subsets of GATA-1 gene targets are bound in vivo by distinct complexes, thus linking specific GATA-1 partners to distinct aspects of its functions. Based on these findings we suggest a model for the different roles of GATA-1 in erythroid differentiation.

Keywords: chromatin/GATA-1/hematopoiesis/repression/transcription factors
Subject category: Chromatin and transcription

Introduction

Hematopoiesis has served as a model for cellular commitment and differentiation. Hematopoietic stem cells (HSCs) commit to a number of divergent yet narrowly defined lineages, each giving rise to a specific type of blood cell. Lineage commitment involves the upregulation of a particular transcription program with the concomitant suppression of "multipotentiality" and transcriptional programs specifying alternative lineages (Orkin, 2000).

GATA-1 is a key regulator of the differentiation of the erythroid, megakaryocytic, eosinophilic and mast cell lineages and is the founding member of the GATA family of zinc-finger factors implicated in the development and differentiation of several cell types . Efforts to understand GATA-1 functions have identified a number of protein interactions with transcription factors, such as TAL-1 (and its associated proteins Ldb1, LMO2 and E2A), EKLF, PU.1 and Sp1 (reviewed by Cantor, 2002). GATA-1 is also reported to interact with chromatin remodeling/ modification proteins, including the CBP/p300 histone acetyltransferases and the SWI/SNF chromatin remodeling complex (Blobel et al., 1998; Kadam and Emerson, 2003). Prominent amongst the GATA-1 interacting partners is FOG-1, originally identified in a yeast two-hybrid screen (Tsang et al., 1997). A direct interaction between the two factors is required for erythroid differentiation (Crispino et al., 1999) and the GATA-1 and FOG-1 knockout phenotypes are very similar (Tsang et al., 1998).

GATA-1 functions as both an activator and a repressor. The GATA-1/FOG-1 complex has been shown to repress some genes, such as GATA-2, and activate others, such as β-globin or the EKLF gene (Anguita et al., 2004; Letting et al., 2004; Pal et al., 2004). GATA-1 has also been linked to the repression of genes with cell proliferation functions, e.g. myc and myb (Rylski et al., 2003), though its protein partners are unknown. The multimeric GATA-1/TAL-1/Ldb1/ E2A/LMO2 complex binds to closely spaced GATA and E-box binding motifs and has been associated with the activation of erythroid genes, such as glycophorin A and the α-globin locus (Anguita et al., 2004; Lahlil et al., 2004). The duality of GATA-1 as an activator and repressor has been reinforced by the recent microarray analysis of terminal erythroid differentiation following induction of GATA-1 expression (Welch et al., 2004). Despite all this information, important questions remain as to how can GATA-1 accommodate all these functions and interactions at the same time in erythroid cells? In addressing this, we have undertaken a biotinylation tagging-proteomics approach to characterize GATA-1 complexes from erythroid cells (de Boer et al., 2003). We show that GATA-1 forms distinct complexes with hematopoietic transcription factors and chromatin remodeling and modification complexes. Our findings provide an explanation for a number of previous observations regarding GATA-1 functions and protein interactions. In addition, we provide evidence for distinct GATA-1 complexes performing specific functions in erythroid cells, thus providing a new framework for future work on GATA-1 and its parallel functions in erythroid differentiation.

Results

Identification of GATA-1 complexes from erythroid cells by biotinylation tagging and mass spectrometry

GATA-1 was tagged by fusing of a small (23aa) peptide sequence to its N-terminus. This tag is efficiently biotinylated by the bacterial BirA biotin ligase which is co-expressed in stably transfected mouse erythroleukemic (MEL) cells, allowing a single step purification of biotinylated GATA-1 using streptavidin beads under mild conditions (de Boer et al., 2003). Proteins co-purified with GATA-1 from MEL cells chemically induced to undergo terminal differentiation, were identified by mass spectrometry, classified according to Gene Ontology terms or by BLAST searches and compared to the background (Suppl. Table 1). Additional experiments employed more stringent conditions and different nuclear extract preparations from induced MEL cells. We

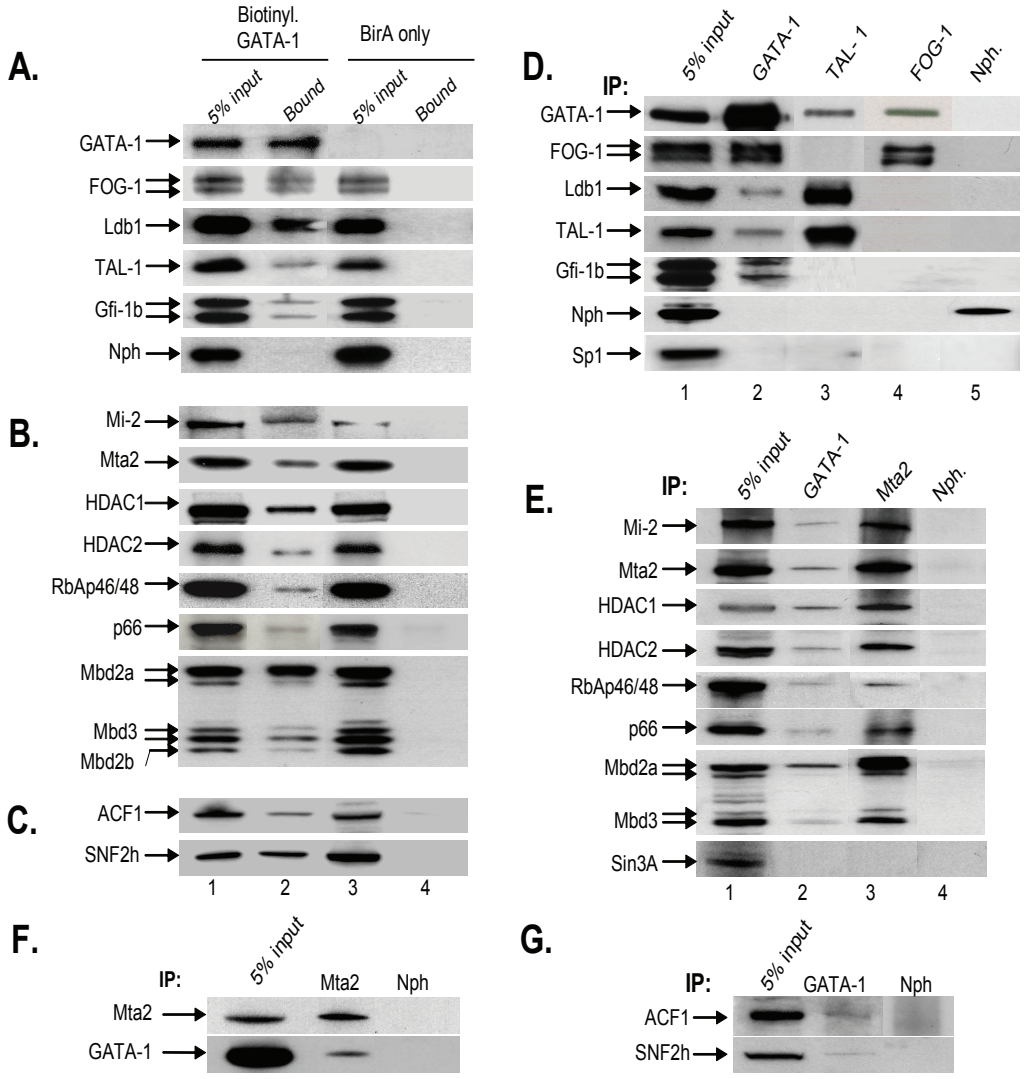


Figure 1: Confirmation by streptavidin pull-downs (A-C) and immunoprecipitations (D-G) of proteins identified co-purifying with GATA-1. (A) Streptavidin pull-downs of transcription factors. Biotinylated GATA-1 (top panel) is detected by streptavidin-HRP and is absent from the BirA only transfected cells. (B) Pull-downs of the MeCP1 complex. (C) Pull-downs of ISWI-containing complexes. SB: streptavidin-bound. (D) Immunoprecipitations (IP) using antibodies against GATA-1, TAL-1, FOG-1 (lanes 2, 3 and 4, respectively) and nucleophosmin as negative control (lane 5). (E) IP of the MeCP1 complex by antibodies against GATA-1 and MTA2 (lanes 2 and 3) and nucleophosmin (lane 4). (F) GATA-1, can be specifically IP'd by an antibody against MTA2 (G) IP of the ACF/WCRF complex by GATA-1 antibodies. Nuclear extract equivalent to 5% used in each pull-down or IP was loaded as control for input material. IP: immunoprecipitating antibody. Arrows show the detecting.

rejected proteins that appeared in the background binding experiments (de Boer et al., 2003), or proteins that belonged to a subnuclear compartment from which GATA-1 is excluded, e.g. the nucleolus (Elefanty et al., 1996). Streptavidin pull-downs of nuclear extracts under more stringent conditions (Fig. 1A-C) and immunoprecipitations of induced non-transfected MEL nuclear extracts provided further validation (Fig. 1D-G). The identities of proteins confirmed in this way as co-purifying with GATA-1 are shown in Table 1.

Finding FOG-1, TAL-1 and Ldb1 co-purifying with GATA-1 (Tsang et al., 1997; Wadman et al., 1997) validated our approach. The Gfi-1b hematopoietic transcription factor was also identified under moderate stringency conditions and verified by immunoprecipitation (Fig. 1D, lane 2) demonstrating an interaction between the two factors. This is in line with the similarities observed in the Gfi-1b and GATA-1 knockout phenotypes which result in differentiation arrest of the erythroid and megakaryocytic lineages (Pevny et al., 1995; Saleque et al., 2002). Chromatin remodeling and modification proteins also co-eluted with GATA-1 (Table 1) including the entire MeCP1 complex (Fig. 1). MeCP1 consists of the methyl-DNA binding protein MBD2 (Feng and

Zhang, 2001), p66/p68 (Feng et al., 2002) and the multi-subunit Mi-2/NuRD complex containing the nucleosome stimulated Mi-2β ATPase, the histone deacetylases HDAC1 and HDAC2 and other subunits of unknown function . The Mi2/NuRD and MeCP1 complexes are associated with epigenetic mechanisms of repression during development (Ahringer, 2000), potentially linking the GATA-1 repressive functions to the MeCP1 complex.

The SNF2h and ACF1 members of mammalian ISWI chromatin remodeling complexes also co-purified with GATA-1 (Table 1, Fig. 1C, G). SNF2h, a homolog of the Drosophila protein ISWI, is the "signature" ATPase of this class of complexes and participates in three distinct complexes in human cells: RSF, hACF/WCRF, hCHRAC (reviewed in (Corona and Tamkun, 2004)). We did not detect by mass spectrometry or immunoprecipitation (not shown) the additional p15 and p17 protein partners present in the hCHRAC complex hence GATA-1 appears to interact with SNF2h/ACF1 in the context of the ACF/WCRF complex (Bochar et al., 2000). ISWI/SNF2h-containing chromatin remodeling complexes have been associated with gene activation and repression (reviewed by (Corona and Tamkun, 2004). The interaction between SNF2h and GATA-1 may help explain the observation that knocking down SNF2h expression in primary hematopoietic progenitor cells blocked erythroid differentiation (Stopka and Skoultchi, 2003).

Further validation for the GATA-1 interactions was provided by reverse immunoprecipitations using antibodies against TAL-1, FOG-1 (Fig. 1D) or MTA2 (Fig. 1E). TAL-1 antibodies specifically immunoprecipitated GATA-1 and Ldb1 (Fig. 1D), as previously observed (Osada et al., 1995; Visvader et al., 1997). LMO2 or E2A were not detected co-purifying with GATA-1 from induced MEL cells, but it cannot be excluded that their absence is due to the very low abundance of the GATA-1/TAL-1/Ldb1/E2A/LMO2 complex (Table 1), in agreement with previous reports of a very small fraction of LMO2 being immunoprecipitated by GATA-1 antibodies (Osada et al., 1995). Interestingly, FOG-1 antibodies immunoprecipitated GATA-1 but not TAL-1, Ldb1, or Gfi-1b (Fig. 1D, lane 4). The converse was also true using TAL-1 antibodies (Fig. 1D, lane 3). Thus, GATA-1 interactions with TAL-1, FOG-1 and Gfi-1b are non-overlapping and must thus occur in distinct complexes. GATA-1 was also immunoprecipitated by MTA2 antibodies (Fig. 1F). By contrast, the Sin3A co-repressor, which interacts with HDACs but not in the MeCP1 complex, was not immunoprecipitated by GATA-1 or MTA2 antibodies (Fig. 1E) further supporting the specificity of the GATA-1 interactions with MeCP1. The participation of GATA-1 in multiple protein interactions is supported by size fractionation experiments of nuclear extracts, which showed that GATA-1 fractionates with a broad profile overlapping the profiles of the partners identified here (Suppl. Fig. 1), also providing evidence for interactions occurring in distinct complexes (for example, fractionation peaks of MeCP1 versus SNF2h/ACF1).

Other abundant chromatin-associated proteins also co-purified with GATA-1 (Table 1), including topoisomerases and Ku autoantigen or ADP ribosyltransferase (PARP). We tested the association of these proteins with DNA and GATA-1 by treating nuclear extracts with DNase I. In contrast to MTA2, there were no topoisomerase I or PARP co-purifying with GATA-1 after DNase I treatment. (Suppl. Fig. 2). Though it remains formally possible that interactions of GATA-1 with topoisomerase I or PARP are relevant and require DNA, on the basis of our DNase I results and on previous evidence by other groups describing topoisomerases as a common contaminant (Eberharter et al., 2001), we did not pursue these further.

GATA-1 forms several distinct complexes

To directly confirm the distinct GATA-1 interactions and to assess how the GATA-1 partners may be partitioned in the GATA-1 complexes, we carried out sequential immunodepletion experiments. First, we used an antibody against one of the GATA-1 partners, i.e. FOG-1, TAL-1 or MTA2 in order to immunodeplete from a nuclear extract the fraction of GATA-1 that is in complex with this factor (Fig. 2A). The remaining GATA-1 in the supernatant was subsequently

immunoprecipitated with a GATA-1 antibody and both immunoprecipitates were tested for the presence or absence of GATA-1 and interacting proteins (Fig. 2A).

We first established that antibodies against GATA-1, TAL-1, FOG-1 and MTA2 were efficient in immunodepleting most of these proteins from nuclear extracts (Fig. 2B). As expected, FOG-1 antibodies immunoprecipitated a fraction of GATA-1 (Fig. 2C, lane 2). Surprisingly, MTA2 was also specifically immunoprecipitated by FOG-1 antibodies (Fig. 2C, lane 2), suggesting an interaction between FOG-1 and the MeCP1 complex. This was confirmed by the reverse immunoprecipitation of FOG-1 by an MTA2 antibody (Fig. 2E). There was no immunoprecipitation of TAL-1, Gfi-1b or ACF1 by FOG-1 antibodies (Fig. 2C, lane 2). Importantly, MTA2 could no longer be detected in the subsequent immunoprecipitation of the supernatant with GATA-1 antibodies, (Fig. 2C, lane 3). Thus the fraction of GATA-1 that is in complex with MTA2 (and MeCP1) was depleted in the first step by the FOG-1 antibodies, leading us to conclude that FOG-1 and GATA-1 interact together in the same complex with MeCP1. Further confirmation was provided by the immunoprecipitation of

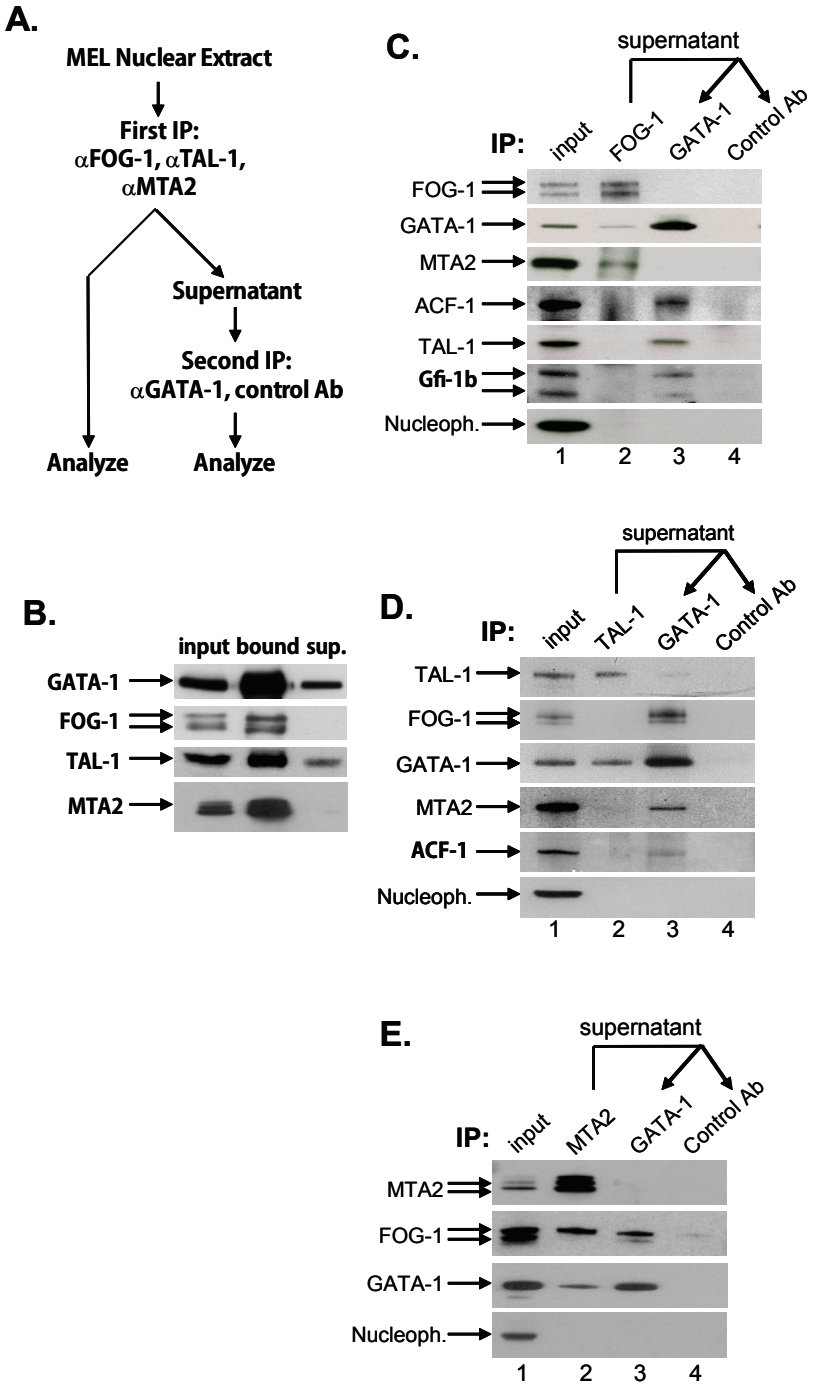


Figure 2: Distinct GATA-1 complexes by sequential immunoprecipitations (IP). (A) Experimental procedure. (B) Efficiency of immunoprecipitating antibodies (also used to detect the immunoprecipitated protein). Sup: supernatant after IP. (C) FOG-1 immunodepletion, FOG-1 IP (lane 2) followed by IP of supernatant with GATA-1 or control antibodies (lanes 3 and 4). (D) TAL-1 immunodepletion, same as (C) using TAL-1 antibodies in first IP. (E) MTA2 immunodepletion, same as (C) and (D) using MTA2 antibodies in first IP. The MTA2 antibody used in panels B and E is different to that used in panels C and D (see Suppl. Methods). IP: immunoprecipitating antibody. Arrows show the detecting antibodies.

MeCP1-associated histone deacetylase (HDAC) activity by GATA-1 and FOG-1 antibodies (Suppl. Fig. 3). Following the FOG-1 immunodepletion, GATA-1 antibodies could still immunoprecipitate TAL-1, Gfi-1b and ACF1 (Fig. 3C, lane 3). This confirms that GATA-1 participates in a complex with FOG-1 and MeCP1 that is distinct from those with TAL-1, Gfi-1b or ACF/WCRF. Using TAL-1 antibodies in the first immunodepletion step, a small fraction of GATA-1, but no MTA2 or ACF1 was immunoprecipitated (Fig. 2D, lane 2).

We also carried out an MTA2 immunodepletion to determine whether the entire fraction of GATA-1 interacting with FOG-1 does so in the context of the MeCP1 complex. Following the immunodepletion of MTA2 (Fig. 2B), an appreciable amount of FOG-1 was subsequently immunoprecipitated by GATA-1 antibodies (Fig. 2E, lane 3). Thus, GATA-1 also interacts with FOG-1 independently of the MeCP1 complex. Interestingly, the MTA2 antibody specifically immunoprecipitated the slower migrating of the two bands detected by the FOG-1 antibody, suggesting differential interaction with one of the two FOG-1 isoforms, while GATA-1 can interact with both FOG-1 isoforms (Fig. 2E, lane 3). Taken together, these experiments show that GATA-1 forms at least five complexes: first with FOG-1 and MeCP1, second with FOG-1 alone, third with TAL-1 (and Ldb1 since it can be almost completely immunodepleted by TAL-1 antibodies [not shown]), fourth with Gfi-1b and fifth with ACF/WCRF.

The GATA-1 zinc fingers mediate differential protein interactions

GATA-1 contains two evolutionarily conserved, closely spaced zinc finger domains. The C-terminal zinc finger (C-ZnF) is essential for DNA binding whereas the N-terminal zinc finger (N-ZnF) is primarily involved in protein-protein interactions, for example with FOG-1, which contribute to the specificity and stability of DNA binding by the C-ZnF (review by (Blobel and Weiss, 2001). Significantly, the C-ZnF is essential for all *in vivo* GATA-1 functions, whereas the N-ZnF is required for definitive, but not primitive, erythropoiesis erythropoiesis (Shimizu et al., 2001). We addressed how the GATA-1 zinc fingers mediated its multiple protein interactions by expressing in MEL cells biotin-tagged mutants lacking the N-ZnF or the C-ZnF followed by streptavidin pull-downs (Fig. 3A-C). As described (Tsang et al., 1998), GATA-1 interaction with

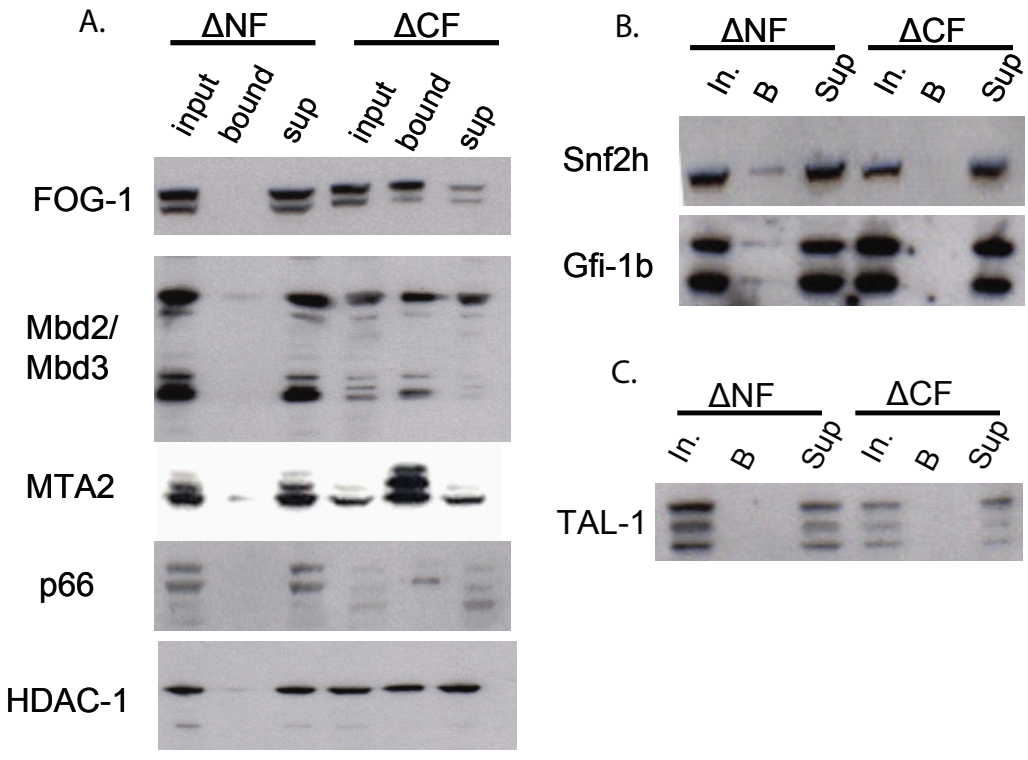


Figure 3: (A-C): Differential interactions mediated by the GATA-1 zinc fingers. GATA-1 zinc finger deletions were expressed as biotin tagged proteins in MEL cells and interactions were assessed by streptavidin pull-downs and Western blots. (A) FOG-1 and the MeCP1 complex require the N-ZnF for interactions. (B) TAL-1 requires both zinc fingers. (C) Gfi-1b and SNF2h require the C-ZnF for interactions with GATA-1. The TAL-1 antibody used is different to that used in Fig. 1.

FOG-1 requires the N-ZnF. Interactions of the MeCP1 members MTA2, Mbd2 and HDAC 1 also occur through the N-ZnF of GATA-1 (Fig. 3A). Interestingly, interactions of GATA-1 with TAL-1 require both zinc fingers (Fig. 3B), whereas interactions with SNF2h or Gfi-1b require only the C-ZnF (Fig. 3C). , We tested by immunoprecipitation using Gfi-1b antibodies whether Gfi-1b and SNF2h were in complex but found no evidence of such an interaction (not shown). Thus the multiple, distinct interactions of GATA-1 are differentially mediated through its zinc finger domains.

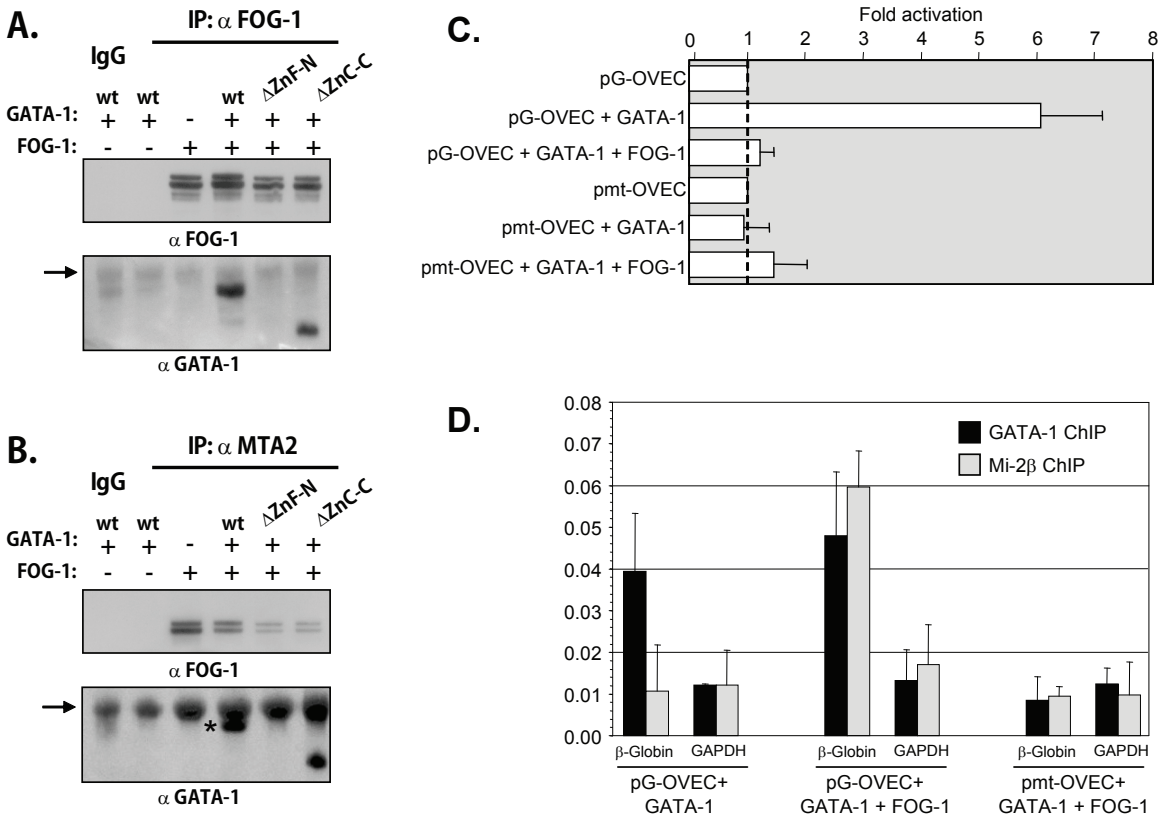


Figure 4: (A-B): FOG-1 bridges GATA-1 and MeCP1 . Nuclear extracts from HeLa cells transfected with the FOG-1 and GATA-1 combinations indicated, were immunoprecipitated with FOG-1 (panel A) or MTA2 antibodies (panel B) and detected with FOG-1 and GATA-1 antibodies. dZn-N and dZn-C: GATA-1 N- and C-terminal zinc-finger deletion mutants. Arrows: cross-reacting IgG. Asterisk (panel B): GATA-1 signal. (C) Real-Time PCR transcription assays in transfected HeLa cells. GATA-1 activates transcription of pG-OVEC, whereas co-transfection of FOG-1 represses to basal levels. (D) Specific recruitment of Mi-2β by co-transfected GATA-1 and FOG-1 by ChIP assays in HeLa cells. Mi-2β recruitment to the repressed gene requires GATA-1 binding to the promoter.

FOG-1 mediates interactions of GATA-1 with the MeCP1 complex in repressing transcription

We next tested whether FOG-1 mediates interactions between GATA-1 and the MeCP1 complex. . GATA-1 was transiently expressed in HeLa cells (which express endogenous MeCP1, but not GATA-1 or FOG-1) with or without FOG-1, followed by immunoprecipitation using FOG-1 or MTA2 antibodies (Fig. 4A-B). We find that the interaction of GATA-1 with MTA2 occurs only in the presence of FOG-1 (Fig. 4B), whereas FOG-1 interacts with MTA2 regardless of the presence or absence of GATA-1 (Fig. 4B, upper panel). Expression of the GATA-1 zinc-finger deletion mutants (Fig. 4B, lower panel) confirmed these observations. We conclude that interaction of GATA-1 with the MeCP1 complex requires interaction with FOG-1, which thus serves as the bridging factor.

We next tested whether the well known GATA-1 and FOG-1 mediated repression is due to the recruitment of the MeCP1 complex to a GATA-dependent promoter. To this end, we used a reporter plasmid containing the rabbit β-globin minimal promoter (pOVEC-1, (Westin et al., 1987), carrying four copies of an optimal GATA-1 binding sequence, or four copies of a mutated

sequence that abolish GATA-1 binding (Whyatt et al., 1993). The GATA-binding promoter was activated more than six-fold by co-transfection of GATA-1 alone (Fig. 4C). As expected, co-transfection of FOG-1 and GATA-1 repressed activation of the GATA-dependent promoter (Fig. 4C). Chromatin immunoprecipitations (ChIP) showed that repression by GATA-1 and FOG-1 was due to the specific recruitment of the MeCP1 complex. Binding of Mi-2β to the repressed gene was specifically enriched in GATA-1 and FOG-1 transfected cells (Fig. 4D), but not in cells transfected with GATA-1 only. The promoter bearing the mutated GATA binding sites does not bind MeCP1, even in the presence of FOG-1 (Fig. 4D). Thus, FOG-1/MeCP1 repression is mediated through GATA-1 binding at its cognate binding sites.

GATA-1, FOG-1, MeCP1 and Gfi-1b are bound to repressed genes *in vivo*

The results above suggest that GATA-1/FOG-1 interactions can tether MeCP1 to repressed GATA-1 target sequences *in vivo*. We therefore employed ChIP assays using the GATA-2 locus, which is repressed by GATA-1 in a FOG-1 dependent manner (Grass et al., 2003; Pal et al., 2004). As seen before, the -2.8kb region upstream of the GATA-2 promoter was enriched for GATA-1 and FOG-1 binding (Fig. 5A). The same sequence was also enriched for Mbd2 binding (Fig. 5A), with similar results obtained with an antibody against Mi-2β (Suppl. Fig. 4A). No binding of TAL-1 or Gfi-1b was observed in the GATA-2 sequences. Interestingly, the –3.4kb, but not the –4.2kb and –2.2kb flanking sequence used as negative controls for GATA-1 binding, was enriched for FOG-1 and Mbd2 binding (Fig. 5A and Suppl. Fig. 4A) suggesting that the FOG-1/MeCP1 binding at -3.4kb may reflect a very localized spreading of these proteins over a few nucleosomes to sequences upstream of the -2.8kb element, or that they were accidentally crosslinked to neighboring DNA. The latter possibility would suggest that the FOG-1/MeCP1 complex is closer to the upstream sequences around the GATA binding sites (see also below). We found no evidence that the binding of the GATA-1/FOG-1/MeCP1 complex to the –2.8kb region was mediated by DNA methylation (Suppl. Fig. 5), however this does not exclude the possibility of highly localized methylation to specific CpG residues elsewhere in the GATA-2 locus. Thus, considering that GATA-1 binding is essential for GATA-2 repression (Pal et al., 2004), our findings strongly suggest that GATA-1, FOG-1 and MeCP1 form the repressive complex responsible for GATA-2 silencing (see below).

Ectopic expression of FOG-1 in eosinophilic cells results in the downregulation of eosinophilic GATA-1 target genes and the reprogramming of these cells towards an earlier, less differentiated cell type which may represent a common progenitor for the erythroid/megakaryocytic and eosinophilic lineages (Querfurth et al., 2000). We thus reasoned that eosinophilic GATA-1 target genes, like the major basic protein (MBP) (Du et al., 2002), which is inactive in erythroid cells (Welch et al., 2004), may be suppressed by the GATA-1/FOG-1/MeCP1 complex. We tested this hypothesis by ChIP in induced MEL cells using as control chromatin from mouse eosinophils where MBP is expressed (Guyot et al., 2004). As expected, the promoter of the MBP gene was bound by GATA-1 in eosinophils (Fig. 5B). Importantly, GATA-1 was also bound to the inactive MBP promoter in induced MEL cells (Fig. 5C). FOG-1 and Mbd2 were also bound to the MBP promoter in MEL cells but not in eosinophils (Fig. 5C), consistent with the prediction above. Similar results were also obtained with an antibody against Mi-2β (Suppl. Fig. 4B). Again, no TAL-1 of Gfi-1b binding was detected in the MBP promoter (Fig. 5C). Strikingly, in MEL cells we again found binding of the FOG-1 and MeCP1 complex, but not of GATA-1, to the +0.6kb sequence located close to the MBP promoter but not to other sequences located further upstream (-1.8kb) or downstream (+1.2kb) of the promoter (Fig. 5C and Suppl. Fig. 4B). This observation is similar to that seen at the GATA-2 –2.8kb element.

We next tested the myc and myb genes which are down-regulated with MEL differentiation (Chen and Bender, 2001; Lachman and Skoultchi, 1984) and references therein). Repression of the myc and myb genes has been linked to the proliferation arrest that accompanies terminal

erythroid differentiation. The myc gene has also been shown to be a GATA-1 target gene in G1E cells (Rylski et al., 2003; Welch et al., 2004). We found GATA-1 binding to both promoters in induced MEL cells but we could not detect binding of FOG-1 or Mbd2 to the same sequences (Fig. 5D and E). By contrast, Gfi-1b (absent from all other genes tested) was found binding to both promoters (Fig. 5D and E), suggesting a role for the GATA-1/Gfi-1b complex in the repression of genes associated with cell proliferation. This may explain the observations of rapidly proliferating Gfi-1b -/-immature erythroid precursors in colony assays (Saleque et al., 2002) and of Gfi-1b overexpression inducing proliferation arrest and differentiation in erythroid progenitors (Garcon et al., 2005).

Finally, we also tested the EKLF gene as an example of a gene that is activated during erythropoiesis. The EKLF enhancer sequence contains a GATA-E-box motif (Anderson et al., 1998) which is bound *in vivo* by GATA-1 independently of FOG-1 (Letting et al., 2004). Strong

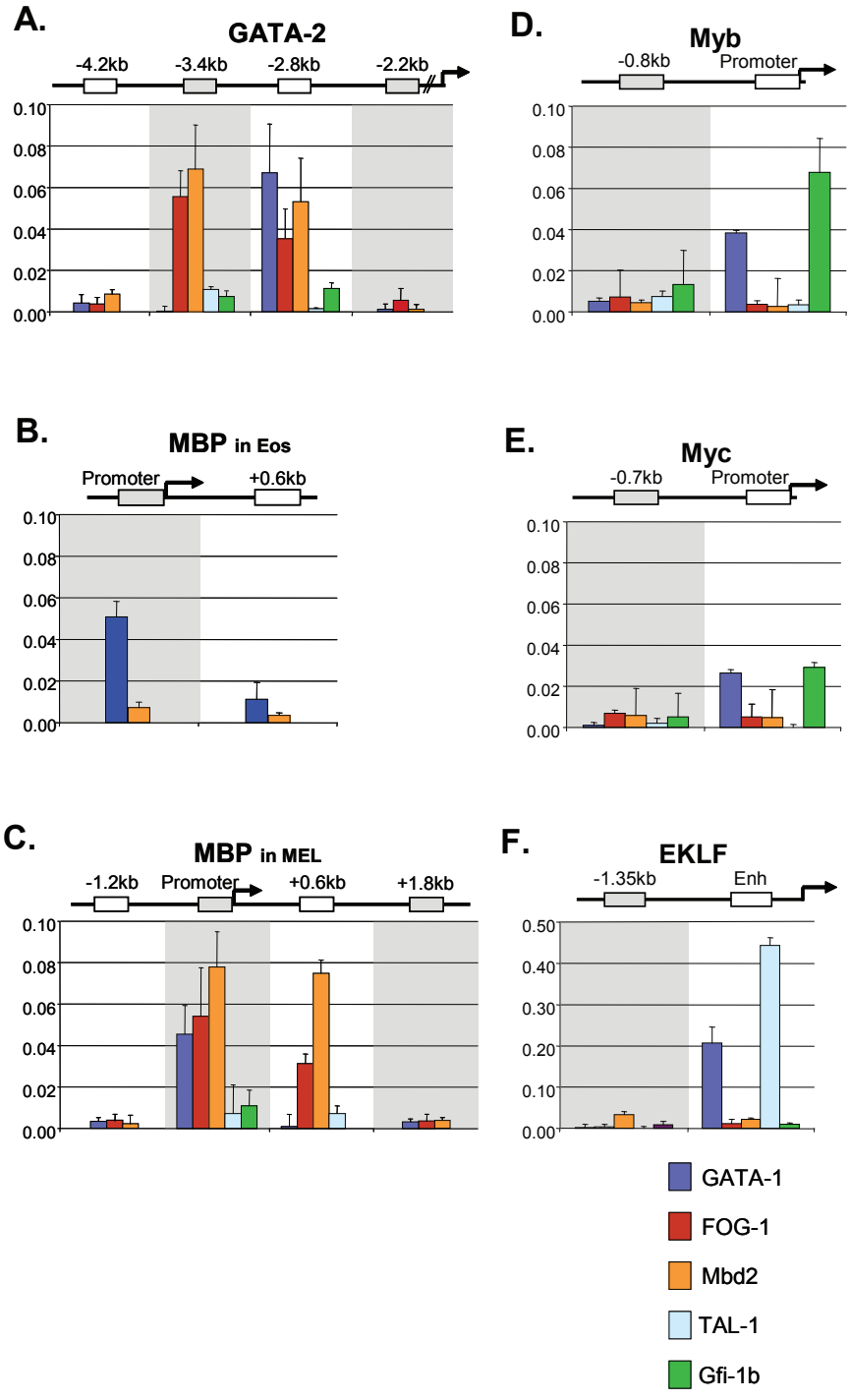


Figure 5: Binding of GATA-1 repressive and activating complexes to target genes by ChIP assays in induced MEL cells. Binding patterns of GATA-1, FOG-1, Mbd2, TAL-1 and Gfi-1b to the –2.8kb element of the GATA-2 locus (A), the MBP promoter in eosinophils (B) and in MEL cells (C), the myb (D) and myc (E) promoters and at the EKLF upstream enhancer (F). Relative enrichment has been normalized to input and corrected for background binding of species- and isotype-matched immunoglobulins. Antibodies: GATA-1, N6 (Santa Cruz); FOG-1 as in Tsang et al. (1997); Mbd2 S923 (Ng et al., 1999); TAL-1 as in Porcher et al. (1996); Gfi-1b, D19 (Santa Cruz).

GATA-1 binding and a clear enrichment for TAL-1 binding was indeed detected at the EKLF enhancer (Fig. 5F), thus providing a clear demonstration for the alternative (activating) GATA-1 complex with TAL-1 binding to a target gene *in vivo*. This may be related to the low level of HDAC activity associated with the TAL-1 immunoprecipitate (Suppl. Fig. 3A). No significant binding of FOG-1, Mbd2 or Gfi-1b could be detected in the EKLF enhancer sequences (Fig. 5F).

Our analysis of the GATA-1/ACF/WCRF complex by ChIP, or any other, assays has been hindered by the quality of ACF/WCRF reagents available to us, hence it is presently not known whether the GATA-1 and ACF/WCRF complex binds to active or repressed genes.

GATA-1 represses GATA-2 expression through the recruitment of FOG-1 and MeCP1

In order to confirm that GATA-2 repression during erythroid differentiation is specifically due to GATA-1 recruiting FOG-1 and the MeCP1 complex, we took advantage of the GATA-1 null G1E proerythroblastic cell line. These cells are derived from in vitro differentiated GATA-1 null ES cells and can undergo terminal differentiation only upon restoration of GATA-1 expression (Weiss et al., 1997). We used two G1E cell lines. The first one expresses wild type GATA-1 fused to an estrogen receptor (ER) ligand binding domain (GATA-1-ER) which can mediate terminal erythroid differentiation upon induction by estradiol (Tsang et al., 1997). The second cell line expresses ER fused to a mutant GATA-1 form bearing a single V205M amino acid substitution in the GATA-1 N-terminal zinc finger. Whilst not affecting GATA-1 DNA binding, this mutant abrogates interaction with FOG-1 and fails to rescue differentiation of G1E cells (Crispino et al., 1999; Nichols et al., 2000). We first determined that repression of the GATA-2 gene in G1E cells was absolutely dependent on GATA-1 being capable of interacting with FOG-1 (Fig. 6A). We next tested by ChIP whether interaction of GATA-1 with FOG-1 binding at -2.8kb was responsible for the recruitment of MeCP1 to this sequence and to the neighboring –3.4kb sequence (Fig. 5A). As control, we also tested the more distal –4.2kb sequence which did not show binding for any of these factors (Fig. 5A). In agreement with the MEL data, GATA-1, FOG-1 and Mi-2β were bound to the –2.8kb and to the –3.4kb sequence (for FOG-1 and Mi-2β) in differentiated GATA-1-ER cells (24 hours after induction with estradiol), albeit at lower levels compared to MEL cells (Fig. 6B, left panels). By contrast, in the GATA-1(V205M)-ER expressing cells GATA-1 was bound to the –2.8kb sequence, but no binding of FOG-1 and Mi-2β to the –2.8kb or –3.4kb sequences was detected (Fig. 6B right panels). We conclude that FOG-1 and the MeCP1 complex are specifically recruited by GATA-1 to the GATA-2 locus and are responsible for GATA-2 repression in terminal erythroid differentiation.

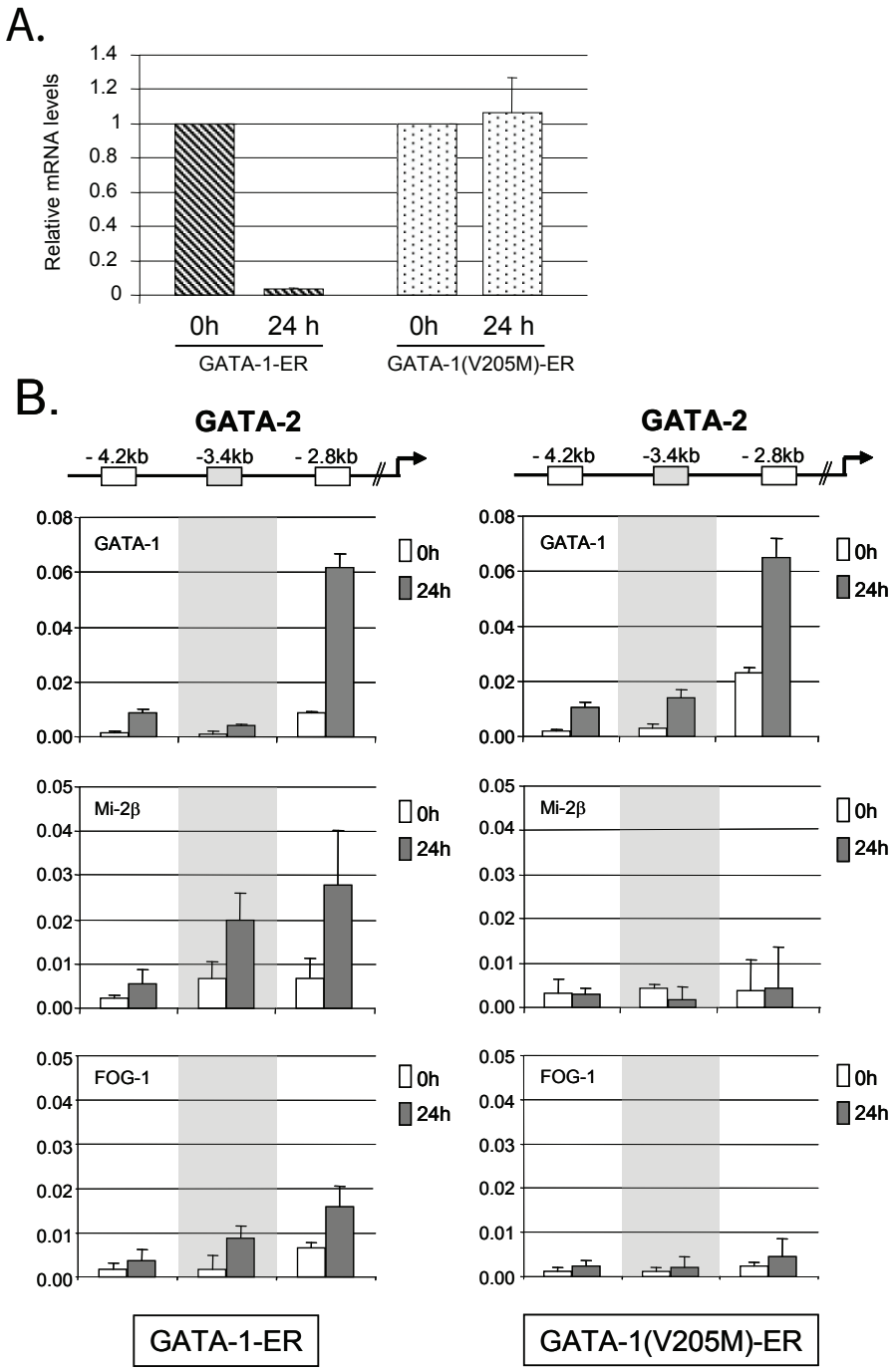


Figure 6: Silencing of GATA-2 requires recruitment of FOG-1 and MeCP1 by GATA-1. (A) The V205M GATA-1 mutation fails to repress GATA-2. GATA-2 mRNA was measured by Real-Time PCR in G1E GATA-1-ER and G1E GATA-1(V205M)-ER cells before (0 hours) and after 24 hours of estradiol induction. Expression at 0 hour was normalized against GAPDH expression and set as 1. (B) ChIP to show binding of GATA-1, Mi-2β and FOG-1 in G1E GATA-1-ER (left panels) and in G1E GATA-1(V205M)-ER (right panels) at time 0 and 24 hours of estradiol induction. Relative enrichment has been normalized to input and corrected for background binding of species- and isotype-matched immunoglobulins. Data represent an of two independent IPs and three PCRs with duplicate samples. Antibodies used were as in the legend of Figure 5, except Mi-2β antibody (Fujita et al., 2003).

Discussion

We describe here the characterization of GATA-1 complexes from erythroid cells by *in vivo* biotinylation tagging and purification by streptavidin beads. This work has led to a number of important findings. First, we identified novel GATA-1 partners, including the essential hematopoietic factor Gfi-1b and the chromatin remodeling and modification complexes MeCP1 and ACF/WCRF, in addition to the known GATA-1 interacting factors FOG-1, TAL-1 and Ldb1. Second, we showed that GATA-1 forms several distinct complexes with FOG-1, with FOG-1 and MeCP1, with TAL-1/Ldb1, with Gfi-1b and with the ACF/WCRF complex. Third, we found that the most abundant of the GATA-1 complexes are those with FOG-1 and with FOG-1 and MeCP1, with FOG-1 serving as the bridging factor between GATA-1 and the MeCP1 complex. Fourth, we showed that the distinct interactions of GATA-1 with its protein partners are differentially

mediated through the two GATA-1 zinc finger domains. Fifth, we show that the known GATA-1 and FOG-1 mediated repression is due to the recruitment of the MeCP1 complex to the repressed gene(s). Sixth, we present evidence for the *in vivo* binding of the repressive GATA-1/FOG-1/MeCP1 complex to silenced hematopoietic genes in erythroid cells and of the activating GATA-1/TAL-1 complex to erythroid specific genes. Significantly, we also showed binding of the GATA-1/Gfi-1b complex to genes associated with cell proliferation functions, which become repressed with erythroid differentiation. Finally, our work demonstrates the utility of biotinylation tagging as an efficient approach for the rapid isolation and identification by mass spectrometry of multiple protein complexes.

Biotinylation tagging and protein complex purification

From our previous work (de Boer et al., 2003) and the work described here, we show that background using biotinylation tagging consists of naturally biotinylated proteins, of abundant nuclear proteins such as splicing factors binding nonspecifically to the beads (de Boer et al., 2003) and, potentially, of abundant chromatin-associated proteins, such as topoisomerase I, which are indirectly pulled-down with the tagged transcription factor (Suppl. Table 1). We have validated a number of the remaining proteins as being true GATA-1 partners, some of which represent low abundance or weaker GATA-1 interactions, e.g. with TAL-1/Ldb1, Gfi-1b and ACF/WCRF. Importantly, purification required a single capture step.

We cannot be certain that we identified all GATA-1 complexes in differentiated MEL cells. Indeed, some of the size-fractionation profiles (Suppl. Fig. 1) suggest that there may be additional protein partners that were not identified perhaps due to their very low abundance or instability. This may be the case for the multimeric GATA-1/TAL-1/Ldb1/E2A/LMO2 complex. Sveral lines of evidence have suggested the presence of this complex in erythroid cells binding to distinct E-box and GATA motifs spatially arranged 9 to 12 nucleotides apart (review by (Lecuyer and Hoang, 2004). Many erythroid genes identified to-date contain such motifs, including GATA-1 itself, EKLF, glycophorin A and 4.2 protein (Lecuyer and Hoang, 2004). Evidence for the multimeric GATA-1/TAL-1 complex binding to erythroid genes *in vivo*, such as α globin and glycophorin A, has been provided recently by ChIP assays (Anguita et al., 2004; Lahlil et al., 2004). Nevertheless, we did not find any co-purification of E2A or LMO2 with GATA-1 from induced MEL cells. The complementary isolation by biotinylation tagging of protein partners, such as TAL-1, will be informative in that respect and may also reveal additional protein partners.

Novel GATA-1 protein partners

We describe here, for the first time, an interaction of GATA-1 with the essential hematopoietic transcription factor Gfi-1b. This factor contains six C-terminal C_2H_2 zinc fingers, which bind a defined DNA consensus sequence, and an N-terminal SNAG domain associated with repression (Doan et al., 2004; Duan and Horwitz, 2003). The Gfi-1b knockout is remarkably similar to that of GATA-1, i.e. it shows embryonic lethality E15 due to the developmental arrest of erythroid and megakaryocytic differentiation in the fetal liver (Saleque et al., 2002). Our data that GATA-1 and Gfi-1b (but not FOG-1 or MeCP1) are bound to the myb and to the myc promoters provide a basis for the similarities in the two knockouts (Pevny et al., 1995; Saleque et al., 2002). These data may also be related to the proliferation defects observed in GATA-1 overexpressing mice (Whyatt et al., 2000). It is important to note that although there are also similarities between the FOG-1 and Gfi-1b knockout phenotypes, we did not find FOG-1 and Gfi-1b to directly interact in induced MEL cells (Fig. 2). Possibly, the two factors regulate common gene targets through distinct complexes and binding sites. Alternatively, the functions of GATA-1 with FOG-1 or Gfi-1b could be separate, e.g. differentiation (FOG-1) versus proliferation arrest (Gfi-1b), with each function being essential for erythropoiesis.

We also describe, for the first time, interactions of GATA-1 with the MeCP1 and ACF/WCRF complexes, linking GATA-1 to repressive functions (with MeCP1) and chromatin structure. Previous evidence linking GATA-1 to chromatin structure involved interactions with the histone acetyltransferases (HATs) CBP and p300 (Blobel et al., 1998) and *in vitro* experiments where GATA-1 co-operated with the SWI/SNF remodeling complex in transcriptional activation (Kadam and Emerson, 2003). However, we did not observe these interactions in our GATA-1 purification from induced MEL cells or in immunoprecipitations (data not shown).

Our observations on the interactions of GATA-1 (and FOG-1) with the MeCP1 complex add to previous reports linking MeCP1 (and the closely related NuRD complex) to transcription factors in hematopoiesis (Hutchins et al., 2002; Kim et al., 1999; O'Neill et al., 2000). The conditional knockout of Mi-2β in thymocytes revealed a requirement in different stages of T cell maturation (Williams et al., 2004). In addition, the characterization of the MTA3 member of NuRD in B lymphocytes showed an interaction with BCL-6, a key repressor of the mature plasma cell transcription program (Fujita et al., 2004). It was thus suggested that MTA3 and the NuRD complex play a role in the maintenance of a population of less differentiated, "poised" B lymphocytes (Fujita et al., 2004). In contrast to these data, our data in erythroid cells suggest that the MeCP1 complex works with tissue-specific transcription factors to effect terminal differentiation by shutting down transcription programs associated with early multipotential ("poised") states.

GATA-1 and FOG-1 interactions

Considerable evidence has linked GATA-1 functions to FOG-1 (reviewed by (Blobel and Weiss, 2001; Cantor and Orkin, 2002). A single amino acid change in the N-terminal zinc finger of GATA-1 which abolishes 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). Our work suggests 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. Clearly, the association of GATA-1 and FOG-1 with the MeCP1 complex provides the molecular basis for the well-documented repressive properties of GATA-1 and FOG-1 interactions (Crispino et al., 1999; Fox et al., 1999; Letting et al., 2004; Pal et al., 2004). Only the slower migrating isoform of FOG-1 (Fig. 3) interacts with the GATA-1/MeCP1 complex, providing a potential mechanism for the selective formation of the GATA-1/FOG-1/MeCP1 complex.

We suggest that the separate GATA-1/FOG-1 complex without MeCP1 links GATA-1 with FOG-1 to transcriptional activation. For example, disruption of GATA-1 and FOG-1 interactions down-regulates erythroid genes such as α and β globin, Band 3, DC11 and HD2 genes (Crispino et al., 1999; Letting et al., 2004). ChIP assays have also shown GATA-1 and FOG-1 to be bound *in vivo* to active genes such as the A globin locus and the GATA-1 gene itself (Anguita et al., 2004; Pal et al., 2004). Significantly, in the α globin locus the GATA-1/FOG-1 complex occupies sites distinct from those occupied by the GATA-1/TAL-1/Ldb1 complex (Anguita et al., 2004), in agreement with our findings of distinct GATA-1 complexes.

Our finding that FOG-1 bridges GATA-1 to the repressive MeCP1 complex partly explains the common features of the GATA-1 and FOG-1 knockouts and the phenotypes caused by the single amino acid change in the N-terminal zinc finger of GATA-1 in mice and patients. In the GATA-1 knockout FOG-1/MeCP1 cannot be tethered to target genes, whereas in the FOG-1 knockout, the interaction between GATA-1 and the MeCP1 complex cannot take place. In patients, the lack of interaction between GATA-1 and FOG-1 would also fail to tether the MeCP1 complex to some of their target genes.

GATA-1 complexes and erythropoiesis

An important aspect in hematopoietic development to a particular lineage is the suppression of alternative "primed" lineage transcription programs and of genes that maintain multipotentiality, while upregulating genes associated with the differentiated cell type (Enver et al., 1998; Orkin, 2000). In addition erythroid terminal differentiation is accompanied by cell cycle arrest. GATA-1 has been implicated in the regulation of most of these aspects (Blobel and Weiss, 2001). In fact, a recent microarray analysis of GATA-1 dependent erythroid terminal maturation revealed an early wave of repression of genes like GATA-2, myc and myb, followed by the upregulation of erythroid specific genes (Welch et al., 2004). Here we identified two GATA-1 repressive complexes acting on distinct sets of genes. Thus we suggest that the GATA-1/Gfi-1b complex acts early and suppresses genes involved in cell proliferation, e.g. myc and myb, while the GATA-1/FOG-1/MeCP1 complex also acts early to suppress genes required to maintain the "primed" multipotential state, e.g GATA-2 and alternative hematopoietic lineage genes, e.g. MBP (Fig. 7). In contrast, the GATA-1/FOG-1 and the GATA-1/TAL-1/Ldb1 complexes would play a major role in the later upregulation of erythroid genes (Fig. 7). The role of the GATA-1/ACF/WCRF complex remains to be established. Thus, GATA-1 provides specific early versus late differentiation functions in the context of distinct complexes (Fig. 7). The model of different GATA-1 complexes executing specific tasks in different stages of erythroid differentiation suggests a dynamic aspect in the GATA-1 complex interactions during differentiation and also raises the prospect of dissecting the contribution of distinct GATA-1 interactions in erythropoiesis (i.e. essential versus dispensable) by selectively manipulating a specific GATA-1 complex at a time.

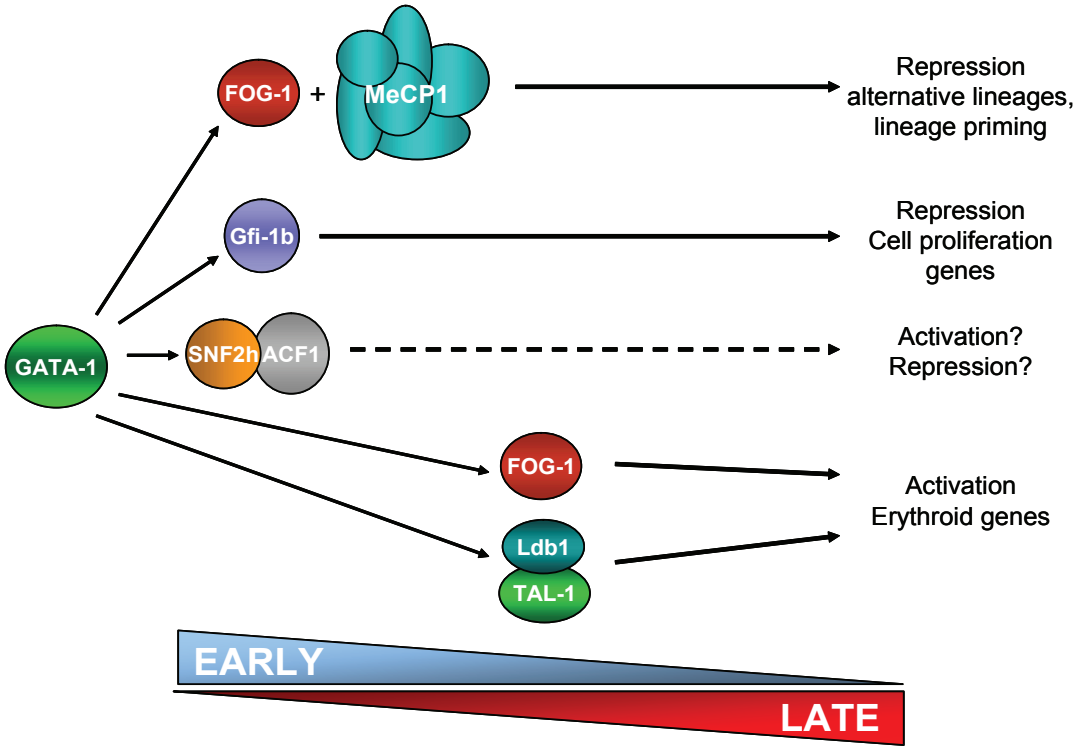


Figure 7: Model for the distinct GATA-1 complexes and their role in erythropoiesis. Broken arrow indicates unknown function and timing. See text for explanation.

Materials and Methods

Constructs, nuclear extract preparation, streptavidin binding, mass spectrometry and immunoblot analysis. Tagged constructs and procedures involving MEL cells, biotinylated proteins and mass spectrometry were previously described (de Boer et al., 2003). The GATA-1 zinc finger deletions have been described (Whyatt et al., 1993). G1E cells and induction were described (Tsang et al., 1997; Weiss et al., 1997).

Superose 6 gel filtration. Size fractionation of protein complexes was done on an AKTA FPLC apparatus with a Superose 6 10/30 column (Amersham Biosciences, Piscataway NJ). Fractions were precipitated with 100% trichloroacetic acid and analyzed by Western immunoblotting, as described (de Boer et al., 2003). Molecular size standards were thyroglobulin (670kDA) and albumin (66kDa) (Amersham Biosciences, Piscataway NJ).

Immunoprecipitations. Nuclear extracts were ple-cleared at 4°C using Protein G sepharose beads and affinity-purified IgG (rat [Santa Cruz, CA, sc-2026], rabbit [Santa Cruz, sc-2027], goat [Santa Cruz, sc-2028]) in HENG150 buffer. GATA-1 and TAL-1 antibodies were crosslinked to beads using dimethyl pimelimidate. Immunoprecipitations were performed in HENG150 / 0.3% NP-40 buffer overnight at 4°C using protein-G Sepharose beads. Washes were done at room temperature in HENG250 / 0.3% NP-40 buffer. Bound material was eluted by boiling in 1x Laemmli buffer.

HeLa transient transfection and transcription assays. GATA-1 and FOG-1 cDNAs cloned in pCDNA 3.1 (Invitrogen, Carlsbad CA) were transiently transfected using 2μg DNA and Lipofectamine 2000 (Invitrogen, Carlsbad CA). Cells were harvested after 24 hours and nuclear extracts were used for immunoprecipitations as above. pEGFP-N1 (Invitrogen, Carlsbad CA) was included as transfection efficiency control. Transcription was assayed by Real-Time PCR with primers for exon 2 of the pOVEC reporter plasmid. ChIP assays were done as below using GATA-1 and Mi-2β antibodies. The endogenous human GAPDH gene was used as control.

ChIP assays. Preparation of crosslinked chromatin ($2x10^7$ induced MEL cells treated with 0.4% formaldehyde for 10 minutes at room temperature), sonication to 300-800 base pair fragments and immunoprecipitations were as described in the Upstate protocol (www.upstate.com). Anti-GATA-1 protein-DNA immunocomplexes were immunoprecipitated in an additional step with an AffiniPure rabbit anti-rat antibody (Jackson ImmunoResearch Laboratories, West Grove, PA). Eosinophilic chromatin was prepared as previously described (Guyot et al., 2004). At least two independent chromatin immunoprecipitations were carried out per experiment. Antibodies used : GATA-1, N6 (Santa Cruz); Mbd2, S923 sheep polyclonal (Ng et al., 1999) and rabbit polyclonal anti Mbd2/3 antibody (Upstate 07-199); FOG-1 rabbit polyclonal (Tsang et al., 1997); TAL-1 rabbit polyclonal (Porcher et al., 1996); Gfi-1b D19 goat polyclonal (Santa Cruz sc-8559).

Real time PCR. Quantitative RealTime PCR (Opticon I, MJ Research) was done using SYBR Green I. PCR primers were designed by Primer Express 2.0 (PE Applied Biosystems). The qPCR Core Kit (Eurogentec, Belgium) was used with 400 nM of each primer under the following cycling conditions: 95^0C for 10 minutes, 40 cycles of 30 seconds at 95^0C, 60 seconds at 60^0C, 15 seconds at 75^0C. Enrichment for a specific DNA sequence was calculated using the comparative C_T method (Litt et al., 2001). PCR primer sequences are provided in Supplementary Materials and **Antibodies.** See Supplementary Information.

Acknowledgements

We are grateful to our many colleagues cited in Supplementary Methods for generously sharing reagents. We are also grateful to Karim Bouazoune and Alexander Brehm (Adolf Butenandt Institut, Munich) for help with the HDAC assays, Mitch Weiss (Children's Hospital of Philadelphia) for G1E cells, Sjaak Philipsen and Marieke von Lindern (Erasmus MC) for comments on the manuscript and Raymond Poot (CSC, London) for helpful discussions. We are grateful to Gerd Blobel (Children's Hospital of Philadelphia) for exchanging pre-publication data. This work has been supported by the Netherlands Research Organization (NWO), the European Union (grant HPRN-CT-2000-00078), the NIH (grant RO1 HL 073455-01) and the Netherlands Proteomics Center.

References

Ahringer, J. (2000) NuRD and SIN3 histone deacetylase complexes in development. *Trends Genet*, **16**, 351-356.

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

Anguita, E., Hughes, J., Heyworth, C., Blobel, G.A., Wood, W.G. and Higgs, D.R. (2004) Globin gene activation during haemopoiesis is driven by protein complexes nucleated by GATA-1 and GATA-2. *Embo J*, **23**, 2841-2852.

Blobel, G.A., Nakajima, T., Eckner, R., Montminy, M. and Orkin, S.H. (1998) CREB-binding protein cooperates with transcription factor GATA-1 and is required for erythroid differentiation. *Proc Natl Acad Sci U S A*, **95**, 2061-2066.

Blobel, G.A. and Weiss, M.J. (2001) Nuclear factors that regulate erythropoiesis. In Steinberg, M.H., Forget, B.G., Higgs, D.R. and Nagel, R.L. (eds.), *Disorders of Hemoglobin: Genetics, Pathophysiology, and Clinical Management.* Cambridge University Press, Cambridge, pp. 72-94.

Bochar, D.A., Savard, J., Wang, W., Lafleur, D.W., Moore, P., Cote, J. and Shiekhattar, R. (2000) A family of chromatin remodeling factors related to Williams syndrome transcription factor. *Proc Natl Acad Sci U S A*, **97**, 1038-1043.

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

Chang, A.N., Cantor, A.B., Fujiwara, Y., Lodish, M.B., Droho, S., Crispino, J.D. and Orkin, S.H. (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**, 9237-9242.

Chen, J. and Bender, T.P. (2001) A novel system to identify Myb target promoters in friend murine erythroleukemia cells. *Blood Cells Mol Dis*, **27**, 429-436.

Corona, D.F. and Tamkun, J.W. (2004) Multiple roles for ISWI in transcription, chromosome organization and DNA replication. *Biochim Biophys Acta*, **1677**, 113-119.

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

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

Doan, L.L., Porter, S.D., Duan, Z., Flubacher, M.M., Montoya, D., Tsichlis, P.N., Horwitz, M., Gilks, C.B. and Grimes, H.L. (2004) Targeted transcriptional repression of Gfi1 by GFI1 and GFI1B in lymphoid cells. *Nucleic Acids Res*, **32**, 2508-2519.

Du, J., Stankiewicz, M.J., Liu, Y., Xi, Q., Schmitz, J.E., Lekstrom-Himes, J.A. and Ackerman, S.J. (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**, 43481-43494.

Duan, Z. and Horwitz, M. (2003) Gfi-1 oncoproteins in hematopoiesis. *Hematology*, **8**, 339-344.

Eberharter, A., Ferrari, S., Langst, G., Straub, T., Imhof, A., Varga-Weisz, P., Wilm, M. and Becker, P.B. (2001) Acf1, the largest subunit of CHRAC, regulates ISWI-induced nucleosome remodelling. *Embo J*, **20**, 3781-3788.

Elefanty, A.G., Antoniou, M., Custodio, N., Carmo-Fonseca, M. and Grosveld, F.G. (1996) GATA transcription factors associate with a novel class of nuclear bodies in erythroblasts and megakaryocytes. *Embo J*, **15**, 319-333.

Enver, T., Heyworth, C.M. and Dexter, T.M. (1998) Do stem cells play dice? *Blood*, **92**, 348-351; discussion 352.

Feng, Q., Cao, R., Xia, L., Erdjument-Bromage, H., Tempst, P. and Zhang, Y. (2002) Identification and functional characterization of the p66/p68 components of the MeCP1 complex. *Mol Cell Biol*, **22**, 536-546.

Feng, Q. and Zhang, Y. (2001) The MeCP1 complex represses transcription through preferential binding, remodeling, and deacetylating methylated nucleosomes. *Genes Dev*, **15**, 827-832.

Fox, A.H., Liew, C., Holmes, M., Kowalski, K., Mackay, J. and Crossley, M. (1999) Transcriptional cofactors of the FOG family interact with GATA proteins by means of multiple zinc fingers. *Embo J*, **18**, 2812-2822.

Fujita, N., Jaye, D.L., Geigerman, C., Akyildiz, A., Mooney, M.R., Boss, J.M. and Wade, P.A. (2004) MTA3 and the Mi-2/NuRD complex regulate cell fate during B lymphocyte differentiation. *Cell*, **119**, 75-86.

Fujita, N., Jaye, D.L., Kajita, M., Geigerman, C., Moreno, C.S. and Wade, P.A. (2003) MTA3, a Mi-2/NuRD complex subunit, regulates an invasive growth pathway in breast cancer. *Cell*, **113**, 207-219.

Garcon, L., Lacout, C., Svinartchouk, F., Le Couedic, J.P., Villeval, J.L., Vainchenker, W. and Dumenil, D. (2005) Gfi-1B plays a critical role in terminal differentiation of normal and transformed erythroid progenitor cells. *Blood*, **105**, 1448-1455.

Grass, J.A., Boyer, M.E., Pal, S., Wu, J., Weiss, M.J. and Bresnick, E.H. (2003) GATA-1-dependent transcriptional repression of GATA-2 via disruption of positive autoregulation and domain-wide chromatin remodeling.

Proc Natl Acad Sci U S A, **100**, 8811-8816.
Guyot, B., Valverde-Garduno, V., Porcher, C. and Vyas, P. (2004) Deletion of the major GATA1 enhancer HS 1 does not affect eosinophil GATA1 expression and eosinophil differentiation. *Blood*, **104**, 89-91.
Hutchins, A.S., Mullen, A.C., Lee, H.W., Sykes, K.J., High, F.A., Hendrich, B.D., Bird, A.P. and Reiner, S.L. (2002) Gene silencing quantitatively controls the function of a developmental trans-activator. *Mol Cell*, **10**, 81-91.
Kadam, S. and Emerson, B.M. (2003) Transcriptional specificity of human SWI/SNF BRG1 and BRM chromatin remodeling complexes. *Mol Cell*, **11**, 377-389.
Kim, J., Sif, S., Jones, B., Jackson, A., Koipally, J., Heller, E., Winandy, S., Viel, A., Sawyer, A., Ikeda, T., Kingston, R. and Georgopoulos, K. (1999) Ikaros DNA-binding proteins direct formation of chromatin remodeling complexes in lymphocytes. *Immunity*, **10**, 345-355.
Lachman, H.M. and Skoultchi, A.I. (1984) Expression of c-myc changes during differentiation of mouse erythroleukaemia cells. *Nature*, **310**, 592-594.
Lahlil, R., Lecuyer, E., Herblot, S. and Hoang, T. (2004) SCL assembles a multifactorial complex that determines glycophorin A expression. *Mol Cell Biol*, **24**, 1439-1452.
Lecuyer, E. and Hoang, T. (2004) SCL: from the origin of hematopoiesis to stem cells and leukemia. *Exp Hematol*, **32**, 11-24.
Letting, D.L., Chen, Y.Y., Rakowski, C., Reedy, S. and Blobel, G.A. (2004) Context-dependent regulation of GATA-1 by friend of GATA-1. *Proc Natl Acad Sci U S A*, **101**, 476-481.
Litt, M.D., Simpson, M., Recillas-Targa, F., Prioleau, M.N. and Felsenfeld, G. (2001) Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. *Embo J*, **20**, 2224-2235.
Ng, H.H., Zhang, Y., Hendrich, B., Johnson, C.A., Turner, B.M., Erdjument-Bromage, H., Tempst, P., Reinberg, D. and Bird, A. (1999) MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nat Genet*, **23**, 58-61.
Nichols, K.E., Crispino, J.D., Poncz, M., White, J.G., Orkin, S.H., Maris, J.M. and Weiss, M.J. (2000) Familial dyserythropoietic anaemia and thrombocytopenia due to an inherited mutation in GATA1. *Nat Genet*, **24**, 266-270.
O'Neill, D.W., Schoetz, S.S., Lopez, R.A., Castle, M., Rabinowitz, L., Shor, E., Krawchuk, D., Goll, M.G., Renz, M., Seelig, H.P., Han, S., Seong, R.H., Park, S.D., Agalioti, T., Munshi, N., Thanos, D., Erdjument-Bromage, H., Tempst, P. and Bank, A. (2000) An ikaros-containing chromatin-remodeling complex in adult-type erythroid cells. *Mol Cell Biol*, **20**, 7572-7582.
Orkin, S.H. (2000) Diversification of haematopoietic stem cells to specific lineages. *Nat Rev Genet*, **1**, 57-64.
Osada, H., Grutz, G., Axelson, H., Forster, A. and Rabbitts, T.H. (1995) Association of erythroid transcription factors: complexes involving the LIM protein RBTN2 and the zinc-finger protein GATA1. *Proc Natl Acad Sci U S A*, **92**, 9585-9589.
Pal, S., Cantor, A.B., Johnson, K.D., Moran, T.B., Boyer, M.E., Orkin, S.H. and Bresnick, E.H. (2004) Coregulator-dependent facilitation of chromatin occupancy by GATA-1. *Proc Natl Acad Sci U S A*, **101**, 980-985.
Pevny, L., Lin, C.S., D'Agati, V., Simon, M.C., Orkin, S.H. and Costantini, F. (1995) Development of hematopoietic cells lacking transcription factor GATA-1. *Development*, **121**, 163-172.
Porcher, C., Swat, W., Rockwell, K., Fujiwara, Y., Alt, F.W. and Orkin, S.H. (1996) The T cell leukemia oncoprotein SCL/tal-1 is essential for development of all hematopoietic lineages. *Cell*, **86**, 47-57.
Querfurth, E., Schuster, M., Kulessa, H., Crispino, J.D., Doderlein, G., Orkin, S.H., Graf, T. and Nerlov, C. (2000) Antagonism between C/EBPbeta and FOG in eosinophil lineage commitment of multipotent hematopoietic progenitors. *Genes Dev*, **14**, 2515-2525.
Rylski, M., Welch, J.J., Chen, Y.Y., Letting, D.L., Diehl, J.A., Chodosh, L.A., Blobel, G.A. and Weiss, M.J. (2003) GATA-1-mediated proliferation arrest during erythroid maturation. *Mol Cell Biol*, **23**, 5031-5042.
Saleque, S., Cameron, S. and Orkin, S.H. (2002) The zinc-finger proto-oncogene Gfi-1b is essential for development of the erythroid and megakaryocytic lineages. *Genes Dev*, **16**, 301-306.
Shimizu, R., Takahashi, S., Ohneda, K., Engel, J.D. and Yamamoto, M. (2001) In vivo requirements for GATA-1 functional domains during primitive and definitive erythropoiesis. *Embo J*, **20**, 5250-5260.
Stopka, T. and Skoultchi, A.I. (2003) The ISWI ATPase Snf2h is required for early mouse development. *Proc Natl Acad Sci U S A*, **100**, 14097-14102.
Taunton, J., Hassig, C.A. and Schreiber, S.L. (1996) A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science*, **272**, 408-411.
Tsang, A.P., Fujiwara, Y., Hom, D.B. and Orkin, S.H. (1998) Failure of megakaryopoiesis and arrested erythropoiesis in mice lacking the GATA-1 transcriptional cofactor FOG. *Genes Dev*, **12**, 1176-1188.
Tsang, A.P., Visvader, J.E., Turner, C.A., Fujiwara, Y., Yu, C., Weiss, M.J., Crossley, M. and Orkin, S.H. (1997) FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. *Cell*, **90**, 109-119.
Visvader, J.E., Mao, X., Fujiwara, Y., Hahm, K. and Orkin, S.H. (1997) The LIM-domain binding protein Ldb1

and its partner LMO2 act as negative regulators of erythroid differentiation. *Proc Natl Acad Sci U S A*, **94**, 13707-13712.

Wadman, I.A., Osada, H., Grutz, G.G., Agulnick, A.D., Westphal, H., Forster, A. and Rabbitts, T.H. (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**, 3145-3157.

Weiss, M.J., Yu, C. and Orkin, S.H. (1997) Erythroid-cell-specific properties of transcription factor GATA-1 revealed by phenotypic rescue of a gene-targeted cell line. *Mol Cell Biol*, **17**, 1642-1651.

Welch, J.J., Watts, J.A., Vakoc, C.R., Yao, Y., Wang, H., Hardison, R.C., Blobel, G.A., Chodosh, L.A. and Weiss, M.J. (2004) Global regulation of erythroid gene expression by transcription factor GATA-1. *Blood*, **104**, 3136-3147.

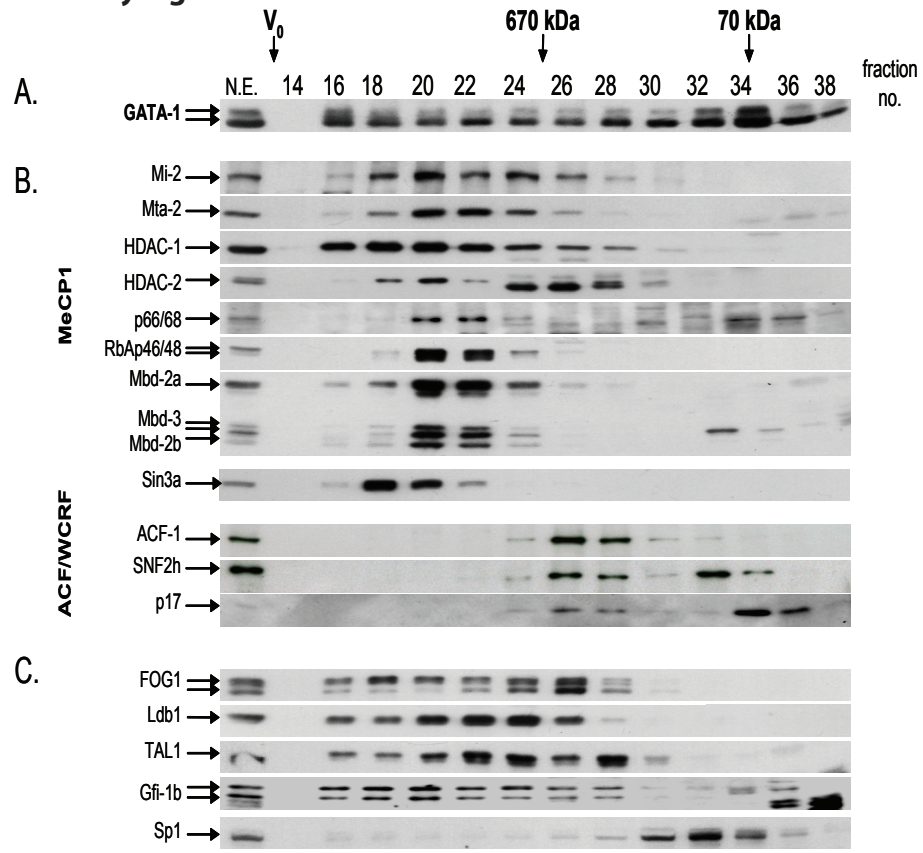
Westin, G., Gerster, T., Muller, M.M., Schaffner, G. and Schaffner, W. (1987) OVEC, a versatile system to study transcription in mammalian cells and cell-free extracts. *Nucleic Acids Res*, **15**, 6787-6798.

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

Whyatt, D.J., deBoer, E. and Grosveld, F. (1993) The two zinc finger-like domains of GATA-1 have different DNA binding specificities. *Embo J*, **12**, 4993-5005.

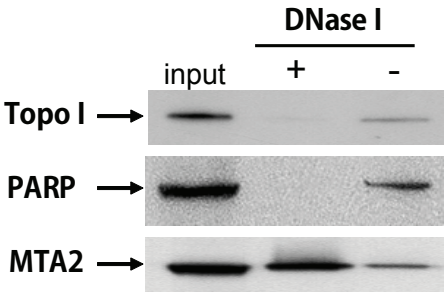
Williams, C.J., Naito, T., Arco, P.G., Seavitt, J.R., Cashman, S.M., De Souza, B., Qi, X., Keables, P., Von Andrian, U.H. and Georgopoulos, K. (2004) The chromatin remodeler Mi-2beta is required for CD4 expression and T cell development. *Immunity*, **20**, 719-733.

Supplementary Figures

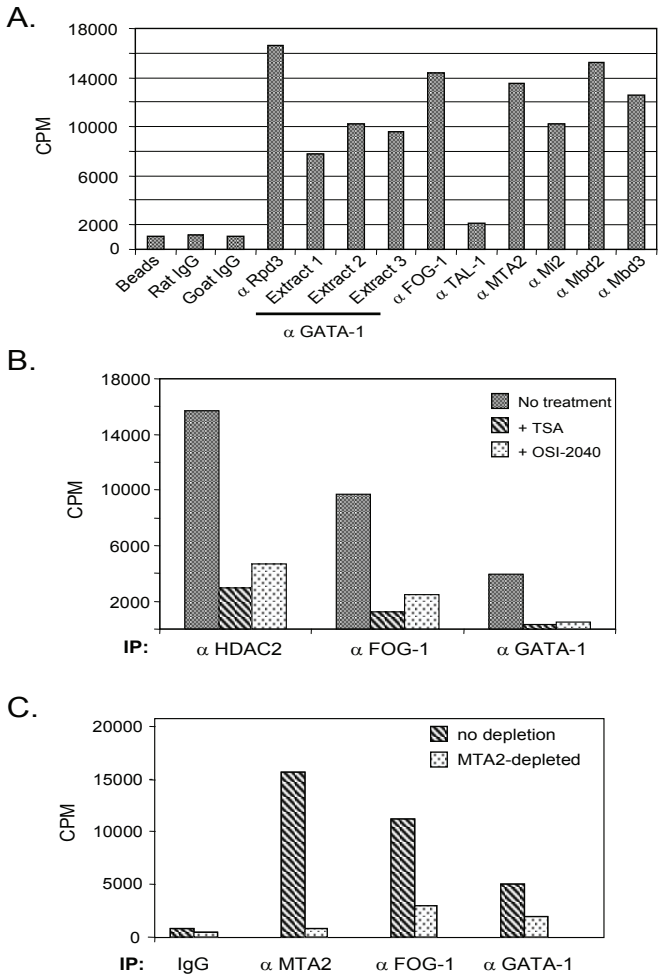


Supplementary Figure 1: Size-fractionation profiles by Superose 6 column of GATA-1 (panel A), members of the MeCP1 and ACF/WCRF complexes (panel B) and transcription factors (panel C). Molecular mass markers are indicated on the top. V_0: void volume. N.E.: input nuclear extract.

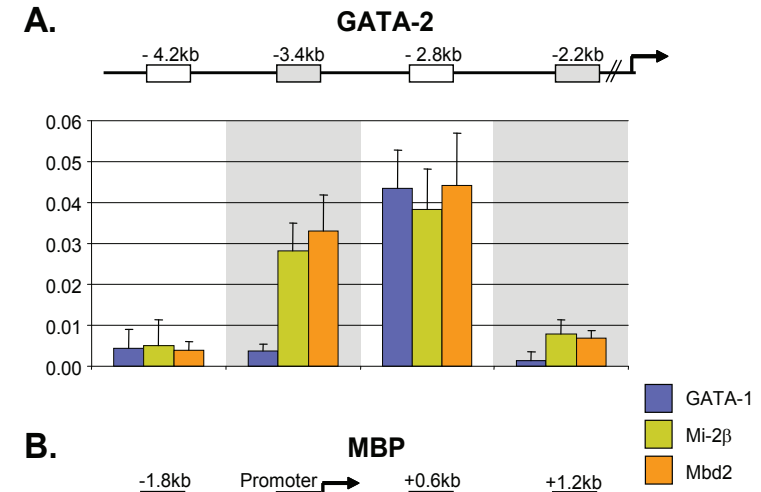
GATA-1 displays a broad fractionation profile with several peaks. The profile of tagged GATA-1 closely follows that of endogenous GATA-1, is stable in salt concentrations up to 1M and is not dependent on the presence of DNA (data not shown). Members of the MeCP1 complex showed overlapping peaks around fractions 20-22 (e.g. MTA2, RbAp46/48, Mbd2/3) in contrast to Sin3A which peaked around fractions 18-20. SNF2h and ACF1 peaked around fractions 26-28 and are distinctly different from those of the MeCP1 complex. p17 elutes in fractions 34-36, further suggesting that GATA-1 interactions with SNF2h and ACF1 occur in the context of the ACF/WCRF complex. These observations also suggest that GATA-1 interactions with MeCP1 and ACF/WCRF occur in distinct complexes. Fractionation profiles of transcription factors FOG-1, Ldb1, TAL-1 and Gfi-1b are largely coincident within the higher molecular weight fractions 16-28, though peaks vary between them. For most of these factors little or no protein is detected in the free protein fractions (i.e. fractions 32-38), in contrast to Sp1 which elutes as free protein.



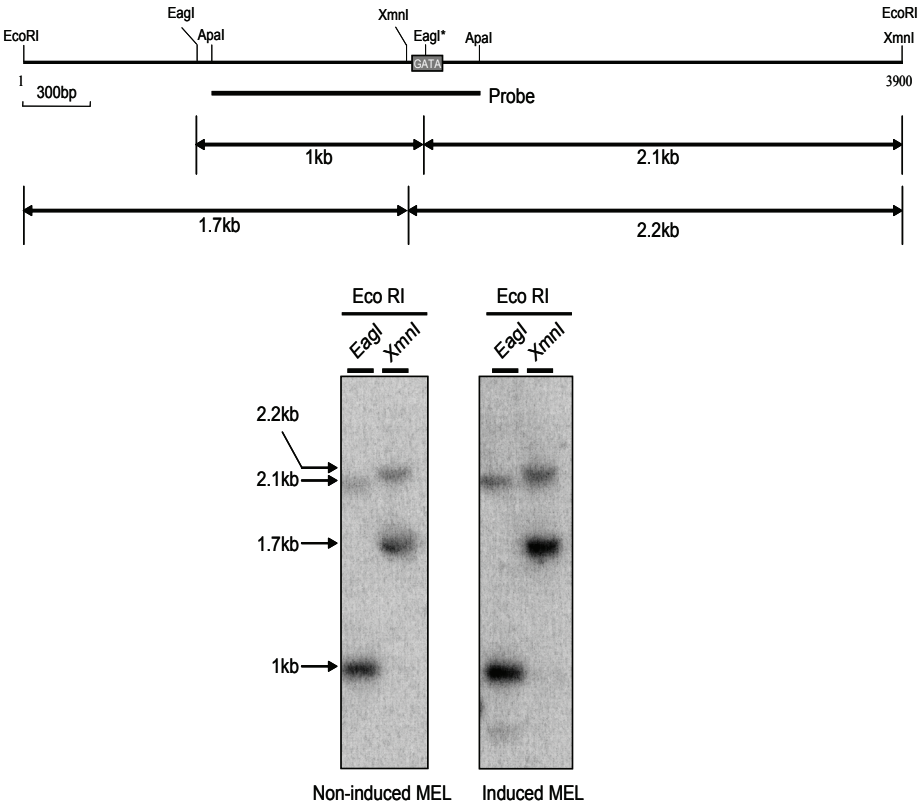
Supplementary Figure 2: DNase I treatment of nuclear extracts shows the indirect co-purification of abundant chromatin associated proteins with GATA-1. Nuclear extracts from MEL cells treated or not treated with DNAse I were immunoprecipitated with GATA-1 antibodies. Co-immunoprecipitation of abundant chromatin associated proteins such as topoisomerase I orPARP, identified by mass spectroscopy as co-purifying with GATA-1, was lost upon DNAse I treatment. By contrast, co-immunoprecipitation of MTA2, a member of the MeCP1 complex co-purified with GATA-1, was unaffected by DNase I treatment thus showing a direct interaction with GATA-1.



Supplementary Figure 3: Histone deacetylase (HDAC) assays. (A) HDAC activity associated with proteins immunoprecipitated by the indicated antibodies. Protein G beads and rat and goat immunoglobulins (IgG) were used as background controls. Three different nuclear extracts were used in the GATA-1 immunoprecipitations (extract 1: non-transfected MEL cells; extracts 2 and 3: biotinylated GATA-1). (B) HDAC activity immunoprecipitated by HDAC2, FOG-1 and GATA-1antibodies is sensitive to the Class I HDAC inhibitors TSA and OSI-2040. (C) Antibodies against the Mi2/NuRD-associated protein MTA2 deplete a considerable part of the HDAC activity associated with FOG-1 and GATA-1 immunoprecipitates. Immunoglobulins (IgG) were used as background control.



Supplementary Figure 4: ChIP assay to show GATA-1, Mbd2 and Mi-2 β binding to the GATA-1 (panel A) and MBP (panel B) loci in induced MEL cells. It is clear that the binding pattern observed for GATA-1 and Mbd2 in these gene loci in Figure 5A and 5C is reproduced by the binding pattern of Mi-2β. Mi-2β antibody: as in (Fujita et al., 2003).



Supplementary Figure 5: DNA methylation assay by restriction enzyme digestion at the –2.8kb element of the GATA-2 locus. Genomic DNA (5-10□g) from non-induced (left panel) and induced (right panel) MEL cells was first digested with Eco RI, which releases a 3.9kb fragment, followed by digestion with Eag I (methylation sensitive, indicated by asterisk), or Xmn I as control for complete digestion of the genomic DNA samples. Eag I maps close to the GATA-1 binding sites in the –2.8kb element of the GATA-2 locus and within the PCR fragment (grey box, not to scale) amplified in the ChIP assays shown in Figures 5 and 6. Digested DNA was blotted and probed with a 1.2kb Apa I fragment (solid line) which detects 1kb and 2.1kb fragments on the Eag I digests and 1.7kb and 2.2kb fragments on the Xmn I digests. For both digests, the larger fragments are weaker due to their limited overlap with the probe. It can be seen that Eag I digests completely in DNA of both induced and non-induced MEL cells, thus suggesting that its recognition site is not affected by DNA methylation.

GO Biological process	GO Cell. component	Total no. of peptides	Remarks
chromatin modeling/ modification	nucleus	289	ATPase-dependent chromatin remodeling, HDACs
transcriptional regulation Unknown		158	putative transcription factors, RNA binding proteins, nucleolar proteins, receptors etc.
mRNA processing	nucleus	139	splicing factors
DNA topological change	nucleus	130	topoisomerases
transcription regulation, DNA dependent	nucleus	105	Transcription factors
Metabolism	mitochondrion	54	carboxylases
protein biosynthesis	cytosol	53	ribosomal proteins
ribosome biogenesis	nucleolus	37	apoptotic chromatin condensation inducer in the nucleus
DNA repair	nucleus	28	XRCC1, Ku autoantigen
nucleolus biogenesis	nucleus	24	
cell proliferation	nucleus	23	Ki-67
DNA replication	nucleus	23	ORC
chromatin assembly	nucleus	21	histones
Cytoskeleton		18	tubulin
structural molecule	nucleus	11	NuMA1
protein targeting	nucleus	7	
Apoptosis	nucleus	4	CDC5
protein chaperone	nucleus	4	
cell cycle	nucleus	3	RCC1
signal transduction		2	GTP binding protein 3
Total # of peptides:		**>1100**	
Gel slices:		**~45**	
MW range:		**15-324kDa**	

Supplementary Table I: Classification by Gene Ontology (GO) criteria of proteins identified by mass spectrometry as co-purifying with GATA-1

PROTEIN IDENTITY	*NUMBER OF PEPTIDES*	*ADDITIONAL PURIFICATIONS*
MeCP1 complex		
Mi-2	82	+
HDAC 1	10	+
HDAC 2	17	+
MTA1	47	+
MTA2	10	+
MTA3	4	-
Mbd2	14	+
Mbd3	9	+
p66	11	+
RbAp46	10	+
RbAp48	8	+
ACF/WCRF complex		
SNF2h	21	+
ACF1	4	-
Transcription factors		
FOG-1 (Hem.)	47	+
TAL-1 (Hem.)	2	+
Gfi-1b (Hem.)	1	-
Ldb1 (Ubiq.)	1	+
DNA repair		
Rfc5	10	+
XRCC1	3	-
Ku70	9	+
PARP	10	+
DNA ligase III-β	1	-
DNA Topological change		
DNA Topo I	34	+
DNA Topo II α	64	+
DNA Topo II β	32	-

Table I: Proteins specifically co-purifying with biotin tagged GATA-1 as compared to the control purification (de Boer et al., 2003). A number of these proteins have been validated by immunoprecipitations and other assays (see text). Hem: hematopoietic transcription factors. Ubiq: ubiquitous transcription factors

Materials and Methods

ChIP PCR primer sequences:

EKLF upstream enhancer forward PCR primer: 5'-CTGGCCCCCCTACCTGAT-3'
EKLF upstream enhancer reverse PCR primer: 5'-GGCTCCCTTTCAGGCATTATC-3'

EKLF -1.35kb forward PCR primer: 5'-TGCTCCCCACTATGATAATGGA-3'
EKLF -1.35kb reverse PCR primer: 5'-GCCACAACCAAAGAAGACATTTT-3'

MBP -1.2kb forward PCR primer: 5'-GGGTCTAATTCCGAGGGTGAGT-3'
MBP -1.2kb reverse PCR primer: 5'-GGCCTGGAAATCACTGAGCTA-3'

MBP promoter forward PCR primer: 5'-CCGCCAAGGTGTCTATAAATGC-3'
MBP promoter reverse PCR primer: 5'-TGGGTCTTGTCAAGTTTGCAAA-3'

MBP +0.6kb forward PCR primer: 5'-GAAGTAGAGGCAGGATAATCAGGAA-3'
MBP +0.6kb reverse PCR primer: 5'-AGGATGAACCAGGGCTAATGC-3'

MBP +1.8kb forward PCR primer: 5'-TGTGACAGACGTGGACCTTCA-3'
MBP +1.8kb reverse PCR primer: 5'-TGCATCCAGAGTCACCCATAAG-3'

GATA2 -4.2kb region forward PCR primer: 5'-GAATTTCCTGCCGGTCCAT-3'
GATA2 -4.2kb region reverse PCR primer: 5'-GACGCGTTGGCTTTGTGTG-3'

GATA-2 -3.4kb forward PCR primer: 5'-TCCATCCAGCAGCTTTAGGAA-3'
GATA-2 -3.4kb region reverse PCR primer: 5'-GGGTTCGAAGCCACTCCAA-3'

GATA-2 -2.8kb region forward PCR primer: 5'-CCGGGCAGATAACGATTGG-3'
GATA-2 -2.8kb region reverse PCR primer: 5'-TTCATCTCGGCCGGCTAAT-3'

GATA-2 -2.2kb region forward PCR primer: 5'-AGGACCCCCTGCTTCTTGTTTC-3'
GATA-2 -2.2kb region reverse PCR primer: 5'-GGCAGTATGAGGCCCAGAATCTT-3'

Myb promoter forward PCR primer: 5'-GGGCGCCAGATTTGG-3'
Myb promoter reverse PCR primer: 5'-GGAGGAAACAGGTTGATATTAAAGT-3'

Myb -0.8kb forward PCR primer: 5'-GTAGGTTTGTCCAGCAAGTGTTTG-3'
Myb -0.8kb reverse PCR primer: 5'-AGGTGCCTACCACGCACTTCT-3'

Myc promoter forward PCR primer: 5'-CCAGACATCGTTTTTCCTGCATA-3'
Myc promoter reverse PCR primer: 5-'CCGCTCAGTGTGTGGAGTGATA-3'

Myc -0.7kb forward PCR primer: 5'-ACACACACATACGAAGGCA-3'
Myc -0.7kb reverse PCR primer: 5'-ACCGTTAACCCCTTCCTCCC-3'

pOVEC exon 2 forward PCR primer: 5'-TCACCTGGACAACCTCAAAGG-3'
pOVEC exon 2 reverse PCR primer: 5'-CAGGATCCACGTGCAGCTT-3'

GAPDH forward PCR primer: 5'-TGAAGGGGAAGCTCAGTCG-3'
GAPDH reverse PCR primer: 5'-TCCACCACCCTGTTGCTGTA-3'

PCR primers for RNA analysis in G1E cells

GATA-2 mRNA primers:
Exon III forward PCR primer: 5'-ACTATGGCAGCAGTCTCTTCCATC-3'
Exon V reverse PCR primer: 5'-AAGGTGGTGGTTGTCGTCTGAC-3'

DNase I treatment of nuclear extracts

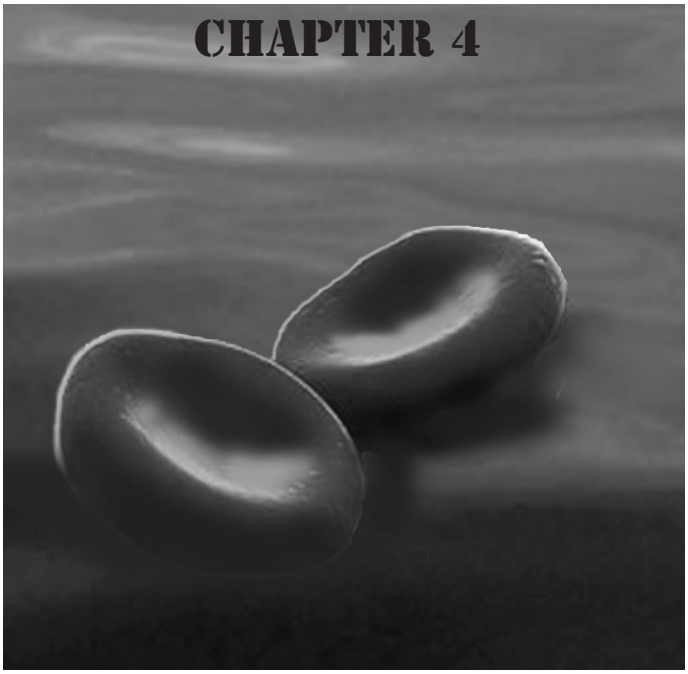
Nuclear extracts (50μg) were diluted to 150 mM KCl with HENG buffer and 5 % glycerol. 10 units of DNaseI (Invitrogen) were added and the extract was incubated overnight at 4°C with the addition of 5 mM $CaCl_2$ and 10 mM $MgCl_2$. Extracts were bound to streptavidin beads and analyzed by western blotting. Antibodies used were:

Topoisomerase I goat polyclonal (Santa Cruz, sc-5342); PARP rabbit polyclonal (Alexis Biochemicals, ALX-210-302).

HDAC assays. Immunoprecipitations for HDAC assays were carried out using 0.3-0.5mg of nuclear extracts in HENG150 / 0.1% NP-40, as described above. Immunoprecipitates were washed once for 10 minutes in HENG150 / 0.3% NP-40 and three times for 10 minutes each in HENG300 / 0.3%NP-40. Beads were resuspended in HENG50 and HDAC assays were done using approximately 32,000cpm of ^{3}H-labelled core histones per reaction, as previously described (Taunton et al., 1996).

Antibodies. The following antibodies were purchased from Santa Cruz Biotechnology (Santa Cruz, CA): N6 GATA-1 rat monoclonal (sc-265); M20 FOG-1 (sc-9361); SNF2h (sc-8760), Gfi-1b D19 (sc-8559) and TAL-1 (sc-12982) were goat polyclonal antibodies; mouse monoclonal antibodies for mSin3A (sc-5299) and HDAC2 (sc-9959); rabbit polyclonal antibodies for HDAC1 (sc-7872) and Sp1 (sc-59). 15G12 RbAp46/48 mouse monoclonal (ab490), MTA2 goat polyclonal (ab9949) and MBD2a rabbit polyclonal (ab3754) antibodies were purchased from Abcam (Cambridge, UK). p66 mouse monoclonal was purchased from Upstate Biotechnology (07-365; Waltham, MA). Rabbit polyclonal antibodies against TAL-1 were a generous gift by Richard Baer (Columbia University, NY) and by Catherine Porcher (WIMM, Oxford). Rabbit polyclonal antibodies against Ldb1 were kindly donated by Gordon N. Gill (Stanford University, CA). Rabbit polyclonal antibodies against FOG-1 were a generous gift by Stuart H. Orkin (Harvard Medical School, MA). Rabbit polyclonal antibodies against MTA2, Mi2 and Mbd2/3 were kindly donated by Paul A. Wade (NIH/NIEHS, NC). Sheep anti-serum S923 against Mbd2 and rabbit polyclonal antibody R593 against Mbd2/Mbd3 were generously provided by Adrian Bird (University of Edinburgh, UK). Rabbit polyclonal antibodies against ACF1 were a generous gift of Patrick Varga-Weisz (Marie Curie Research Institute, Surrey, UK). Rpd3 rabbit polyclonal antibodies were donated by Alexander Brehm (Adolf Butenandt Institut, Germany). Mouse monoclonal antibody against B23 nucleophosmin was a kind gift by Pui K. Chan (Baylor College of Medicine, TX). Secondary antibodies conjugated to horseradish peroxidase were purchased from Dako (DakoCytomation, Denmark) and Amersham Biosciences and developed by chemiluminescence using the ECL-PLUS kit (Amersham).

CHARACTERIZATION OF TR2/TR4 COMPLEXES IN MOUSE ERYTHROLEUKEMIC (MEL) CELLS BY in vivo BIOTINYLATION TAGGING AND MASS SPECTROMETRY



Manuscript in preparation

CHARACTERIZATION OF TR2/TR4 COMPLEXES IN MOUSE ERYTHROLEUKEMIC (MEL) CELLS BY in vivo BIOTINYLATION TAGGING AND MASS SPECTROMETRY

Katarzyna E. Kolodziej[1], Osamu Tanabe[2], Jeroen Demmers[1], Frank Grosveld[1], James Douglas Engel[2] and John Strouboulis [1,3]

[1] Department of Cell Biology, Erasmus MC. Dr Molewaterplein 50, 3015 GE Rotterdam, the Netherlands
[2] Depratment of Cellular and Developmental Biology, University of Michigan, 109 Zina Pitcher Place, 2620 BSRB, Ann Arbor, MI 48109-2200
[3] Present address: Institute of Molecular Oncology BSRC "Alexander Fleming" P.O. Box 74145, 16602 Varkiza Greece

Abstract

The TR2 and TR4 orphan nuclear receptors are the core components of the DRED complex previously implicated in ε and γ globin gene silencing in the adult (92, 93). Initially, DRED was purified as 540kDa protein complex, however only TR2 and TR4 were shown to bind to DR1 regulatory elements present in the ε and γ globin gene promoters (92). In order to identify additional DRED complex components other than TR2/TR4 we applied in vivo biotinylation tagging (13) of TR2 and TR4 in erythroid cells for the efficient isolation and characterization of protein complexes. Here we present biochemical evidence for TR2/TR4 interactions with the DNA methyltrasferase 1, the histone demethylase Lsd1, with histone deacetylases such as HDAC3 and transcriptional co-repressors, such as TIF1β and RbAp46/48. The possible function of these factors in transcriptional repression are in agreement with previous observations of the DRED complex being responsible for ε and γ globin gene silencing (70, 92, 93, 95). Our collected data suggest that TR2/TR4 form two distinct complexes that may have distinct functions. However, further investigation is required to confirm this model and to further relate these interactions to the initially identified DRED complex and its functions.

Introduction

The CAC and CCAAT box is one of the most conserved common motifs found in the globin promoters. The distal CCAAT box in the human γ-globin promoter contains an approximate direct repeat (DR1) sequence TGACCAATAGCC. Importantly the CAAT box regions of the human ε- and ζ –globin genes as well as of the murine εy- and βh1-globin genes also contain a DR-1 type sequence, as shown in Table 1 (24).

Table1. After (24) and (93). Alignment of the human and mouse globin genes promoter sequence. Potential DR1 sites indicated by bold fonts.

	Distal DR1 **AGGTCAnAGGTCA**	Proximal DR1 **TGACCTnTGACCT**
Human ε	**AGGaCA**.c**AGGTCA**gc	**TGACCa**a**TGACt**T
Mouse εy	**AGGaC**cac**GGGTCA**gg	**TGACCa**a**TGgCt**T
Human γ	ggttggccagccttg	**TGACCa**a**TagCCT**
Mouse βh1	tcttgcccagactct	**TGACCa**a**TagCCT**
Human β	ggagccacaccct	**TGgCCa**a**TctaCT**
Mouse βmaj	agagccacaccctg	**gGgCCa**a**TctgCT**
Aγ -117 HPFH		**TaACCA**a**TAGCCT**

Direct repeat (DR) sites are direct binding targets for nuclear receptors (63). Interestingly, the hereditary persistence of fetal haemoglobin (HPFH) an inherited human condition resulting in continued γ-globin gene expression in adult life (73) has been associated with deletions/mutations in distinct sites in the γ-globin promoter, including DR1 sites such as in the -117 Greek HPFH (11). A growing list of factors including GATA-1, NF-E3, COUP-TF1, EKLF, Sp1, CP1/NF-Y and CDP have been suggested to bind to the CAC/CCAAT motifs in globin promoters (24, 95).

DNA methylation has also been implicated in γ globin silencing (17, 18, 30, 90). A high local concentration of CpG residues is found in the 5'γ globin promoter (five CpGs in the first 105bp) that is absent from the other β-like globin gene promoters (83). Previous observations suggested that there is a correlation between γ-globin gene expression and the lack of methylation of CpGs in the γ-globin promoter region (64, 83, 98). Moreover, phylogenetic footprinting studies demonstrated that acquisition of CpG residues in the γ globin promoter coincided with recruitment of the γ-globin gene to fetal expression in primates (83, 91). In addition, 5-azacytidine and its derivatives, which are known DNA methyltrasferase inhibitors, have been tested on baboons and human patients for their potential to re-activate γ-globin expression in adults, with very promising results (22, 33, 55, 56). However, concern about potential carcinogenic effects have halted the development of 5-azacytidine derivatives as therapeutic agents for patients with β-thalassemia and sickle cell disease (20).

Initial studies from the Engel laboratory showed that the suppression of the ε-globin transcription during definitive erythropoiesis is mediated by the binding of a repressor that prevents EKLF from activating the ε-globin gene via the CAC element (95). That repressor activity was further attributed to a novel protein complex called DRED (Direct Repeat Erythroid Definitive) as detected by electrophoretic mobility shift assays (95). Subsequently the Engel laboratory showed that the DNA binding core activity of the DRED complex consists of two orphan nuclear receptors TR2 and TR4. The DRED complex binds to DR1 elements in ε- and γ- globin gene promoters to mediate stage-specific globin switching (92, 95). DNA-affinity purification of the DRED complex revealed five proteins: CDP, USF-1, LBP-1a, and the orphan nuclear receptors TR2 and TR4. However, DRED DNA binding activity was confirmed by EMSA only for TR2 and TR4 (92). TR2 and TR4 belong to the subfamily of orphan nuclear receptors and are expressed in many tissues (including hematopoietic) at all developmental stages (51). Although TR2 and TR4 preferentially

heterodimerize, they may form homodimers (49). Significantly, initial gel filtration studies showed that the DRED repressor complex migrates as a much larger species than the TR2:TR4 heterodimer (92). In order to identify additional TR2 and TR4 interactors we applied an in vivo biotinylation tagging approach in order to purify TR2/TR4 complexes from mouse erythroleukemic (MEL) cells (13). To this end, we generated stably transfected MEL cell lines expressing the BirA biotin ligase (13) and biotin-tagged versions of both TR2 and TR4 together.

In vivo biotinylation was previously successfully applied by our laboratory to study the interacting partners of GATA-1 and Ldb1 in erythroid cells. Previous work ((13, 31, 74) and discussed in greater detail elsewhere in Chapters 2 and 3) showed that biotin tagging offered a powerful tool for the study of protein complexes in erythroid (and other mammalian) cells.

Here we show that we could express biotin tagged versions of TR2 and TR4 in MEL cells. The biotin-tagged TR2/TR4 could be efficiently bound to magnetic M280 Dynabeads. However, initial experiments with the TR2/TR4 complex purification indicated that significant optimization was needed in order to enrich for TR2 and TR4 complexes. Here we present our successful attempts at optimizing TR2/TR4 complex purification.In addition, we present the biochemical identification and analysis of the mass spectrometry results of the TR2/TR4 complexes. We identified a number of chromatin associated co-repressors, including DNA methyltransferase (Dnmt1), members of the NuRD complex and TIF1β co-purifying with TR2/TR4. Interestingly, we also identified the Lsd1 (AOF 2) Lysine Specific Histone Demethylase, which has recently been reported to interact with hematopoietic transcription factors Gfi1b and GATA-2 ((80) and H. Braun unpublished observations). Significantly, the most prominent interacting partner co-purifying with TR2/TR4 is the DNA methyltransferase Dnmt1. In this chapter, we also present results suggesting that the mechanism of 5-azacytidine mediated γ globin re-activation involves the specific degradation of Dnmt1, TR2 and TR4. These findings provide new leads for investigating how these co-repressors may contribute to the repression the ε- and γ-globin genes.

Material and methods:

DNA Constructs: N-terminally tagged TR2 and TR4 were constructed by cloning the 23-amino acid biotin tag in the NcoI site overlapping the start codon of both proteins. Tagged TR2 and TR4 cDNAs were then cloned in the pEV-Neo β-globin based erythroidexpression vector. DNA constructs were made by Dr Shoko Kobayashi (the Engel lab, University of Michigan).

Cell culture: MEL cells were cultured in DMEM medium supplemented with 10%FCS and Pen/Strep and were induced by adding DMSO (Sigma) to a final concentration of 2%. Stable transfectants (Figure 1) were obtained as previously described (3, 13). Nuclear extracts were prepared as described in (2, 74).

Antibodies: TR2 and TR4 rabbit polyclonal antibodies were from Santa Cruz (Cat No sc-9087 and ac 9086) or generated by Engel's laboratory. Anti-Dnmt1 antibodies were mouse monoclonal ab13537 (Abcam), rabbit polyclonal ab 19905 (Abcam) and mouse monoclonal IMG-261A (Imgenex). TIF1α, -β, and -γ antibodies were generously donated by Dr. Florence Cammas (IGBMG, Strasbourg) (7). The actin antibody was from Abcam (ab 8226) and HDAC-1 antibodies were from Santa Cruz (mouse monoclonal sc- 8410) and Abcam (rabbit polyclonal ab 7028). HDAC 3 antibody was from Abcam (rabbit polyclonal; ab 7030). Streptavidin HRP was from Perkin Elmer (NEL 750).

Binding to strepatvidin beads: was done as described in Chapter 2 and 3 (74, 75). Benzonase (Novagen) treated samples were supplemented with 10mM (final concentration) of CaCl2, and benzonase was used in a ratio of 500 U per 1mg total protein.

Ammonium sulphate precipitation: was carried out as described in the initial DRED isolation (92).

Sucrose gradient: different concentrations of sucrose (Sigma) varying from 70% to 10%

were prepared in HEPES buffer (pH 7.9) buffer and stored at 4ºC (swing out rotor). The gradient was assembled in 5mL polyallomer tubes (Beckman), (Figure 6A) and ultracentrifuged 40000 rpm over-night at 4ºC. Fractions (500µl each) were collected from the top of the gradient and injected onto the analytical superpose 6 gel filtration column. Fractionation profiles were analysed by Western immunoblotting.

Size fractionation by Superose 6 column: size fractionation using the analytical Superose 6 column and processing of fractions for analysis were carried out as described in (75) Appendix 1. Fractionation using the preparative column was done as follows. 5 mL of nuclear extract, equivalent to 40 mg of total protein was injected onto a preparative grade Superose 6 XK50/600 column connected to the AKTA FPLC system (Amersham Biosciences UK), equilibrated in running buffer (20mM HEPES pH 7.9, 0.5mM EGTA, 1mM MgCl2, 200mM KCL, 10% glycerol). The column was ran at 4º C. 10mL fractions (numbers 18 to 38) were pooled together and binding to streptavidin beads was carried out as described in (74) Chapter 2 and 3. Eluted proteins were resolved by SDS-PAGE and subsequently analysed by mass spectrometry. Alternatively, proteins bound to the streptavidin beads were directly subjected to trypsinization, as described below and subsequently analysed by mass spectrometry.

On beads trypsinization: was done as described in (78, 79). After binding and washing, M280 paramagnetic beads were resuspended in 50mM ammonium bicarbonate and trypsin was added (60ng of trypsin per mg of total protein used for binding). Digestion was carried out overnight at 37º C. Beads were removed using the magnetic rack and the supernatant containing trypsinized peptides was collected and immediately injected onto a mass spectrometer.

Mass spectrometry analysis: was done as previously described in (65, 74, 75, 81). SDS-PAGE gel lanes were cut into slices using an automatic gel slicer and subjected to in-gel trypsinization (Promega, sequencing grade), essentially as described (86, 104). Alternatively, after stringent washing the streptavidin beads were incubated overnight at 37ºC with trypsin in ammonium bicarbonate buffer for on-beads digestion as described in (78, 79). NanoLC-MS/MS was performed on either a CapLC system (Waters, Manchester, UK) coupled to a Q-ToF Ultima mass spectrometer (Waters, Manchester, UK), operating in positive mode and equipped with a Z-spray source, or an 1100 series capillary LC system (Agilent Technologies) coupled to an LTQ mass spectrometer (Thermo) operating in positive mode and equipped with a nanospray source. Peptide mixtures were trapped on a Jupiter™ C18 reversed phase column (Phenomenex). Peptide separation was performed on Jupiter™ C18 reversed phase column (Phenomenex) using a linear gradient from 0 to 80% B (A = 0.1 M acetic acid; B = 80% (v/v) acetonitrile, 0.1 M acetic acid). The column eluent was directly sprayed into the ESI source of the mass spectrometer. Mass spectra were acquired in continuum mode; fragmentation of the peptides was performed in data-dependent mode.

Data analysis and protein identification: peak lists were automatically created from raw data files using the ProteinLynx Global Server software (version 2.0; Waters, Manchester, UK) for Q-ToF spectra and the Mascot Distiller software (version 2.0; MatrixScience, London, UK) for LTQ spectra. The Mascot search algorithm (version 2.0, MatrixScience, London, UK) was used for searching against the NCBInr database (release data: 3rd March 2006; taxonomy: M. musculus). The Mascot score cut-off value for a positive protein hit was set to 50. Individual peptide MS/MS spectra with Mowse scores below 40 were checked manually and either interpreted as valid identifications or discarded.

Identified proteins listed as NCBInr database entries, were screened to identify proteins that were also identified streptavidin pull-down experiments with extracts from “BirA-only” MEL cells (6). These were removed as background binding proteins. Remaining proteins were classified according to Gene Ontology criteria (http://www.informatics.jax.org/), and grouped according to Biological Process and/or Molecular Function.

Chromatin Immunoprecipitations: were done according to the ChIP kit Upstate/Milipore protocol, with minor modifications. We used 1% gelatine from cold water fish skin for blocking the beads, and no-SDS lysis buffer (Chapter 5). Double cross-linking using Di(N0succinimidyl) glutarate (Sigma), according to Porro and Perini protocol (72). Primers were designed with Primer express 2, in addition to primers previously published in (94). Primers used for PCR are listed in Table 1.

Results

Optimization of TR2 and TR4 high molecular weight complexes from erythroid cells

We obtained stably transfected MEL cell clones expressing bio-TR2 or bio-TR4 or both double bio-TR2/bio-TR4 (Figure 1). Western blot analysis with TR2 and TR4 antibodies showed that the tagged forms were expressed at levels 2-3 fold higher than the endogenous protein levels, particularly in induced MEL cells where the tagged protein expresses at high levels as a result of the β-globin promoter of the pEV-Neo vector (Figure 2 input lane). All selected clones expressed tagged TR2/TR4 in both non-induced and DMSO induced MEL cells.

The efficiency of biotinylation was tested in all three stable transfectants by direct binding

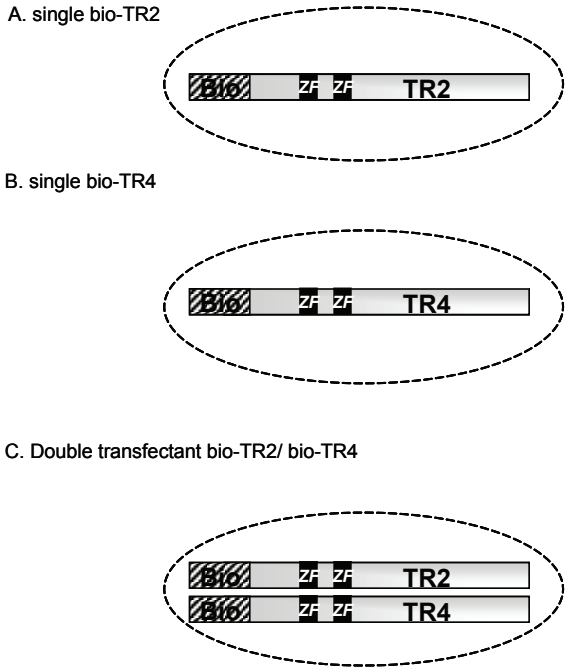


Figure 1: Scheme depicting biotinylation tagging strategy of TR2 and TR4 in order to obtain single (panel A and B) or double (panel C) MEL cells transfectants expressing biotin-tagged TR2 and TR4. cell lines, expressing (next to endogenous protein) either bio-TR2 only (A), bioTR4 only (B), both bio-TR2 and bioTR4 (C)

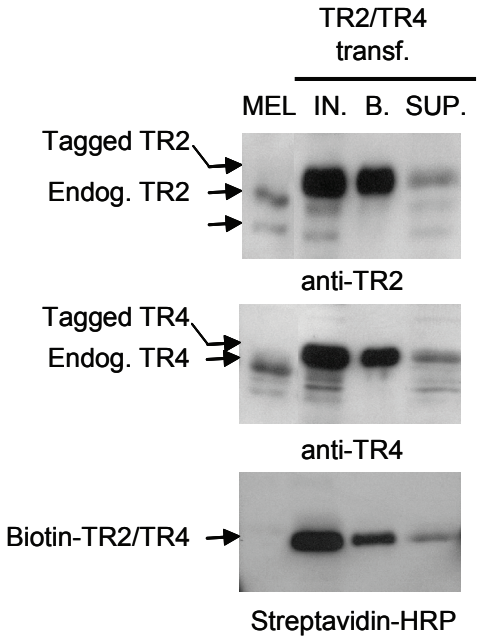


Figure 2: Efficient biotinylation of tagged TR2 and TR4 in stably transfected MEL cell lines. Streptavidin pull-down of nuclear extracts from double transfectant cell line shows that the majority (~90%) of bioTR2 (top panel) and bio-TR4 (middle panel) are bound to the beads, with little tagged protein left in the supernatant. This is also shown with streptavidin–HRP detection (bottom panel). Non-transfected MEL and input (IN.) lanes indicate the relative amounts of endogenous protein and biotin-tagged TR2 and TR4. Input and supernatant (SUP.) lanes were loaded with 10% of total material, bound (B.) lane was loaded with the equivalent of 1% of total protein used for binding. Detection using anti TR2 and anti TR4 antibodies (upper and middle panels) and streptavidin-HRP (bottom panel).

to streptavidin beads, which showed that 80-90% of the tagged protein is biotinylated in all transfectants. We decided to focus our attention on the analysis of the double bio-TR2/bio-TR4 transfectant (Figure 2) since the initial analysis of the DRED complex revealed that the TR2/TR4 heterodimer has the highest affinity toward the DR1 repeats in the human ε- and γ- globin gene promoters (92, 93). Furthermore, a previous study demonstrated that TR2 and TR4 preferentially heterodimerize (49).

To determine whether bio-TR2/bio-TR4 form complexes in MEL cells we tested nuclear extracts from the double transfected cells by size fractionation using an analytical gel filtration superose 6 column. As a control we analyzed nuclear extracts from non-trasfected MEL cells under identical conditions. The elution profiles of bio-TR2 and bio-TR4 and of the endogenous proteins in nontransfected MEL cells were very similar (Figure 3). A proportion of TR2 and TR4

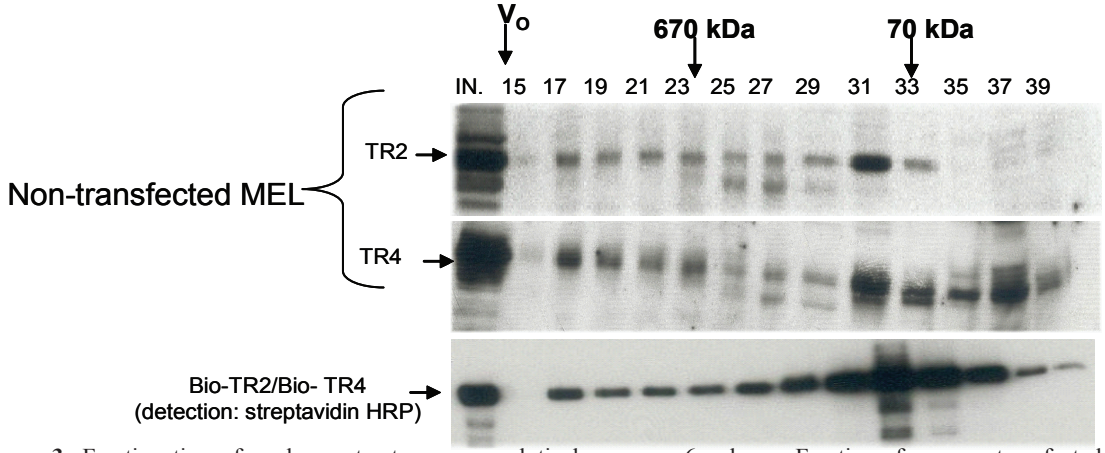


Figure 3: Fractionation of nuclear extracts on an analytical superose 6 column. Fractions from non-transfected MEL cells were analysed with anti-TR2 (top panel) and anti-TR4 (middle panel) antibodies. Fractions from double transfected bio-TR2/bio-TR4 cells were analyzed with streptavidin HRP (bottom panel). The void volume (V_0) and the elution of two molecular weight markers are indicated by arrows.

protein elutes in high molecular weight (HMW) fraction, however the majority of TR2 and TR4 proteins are present in low molecular weight (LMW) fractions. This indicates that a minority of TR2 and TR4 form large protein complexes under the extraction conditions employed. Indeed, SDS-PAGE analysis of bio-TR2/bio-TR4 nuclear extracts bound to streptavidin beads demonstrated the presence of large amounts of TR2 and TR4, with potential interacting protein present in sub-stoichiometric amounts (Figure 4). Mass spectrometry analysis of the gel slices resulted in the identification of predominantly TR2/TR4 peptides with little information on the identity of other co-eluted proteins. To enrich for the isolation of TR2/TR4 HMW complexes, we tested ammonium sulphate precipitation, sucrose gradient centrifugation and superose 6 gel filtration. Of these three

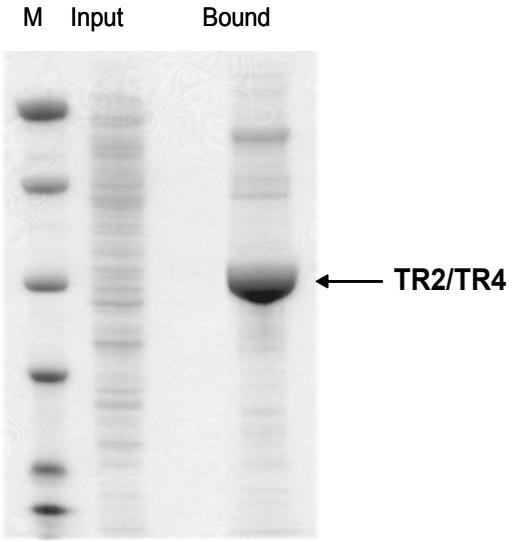
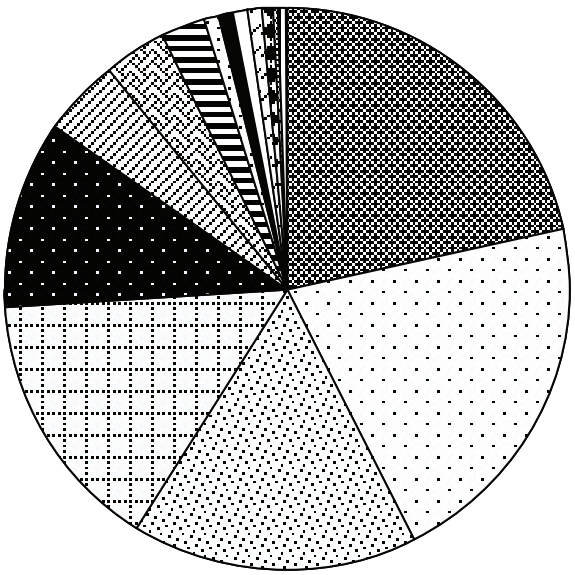


Figure 4: Coomasie blue staining of an SDS-PAGE gel after streptavidin pull-down of nuclear extract from bio-TR2/bio-TR4 MEL cells. The arrow indicates a strongly staining band where TR2 and TR4 migrate. TR2 and TR4 have similar molecular weight (~67kDa) and run as one band on the gel.

methods, only size fractionation on superose 6 columns yielded reproducible enrichment of HMW TR2/TR4 complexes. We therefore used a preparative grade superose 6 gel filtration column with a matrix bed volume of 550 ml allowing us to load over 5 ml of nuclear extract totaling more than 40mg of protein per run. We pooled 20 fractions, spanning a size range of ~200-2000 kDa, for streptavidin pull-down (74). Before mass spectrometry analysis, streptavidin bound proteins were either first separated on SDS-PAGE followed by trypsin digestion of gel slices, or were digested directly on the streptavidin beads. Both methods gave very similar results. We performed a total of five mass spec analyses of different nuclear extract preparations from induced MEL cells expressing biotin tagged TR2 and TR4, yielding similar results. Proteins potentially matching the peptides were identified using the NCBI database. Proteins that also appeared in streptavidin pull-down experiments of BirA-only MEL nuclear extracts were rejected (13, 74). The list of remaining proteins was manually curated and grouped according to Gene Ontology criteria (Biological Process, Molecular Function, Cellular Component) using the Mouse Genome Informatics database (Table 2 and Table 3), and plotted in a pie chart to provide a general overview (Figure 5).

Analysis of TR2/TR4 interacting proteins



	Peptides	%age
orphan nuclear receptor	100	21,70%
cell cycle	96	20,80%
heat shock	76	16,50%
ubiquitin-proteasome pathway	68	14,75%
Dnmt1	50	10,80%
co-repressor	21	4,55%
NuRD	17	3,70%
transcription factor	11	2,40%
signalosome complex	4	0,90%
signal transduction	4	0,90%
ISWI	4	0,90%
Pc-G	3	0,65%
DNA replication	3	0,65%
SWI/SNF	2	0,40%
FACT	2	0,40%
Total	461	100,00%

Figure 5: Pie chart representing the most abundant proteins identified by mass spectrometry. Proteins have been grouped according to the Gene Ontology, the number of peptides and %age of total peptides is shown below to the pie chart.

The number of peptides identified per protein is an indication of its relative abundance in the sample. TR2 and TR4 are the most abundant proteins identified, followed by proteins involved in cell cycle, heat shock and the ubiquitin-proteasome pathway. In addition, we identified the DNA methyltransferase Dnmt1, and members of the NuRD complex, and the TIF1β and TIF1γ co-repressors that are known to interact with nuclear receptors (45).

In order to validate these results, we analyzed by western blotting whether these proteins were pulled down with TR2 and TR4 (Figure 6). This confirmed that bio-TR2/bio-TR4 specifically pulled down Dnmt1, TIF1β and the NuRD components RbAp46/48, Mta1 and Mi2β. We also confirmed interactions with HDAC3, a protein previously reported to interact directly with TR2 (25, 57).

To confirm these interactions with an independent method, we performed co-immunoprecipitations (co-IPs) using antibodies against Smc3, Hsp 90, TIF1β, TIF1γ, HDAC3, Mi2β, HDAC1, RbAp46/48, with mouse- and rabbit IgGs as negative controls. We used nuclear extracts from double (Figure 7A) and single (Figure 7B and 7C) transfectants expressing biotin-TR2 and/or biotin-TR4. Immunoprecipitates were assayed by Western blotting using streptavidin HRP which detects biotin-tagged TR2 and TR4. This confirmed that Dnmt1, TIF1β, HDAC-3, Mi2β and Hsp90 interact with bio-TR2/bio-TR4. Smc3 antibodies did not co-IP bio-TR2/bio-TR4. In addition,

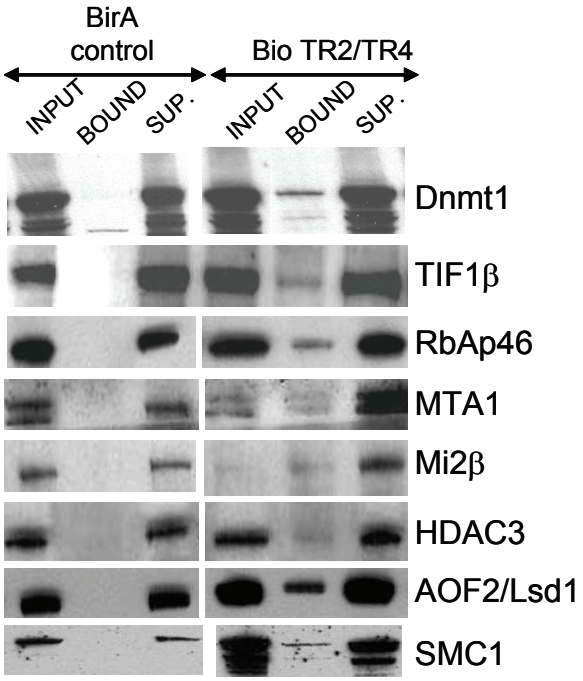


Figure 6: Confirmation of TR2/TR4 co-purified proteins by binding nuclear extracts from bio-TR2/bio-TR4 and BirA-only control MEL cells to streptavidin beads followed by Western immunoblot analysis using antibodies specific for TR2/TR4 co-purified proteins. The antibodies used are listed on the right. The BirA-only control shows that none of the proteins analyzed bound as background to the streptavidin beads. Loading: Input (IN), bound to streptavidin beads (B.) and supernatant (SUP).

TIF1γ, HDAC-1, Mi2β and Dnmt1 antibodies co-immunoprecipitated bio-TR2. We failed to confirm interaction between bio-TR2 and RbAp46/48, despite the fact that we could detect this protein in the double transfectant pull-downs and in co-IP.

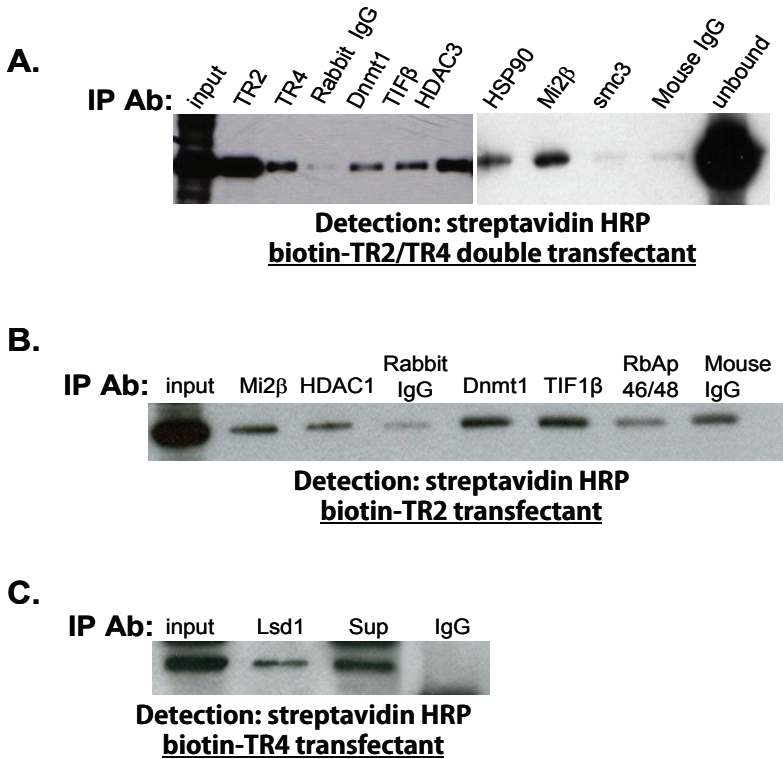


Figure 7: A) Co-IP experiment using TR2, TR4, Dnmt1, TIF1β, Hsp90, Mi2β, smc3, Hdac3 and control mouse and rabbit IgG antibodies. Nuclear extract from bio-TR2/bio-TR4 double transfectant MEL cells was used for co-IP and TR2 and TR4 are detected with streptavidin-HRP. B) Co-IP using anti Mi2β, HDAC1, Dnmt1, TIF1β, RbAp46/48, and control rabbit and mouse IgG antibodies on nuclear extract from bio-TR2 single transfectant MEL cells. Bio-tagged TR2 was detected with streptavidin HRP.

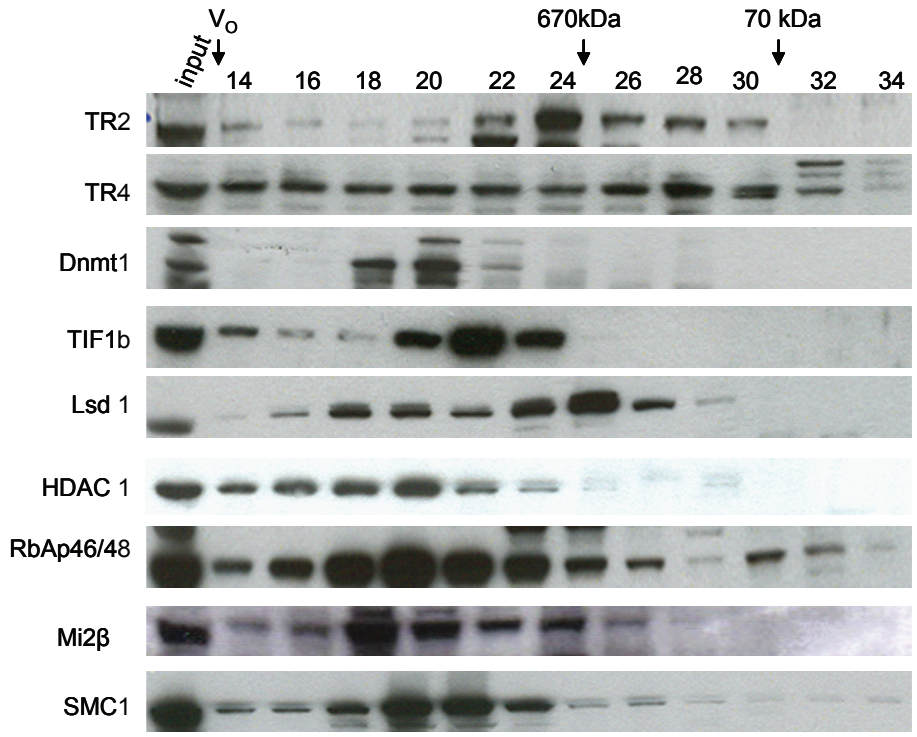


Figure 8: Nuclear extract from induced, non-transfected MEL cells was run on the analytical superose 6 gel filtration column. Western blots detecting TR2, TR4 and potential interacting partners (Dnmt1, TIF1β, HDAC1, RbAp46/48, Mi2β, SMC1), showing the fractionation profiles of these proteins. The elution of molecular weight markers and the void volume are indicated by arrows on the top.

In summary, our results indicate that bio-TR2/bio-TR4 interacts with a number of transcriptional co-repressors. To assess whether these interactions occurred in distinct or overlapping complexes, we compared the superose 6 size fractionation profiles of the potential interacting proteins. Overlapping elution profiles would indicate that the proteins are components of the same complexes (Figure 8). We find that most proteins fractionate with peaks around fractions 18 to 24, including Dnmt1, TIF1β, HDAC1, Mi2β, RbAp46/48 and SMC1. Interpretation of these results with regards to possible interactions in distinct or overlapping complexes is difficult due to the overlapping fractionation profiles of most proteins tested. One other complicating factor is that the peaks of the TR2/TR4 fractionation profiles do not clearly overlap with those of the other proteins tested.

Since these co-elution results were inconclusive, we performed co-IPs with antibodies against Dnmt1, TIF1β and HDAC1 (Figure 9). We find that the Dnmt1 antibody co-immunoprecipitates TIF1β (Fig. 9A). The interactions of Dnmt1 with TIF1β are also confirmed by co-IPs using TIF1β antibodies as the primary antibody (Fig. 9B and C). In addition, we find that Dnmt1 interacts with HDAC3 (Fig. 9A). Interestingly, HDAC1 antibodies could not immunoprecipitate detectable amounts of TIF1β (Fig. 9). These results are summarized in Figure 10.

TR2/TR4 interact with Dnmt1

As we have shown above (Figure 5 and 6 and Table 2), we could isolate Dnmt1 in our mass spectrometry results as a potentially TR2/TR4-interacting protein. As unrelated proteins may be bound to the same piece of DNA and therefore appear to be part of the same complex, we tested whether the interaction of TR2/TR4 with Dnmt1 is due to nucleic acid contamination. To eliminate nucleic acids we treated nuclear extracts with benzonase, an enzyme that hydrolyses both DNA and RNA. As shown in Figure 11, Dnmt1 was co-purified with TR2/TR4 in both benzonase- and mock- treated samples. We conclude that the interaction of TR2/TR4 and Dnmt1 is not dependent on the presence of DNA.

Affinity tags may affect the yield, solubility and even the folding of their fusion partners (101). This may create non-physiological interactions with other proteins. To test whether the TR2/TR4/Dnmt1 interaction is present in non-transfected MEL cells we performed immune

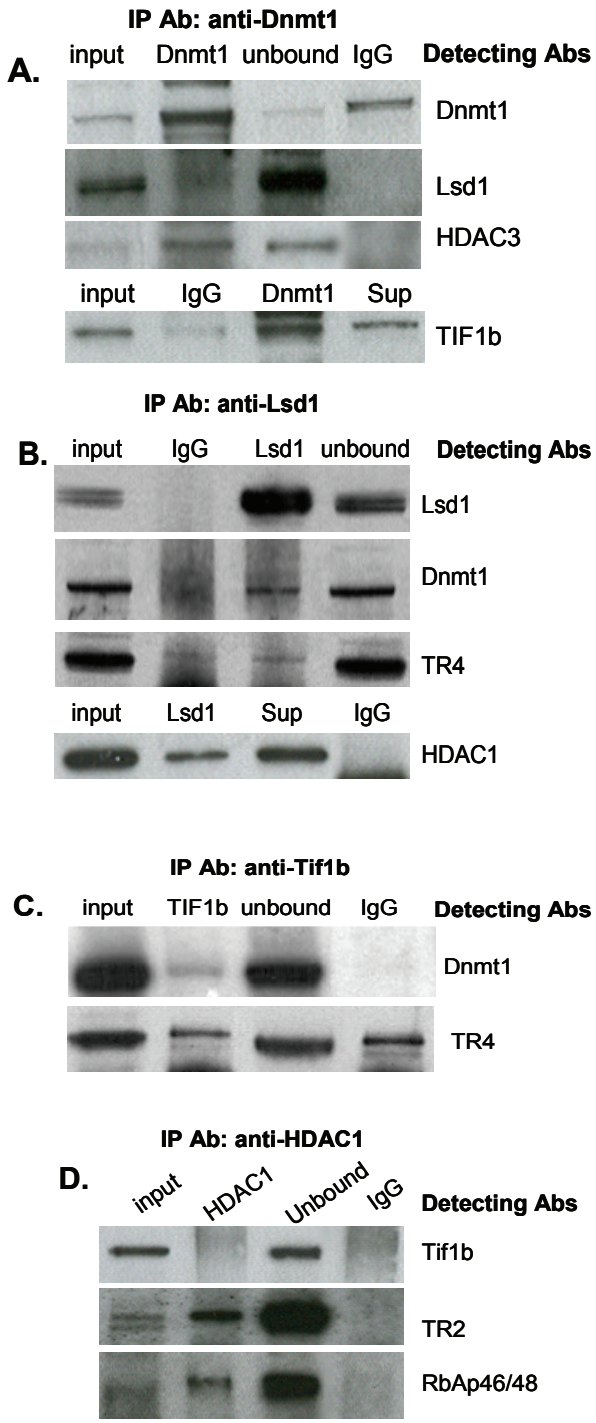


Figure 9: Co-IPs using antibodies against potential TR2/TR4 interacting proteins and nuclear extracts from induced non-transfected MEL cells. The order of loading of input supernatant (unbound) and bound (immunoprecipitated) fractions is indicated. Input and supernatant loading represents approx. 1% of starting material, while bound fraction loading represents approx 10-15 % of starting protein A) Co-IP experiment using Dnmt1 antibody, using mouse IgG as a control. C) IP using anti HDAC1 antibodies and rabbit IgG control.

precipitation experiments using TR2 and TR4 antibodies. Although signals are not very strong, we could clearly demonstrate that Dnmt1 precipitated with TR2 and TR4 antibodies (Fig. 11B).

Developmental interaction of TR2/TR4 complexes to the globin promoters in erythroid cells

Previous work showed that TR2 and TR4 are the core components of the DRED complex that binds to the DR1 sites in the γ- and ε- globin promoters (92). Such DR1 sites are also present in the mouse embryonic εy and βh1 globin promoters (24, 95). In order to test the functional relevance of the TR2/TR4/Dnmt1 interaction in globin gene regulation we performed Chromatin Immunoprecipitation (ChIP) experiments using streptavidin beads and specific antibodies. Firstly, we tested whether we could ChIP TR2/TR4 bound to the ε- and γ-globin promoters in non-induced MEL cells and 14.5 dpc fetal livers of transgenic mice carrying a human β-globin locus. Fetal

	Immunoprecipitating antibodies			
Immunoprecipitated proteins	***α-Dnmt1***	***α-Lsd1***	***α-Tif1β***	***α-HDAC1***
Dnmt1		**yes**	**yes**	n.t.*
Lsd1	**yes**		n.t.	n.t.*
Tif1β	**yes**	n.t.		**no**
HDAC1	n.t.*	**yes**	n.t.	
HDAC3	**yes**	n.t.	n.t.	n.t.*
RbAp46/48	n.t.	n.t.	n.t.	**yes**
TR2/TR4	**yes**	**yes**	**yes**	**yes**

n.t.: not tested
* : previously reported in literature

Figure 10: Potential cross-interactions of the TR2/TR4 proteins and interacting partners

livers at this stage are developmentally similar to induced MEL cells, and represent the adult stage of erythropoiesis (Fig 12) However, in none of these experiments did we were able detect binding of TR2 or TR4 to the εy- and βh1 promoters or to the human ε- or γ-globin promoters. Instead, much to our surprise, initially biotin tagged TR2/TR4 double transfectant (data not shown) and

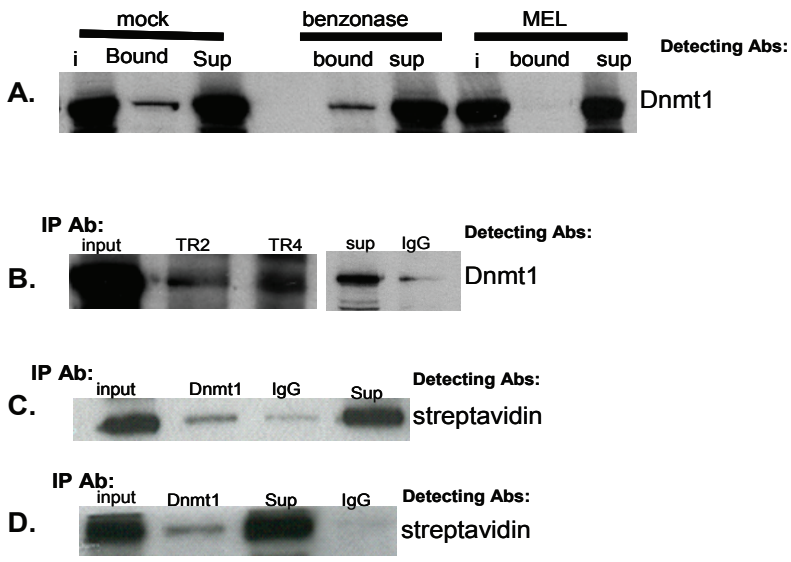


Figure 11: Confirmation of TR2/TR4 and Dnmt1 interactions by immunoprecipitation A) Nuclear extracts of bio-TR2/bio-TR4 were either mock or benzonase treated and substequently bound to the streptavidin beads. Nuclear extracts from BirA–only trasfected cells were used as a negative control. The western blot was detected with a Dnmt1 antibody. B) IP experiment using induced non-transfected MEL cell nuclear extract. Antibodies against TR2 and TR4 could specifically immunoprecipitate Dnmt1. C) Dnmt1 can specifically immunoprecipitate bio-TR2, as detected by streptavidin HRP D) reverse IP using Dnmt1 antibody and relative IgG negative control and nuclear extracts from single transfectant bio-TR4 MEL cells. Dnmt1 can specifically immunoprecipitate bio-TR2, as detected by streptavidin HRP.

later on single biotin-tagged TR4 and anti-TR4 antibodies showed a marked enrichment for mouse βmajor (β_{maj}) promoter (Fig.12). In addition, using a Dnmt1 antibody previously reported to work in ChIP assays (54), we were unable to detect any promoter-specific enrichment of Dnmt1 in induced MEL cells, even after using double crosslinking (first protein-protein interactions with DSG subsequently protein-DNA with formaldehyde) (Fig. 12F and data not shown). It is possible that the anti-Dnmt1 antibody did not work efficiently enough, and this problem may be overcome in future by using chromatin from biotin-tagged Dnmt1 MEL cells (K.Kolodziej, unpublished). This could be also due to the fact that the Dnmt1 antibody binds to the chromatin indiscriminately. We also tested chromatin prepared from 13.5 dpc fetal livers. We found enrichment of the β_{maj} promoter only with anti-TR4 but not with anti-TR2 or anti-Dnmt1 antibody (Fig. 12E). The only potentially TR4-interacting factors that showed similar tendency in ChIP experiments (i.e. bound β_{maj} globin) was RbAp46/48 (Fig 12G). In summary our ChIP results suggest that only TR4

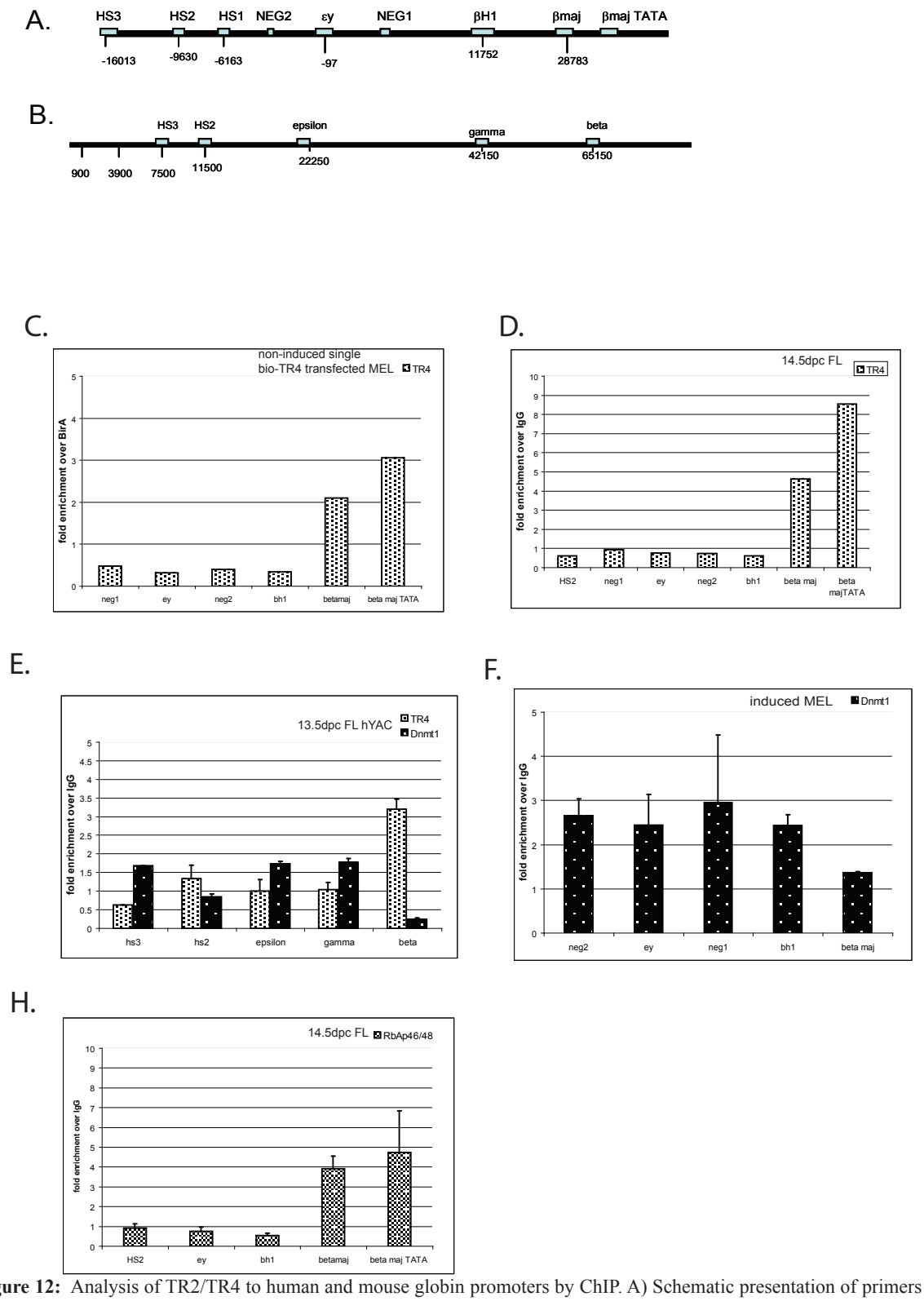


Figure 12: Analysis of TR2/TR4 to human and mouse globin promoters by ChIP. A) Schematic presentation of primers used for mouse globin locus ChIP. B) Schematic presentation of primers used in human β- globin locus ChIP. C) ChIP using non-induced MEL cells chromatin from single bio-TR4 transfectant. Fold enrichment calculated over the BirA negative control experiment (by division) that was done in parallel. D) ChIP using chromatin from 14.5 dpc fetal liver cells. E) anti-TR4 and anti-Dnmt1 antibodies ChIP using chromatin from 13.5 dpc fetal livers of the transgenic mice carrying human β-globin locus YAC and. F) anti-Dnmt1 antibody ChIP using chromatin from induce MEL cells G) anti-RbAp 46/48 antibodies ChIP using the chromatin from 14.5 dpc fetal livers.

homodimers bind to the active β_{maj} globin promoter *in vivo*.

TR4 binds to the DR4 element in β_{maj} promoter

The β_{maj} ChIP primers overlap with a potential DR4 site which is 28bp downstream of the CAAT box that was tested in band shift experiments in (92) (Fig. 13 A). Considering that TR2 and TR4 bind to DR4 response elements (9, 50), we propose that the β_{maj} promoter contains a potential DR4 response element that has not been reported before. To further investigate whether this potential DR4 element can be bound by TR2 and TR4 we performed band shift assays with nuclear extracts from MEL cells (Fig. 13 B). Consistent with previous data, we find the DRED complex binding to DR1 in the ε-globin promoter (Fig. 5, lane 2). We used competitor oligonucleotides that were published previously (92): EPSI which strongly competed for DRED binding (Fig. 13B, lanes 2 and 3) and γCAAT which was less efficient in competing the DRED complex. As a negative control we used βCAAT oligonucleotides, which do not compete for DRED binding (Fig. 13B, lanes 7 and 8). In comparison the oligonucleotide corresponding to the potential DR4 response element competed for DRED complex binding almost as efficiently as the γCAAT oligonucleotide. We conclude that TR2/TR4 may bind to the potential DR4 response element in the β_{maj} globin gene promoter *in vivo*.

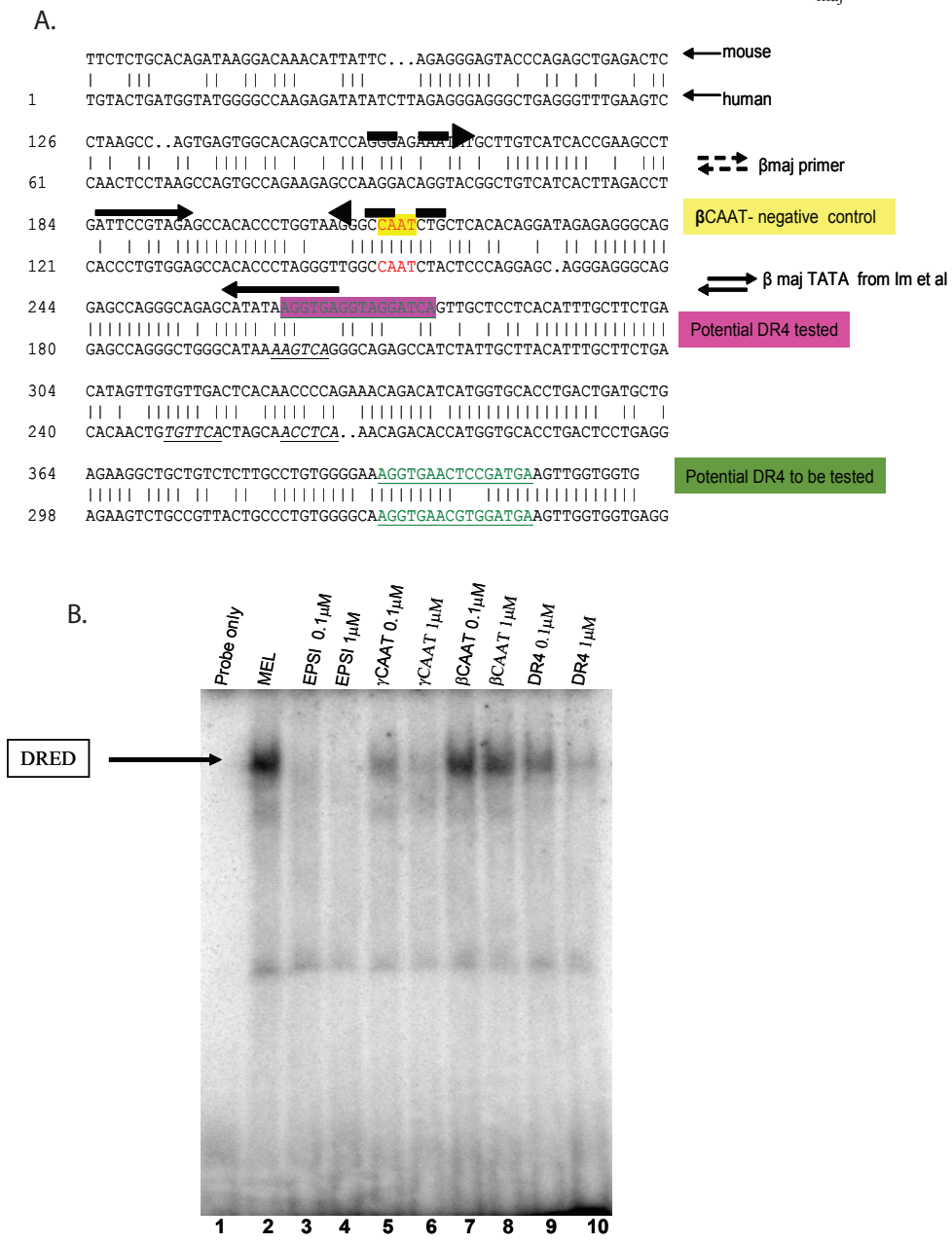


Figure 13: A) Alignment of sequences upstream of the human (lower sequence) and mouse (upper sequence) β-globin gene promoters. The CAAT box previously tested and used as a negative control for band shift assay (92) is marked in yellow (βCAAT negative control). Broken line arrows mark the βmaj primer set used in ChIP. Solid line arrows mark βmaj TATA primer set used in ChIP. The potential DR4 element that was tested by EMSA is highlighted in pink. Another potential DR4 element that is shared by both species and which should be further tested, is marked by underlined green font.
B). Band shift assays using MEL nuclear extract. EPSI probe (from (92)) was labeled with P^{32} according to the protocol in (92). Cold probes (EPSI, γ CAAT, potential β-globin DR4 element, and βCAAT negative control) were used as competitors in two concentrations: 0.1μM and 1 μM. In lane 1 only labelled probe was used. DRED complex formation is indicated by an arrow.

The alignment of mouse and human beta globin gene promoters (Figure 13A) revealed that there is an additional potential DR4 response element further downstream from the CCAAT box in both mouse and human. Whether it can be bound by TR2 and TR4 remains to be tested. It is possible that the enrichments we observe by ChIP using chromatin from mouse cells are the result of TR2/TR4 binding to both DR4 elements, and in the case of human cells the enrichment in ChIP is the result of binding of TR2/TR4 to only one DR4 element.

The observation that TR2/TR4 binding to the β-like globin promoters is a very intriguing

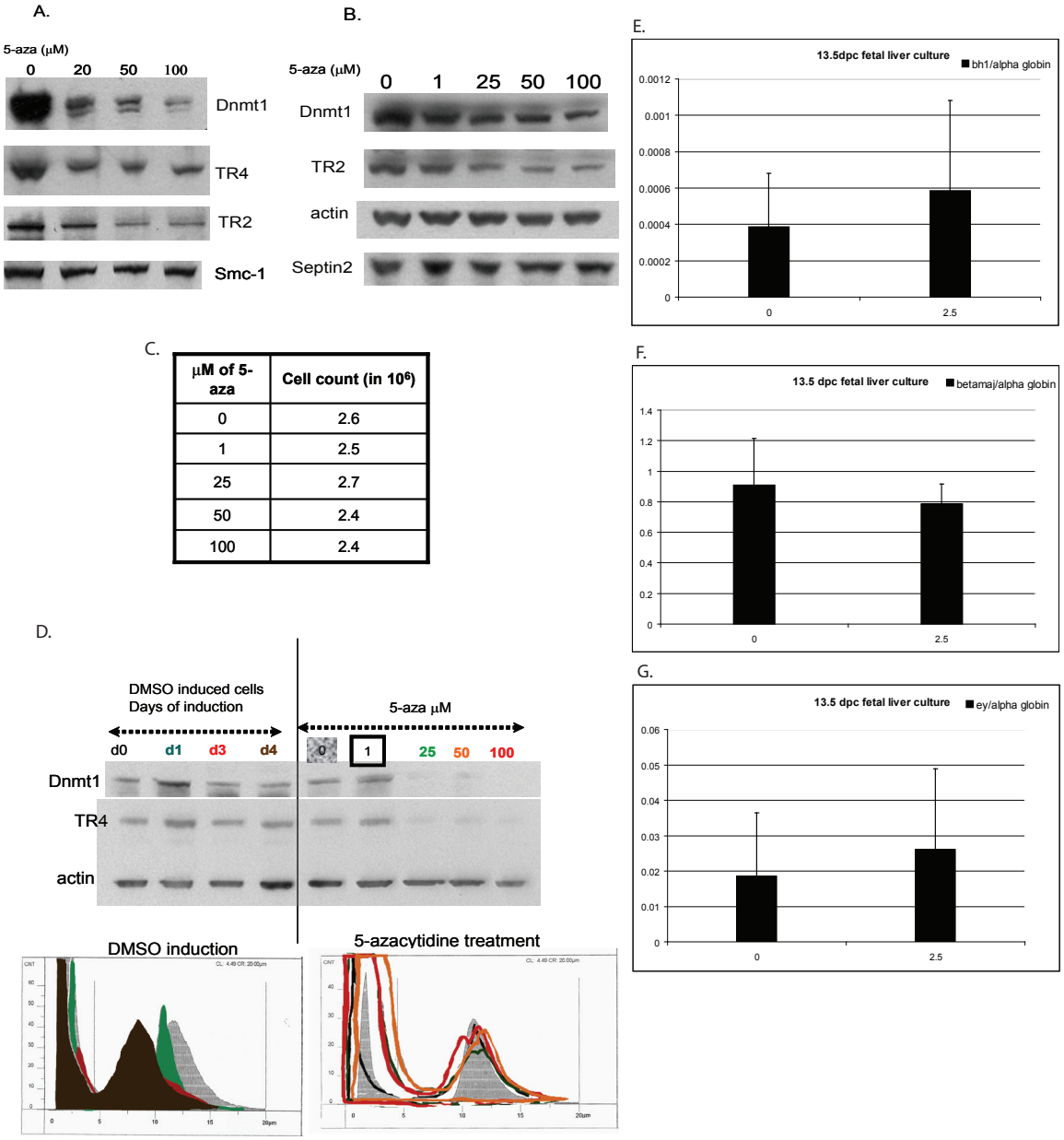


Figure 14: 5-azacytidine treatment causes specific protein degradation of Dnmt1 and TR2/TR4 in fetal liver and MEL cells. Cells were treated for 12-15 hours with different concentrations of 5-azaC (as indicated above the panels).). Whole cell extracts were tested by Western immunoblot analysis using antibodies against Dnmt1, TR2, TR4 and control antibodies against Smc1, actin and septin2. A) Primary cultured 14.5 fetal liver cells B) MEL cells C) Table showing the results of the cell count after 5-azaC treatment. Non-treated cells were used as a control. D) MEL cells were either induced with 2% DMSO or treated with increasing amount of 5-azaC. The cell number and size was monitored using CASY cell counter. The induction with DMSO has no apparent effect on the levels of Dnmt1 or TR2/TR4 proteins, whereas 5-azaC induces their specific degradation. E-G: RNA levels of globin genes upom treatment with 5-azacytidine.13.5 dpc fetal livers were collected for primary culture and treated with 2.5 μM 5-azacytidine for 48hours, when RNA was collected. 13.5 dpc fetal livers were collected for primary culture and treated with 2.5 μM 5-azacytidine for 48hours, before RNA was extracted. The εy (panel E), βmaj (panel F) and βh1 Panel G) globin expression levels were determined by quantitative RT-PCR and normalized to α globin expression levels.

one and the function(s) of TR2/TR4 in regulating β-globin remains to be discovered.

5-azacytidine induces degradation of Dntm1 and TR2/TR4

5-azacytidine is a potent DNA methyltransferase inhibitor that was previously tested in tissue culture but also in baboons and human patients initially as an anticancer drug (90). Recent studies showed that one of the possible effects of 5-azacytidine in inhibiting DNA methyltransferases is by specific Dnmt1 protein degradation through the proteasome pathway (28). In order to test whether 5-azacytidine would have any effect on the TR2/TR4/Dnmt1 complex stability in erythroid cells we treated cultured 14.5 dpc fetal liver cells (Fig. 14A) and MEL cells (Fig. 9B) with 5-azacytidine. Increasing concentrations of 5-azacytidine resulted in decreasing amount of TR2/TR4 and Dnmt1 proteins (Fig.14). This effect is specific as the levels of control proteins such as Smc-1, Septin or Actin did not change upon 5-azacytidine treatment (Fig 14A and B). Furthermore, the effect of 5-azacytidine on TR2/TR4 and Dnmt1 cannot be explained solely by increased cytotoxity as the numbers of MEL cells do not significantly change before and after treatment (Fig. 14C).

In addition to its inhibitory and pro-apoptotic functions, 5-azacytidine has been previously shown to induce differentiation in many cellular systems (59, 82, 100). We therefore tested if the decreasing levels of Dnmt1 and orphan receptors observed after 5-AzaC treatment of MEL cells are a physiological consequence of cellular differentiation. To this end, we divided MEL cells into two groups: one was induced to differentiate with the very potent inducer DMSO, the other group was incubated with various concentrations of 5-azacytidine. Cells were subsequently counted and lysed for proteins to be analyzed by western blotting (Fig. 14D). The results show that in cells induced with DMSO the levels of TR2/TR4 or Dnmt1 protein did not change significantly (Fig. 14 D, left panel). In contrast, again the levels of TR4 and Dnmt1 protein again decreased with increasing amounts of 5-azacytidine (Fig. 14 D, right panel). In addition, 5-azacytidine appeared to only slightly induce cellular differentiation (see small additional peak on the red and green graph cells treated with 25 and 100 μM 5-azacytidine (Figure 14 D, right panel)).

In order to investigate whether in the conditions we used 5-azacytidine could induce the embryonic and fetal globin expression, we cultured 13.5 dpc fetal liver cells in the presence of 2.5 μM 5-azacytidine for 72 hours. A slight increase in the expression of the εy- and βh1- mouse embryonic globins was detected (Fig. 14).

Discussion

The TR2/TR4 orphan nuclear receptors have been identified as components of DRED complex involved in developmental silencing of embryonic and fetal globin genes (70, 92, 93). In order to identify additional components of the DRED complex we employed biotinylation tagging of TR2/TR4 in MEL cells, by size fractionating nuclear extracts followed by streptavidin binding and mass spectrometry. Apart from TR2 and TR4 the most abundant proteins identified in this way were proteins involved in the cell cycle, heat shock and the ubiquitin-proteasome pathway followed by chromatin associated proteins involved in transcriptional repression (Tables 2 and 3). Although we did confirmed the interaction of TR2/TR4 with Hsp90 (Fig. 7) we did not pursue any further our observations on the co-purification of cell-cycle, heat shock and ubiquitin proteasome proteins with TR2/TR4. Interactions of nuclear receptors with heat shock proteins acting as chaperones are well known and have been implicated in the refolding of nuclear receptors and their nuclear translocation (67). There is also increasing evidence that components of the ubiquitin-proteasome pathway interact with nuclear receptors in modulating their transcriptional activities (36, 42). The co-purification of several cell cycle associated proteins with TR2/TR4 may be related to a wealth of evidence implicating nuclear receptors in cell cycle regulation in erythropoiesis (for example (4)). Significantly, overexpression of TR4 in myeloid cell lines led to increased cell proliferation (43), though we did not observe any overt effects on MEL cell differentiation upon overexpression

of biotin tagged TR4 and/or TR2. We do not presently know whether these TR2/TR4 protein interaction/complexes are distinct or overlap with other complexes that were identified in our studies. It is possible that they are specifically enriched in the size fractionation protocol that we used. In that respect, it would be interesting to determine the fractionation profiles for several representative members of cell cycle, heat shock and ubiquitin-proteasome proteins and compare them to the TR2/TR4 profiles. Interactions of nuclear receptors with heat shock proteins acting as chaperones are well known and have been implicated in the refolding of nuclear receptors and their nuclear translocation (67). There is also increasing evidence that components of the ubiquitin-proteasome pathway interact with nuclear receptors in modulating their transcriptional activities (36, 42). The co-purification of several cell cycle associated proteins with TR2/TR4 may be related to a wealth of evidence implicating nuclear receptors in cell cycle regulation in erythropoiesis (for example (4)). Significantly, overexpression of TR4 in myeloid cell lines led to increased cell proliferation (43), though we did not observe any overt effects on MEL cell differentiation upon overexpression of biotin tagged TR4 and/or TR2.

Interestingly, we also identified and confirmed interactions of TR2/TR4 with the Dnmt1 DNA methyltransferase, the Lsd1 histone demethylase, the Tif1β and γ co-repressors and members of the NuRD complex (mainly HDAC1, M2β, RbAp46/48). Most of these interactions have also been confirmed by reverse IPs (7 and 9). Due to the involvement of TR2/TR4 in the DRED complex and in globin gene repression, we focused on this group of chromatin associated repressive proteins. Most importantly, we show for the first time that Lsd1 and Dnmt1 interact together (Fig. 9). We do not presently know whether these interactions occur in addition or exclusively in the context of their interactions with TR2/TR4. We also find that TR2/TR4 interact with HDAC3, as has been previously reported for TR2 (25).

Chromatin remodeling protein Mi2β interacts together with HDAC1 and RbAp46/48 in the NuRD complex that contains additionally Mbd3, Mta1, Mta2 and p66 proteins (105) and, as shown in Figure 4, they share overlapping fractionation profiles. Thus, NuRD combines in the same complex histone deacetylation and chromatin remodeling ATPase activities to methyl DNA binding. Importantly, members of the NurD complex have been reported to participate in the PYR complex which binds to the intergenic region between the human δ and γ globin genes (69) whereas the MeCP1 complex has been reported to bind to HS2 of the human β-globin LCR (62). Lsd1 is a lysine specific histone demethylase (87) that is associated with HDACs 1 and 2 in the context of the coREST co-repressor complex (44, 88) (but not with the NuRD complex). Consistent with this, it has been previously shown that the preferred Lsd1 substrates are hypoacetylated histones and treatment with the HDAC inhibitor trichostatin (TSA) causes Lsd1 target gene de-repression (87, 88). Lsd1 has not been previously reported to interact with Dnmt1 or with Tif1β. Intriguingly, interactions of Lsd1 with nuclear receptors have been previously reported to result in gene activation (71) (also see Discussion Chapter).

TIF1 (transcriptional intermediary factor 1) β and γ are members of a family of transcriptional co-repressors interacting with nuclear factors (8, 48, 99). In addition, TIF1β is a universal co-repressor for KRAB-domain zinc-finger proteins (26, 41, 66). Importantly, TIF1β has been reported to co-purify with the N-CoR1 complex which contains HDAC3, but not HDAC1 or 2 (44). It has also been reported to interact with Mi2α, HDAC1 and RbAp48 (85, 97) which are members of the NuRD complex. TIF1β has also been shown to interact directly with histone methyltrasnferase SETB1 (84) and HP1 (45, 46, 68, 77). Our results clearly show that TIF1β can interact with TR2/TR4, but it can also be immunoprecipitated with Dnmt1 antibodies and that this interaction is reciprocal (see Figure 9A and C). However, we have also clearly shown that there is no detectable interaction between Tif1β and HDAC1 in MEL cells (Fig.9D). This suggests that interactions of TR2/TR4 (and of Dnmt1) with Tif1β may be taking place in a complex that is distinct to TR2/TR4 complexes containing HDAC1. It is not clear from our data whether Tif1β

interactions with TR2/TR4 also involve HDAC3, however, Tif1β has been reported to interact with HDAC3 in the context of the NCoR1 complex (44) and we show here that Dnmt1 can also immunoprecipitate HDAC3 (Fig. 9A). It is also not clear from our data whether Lsd1 interacts with Tif1β, however, Lsd1 has not been previously reported to interact with any of the reported Tif1β interacting proteins (except HDAC1). By putting our data together and by extrapolating from previous results from other labs, we suggest that TR2/TR4 possibly form a distinct complex with Tif1β, Dnmt1 and HDAC3 (complex 1, Figure 15). In order to provide evidence for a role for Tif1β together with TR2/TR4 in the regulation of globin genes, we employed ChIP assays but failed to detect any binding of Tif1β to the mouse β-globin locus in 14.5 dpc fetal livers (data not shown). TIF1α and TIF1β have been previously shown to interact with the LBD of nuclear receptors (NR), including estrogen receptor, thyroid hormone receptor, vitamin D receptor and retinoic acid receptor (45, 47, 48). The TIF1 family of proteins can interact with NRs in the absence of activating ligands, or compete with the activating ligands for binding to the NRs. TIF1 bound to NRs can result in a change in the conformation of the NR protein and as result block DNA binding by the nuclear receptor causing transcriptional repression (e.g. by stalling the RNA polymerase or by outcompeting the transcriptional co-activators) (89). Thus, a possible explanation for the fact that we could not detect by conventional ChIP Tif1β bound to globin sequences, would be that when in complex it would render TR2/TR4 incapable of binding to DNA. An erythroid specific TIF1β knock-out (currentlu under way in the Engel lab) would help bring further insights into the functional relevance of Tif1β and TR2/TR4 interaction in globin gene regulation.

The Lsd1 immunoprecipitations reveals interactions with Dnmt1, HDAC1 and TR2/TR4 (Fig. 6 and 9B). Whereas it is well documented that Lsd1 and Dnmt1 can interact with HDAC1, it is not clear from our data, or from previous work, whether they interact with the Mi2β and RbAp46/48 members of the NuRD complex. Nevertheless, for this data we would like to suggest that these proteins interact to form a second, possibly distinct, TR2/TR4 complex in MEL cells (complex 2, Figure 15).

By means of IPs we were able to demonstrate that HDAC1 can bring down TR2 (Fig. 7B and 9B). HDAC-1 is part of NuRD complex and therefore in this context co-exists together with RbAp 46/48 (14, 105), and has also been previously shown- interacts with Lsd1 (88). Histone deacteylation is an important repressive chromatin mark and has also been shown to be required for histone methylation (88). Histone acetylation occurs during chromatin remodeling which is also connected to the cell cycle progression (32). It has been previously suggested that specific histone H4 acetylation may recruit basal transcription factor TFIID to activate the β major promoter (1). Also acetylation of histones H3 and H4 play an important role in the stage specific regulation of α and γ globin gene expression (21). The histone deacetylase inhibitors sodium butyrate and TSA have been shown to activate globin genes (23, 35, 103) though the exact function of butyrate on the chromatin structure of globin genes is not yet well understood (21). In summary, there has long been evidence for histone acetylation playing a major role in epigenetic regulation of γ globin silencing. It is plausible to hypothesise that DRED, as a repressive, stage specific complex could contain or recruit histone deacetylase activity to ε and γ globin gene promoters.

Previous work by the Engel laboratory showed by band shift assays TR2 and TR4 to be the core components of the DRED complex that binds to the DR1 sites in the γ- and ε- globin promoters (92). In order to test the functional relevance of TR2/TR4/Dnmt1 interaction we performed Chromatin Immunoprecipitation (ChIP) experiments using streptavidin pull-downs and specific antibodies. However, we failed to detect the binding of TR2 or TR4 to the mouse εy- and βh1 promoters in MEL cells or to the human ε- and γ-globin in 14.5 fetal liver cells in transgenic mice (Fig. 12). This is in contrast to what we would have predicted by the previous data on DRED binding to the embryonic and fetal globin promoters. It is possible that the epigenetic marks on the chromatin of the embryonic and fetal ε- and γ- globin gene promoters are already set in a manner

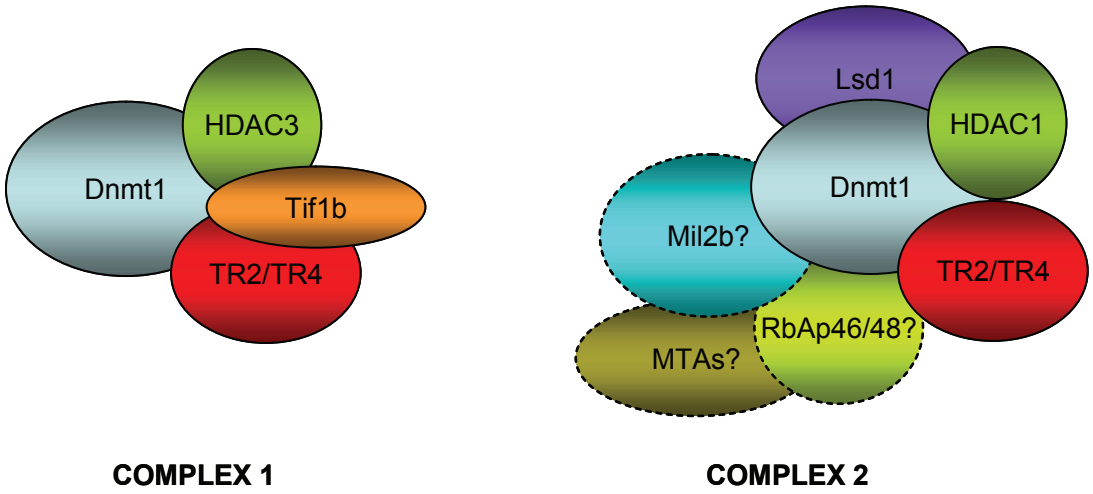


Figure 15: Working model presenting two possible distinct complexes that interact with TR2/TR4. We hypothesize that Complex 1 consists of TR2/TR4, Dnmt1, HDAC3 and TIF1β, whereas Complex 2 consists of TR2/TR4, HDAC1, Dnmt1, Lsd1 and possibly NuRD components RbAp46/48, Mi2β, MTA1. Further experiments should be conducted to test these models.

that does not require continued DRED binding anymore. We also failed to detect any enrichment at the globin promoters using anti Dnmt1 antibody. This may be a technical problem which can be overcome by using bio-Dnmt1 MEL cells chromatin for streptavidin ChiP. It is also possible that Dnmt1, which confers maintenance methylation function is spread across the silenced ε and γ globin genes making it difficult to show local enrichments at particular sites of these promoters.

Instead, to our surprise, the anti-TR2 (data not shown) and anti-TR4 ChIP showed consistent and specific enrichment of the adult mouse β major in MEL cells or the human β-globin promoters in transgenic mice. Considering that TR2 and TR4 have been reported before to bind to DR4 response elements (9, 50) we propose that it is plausible that β major and β-globin contain a functional DR4 response element that has been never reported before. Moreover the alignment of the mouse and human β-globin gene promoters (Fig. 13A) revealed that there is an additional potential DR4 response element even further downstream of the CAAT box and which is conserved between the two species. Whether it can be bound by TR2 and TR4 remains to be tested.

It should be noted though that the binding affinity of nuclear receptors for DR4 is lower than that for DR1 response elements: DR1>DR2: DR5, DR4, DR6> DR3 (58). In order to investigate whether the enrichment of the β major with anti-TR4 antibodies is genuine or an artefact caused by conformation or "stickiness" of chromatin of this particular promoter region (enrichment of IgGs negative control is usually higher with those primers as well data not shown), we performed band shift assays as previously documented for the binding of the DRED complex to the ε- and γ- globin promoters (92). On the basis of these experiments we conclude that TR2/TR4 can bind to the potential DR4 response element proximal to the TATA box in β major globin gene promoter. During chromatin preparation DNA is randomly sheared in 300-800 bp fragments, while for band shift assays short 20-30 bp oligos are used. This means that in ChIP experiments we can test binding to a much longer DNA string. Also, orphan nuclear receptors allow certain variations in their consensus binding sites thus escaping standardized bioinformatics tools searching for transcription factor binding motifs in promoters. Hence, this potential DR4 site was never spotted before. It has to be stressed that the function of binding of TR2/TR4 the β major globin promoter remains to be discovered.

Dnmt1 plays an important role during cell cycle (16, 53, 76) when it is targeted to the replication foci maintaining the methylation pattern following the DNA replication (6, 27, 34, 38). Dnmt1 has been shown to interact with the histone methylatransferases G9a and SUV39H1during the cell cycle to ensure the re-establishing of the epigenetic marks after cell division (19). We do not know whether interaction between Lsd1 and Dnmt1 and possibly TR2/TR4 are also cell-cycle

related. We can not exclude the possibility that this interaction has a distinct function in cell cycle progression or, possibly, is regulated by the cell cycle. This hypothesis could account for the cell-cycle related proteins which were found in our mass spec results.

5-azacytidine (5-azaC) and its derivatives (such as 5-aza- deoxycitidine, decytabine) have been used for cancer therapy (5, 10, 37, 40, 52) and in treating hemoglobinopathies since they can induce γ-globin reactivation (82, 90, 102). Despite many efforts in the elucidation of the molecular mechanism that these potent DNA methyltraserase inhibitors employ, details regarding the exact mechanism of their action remain to be determined. 5-azaC is at first converted to a nucleoside triphosphate, which can be subsequently incorporated into DNA and RNA and consequently alter protein synthesis. Importantly, after incorporation of 5-azaC into DNA methylatransferases, Dnmts irreversibly bind to DNA and can no longer methylate it (12, 28, 39, 96, 106).

However, it has also been suggested that the formation of tight covalent complexes between Dnmts and 5-aza-C substituted DNA alone cannot explain many aspects of the action of these drugs (28). Firstly, Dnmt1 activity decreases faster than incorporation of 5-azaC to DNA (12, 28). Secondly, microarray expression profiling analysis of colon cancer (HCT116) and Dnmt1 -/- cell lines showed that the expression changes induced by 5-azaC were not dependent on cell proliferation or apoptosis (29). These results did not fit with the view that 5-azaC acts solely by incorporation into DNA during one cell division, followed by loss of methylation during the subsequent round of replication (28, 29). Further, it has been suggested that 5-azaC specifically targets Dnmt1 for proteosomal degradation (28). We tested whether we could observe a similar effect of 5-azaC in MEL cells. As the amount of 5-azaC required for inducing this process depends on the level of Dnmt1 in the cell line (28), we tested different concentrations and we could confirm that 5-azaC treatment did indeed cause a decrease in Dnmt1 protein and also in TR2 and TR4 protein in MEL cells (Fig. 14). We did not confirm however whether the degradation is specifically mediated by proteasome (by using Mg132) (our initial attempts were so far unsuccessful, further optimization is required; K.K observations) nevertheless it seems very likely that this is the case.

It may be of interest to establish whether the specific degradation of TR2/TR4 and Dnmt1 upon 5-azaC treatment is simultaneous, or whether the degradation of orphan receptors is a signal for Dnmt1 degradation (or the other way around). The understanding of these processes would significantly enhance the understanding of epigenetic regulation of globin promoters by DNA methylation particularly in the light of recent reports showing that the Dnmt1 knock-down by RNAi which results in a decrease of γ-globin promoter methylation (to similar levels seen upon 5-azaC treatement) was not a primary cause of induction of γ –globin expression (60). This suggests that additional effects of 5-azaC, next to the inhibition of DNA methylation, are responsible for the activation of HbF observed in previous studies (15, 60, 61).

Reference:

1. **Amrolia PJ, Ramamurthy L, Saluja D, Tanese N, Jane SM, and Cunningham JM**. The activation domain of the enhancer binding protein p45NF-E2 interacts with TAFII130 and mediates long-range activation of the alpha- and beta-globin gene loci in an erythroid cell line. *Proceedings of the National Academy of Sciences of the United States of America* 94: 10051-10056, 1997.

2. **Andrews NC, and Faller DV**. A rapid micropreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. *Nucleic acids research* 19: 2499, 1991.

3. **Antoniou M** editor. *Gene Transfer and Expression Protocols*. 1991, p. 421-434.

4. **Bauer A, Mikulits W, Lagger G, Stengl G, Brosch G, and Beug H**. The thyroid hormone receptor functions as a ligand-operated developmental switch between proliferation and differentiation of erythroid progenitors. *The EMBO journal* 17: 4291-4303, 1998.

5. **Baylin SB**. Reversal of gene silencing as a therapeutic target for cancer--roles for DNA methylation and its interdigitation with chromatin. *Novartis Foundation symposium* 259: 226-233; discussion 234-227, 285-228, 2004.

6. **Bestor TH**. The DNA methyltransferases of mammals. *Human molecular genetics* 9: 2395-2402, 2000.

7. **Cammas F, Herzog M, Lerouge T, Chambon P, and Losson R**. Association of the transcriptional corepressor TIF1beta with heterochromatin protein 1 (HP1): an essential role for progression through differentiation.

Table 2. Mass spectrometry results

Genes & development 18: 2147-2160, 2004.

8. **Cammas F, Mark M, Dolle P, Dierich A, Chambon P, and Losson R**. Mice lacking the transcriptional corepressor TIF1beta are defective in early postimplantation development. *Development (Cambridge, England)* 127: 2955-2963, 2000.

9. **Chang C, and Pan HJ**. Thyroid hormone direct repeat 4 response element is a positive regulatory element for the human TR2 orphan receptor, a member of steroid receptor superfamily. *Molecular and cellular biochemistry* 189: 195-200, 1998.

10. **Christman JK**. 5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. *Oncogene* 21: 5483-5495, 2002.

11. **Collins FS, Metherall JE, Yamakawa M, Pan J, Weissman SM, and Forget BG**. A point mutation in the A gamma-globin gene promoter in Greek hereditary persistence of fetal haemoglobin. *Nature* 313: 325-326, 1985.

12. **Creusot F, Acs G, and Christman JK**. Inhibition of DNA methyltransferase and induction of Friend erythroleukemia cell differentiation by 5-azacytidine and 5-aza-2'-deoxycytidine. *The Journal of biological chemistry* 257: 2041-2048, 1982.

13. **de Boer E, Rodriguez P, Bonte E, Krijgsveld J, Katsantoni E, Heck A, Grosveld F, and Strouboulis J**. Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. *Proceedings of the National Academy of Sciences of the United States of America* 100: 7480-7485, 2003.

14. **Denslow SA, and Wade PA**. The human Mi-2/NuRD complex and gene regulation. *Oncogene* 26: 5433-5438, 2007.

15. **DeSimone J, Heller P, Hall L, and Zwiers D**. 5-Azacytidine stimulates fetal hemoglobin synthesis in anemic baboons. *Proceedings of the National Academy of Sciences of the United States of America* 79: 4428-4431, 1982.

16. **Easwaran HP, Schermelleh L, Leonhardt H, and Cardoso MC**. Replication-independent chromatin loading of Dnmt1 during G2 and M phases. *EMBO reports* 5: 1181-1186, 2004.

17. **Engel JD, and Tanimoto K**. Looping, linking, and chromatin activity: new insights into beta-globin locus regulation. *Cell* 100: 499-502, 2000.

18. **Enver T, Zhang JW, Papayannopoulou T, and Stamatoyannopoulos G**. DNA methylation: a secondary event in globin gene switching? *Genes & development* 2: 698-706, 1988.

19. **Esteve PO, Chin HG, Smallwood A, Feehery GR, Gangisetty O, Karpf AR, Carey MF, and Pradhan S**. Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. *Genes & development* 20: 3089-3103, 2006.

20. **Fathallah H, and Atweh GF**. DNA hypomethylation therapy for hemoglobin disorders: molecular mechanisms and clinical applications. *Blood reviews* 20: 227-234, 2006.

21. **Fathallah H, Portnoy G, and Atweh GF**. Epigenetic analysis of the human alpha- and beta-globin gene clusters. *Blood cells, molecules & diseases* 2007.

22. **Fathallah H, Sutton M, and Atweh GF**. Pharmacological induction of fetal hemoglobin: Why haven't we been more successful in thalassemia? *Annals of the New York Academy of Sciences* 1054: 228-237, 2005.

23. **Fathallah H, Weinberg RS, Galperin Y, Sutton M, and Atweh GF**. Role of epigenetic modifications in normal globin gene regulation and butyrate-mediated induction of fetal hemoglobin. *Blood* 110: 3391-3397, 2007.

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

25. **Franco PJ, Farooqui M, Seto E, and Wei LN**. The orphan nuclear receptor TR2 interacts directly with both class I and class II histone deacetylases. *Molecular endocrinology (Baltimore, Md* 15: 1318-1328, 2001.

26. **Friedman JR, Fredericks WJ, Jensen DE, Speicher DW, Huang XP, Neilson EG, and Rauscher FJ, 3rd**. KAP-1, a novel corepressor for the highly conserved KRAB repression domain. *Genes & development* 10: 2067-2078, 1996.

27. **Fuks F, Burgers WA, Brehm A, Hughes-Davies L, and Kouzarides T**. DNA methyltransferase Dnmt1 associates with histone deacetylase activity. *Nature genetics* 24: 88-91, 2000.

28. **Ghoshal K, Datta J, Majumder S, Bai S, Kutay H, Motiwala T, and Jacob ST**. 5-Aza-Deoxycytidine Induces Selective Degradation of DNA Methyltransferase 1 by a Proteasomal Pathway That Requires the KEN Box, Bromo-Adjacent Homology Domain, and Nuclear Localization Signal. 2005, p. 4727-4741.

29. **Gius D, Cui H, Bradbury CM, Cook J, Smart DK, Zhao S, Young L, Brandenburg SA, Hu Y, Bisht KS, Ho AS, Mattson D, Sun L, Munson PJ, Chuang EY, Mitchell JB, and Feinberg AP**. Distinct effects on gene expression of chemical and genetic manipulation of the cancer epigenome revealed by a multimodality approach. *Cancer cell* 6: 361-371, 2004.

30. **Goren A, Simchen G, Fibach E, Szabo PE, Tanimoto K, Chakalova L, Pfeifer GP, Fraser PJ, Engel JD, and Cedar H**. Fine tuning of globin gene expression by DNA methylation. *PLoS ONE* 1: e46, 2006.

31. **Grosveld F, Rodriguez P, Meier N, Krpic S, Pourfarzad F, Papadopoulos P, Kolodziej K, Patrinos GP, Hostert A, and Strouboulis J**. Isolation and characterization of hematopoietic transcription factor complexes by in vivo biotinylation tagging and mass spectrometry. *Annals of the New York Academy of Sciences* 1054: 55-67, 2005.
32. **Harju S, McQueen KJ, and Peterson KR**. Chromatin structure and control of beta-like globin gene switching. *Experimental biology and medicine (Maywood, NJ* 227: 683-700, 2002.
33. **Heller P, and DeSimone J**. 5-Azacytidine and fetal hemoglobin. *American journal of hematology* 17: 439-447, 1984.
34. **Iida T, Suetake I, Tajima S, Morioka H, Ohta S, Obuse C, and Tsurimoto T**. PCNA clamp facilitates action of DNA cytosine methyltransferase 1 on hemimethylated DNA. *Genes Cells* 7: 997-1007, 2002.
35. **Ishiguro K, and Sartorelli AC**. Coinduction of embryonic and adult-type globin mRNAs by sodium butyrate and trichostatin A in two murine interleukin-3-dependent bone marrow-derived cell lines. *Blood* 92: 4383-4393, 1998.
36. **Ismail A, and Nawaz Z**. Nuclear hormone receptor degradation and gene transcription: an update. *IUBMB life* 57: 483-490, 2005.
37. **Issa JP**. Decitabine. *Current opinion in oncology* 15: 446-451, 2003.
38. **Jones PL, and Wolffe AP**. Relationships between chromatin organization and DNA methylation in determining gene expression. *Seminars in cancer biology* 9: 339-347, 1999.
39. **Juttermann R, Li E, and Jaenisch R**. Toxicity of 5-aza-2'-deoxycytidine to mammalian cells is mediated primarily by covalent trapping of DNA methyltransferase rather than DNA demethylation. *Proceedings of the National Academy of Sciences of the United States of America* 91: 11797-11801, 1994.
40. **Kantarjian HM, O'Brien S, Cortes J, Giles FJ, Faderl S, Issa JP, Garcia-Manero G, Rios MB, Shan J, Andreeff M, Keating M, and Talpaz M**. Results of decitabine (5-aza-2'deoxycytidine) therapy in 130 patients with chronic myelogenous leukemia. *Cancer* 98: 522-528, 2003.
41. **Kim SS, Chen YM, O'Leary E, Witzgall R, Vidal M, and Bonventre JV**. A novel member of the RING finger family, KRIP-1, associates with the KRAB-A transcriptional repressor domain of zinc finger proteins. *Proceedings of the National Academy of Sciences of the United States of America* 93: 15299-15304, 1996.
42. **Kinyamu HK, Chen J, and Archer TK**. Linking the ubiquitin-proteasome pathway to chromatin remodeling/modification by nuclear receptors. *Journal of molecular endocrinology* 34: 281-297, 2005.
43. **Koritschoner NP, Madruga J, Knespel S, Blendinger G, Anzinger B, Otto A, Zenke M, and Bartunek P**. The Nuclear Orphan Receptor TR4 Promotes Proliferation of Myeloid Progenitor Cells. 2001, p. 563-572.
44. **Kuppuswamy M, Vijayalingam S, Zhao LJ, Zhou Y, Subramanian T, Ryerse J, and Chinnadurai G**. Role of the PLDLS-binding cleft region of CtBP1 in recruitment of core and auxiliary components of the corepressor complex. *Molecular and cellular biology* 28: 269-281, 2008.
45. **Le Douarin B, Nielsen AL, Garnier JM, Ichinose H, Jeanmougin F, Losson R, and Chambon P**. A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic control of transcription by nuclear receptors. *The EMBO journal* 15: 6701-6715, 1996.
46. **Le Douarin B, vom Baur E, Zechel C, Heery D, Heine M, Vivat V, Gronemeyer H, Losson R, and Chambon P**. Ligand-dependent interaction of nuclear receptors with potential transcriptional intermediary factors (mediators). *Philosophical transactions of the Royal Society of London* 351: 569-578, 1996.
47. **Le Douarin B, You J, Nielsen AL, Chambon P, and Losson R**. TIF1alpha: a possible link between KRAB zinc finger proteins and nuclear receptors. *The Journal of steroid biochemistry and molecular biology* 65: 43-50, 1998.
48. **Le Douarin B, Zechel C, Garnier JM, Lutz Y, Tora L, Pierrat P, Heery D, Gronemeyer H, Chambon P, and Losson R**. The N-terminal part of TIF1, a putative mediator of the ligand-dependent activation function (AF-2) of nuclear receptors, is fused to B-raf in the oncogenic protein T18. *The EMBO journal* 14: 2020-2033, 1995.
49. **Lee C-H, Chinpaisal C, and Wei L-N**. A Novel Nuclear Receptor Heterodimerization Pathway Mediated by Orphan Receptors TR2 and TR4. 1998, p. 25209-25215.
50. **Lee Y-F, Pan H-J, Burbach JPH, Morkin E, and Chang C**. Identification of Direct Repeat 4 as a Positive Regulatory Element for the Human TR4 Orphan Receptor. A MODULATOR FOR THE THYROID HORMONE TARGET GENES. 1997, p. 12215-12220.
51. **Lee YF, Lee HJ, and Chang C**. Recent advances in the TR2 and TR4 orphan receptors of the nuclear receptor superfamily. *The Journal of steroid biochemistry and molecular biology* 81: 291-308, 2002.
52. **Leone G, Teofili L, Voso MT, and Lubbert M**. DNA methylation and demethylating drugs in myelodysplastic syndromes and secondary leukemias. *Haematologica* 87: 1324-1341, 2002.
53. **Leonhardt H, Page AW, Weier HU, and Bestor TH**. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* 71: 865-873, 1992.
54. **Leu YW, Yan PS, Fan M, Jin VX, Liu JC, Curran EM, Welshons WV, Wei SH, Davuluri RV, Plass C, Nephew KP, and Huang TH**. Loss of estrogen receptor signaling triggers epigenetic silencing of downstream targets in breast cancer. *Cancer research* 64: 8184-8192, 2004.

55. **Ley TJ, Anagnou NP, Young NS, Nienhuis AW, DeSimone J, and Heller P**. 5-azacytidine for beta thalassaemia? *Lancet* 1: 467, 1983.
56. **Ley TJ, DeSimone J, Noguchi CT, Turner PH, Schechter AN, Heller P, and Nienhuis AW**. 5-Azacytidine increases gamma-globin synthesis and reduces the proportion of dense cells in patients with sickle cell anemia. *Blood* 62: 370-380, 1983.
57. **Li G, Franco PJ, and Wei LN**. Identification of histone deacetylase-3 domains that interact with the orphan nuclear receptor TR2. *Biochem Biophys Res Commun* 310: 384-390, 2003.
58. **Lin TM, Young WJ, and Chang C**. Multiple functions of the TR2-11 orphan receptor in modulating activation of two key cis-acting elements involved in the retinoic acid signal transduction system. *The Journal of biological chemistry* 270: 30121-30128, 1995.
59. **Liu Y, Song J, Liu W, Wan Y, Chen X, and Hu C**. Growth and differentiation of rat bone marrow stromal cells: does 5-azacytidine trigger their cardiomyogenic differentiation? *Cardiovascular research* 58: 460-468, 2003.
60. **Mabaera R, Greene MR, Richardson CA, Conine SJ, Kozul CD, and Lowrey CH**. Neither DNA hypomethylation nor changes in the kinetics of erythroid differentiation explain 5-azacytidine's ability to induce human fetal hemoglobin. 2007, p. blood-2007-2006-093948.
61. **Mabaera R, Richardson CA, Johnson K, Hsu M, Fiering S, and Lowrey CH**. Developmental- and differentiation-specific patterns of human gamma- and beta-globin promoter DNA methylation. *Blood* 110: 1343-1352, 2007.
62. **Mahajan MC, Narlikar GJ, Boyapaty G, Kingston RE, and Weissman SM**. Heterogeneous nuclear ribonucleoprotein C1/C2, MeCP1, and SWI/SNF form a chromatin remodeling complex at the beta-globin locus control region. *Proceedings of the National Academy of Sciences of the United States of America* 102: 15012-15017, 2005.
63. **Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schutz G, Umesono K, Blumberg B, Kastner P, Mark M, Chambon P, and Evans RM**. The nuclear receptor superfamily: the second decade. *Cell* 83: 835-839, 1995.
64. **Mavilio F, Giampaolo A, Care A, Migliaccio G, Calandrini M, Russo G, Pagliardi GL, Mastroberardino G, Marinucci M, and Peschle C**. Molecular mechanisms of human hemoglobin switching: selective undermethylation and expression of globin genes in embryonic, fetal, and adult erythroblasts. *Proceedings of the National Academy of Sciences of the United States of America* 80: 6907-6911, 1983.
65. **Meier N, Krpic S, Rodriguez P, Strouboulis J, Monti M, Krijgsveld J, Gering M, Patient R, Hostert A, and Grosveld F**. Novel binding partners of Ldb1 are required for haematopoietic development. *Development (Cambridge, England)* 133: 4913-4923, 2006.
66. **Moosmann P, Georgiev O, Le Douarin B, Bourquin JP, and Schaffner W**. Transcriptional repression by RING finger protein TIF1 beta that interacts with the KRAB repressor domain of KOX1. *Nucleic acids research* 24: 4859-4867, 1996.
67. **Morimoto RI**. Dynamic remodeling of transcription complexes by molecular chaperones. *Cell* 110: 281-284, 2002.
68. **Nielsen AL, Ortiz JA, You J, Oulad-Abdelghani M, Khechumian R, Gansmuller A, Chambon P, and Losson R**. Interaction with members of the heterochromatin protein 1 (HP1) family and histone deacetylation are differentially involved in transcriptional silencing by members of the TIF1 family. *The EMBO journal* 18: 6385-6395, 1999.
69. **O'Neill DW, Schoetz SS, Lopez RA, Castle M, Rabinowitz L, Shor E, Krawchuk D, Goll MG, Renz M, Seelig HP, Han S, Seong RH, Park SD, Agalioti T, Munshi N, Thanos D, Erdjument-Bromage H, Tempst P, and Bank A**. An ikaros-containing chromatin-remodeling complex in adult-type erythroid cells. *Molecular and cellular biology* 20: 7572-7582, 2000.
70. **Omori A, Tanabe O, Engel JD, Fukamizu A, and Tanimoto K**. Adult stage gamma-globin silencing is mediated by a promoter direct repeat element. *Molecular and cellular biology* 25: 3443-3451, 2005.
71. **Perillo B, Ombra MN, Bertoni A, Cuozzo C, Sacchetti S, Sasso A, Chiariotti L, Malorni A, Abbondanza C, and Avvedimento EV**. DNA oxidation as triggered by H3K9me2 demethylation drives estrogen-induced gene expression. *Science (New York, NY* 319: 202-206, 2008.
72. **Perini G, Diolaiti D, Porro A, and Della Valle G**. In vivo transcriptional regulation of N-Myc target genes is controlled by E-box methylation. *Proceedings of the National Academy of Sciences of the United States of America* 102: 12117-12122, 2005.
73. **Poncz M, Henthorn P, Stoeckert C, and Surrey S**. Globin gene expression in hereditary persistence of fetal haemoglobin and (delta beta) (0)-thalassaemia. *Oxford surveys on eukaryotic genes* 5: 163-203, 1988.
74. **Rodriguez P, Bonte E, Krijgsveld J, Kolodziej KE, Guyot B, Heck AJ, Vyas P, de Boer E, Grosveld F, and Strouboulis J**. GATA-1 forms distinct activating and repressive complexes in erythroid cells. *The EMBO journal* 24: 2354-2366, 2005.
75. **Rodriguez P, Braun H, Kolodziej KE, de Boer E, Campbell J, Bonte E, Grosveld F, Philipsen S, and**

Strouboulis J. Isolation of transcription factor complexes by in vivo biotinylation tagging and direct binding to streptavidin beads. *Methods in molecular biology (Clifton, NJ* 338: 305-323, 2006.

76. **Rountree MR, Bachman KE, and Baylin SB**. DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nature genetics* 25: 269-277, 2000.

77. **Ryan RF, Schultz DC, Ayyanathan K, Singh PB, Friedman JR, Fredericks WJ, and Rauscher FJ, 3rd**. KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: a potential role for Kruppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing. *Molecular and cellular biology* 19: 4366-4378, 1999.

78. **Rybak JN, Ettorre A, Kaissling B, Giavazzi R, Neri D, and Elia G**. In vivo protein biotinylation for identification of organ-specific antigens accessible from the vasculature. *Nature methods* 2: 291-298, 2005.

79. **Rybak JN, Scheurer SB, Neri D, and Elia G**. Purification of biotinylated proteins on streptavidin resin: a protocol for quantitative elution. *Proteomics* 4: 2296-2299, 2004.

80. **Saleque S, Kim J, Rooke HM, and Orkin SH**. Epigenetic Regulation of Hematopoietic Differentiation by Gfi-1 and Gfi-1b Is Mediated by the Cofactors CoREST and LSD1. *Molecular Cell* 27: 562-572, 2007.

81. **Sanchez C, Sanchez I, Demmers JA, Rodriguez P, Strouboulis J, and Vidal M**. Proteomics analysis of Ring1B/Rnf2 interactors identifies a novel complex with the Fbxl10/Jhdm1B histone demethylase and the Bcl6 interacting corepressor. *Mol Cell Proteomics* 6: 820-834, 2007.

82. **Sasaki R, and Miura Y**. Effect of 5-azacytidine on the differentiation of human leukemia K-562 cells. *International journal of cell cloning* 4: 192-202, 1986.

83. **Saunthararajah Y, Lavelle D, and DeSimone J**. DNA hypo-methylating agents and sickle cell disease. *British journal of haematology* 126: 629-636, 2004.

84. **Schultz DC, Ayyanathan K, Negorev D, Maul GG, and Rauscher FJ, 3rd**. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes & development* 16: 919-932, 2002.

85. **Schultz DC, Friedman JR, and Rauscher FJ, III**. Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2{alpha} subunit of NuRD. 2001, p. 428-443.

86. **Shevchenko A, Wilm M, Vorm O, and Mann M**. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Analytical chemistry* 68: 850-858, 1996.

87. **Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA, and Shi Y**. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 119: 941-953, 2004.

88. **Shi YJ, Matson C, Lan F, Iwase S, Baba T, and Shi Y**. Regulation of LSD1 histone demethylase activity by its associated factors. *Mol Cell* 19: 857-864, 2005.

89. **Shibata H, Spencer TE, Onate SA, Jenster G, Tsai SY, Tsai MJ, and O'Malley BW**. Role of co-activators and co-repressors in the mechanism of steroid/thyroid receptor action. *Recent progress in hormone research* 52: 141-164; discussion 164-145, 1997.

90. **Stamatoyannopoulos G**. Prospects for developing a molecular cure for thalassemia. *Hematology (Amsterdam, Netherlands)* 10 Suppl 1: 255-257, 2005.

91. **Tagle DA, Koop BF, Goodman M, Slightom JL, Hess DL, and Jones RT**. Embryonic epsilon and gamma globin genes of a prosimian primate (Galago crassicaudatus). Nucleotide and amino acid sequences, developmental regulation and phylogenetic footprints. *Journal of molecular biology* 203: 439-455, 1988.

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

93. **Tanabe O, McPhee D, Kobayashi S, Shen Y, Brandt W, Jiang X, Campbell AD, Chen YT, Chang C, Yamamoto M, Tanimoto K, and Engel JD**. Embryonic and fetal beta-globin gene repression by the orphan nuclear receptors, TR2 and TR4. *The EMBO journal* 26: 2295-2306, 2007.

94. **Tanabe O, Shen Y, Liu Q, Campbell AD, Kuroha T, Yamamoto M, and Engel JD**. The TR2 and TR4 orphan nuclear receptors repress Gata1 transcription. *Genes & development* 21: 2832-2844, 2007.

95. **Tanimoto K, Liu Q, Grosveld F, Bungert J, and Engel JD**. Context-dependent EKLF responsiveness defines the developmental specificity of the human epsilon-globin gene in erythroid cells of YAC transgenic mice. *Genes & development* 14: 2778-2794, 2000.

96. **Taylor SM, and Jones PA**. Multiple new phenotypes induced in 10T1/2 and 3T3 cells treated with 5-azacytidine. *Cell* 17: 771-779, 1979.

97. **Underhill C, Qutob MS, Yee S-P, and Torchia J**. A Novel Nuclear Receptor Corepressor Complex, N-CoR, Contains Components of the Mammalian SWI/SNF Complex and the Corepressor KAP-1. 2000, p. 40463-40470.

98. **van der Ploeg LH, and Flavell RA**. DNA methylation in the human gamma delta beta-globin locus in erythroid and nonerythroid tissues. *Cell* 19: 947-958, 1980.

99. **Venturini L, You J, Stadler M, Galien R, Lallemand V, Koken MH, Mattei MG, Ganser A, Chambon P,**

Losson R, and de The H. TIF1gamma, a novel member of the transcriptional intermediary factor 1 family. *Oncogene* 18: 1209-1217, 1999.
100. **Wang XM, Wang X, Li J, and Evers BM**. Effects of 5-azacytidine and butyrate on differentiation and apoptosis of hepatic cancer cell lines. *Annals of surgery* 227: 922-931, 1998.
101. **Waugh DS**. Making the most of affinity tags. *Trends in biotechnology* 23: 316-320, 2005.
102. **Weatherall DJ**. ABC of clinical haematology. The hereditary anaemias. *BMJ (Clinical research ed* 314: 492-496, 1997.
103. **Weinberg RS, Ji X, Sutton M, Perrine S, Galperin Y, Li Q, Liebhaber SA, Stamatoyannopoulos G, and Atweh GF**. Butyrate increases the efficiency of translation of gamma-globin mRNA. *Blood* 105: 1807-1809, 2005.
104. **Wilm M, Shevchenko A, Houthaeve T, Breit S, Schweigerer L, Fotsis T, and Mann M**. Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* 379: 466-469, 1996.
105. **Xue Y, Wong J, Moreno GT, Young MK, Cote J, and Wang W**. NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. *Mol Cell* 2: 851-861, 1998.
106. **Yoder JA, Soman NS, Verdine GL, and Bestor TH**. DNA (cytosine-5)-methyltransferases in mouse cells and tissues. Studies with a mechanism-based probe. *Journal of molecular biology* 270: 385-395, 1997.

Table 2.

NCBI ID	Description	GO Classification*	Peptides
gi|20072056	Brg1	SWI/SNF, chromatin remodeling	2
gi|40254165	SMC hinge domain containing 1	cell cycle?	18
gi|6678832	mcm6	cell cycle	11
gi|6678830	mcm5, cdc46	cell cycle	7
gi|6754816	septin 2	cell cycle	7
gi|51980515	cdk9	cell cycle	6
gi|53551	mcm3	cell cycle	6
gi|10242373	mcm7	cell cycle	6
gi|36031035	SMC3	cell cycle	5
gi|50510569	SCC-112 novel cell cycle-regulated molecule	cell cycle?	4
gi|9957546	Septin 8	cell cycle	4
gi|28204888	septin 9	cell cycle	4
gi|2183319	mcm2	cell cycle	3
gi|9790237	SMC1	cell cycle?	3
gi|1293686	HCF1	cell cycle, transcription factor	3
gi|26354124	cdc1, septin 7	cell cycle	2
gi|12963599	cyclin H	cell cycle	2
gi|19527092	RCC1	cell cycle	2
gi|50360	Cdc2a, cdc2, cdk1	cell cycle	1
gi|22779899	cdc5l	cell cycle	1
gi|4099131	Mad2l1	cell cycle	1
		Total	**96**
gi|37360250	TIF1gamma	co-repressor	3
gi|6755787	TIF 1 beta	co-repressor	17
gi|3452507	CtBP1 protein	co-repressor	1
		Total	**21**
gi|1363105	replication licensing factor MCM4	DNA replication	3
gi|31419356	Dnmt1	Dnmt1	50
gi|33695134	structure specific recognition protein 1‡	FACT	2
gi|14028669	DNA-dependent ATPase SNF2H ‡	ISWI	4
gi|39204553	Mi2 beta	NuRD	4
gi|16905113	MTA1	NuRD	11
gi|2347180	HDAC1	NuRD	2
		Total	**17**
gi|1144348	TR4	orphan nuclear receptor	58
gi|1173533	TR2	orphan nuclear receptor	42
		Total	**100**
gi|2239142	polycomb-M33 interacting protein Ring1B	Pc-G complex, ubiquitin ligase	3
gi|1262300	casein kinase II alpha subunit	signal transduction	3
gi|46237616	casein kinase 2, beta subunit	signal transduction	1
		Total	**4**
gi|2895588	TFII-I protein long form; p128	transcription factor	2

gi|19527168	Psip1	transcription factor	5
gi|110759	nuclear factor I (clone B5)	transcription factor	1
gi|26344676	Pur alpha	transcription factor	1
gi|35505176	CAND1	transcription factor, ubiquitin cycle	2
		Total	**11**
gi|91870	polyubiquitin (clone arf3)	ubiquitin-proteasome pathway	2
gi|12848428	Psmd14	ubiquitin-proteasome pathway	2
gi|30519947	Psmd11	ubiquitin-proteasome pathway	4
gi|12845617	Psmd6	ubiquitin-proteasome pathway	5
gi|12852148	Psmc6	ubiquitin-proteasome pathway	8
gi|7110703	Psmc5	ubiquitin-proteasome pathway	1
gi|6679503	Psmc3	ubiquitin-proteasome pathway	3
gi|13542777	Psmd2	ubiquitin-proteasome pathway	3
gi|6755210	Psmd13	ubiquitin-proteasome pathway	6
gi|6679503	Psmc3	ubiquitin-proteasome pathway	1
gi|20988514	Psmd11	ubiquitin-proteasome pathway	1
gi|26347233	Psmd5	ubiquitin-proteasome pathway	3
gi|2137840	Psmd3	ubiquitin-proteasome pathway	7
gi|1698570	Psme1	ubiquitin-proteasome pathway	2
gi|6754724	Psmd7	ubiquitin-proteasome pathway	5
gi|55217	murine valosin-containing protein	ubiquitin-proteasome pathway	15
		Total	**68**
gi|6753486	Cops2	signalosome complex, cell proliferation	1
gi|33563284	Cops6	signalosome complex	2
gi|6753490	Cops4	signalosome complex	1
		Total	**4**
gi|194027	heat-shock protein hsp84	heat shock	65
gi|1661134	heat shock 70 protein	heat shock	11
		Total	**76**
		Grand Total	**461**

* : most proteins that were identified have multiple GO entries. For practical reasons, we have chosen to present on GO term that we consider to be the most descriptive.

‡ : proteins routinely identified in mass spec experiments be streptavidin pulldown of biotin tagged unrelated chromatin associated factors. In the absence of any other FACT or ISWI complex partners, we tend to consider these proteins as chromatin associated background

Table 2: Classification by Gene Ontology (GO) criteria of proteins identified by mass spectrometry as co-purifying with TR2/TR4

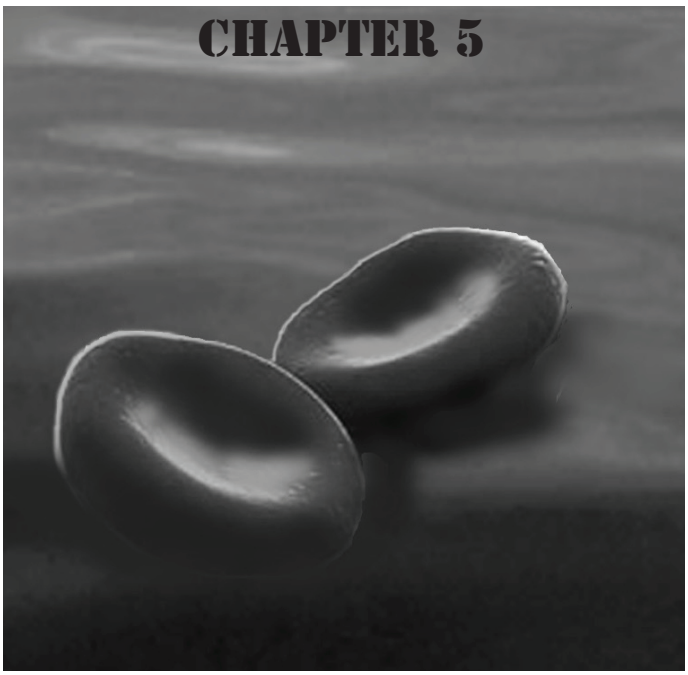
Table 3

		MS1	MS2	MS3	MS4	MS5
Cell Cycle	Cdc2a, cdc2, cdk1	yes				
	cdc5l	yes	yes	yes		yes
	cdk9	yes				
	cyclin H	yes				
	HCF1	yes				yes
	Mad2l1	yes				
	mcm2	yes		yes		
	mcm3	yes				
	mcm4	yes		yes		
	mcm5, cdc46	yes		yes		
	mcm6	yes		yes		
	mcm7	yes	yes	yes		
	RCC1	yes				
	SCC-112 novel cell cycle-regulated molecule	yes		yes		
	septin 2	yes		yes	yes	yes
	Septin 8	yes				
	cdc10, septin 7	yes		yes	yes	yes
	septin 11			yes		yes
	septin 9	yes				
	SMC hinge domain containing 1	yes				
	SMC1	yes				yes
	SMC3	yes	yes			yes
	SMC4		yes	yes		yes
	Smc2l1			yes		
	Smc1l1			yes		
Co-repressors	TIF1 gamma	yes		yes		
	TIF1 beta	yes	yes	yes	yes	
	HP1 gamma					yes
	CtBP1 protein [Mus musculus]*	yes				
	IRA1			yes		
	Sin3A					yes
FACT	Ssrp1	yes		yes		
	suppressor of Ty 5 homolog			yes		
	suppressor of Ty 16 homolog			yes		yes
NuRD	Mi2 beta	yes		yes	yes	yes
	MTA1	yes		yes		yes
	HDAC 1	yes	yes	yes		yes
	mRbAp48		yes	yes	yes	yes
	Mi2 alpha			yes		
	Chd5			yes		
	HDAC 2					yes
	retinoblastoma binding protein 5					yes
signal transduction	casein kinase II alpha subunit	yes	yes	yes		
	casein kinase II, beta subunit	yes		yes		
ubiquitin cycle	polyubiquitin (clone arf3)	yes				
proteasome	Psmd14	yes				
	Psmd11	yes				yes

	Psmd6	yes				
	Psmc6	yes		yes		yes
	Psmc5	yes				yes
	Psmc3	yes				
	Psmd2	yes				
	Psmd13	yes				
	Psmc3	yes				
	Psmd5	yes				
	Psmd3	yes	yes			
	Psme1	yes				
	Psmd7	yes				
	valosin-containing protein	yes	yes	yes	yes	yes
	Cops2, COP9	yes				
	Cops6, COP9	yes				
	Cops4, COP9	yes				
	CAND1	yes		yes	yes	yes
	Psmc1		yes			
	Psmd12					yes
	Psmc2					yes
	Prp19					yes
	FANCL					yes
DNA methyltransferase	Dnmt1	yes		yes	yes	yes

Table 3: Presence of potential interacting proteins in different mass spectrometry analyses. For MS4 and MS5, benzonase was used in strptavidin binding experiments

OPTIMISATION OF ChIP ASSAYS USING in vivo BIOTIN TAGGED TRANSCRIPTION FACTORS



Manuscript in preparation

Provisional title: Optimization of ChIP assays using in vivo biotin tagged transcription factors

Katarzyna E. Kolodziej, Farzin Poufarzad, Ernie de Boer, Frank Grosveld* and John Strouboulis*[1].

Department of Cell Biology, ErasmusMC, Dr Molewaterplein 50, 3015GE Rotterdam, the Netherlands

[1]Present address: Institute of Molecular Oncology BSRC "Alexander Fleming" P.O. Box 74145, 16602 Varkiza Greece

* Corresponding authors: f.grosveld@erasmusmc.nl, strouboulis@fleming.gr

Abstract:

In vivo biotinylation tagging has been used by virtue of its very tight binding to (immobilised) streptavidin as a high affinity approach for purifying nuclear protein complexes from mammalian cells (25) and has also been also shown to work well in combination with other tags (37). Here we describe the optimization of biotin mediated ChIP assays using the biotin-tagged GATA-1 hematopoietic transcription factor in erythroid cells. We show that the choice of beads is crucial for the efficient enrichment of target genes in chromatin binding. In addition, we show that using fish skin gelatin in blocking the beads and omitting SDS when sonicating chromatin, significantly improve enrichments in streptavidin-ChIP. Together these steps provide an optimized approach for the high affinity, low background application of biotinylation tagging in ChIP assays. Finally we show that under these optimized conditions the V5 peptide tag performs equally well in antibody mediated ChIPs.

Introduction

Affinity tagging is used as a powerful and elegant tool for the recovery and purification of recombinant proteins. Among a number of affinity-based purification technologies, the use of the biotin/(strept)avidin system occupies a prominent place due to its unique characteristics (14), which include: (a) the very tight and specific binding of biotin by avidin (or streptavidin) which, with a K_d of 10^{15} L*mol $^{-1}$, is one of the highest non covalent interactions known in nature, close to almost $10^3 - 10^6$ times greater than the interaction of epitopes with their specific antibodies. Once formed, the biotin-streptavidin complex is not disturbed by changes in pH, introduction of detergents or high salt concentration, thus remaining stable even under very stringent washing conditions; (b) biotin is a very small molecule and is not known to affect the biological activity of tagged proteins (11, 21); (c) there are few (mostly cytoplasmic) naturally biotinylated proteins in mammalian cells. As a result the non-specific background binding of nuclear extract is low (10).

Early experiments applying biotinylation tagging of specific proteins in E.coli and yeast (9) combined the use of the bacterial BirA biotin ligase together with a 75 amino-acid long tag derived from naturally biotinylated proteins such as carboxylases. More recently, it was shown that short artificial tags obtained by screening a combinatorial synthetic peptide library can also be very efficiently biotinylated by BirA and have been subsequently used as the tags of choice (4, 10, 28).

Previous work in our lab has used one such short (23aa) tag for the purification of GATA-1 complexes from nuclear extracts of erythroid cells (10, 25) GATA-1 is a DNA sequence-specific zinc finger transcription factor that is essential for the differentiation of erythroid, megakaryocytic, eosinophil and mast cell lineages (6, 13). N-terminally tagged GATA-1 was co-expressed with the E.coli BirA biotin ligase in mouse erythroleukemic (MEL) cells and subsequently purified from nuclear extracts together with interacting proteins by binding to streptavidin beads (10). In this way, a number of known and novel GATA-1 protein partners were identified (25). We also provided preliminary evidence that biotinylation tagging could be applied in place of antibodies in ChIPs of GATA-1 target genes (10, 18).

Several groups, including ours are testing different tags to search for the most optimal for the isolation of transcription factor complexes and their target genes by ChiP-on-chip or ChIP-sequencing (ChIP-seq). In this selection three criteria are important: (a) tags must have high binding affinity; (b) tag binding must be insensitive to formaldehyde fixation; (c) tags should be preferably small and not strongly charged so as to minimize possible interference with transcription factor function. These different (and equally efficient) tags could be used in tandem, separated by a protease cleavage site to allow an extra affinity purification step for lowering the background of non-specific proteins. At the same time this approach can greatly enhance the ability to purify the complexes to homogeneity (by using several factors of the same complex differentially tagged and co-expressed) for other applications, e.g. for crystallisation and structure determination.

Here we present the optimization and comparison of various ChIP protocols using tagged GATA-1 in combination with known target genes (25) as an example. We first show that different streptavidin beads are not equally efficient in ChIP assays. We also show that effective blocking with fish skin gelatin and omission of SDS during chromatin sonication are important factors in reducing background signals, which is a major concern in ChIP using complex chromatin from mammalian cells. Further, we present that these optimal conditions can be efficiently applied for the V5 tag and antibodies mediated ChIP. Finally, it is also important to note that the most frequent formaldehyde cross-link is formed at the end of the side-chain of lysine (K) and histidine (H) and hence a number of tags containing a lysine/histidine in the recognised part of the tag may have very poor yields. To underline this importance we also present evidence that formaldehyde crosslinking can affect the efficiency of the biotin-tag, more than V5–tag binding.

Materials and methods

Cells and constructs. MEL cells were cultured as previously described (2). Constructs and stably transfected cell lines were described previously (10).

Chromatin crosslinking. Approximately $2x10^7$ induced MEL cells were crosslinked with 1% formaldehyde for 10 minutes at room temperature and processed for sonication essentially as described in the Upstate protocol. Chromatin was sonicated on ice with a Sanyo Soniprep 150 sonicator at amplitude 6 using 10 cycles of 15 sec "on" and 45 sec "off" to a DNA fragment size in the range of 300 to 800 nucleotides. The alternative "no SDS" sonication buffer is: 10mM Tris, 1mM EDTA and 0.5mM EGTA. All buffers were supplemented with Complete Protease Inhibitor Cocktail (Roche). Aliquots of sonicated chromatin of $10x10^6$ cells were stored at –80°C.

Streptavidin ChIP. Chromatin pull-downs with streptavidin beads were carried out overnight using 20 µl of streptavidin Dynabeads M280 (Invitrogen) or 20µl of UltraLink Immobilized NeutrAvidin Protein (Pierce) per chromatin aliquot. For streptavidin agarose (Sigma) or Streptavidin mutein (Roche) chromatin pull-downs, 60µl of agarose slurry or beads were used per aliquot. All the beads/slurry were blocked with 400µg sonicated salmon sperm DNA for 1h at 4°C. Pre-clearing of chromatin prior to binding to M280 streptavidin beads was done using 20µl of Protein G Dynabeads (Invitrogen) preblocked with salmon sperm DNA). Chromatin incubation with beads was carried out in a total reaction volume of 1mL supplemented with Complete Protease Inhibitor, at 4°C overnight on a rotating wheel. After binding, beads were washed with 1mL of low salt, high salt, LiCl and TE (10mM Tris-HCl pH 8.0, 1mM ETDA) wash buffers, 3-5 minutes each, as described in Upstate protocol. An additional urea wash (5M urea/2M thiourea/1% TritonX100) was carried out after LiCl buffer wash and before the TE washes (see Results). After the washes, bound chromatin was eluted by resuspending Neutravidin, Mutein, streptavidin agarose and M280 beads in 500 µl 0.1M sodium carbonate, 1%SDS, 0.2M NaCl elution buffer transferring to a fresh tube and decrosslinking with shaking at 65°C for at least 5h. Thermal elution of chromatin from M280 Dynabeads was carried out by resuspending beads in 500 µl 95% formamide, 0.1M sodium bicarbonate and boiling at 95ºC for 10 min. M280 Dynabeads were subsequently separated from the buffer using a magnetic rack. Eluted chromatin was transferred to a fresh tube and decrosslinking was carried out as described above. Decrosslinked samples were deproteinized as described by the Upstate protocol. DNA was recovered by phenol/chloroform/isoamyl alcohol extraction and isopropanol precipitation using 20µg glycogen as carrier.

Antibody ChIP. GATA-1 ChIP has been previously described (25). Anti-GATA-1 antibodies used were N6 and M20 (Santa Cruz). Anti GATA-1 antibody immunoprecipitates were eluted from the beads by incubation in elution buffer (0.1M sodium carbonate, 1%SDS) twice for 15 minutes each at room temperature.

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

Real time PCR. This was done in an Opticon I (MJ Research) thermal cycler using SYBR Green and Platinum Taq polymerase (Invitrogen) as described previously (25). Primers (listed in Table1) were designed using Primer Express 2.0 (PE Applied Biosystems). For each experiment at least two runs were done with each sample loaded in duplicate. PCR conditions: 95°C for 10 min, 40 cycles of 30 sec at 95 °C, 60 sec at 60°C, 15 sec at 75°C. Enrichment for a specific DNA sequence was calculated using the comparative C_T method, as previously described (22). The enrichment of bound DNA over input is calculated using the formula $2^{Ct(IP)-Ct(Ref)}$. Enrichment over the negative primer set or negative control chromatin from MEL cells expressing BirA only was subsequently calculated by dividing.

Nuclear extracts, and immunoblotting. These were done as previously described (10).

Results and Discussion

GATA-1 is a key transcriptional regulator of erythroid genes. In addition, a number of genes involved in cell cycle regulation or cell survival and proliferation, such as Bcl-2, c-myc and c-myb, have also been identified as GATA-1 targets in erythroid cells (5, 15, 26, 30). We have previously shown that GATA-1 binds together with Gfi1b to the repressed c-myb promoter in induced MEL cells, thus suggesting a role for this complex in proliferation arrest during terminal

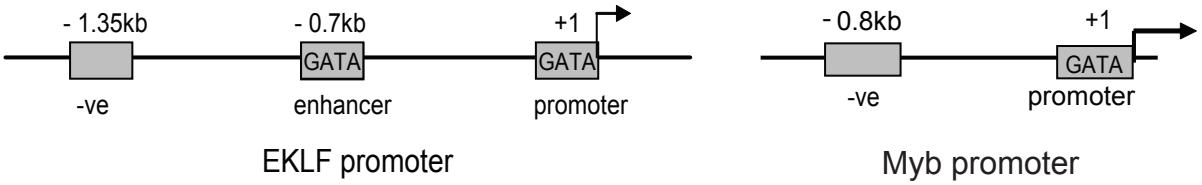


Figure1. Location of the ChIP primers in the EKLF and c-myb promoters. "GATA" boxes indicate GATA-1 binding sites.

erythroid differentiation (16, 17, 25). GATA-1 also regulates expression of key hematopoietic transcription factor genes in erythroid differentiation, such as GATA-2 and EKLF. For example, binding of GATA-1 to the EKLF promoter and enhancer is essential for its transcriptional activation (1, 25), with forced GATA-1 expression activating the proximal EKLF promoter in non-erythroid cells (7).

We previously showed using biotin-tagged GATA-1 that a streptavidin binding of crosslinked chromatin could substitute for antibodies in enriching for GATA-1 target genes in ChIP assays (10). Due to the potential advantages offered by the very high affinity of streptavidin for binding to biotin and the importance of having multiple tags that can be used in the same cell, we wanted to extended these observations in developing an optimized protocol for the streptavidin binding of chromatin from cells expressing biotin-tagged GATA-1. In doing so, we used the EKLF and c-myb promoters as examples of GATA-1 gene targets that are upregulated or repressed, respectively, in erythroid cells. Figure 1 shows the location of primers used for the EKLF and myb promoters and negative control sequences. Primer sequences are listed in Table 1.

Name of the primer	Sequence
Myb prom FOR	ACTGCAGGGGCGCCAGATTT
Myb prom REV	GGAGAAAGGGGAGGAGAAGGAGGTA
Neg myb FOR	GAAGTAGAGGCAGGATAATCAGGAA
Neg myb REV	AGGATGAACCAGGGCTAATGC
EKLF Upstream FOR	CTGGCCCCCCTACCTGAT
EKLF Upstream REV	GGCTCCCTTTCAGGCATTATC
EKLF Promoter FOR	TATCGCACACACCCCTCCTT
EKLF Promoter REV	CCCACATCTGATTGGCTGTCT
Neg EKLF FOR	TGCTCCCCACTATGATAATGGA
Neg EKLF REV	GCCACAACCAAAGAAGACATTTT

Comparison of different types of streptavidin beads

Biotinylation of biological substrates is frequently used in a variety of different applications and hence many manufacturers offer a wide range of immobilized streptavidin matrices. We have previously used paramagnetic M280 streptavidin Dynabeads for the isolation of protein complexes (10, 12, 24, 25, 27). We also tested the performance of M280 beads in chromatin immunoprecipitations

and compared them to three other available products: streptavidin agarose, streptavidin mutein and NeutrAvidin. Streptavidin agarose was used to test whether an immobilization matrix different to that of paramagnetic particles would give better yields. Neutravidin is a streptavidin derivative without carbohydrate side chains which is predicted to reduce background binding. Streptavidin mutein is a mutated recombinant streptavidin which binds biotin with a lower affinity thus allowing

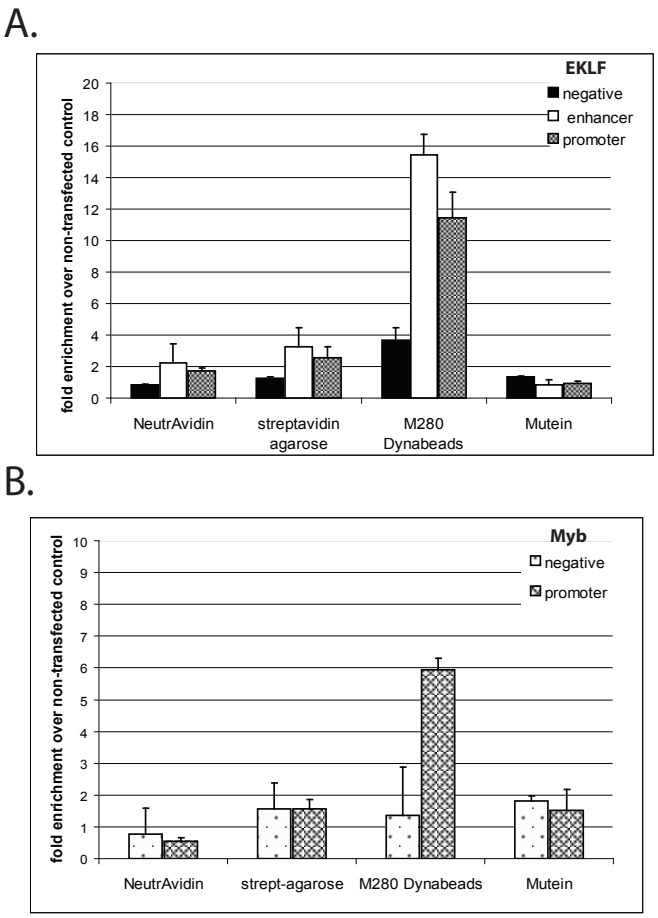


Figure 2. Comparison of different derivatives of immobilized streptavidin: NeutrAvidin, streptavidin agarose, streptavidin mutein and M280 Dynabeads. Relative enrichment for EKLF (panel **A)** and c-myb (panel **B**) promoters was calculated over negative control chromatin isolated from cells expressing BirA biotin ligase, but not tagged GATA-1. The difference in the scales between the two genes indicates a lower enrichment for the c-myb promoter presumably due to the "closed" nature of the chromatin of this repressed gene.

elution of bound material under gentler conditions by using biotin. Chromatin from biotin-tagged GATA-1 cells and non-transfected cells (expressing BirA ligase only) was bound to different types of beads, under identical conditions (overnight binding and subsequent washes). Biotin-tagged GATA-1 was eluted from the beads by decrosslinkig, except for the mutein beads where we used biotin for elution. The results shown in Figure 2A and B clearly show that the M280 streptavidin Dynabeads are the most efficient in capturing biotin-tagged GATA-1 bound to the EKLF and myb promoters. As a result, the M280 Dynabeads were used in all subsequent experiments.

Pre-clearing chromatin

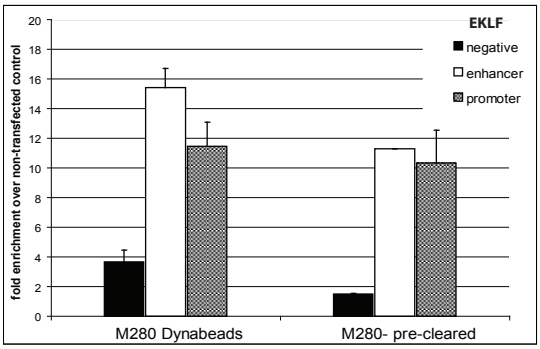
Pre-clearing of chromatin is one of the methods used to decrease background binding in ChIP assays using antibodies. We tested this by preclearing chromatin with Protein G paramagnetic beads (Dynal) for 1 hour at 4°C. As shown in Figure 3, this resulted in lower background and improved enrichment (Fig. 3) of EKLF sequences bound by biotin tagged GATA-1. Similar results were also obtained with the c-myb promoter (not shown).

Blocking with fish skin gelatin

Among various blocking compounds (e.g. BSA, Chicken Egg Albumin etc.) fish skin gelatin (FGEL) has been shown to be very effective for blocking in Western and ELISA experiments (20, 35). We therefore investigated whether the use of FGEL for blocking would help improve the performance of M280 beads in streptavidin ChIP. Figure 4 shows that addition of as little as 0.5%

(final concentration) of FGEL (together with salmon sperm DNA) can significantly improve the yield of EKLF target sequences bound by GATA-1. Similar results were also observed with the c-myb promoter (data not shown). Thus blocking the beads for 1 hour with FGEL and salmon sperm DNA reduces the background compared to blocking with salmon sperm DNA only. In addition to blocking the beads, we also used 1%FGEL by adding it to chromatin during binding to the beads with similar results to those obtained when only blocking the beads with FGEL (not shown). As a result, we have included 1%FGEL in blocking the beads in all subsequent experiments.

A.



B.

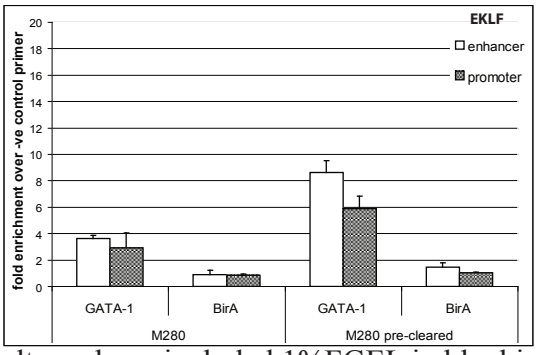


Figure3. Preclearing chromatin with Protein G Dynabeads. **A**) Relative enrichment of EKLF sequences calculated over chromatin from control cells. The enrichment of the specific EKLF promoter elements appears lower after preclearing presumably due to some loss of chromatin in the additional preclearing step. **B**) Relative enrichment of biotin-tagged GATA-1 binding at EKLF promoter and enhancer calculated over the negative control sequence for biotin-tagged GATA-1 chromatin and chromatin from BirA expressing cells as negative control.

It has been shown previously that biotin tagging allows more stringent washes (containing up to 3% SDS) compared to other affinity tags (like TAP tag) (34). For example, urea and thiourea are reagents widely used in proteomics to resuspend hydrophobic proteins. We therefore tested whether the background binding of hydrophobic proteins can be reduced by washing in urea/thiourea/SDS. We found that the additional wash did not significantly lower the background or increase the specific binding signals of the EKLF or c-myb promoters (data not shown) and this parameter was not investigated further.

The non-covalent binding between of biotin to streptavidin is one of the strongest known in nature (3, 38). However this presents a drawback when eluting bound chromatin from the beads as

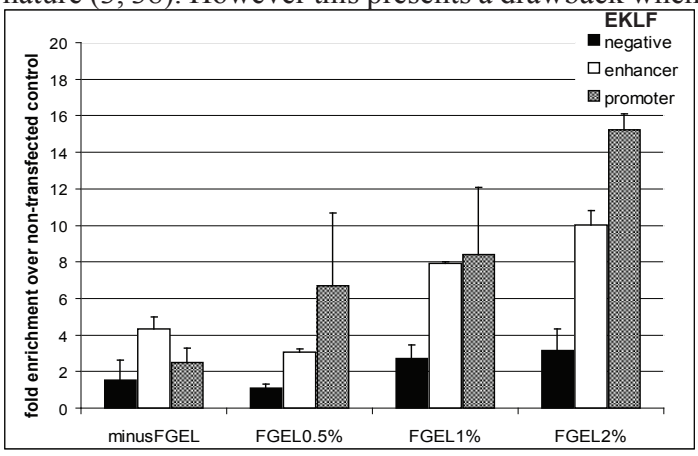


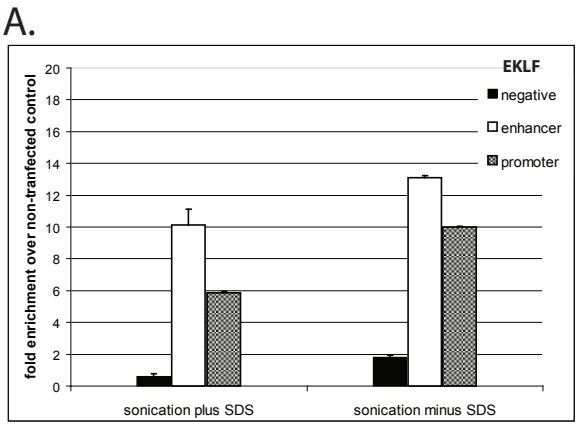
Figure 4. The effect of using different concentrations of FGEL in blocking the beads. Relative enrichment was calculated for EKLF sequences over chromatin from control cells.

is usually done in ChIP using antibodies. We tried to overcome this problem by "thermal elution" i.e. by boiling the beads after the binding (see Materials and Methods). We found this method to be very inefficient and did not pursue it any further.

To possibly overcome this problem, elution from the beads by biotin (similar to mutein) could be applied, although in our hands this did not prove to be very efficient. A better alternative would be inclusion of protease (TEV or Precission) cleavage sites (8, 19, 23, 36)

Sonication without SDS

Most ChIP protocols, including those used in our laboratory, are based on the Upstate ChIP protocol which includes sonicating chromatin in a buffer containing 1% SDS. Addition of SDS introduces stringent conditions and helps prevent the aggregation of insoluble protein complexes. However, high SDS concentration may affect optimal binding of chromatin by the antibody or



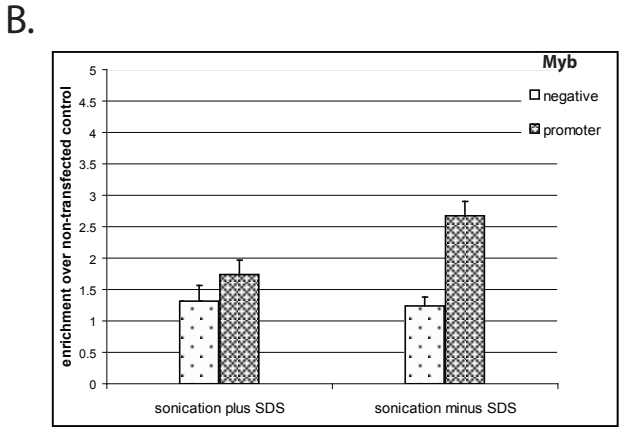


Figure 5. Comparison of sonication buffers with or without the addition of SDS. **A**) Relative enrichment of EKLF sequences calculated over chromatin from control cells. **B**) Relative enrichment of c-myb promoter sequences calculated over chromatin from control cells.

beads and in some approaches inclusion of SDS is not compatible with further experimental procedures, for example in chromatin fractionation by CsCl gradient centrifugation (29, 31).

We therefore tested whether sonicating chromatin without SDS would improve the efficiency of a streptavidin ChIP. Omission of SDS did not affect the efficiency of DNA shearing. As shown in Figure 5, sonicating chromatin without SDS resulted in higher enrichment of both EKLF and c-myb promoter sequences, albeit with a small increase in the background binding of the EKLF negative control sequence. Enrichment of the negative control primer in the c-myb promoter is similar in both conditions. Similar results were also obtained with other biotin-tagged proteins (KK and FP unpublished observations). Thus, omitting SDS from the sonication buffer improves the yield of a streptavidin pull-down significantly.

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

All the experiments described above were carried out with an N-terminally biotin-tagged

GATA-1 (10). We also generated a second construct containing a tandem affinity tag created by fusing the short (14 aa) biotin tag (4) with the 14aa V5 tag to the C-terminus of GATA-1(Figure 6A). V5 is a short peptide sequence derived from the C-terminus of the P and V proteins of Simian virus 5 (33, 39). This construct can be used in the two-step affinity purification of tagged protein complexes, thus reducing background binding. This is an important consideration in

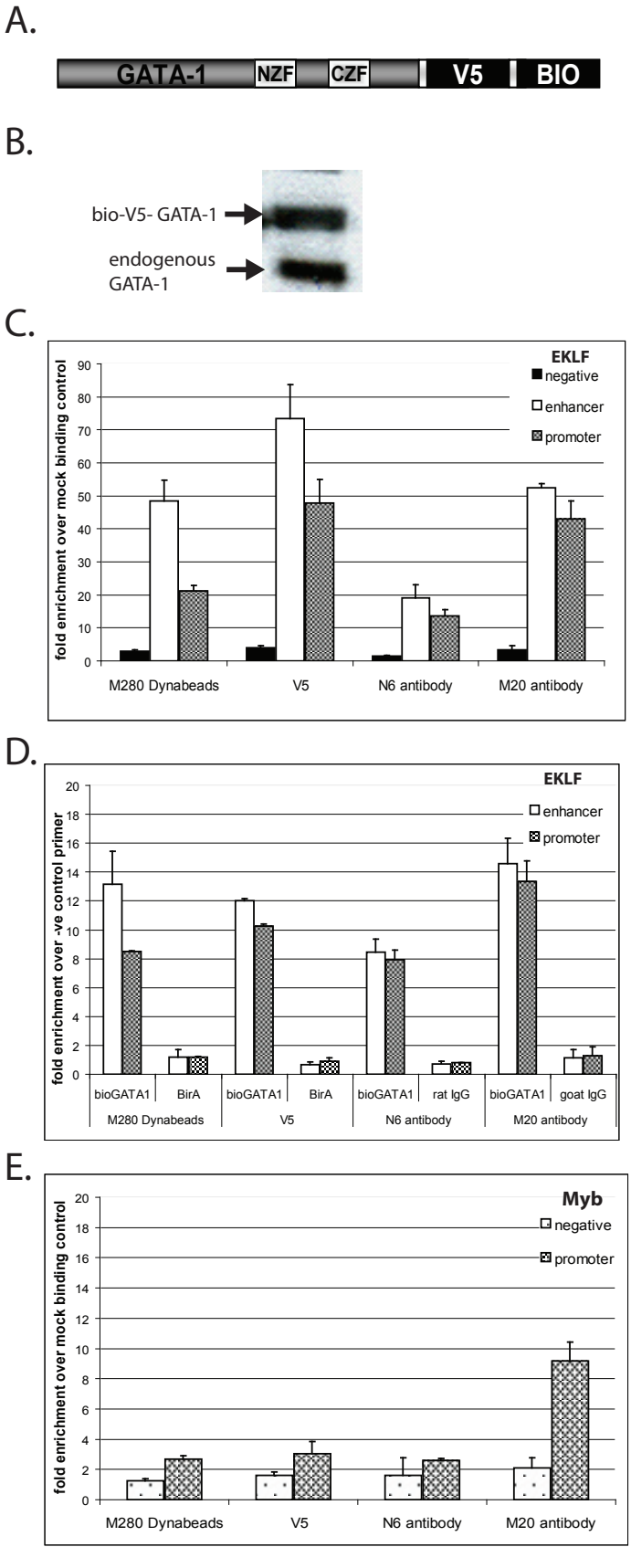


Figure 6.

A) Schematic representation of the biotin-V5-tagged GATA-1

B) Western blot showing the relative amounts of biotin-V5 –GATA 1 and endogenous GATA-1 in the extract used for ChIP. The biotin-V5-GATA-1 was detected with N6 anti-GATA-1 antibody.

C) Comparison of V5, M280 and two different anti-GATA-1 antibodies, N6 and M20, tested for the binding of GATA-1 to the EKLF promoter. The enrichment is calculated over a non-transfected BirA cells control or over IgG negative control, respectively. V5 ChIP gives the highest yield in the EKLF enhancer and promoter elements, streptavidin ChIP with M280 Dynabeads gives comparable yield in the upstream enhancer element as M20 anti-GATA-1 antibody. The M20 antibody can enrich for more GATA-1 bound to the EKLF promoter than the M280 Dynabeads. The N6 antibody precipitates the least amount of GATA-1 bound to EKLF promoter elements.

D) Comparison of V5, M280 and two different anti-GATA-1 antibodies N6 and M20, tested for the binding of GATA-1 to the EKLF promoter. Enrichment of the specific binding to EKLF promoter and enhancer was calculated over the negative primer set (-1.35kb element in EKLF promoter). V5agarose and M280 Dynabeads bring down comparable amounts of GATA-1 bound to EKLF enhancer and promoter sequences. The M20 antibody enriches the most for GATA-1 bound to the EKLF upstream enhancer. Rat and goat IgGs as well as BirA control show similarly low enrichments of specific primer sets.

E) Comparison of V5, M280 and two different anti-GATA-1 antibodies N6 and M20, tested as binding of GATA-1 to myb promoter.

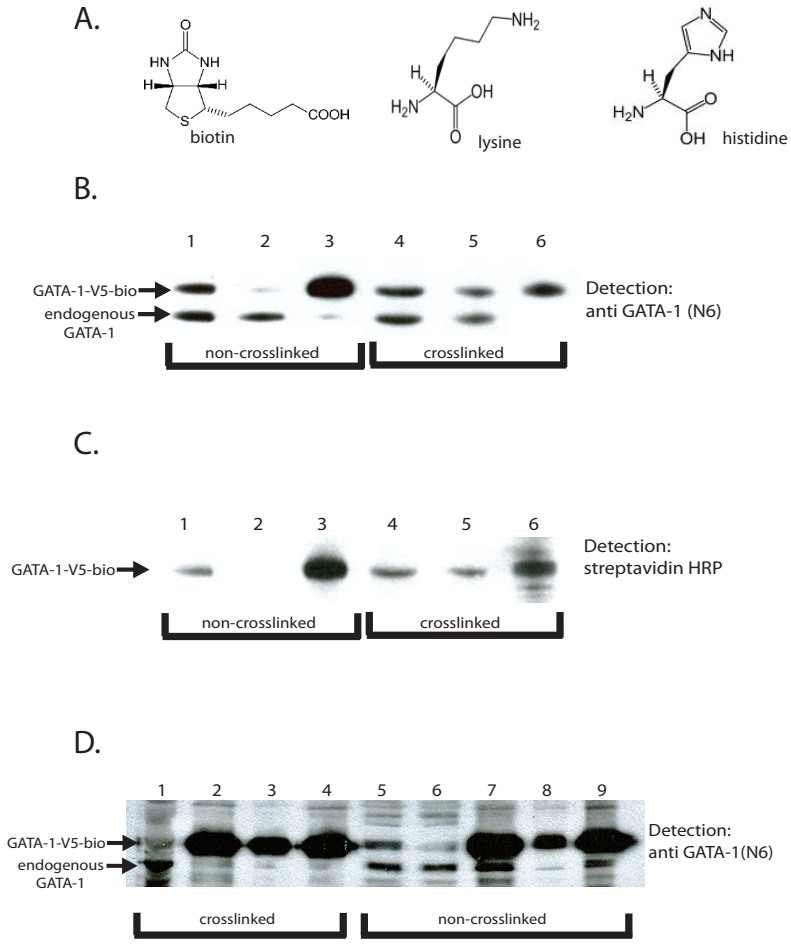


Figure 7. A) The structure of biotin lysine and histidine **B)** Western blot analysis of two binding experiments where non-crosslinked nuclear extract (lane 1,2,3) and formaldehyde crosslinked chromatin (lane 4,5,6) were tested. Input and supernatant fractions (lanes 1 and 2, 4 and 5 respectively) represent 1% of total material, while bound lane (3 and 6) represent 25% of total material. The biotin-V5-tagged GATA-1 was detected with anti-GATA-1 antibody. **C)** After stripping the same membrane was incubated with streptavidin-HRP. **D)** Western blot showing comparison of binding of two different extracts: non-crosslinked nuclear extract (lane 5-9) and crosslinked chromatin (lane 1-5) precipitation by V5 agarose. The biotin-V5- tagged GATA-1was detected with N6 antibody. Lanes 1= input crosslinked loaded 1% of total protein, 2= bound (after washing, before elution), 3= gata-1 eluted from the beads loaded 10% of total material, 4= leftover on the beads after elution loaded 10% of total material, 5= input non- crosslinked loaded 1% of total material, 6= unbound supernatant after binding loaded 1% of total protein, 7= bound (after washing, before elution), 8=GATA-1 eluted from the beads loaded 10% of total protein, 9= leftover on the beads after elution loaded 10% of total protein,

ChIP experiments particularly in applications involving ChIP-on-chip or ChIP sequencing. This construct also allows comparison with the streptavidin-ChIP results obtained with the biotin tag fused to the N-terminus of GATA-1.

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

extract used in this experiment are similar (Fig6B). This means that the enrichments obtained with antibodies should be hypothetically approximately twice as high as with affinity tags pull-downs. Thus, the results obtained with the M20 antibody are an over estimate when compared to the tag based ChIPs (in particular see figure 6D).

The V5 tag appears to work at least as well as streptavidin binding in ChIP (Fig 6C). When normalising to control cells expressing only BirA, the V5 ChIP gives actually a slightly better enrichment compared to the streptavidin ChIP (Fig. 6C), albeit with a slightly higher background binding to the negative control sequence (Fig. 6C and D). One explanation for the difference observed between the V5 ChIP and the streptavidin ChIP may be that the elution from the anti-V5 agarose beads is more efficient than that from the M280 streptavidin beads.

Remarkably, when we tested the c-myb promoter sequences under the same conditions, we found that the M20 antibody (a polyclonal antibody) gave much higher enrichment (Figure 6D). This suggests that the epitope(s) of GATA-1 recognized by M-20 antibody is more accessible for binding by antibodies in the crosslinked "closed" chromatin environment of the repressed c-myb gene, when compared with N-terminus and the very C-terminus of the protein.

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

Formaldehyde cross-linking as first introduced by Salomon et al in *Drosophila* (32), has been widely used to study the binding of proteins to DNA elements in intact cells. We tested whether the structure of the biotin and V5 tags or of the GATA-1 protein is affected by formaldehyde cross linking. The V5 aa sequence (GKPIPNPLLGLDST) contains one lysine (K) residue, whilst the biotin (with chemical formula $C_{10}H_{16}N_2O_3S$, see the structure in Figure 7A) is potentially more sensitive to formaldehyde cross linking. To this end, the extracts containing equal amounts of GATA-1-V5-bio (see input lanes 1 and 4 in figure 7A) were bound under identical conditions. Input, bound and unbound fractions we loaded on an SDS-PAGE gel and the western blots were detected with an anti GATA-1 (N6) antibody and with streptavidin-HRP. The results showed that the binding of crosslinked material to the M280 Dynabeads was less efficient, as there was much more GATA-V5-bio found in unbound fraction in comparison to the binding of non-crosslinked material where there was hardly any GATA-V5-bio left in the unbound fraction (compare the amounts of GATA-1 in lane 2 and lane 5 of unbound fractions 7A and B). In addition, anti-GATA-1 antibody as well as streptavidin–HRP detection (Fig 7A and B) showed that there was significantly less GATA-1-V5-bio bound to the M280 Dynabeads in cross-linked extract pull-down than in non-crosslinked material (compare lane 3 with 6 in figures 7A and 7B). We subsequently tested whether crosslinking affects binding to the anti-V5 agarose. The western blot data presented in Figure 7C suggests that crossilinked and non-crosslinked material can be equally efficiently bound by anti-V5 beads (Figure 8, compare line 2 with 7 and 3 with 8). This observation likely explains the slightly higher V5 enrichments in the graphs in Figure 6. Although more sensitive to formaldehyde, it should be emphasized that despite its sensitivity to formaldehyde the biotin-tag performed very well in all our ChIP experiments.

In summary, we show that pre-clearing chromatin and blocking streptavidin beads with fish skin gelatin reduces the background binding. In addition avoiding SDS in the sonication buffer appears to increase enrichment for bio-GATA-1 binding to specific DNA sequences. Lastly, we provide evidence that a tandem affinity tag composed of the biotin tag and the V5 epitope tag works very efficiently in ChIP, either when used independently or in combination. Our preliminary ChIP-sequencing data (E. de Boer and E. Soler, unpublished) show that biotin and V5 tagging can be very efficiently used for transcription factor target sequence mapping and that the background seen in these experiments contains mostly non-specific DNA fragments that can be easily distinguished from specific target sites.

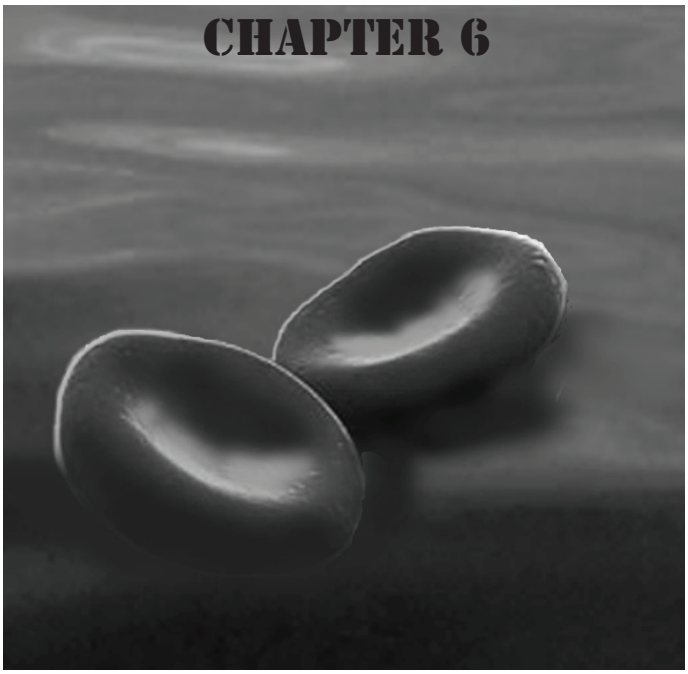
Thus, affinity tagging of transcription factors offers an easy, rapid and effective way for comparative and functional studies of different transcription complexes.

References:

1. **Anderson KP, Crable SC, and Lingrel JB**. Multiple proteins binding to a GATA-E box-GATA motif regulate the erythroid Kruppel-like factor (EKLF) gene. *The Journal of biological chemistry* 273: 14347-14354, 1998.
2. **Antoniou M** editor. *Gene Transfer and Expression Protocols*. 1991, p. 421-434.
3. **Bayer EA, and Wilchek M**. Application of avidin-biotin technology to affinity-based separations. *Journal of chromatography* 510: 3-11, 1990.
4. **Beckett D, Kovaleva E, and Schatz PJ**. A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Sci* 8: 921-929, 1999.
5. **Blobel GA, and Orkin SH**. Estrogen-induced apoptosis by inhibition of the erythroid transcription factor GATA-1. *Molecular and cellular biology* 16: 1687-1694, 1996.
6. **Cantor AB, and Orkin SH**. Hematopoietic development: a balancing act. *Current opinion in genetics & development* 11: 513-519, 2001.
7. **Chen X, Reitman M, and Bieker JJ**. Chromatin structure and transcriptional control elements of the erythroid Kruppel-like factor (EKLF) gene. *The Journal of biological chemistry* 273: 25031-25040, 1998.
8. **Cordingley MG, Callahan PL, Sardana VV, Garsky VM, and Colonno RJ**. Substrate requirements of human rhinovirus 3C protease for peptide cleavage in vitro. *The Journal of biological chemistry* 265: 9062-9065, 1990.
9. **Cronan JE, Jr.** Biotination of proteins in vivo. A post-translational modification to label, purify, and study proteins. *The Journal of biological chemistry* 265: 10327-10333, 1990.
10. **de Boer E, Rodriguez P, Bonte E, Krijgsveld J, Katsantoni E, Heck A, Grosveld F, and Strouboulis J**. Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. *Proceedings of the National Academy of Sciences of the United States of America* 100: 7480-7485, 2003.
11. **Diamandis EP, and Christopoulos TK**. The biotin-(strept)avidin system: principles and applications in biotechnology. *Clinical chemistry* 37: 625-636, 1991.
12. **Driegen S, Ferreira R, van Zon A, Strouboulis J, Jaegle M, Grosveld F, Philipsen S, and Meijer D**. A generic tool for biotinylation of tagged proteins in transgenic mice. *Transgenic research* 14: 477-482, 2005.
13. **Ferreira R, Ohneda K, Yamamoto M, and Philipsen S**. GATA1 function, a paradigm for transcription factors in hematopoiesis. *Molecular and cellular biology* 25: 1215-1227, 2005.
14. **Ford CF, Suominen I, and Glatz CE**. Fusion tails for the recovery and purification of recombinant proteins. *Protein expression and purification* 2: 95-107, 1991.
15. **Gregory T, Yu C, Ma A, Orkin SH, Blobel GA, and Weiss MJ**. GATA-1 and erythropoietin cooperate to promote erythroid cell survival by regulating bcl-xL expression. *Blood* 94: 87-96, 1999.
16. **Hock H, Hamblen MJ, Rooke HM, Schindler JW, Saleque S, Fujiwara Y, and Orkin SH**. Gfi-1 restricts proliferation and preserves functional integrity of haematopoietic stem cells. *Nature* 431: 1002-1007, 2004.
17. **Hock H, and Orkin SH**. Zinc-finger transcription factor Gfi-1: versatile regulator of lymphocytes, neutrophils and hematopoietic stem cells. *Current opinion in hematology* 13: 1-6, 2006.
18. **Horsman S, Moorhouse MJ, de Jager VC, van der Spek P, Grosveld F, Strouboulis J, and Katsantoni EZ**. TF Target Mapper: a BLAST search tool for the identification of Transcription Factor target genes. *BMC bioinformatics* 7: 120, 2006.
19. **Kapust RB, and Waugh DS**. Controlled intracellular processing of fusion proteins by TEV protease. *Protein expression and purification* 19: 312-318, 2000.
20. **Kaur R, Dikshit KL, and Raje M**. Optimization of immunogold labeling TEM: an ELISA-based method for evaluation of blocking agents for quantitative detection of antigen. *J Histochem Cytochem* 50: 863-873, 2002.
21. **Laitinen OH, Nordlund HR, Hytonen VP, and Kulomaa MS**. Brave new (strept)avidins in biotechnology. *Trends in biotechnology* 25: 269-277, 2007.
22. **Litt MD, Simpson M, Recillas-Targa F, Prioleau MN, and Felsenfeld G**. Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. *The EMBO journal* 20: 2224-2235, 2001.
23. **Lucast LJ, Batey RT, and Doudna JA**. Large-scale purification of a stable form of recombinant tobacco etch virus protease. *BioTechniques* 30: 544-546, 548, 550 passim, 2001.
24. **Meier N, Krpic S, Rodriguez P, Strouboulis J, Monti M, Krijgsveld J, Gering M, Patient R, Hostert A, and Grosveld F**. Novel binding partners of Ldb1 are required for haematopoietic development. *Development (Cambridge, England)* 133: 4913-4923, 2006.
25. **Rodriguez P, Bonte E, Krijgsveld J, Kolodziej KE, Guyot B, Heck AJ, Vyas P, de Boer E, Grosveld F, and Strouboulis J**. GATA-1 forms distinct activating and repressive complexes in erythroid cells. *The EMBO journal* 24: 2354-2366, 2005.
26. **Rylski M, Welch JJ, Chen YY, Letting DL, Diehl JA, Chodosh LA, Blobel GA, and Weiss MJ**. GATA-1-mediated proliferation arrest during erythroid maturation. *Molecular and cellular biology* 23: 5031-5042, 2003.
27. **Sanchez C, Sanchez I, Demmers JA, Rodriguez P, Strouboulis J, and Vidal M**. Proteomics analysis of Ring1B/Rnf2 interactors identifies a novel complex with the Fbxl10/Jhdm1B histone demethylase and the Bcl6 interacting corepressor. *Mol Cell Proteomics* 6: 820-834, 2007.
28. **Schatz PJ**. Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: a 13 residue consensus peptide specifies biotinylation in Escherichia coli. *Bio/technology (Nature Publishing Company)* 11: 1138-1143, 1993.
29. **Schwartz YB, Kahn TG, and Pirrotta V**. Characteristic low density and shear sensitivity of cross-linked chromatin containing polycomb complexes. *Molecular and cellular biology* 25: 432-439, 2005.
30. **Shimamoto T, Nakamura S, Bollekens J, Ruddle FH, and Takeshita K**. Inhibition of DLX-7 homeobox gene causes decreased expression of GATA-1 and c-myc genes and apoptosis. *Proceedings of the National Academy of Sciences of the United*

States of America 94: 3245-3249, 1997.
31. **Solomon MJ, Larsen PL, and Varshavsky A**. Mapping protein-DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. *Cell* 53: 937-947, 1988.
32. **Solomon MJ, and Varshavsky A**. Formaldehyde-mediated DNA-protein crosslinking: a probe for in vivo chromatin structures. *Proceedings of the National Academy of Sciences of the United States of America* 82: 6470-6474, 1985.
33. **Southern JA, Young DF, Heaney F, Baumgartner WK, and Randall RE**. Identification of an epitope on the P and V proteins of simian virus 5 that distinguishes between two isolates with different biological characteristics. *The Journal of general virology* 72 (Pt 7): 1551-1557, 1991.
34. **van Werven FJ, and Timmers HT**. The use of biotin tagging in Saccharomyces cerevisiae improves the sensitivity of chromatin immunoprecipitation. *Nucleic acids research* 34: e33, 2006.
35. **Vogt RF, Jr., Phillips DL, Henderson LO, Whitfield W, and Spierto FW**. Quantitative differences among various proteins as blocking agents for ELISA microtiter plates. *Journal of immunological methods* 101: 43-50, 1987.
36. **Walker PA, Leong LE, Ng PW, Tan SH, Waller S, Murphy D, and Porter AG**. Efficient and rapid affinity purification of proteins using recombinant fusion proteases. *Bio/technology (Nature Publishing Company)* 12: 601-605, 1994.
37. **Wang J, Rao S, Chu J, Shen X, Levasseur DN, Theunissen TW, and Orkin SH**. A protein interaction network for pluripotency of embryonic stem cells. *Nature* 444: 364-368, 2006.
38. **Wilchek M, Bayer EA, and Livnah O**. Essentials of biorecognition: the (strept)avidin-biotin system as a model for protein-protein and protein-ligand interaction. *Immunology letters* 103: 27-32, 2006.
39. **Zhou W, Dong L, Ginsburg D, Bouhassira EE, and Tsai HM**. Enzymatically active ADAMTS13 variants are not inhibited by anti-ADAMTS13 autoantibodies: a novel therapeutic strategy? *The Journal of biological chemistry* 280: 39934-39941, 2005.

DISCUSSION



DISCUSSION

During hematopoiesis, the controlled expression of lineage-specific genes is crucial for proliferation and differentiation cues. Many transcription factors are essential for the development of a given hematopoietic lineage, for example GATA-1 which is required for generation of the erythroid and megakaryocytic lineages (Cantor and Orkin 2002; Ferreira, Ohneda et al. 2005; Rodriguez, Bonte et al. 2005). The ability of such factors to influence cell fate is related to their ability to recognize specific DNA sequences via their DNA-binding domains, and recruit to the regulatory elements of genes (via protein interaction domains) co-activators or co-repressors (Rice, Hormaeche et al. 2007). The aim of this thesis was to gain insight into transcription factor function in regulating erythroid gene expression in general and developmental switches in globin gene expression, specifically. We undertook a biochemical approach using *in vivo* biotinylation tagging to isolate and characterize transcription factor complexes of GATA-1 and of the TR2/TR4 orphan nuclear receptors from erythroid cells

- **Biotinylation tagging**

In this thesis we show that *in vivo*-biotin tagging is a very efficient approach to isolate and characterize protein complexes (Chapter 2 and 3). We applied this method to show that GATA-1 can form distinct protein complexes, with either repressive or activating functions ((Rodriguez, Bonte et al. 2005) and Chapter 3). By using biotin-tagged deletion mutants of GATA-1, we could also show that distinct domains of the GATA-1 protein confer specific protein-protein interactions ((Rodriguez, Bonte et al. 2005) and Chapter 3).

Moreover *in vivo* biotinylation tagging can be very efficiently used in ChIP experiments instead of antibodies (Chapter 5). Using GATA-1 as an example, investigated the optimal conditions for streptavidin ChIP assays were investigated. The results in Chapter 5 show that the type of streptavidin beads used for chromatin binding is critical, as in our experience the M280 Dynabeads performed best. Blocking the streptavidin beads with gelatin from cold water fish skin (FGEL) and using a sonication buffer without SDS in the chromatin preparation step significantly reduces background binding.

In addition, the work described here shows that under these optimized conditions antibody based tags can also be efficiently used in ChIP experiments as we were also able to show that the V5 tag performs as efficiently as the biotin-tag in ChIP experiments. This provides a uniform integrated platform for identifying protein interactions and building interactome networks between the hematopoietic transcription factors, as well as for identifying transcription factor binding sites in a genome wide manner for example, in ChIP-on-chip experiments or for ChIP sequencing using massive parallel sequencing platforms (E. de Boer and E. Soler, unpublished).

I will limit the further discussion to TR2/TR4 because the results obtained for GATA-1 were extensively discussed in the thesis of Patrick Rodriguez (Rodriguez 2006)

- **Identification of TR2/TR4 interacting proteins**

Chapter 4 of this thesis describes the efficient expression of biotin-tagged TR2 and TR4 in doubly transfected MEL cells. Both tagged and endogenous TR2 and TR4 have similar fractionation profiles on an analytical gel filtration column. However, only a small portion of TR2/TR4 proteins was present in high molecular weight fractions, with the majority of TR2/TR4 protein present in low molecular weight fractions (i.e. migrating essentially with the mobility of TR2/TR4 heterodimers and TR2 and TR4 monomers alone). As a consequence we had to develop optimization steps in order to enrich for the TR2/TR4 high molecular weight complexes in our streptavidin binding experiments. After testing a number of approaches, a preparative grade superose 6 gel filtration column followed by streptavidin binding and in-gel or on-beads trypsinization was determined to

be the most efficient approach (see Chapter 4). This procedure, followed by mass spectrometry analysis enabled us to identify a long list of potential interacting proteins, some of which we were able to confirm by western blot and co-immunoprecipitation (co-IP). We show that biotinylated-TR2/TR4 reproducibly and specifically binds Dnmt1, HDAC3, TIF1β, Lsd1 and components of the NuRD complex (including HDAC1, Mta-1, RbAp46/RbAp48 and Mi2β) by streptavidin binding and IPs. We also showed that those interactions are reciprocal, i.e. antibodies that recognize (some of) these factors can, in turn, bring down TR2/TR4 in MEL nuclear extracts (Chapter 4).

In our mass spectrometric analysis we did not recover any of the initial proteins that were reported to co-purify with the DRED complex: CDP-1, USF1, and Lbp1a. We also failed to confirm these interactions by IPs (data not shown) in agreement with the data of Tanabe and Engel (unpublished). We conclude that these factors represent contaminants that were bound to the DNA on the DNA affinity column that was used in the original purification of the DRED complex from MEL cells (Tanabe, Katsuoka et al. 2002). Moreover, of the previously published TR2 and TR4 interacting factors, we could only confirm in our studies the co-purification of HDACs (Franco, Farooqui et al. 2001; Zhang and Dufau 2003). We failed to detect previously reported TR2 and TR4 co-factors such as RIP140, TIP27 or TRA16 (Lee, Chinpaisal et al. 1998; Yang, Wang et al. 2002; Nakajima, Fujino et al. 2004). This may result from differences in the assays and cell lines used: TIP27, TRA16 and RIP140 are co-factors that were identified in yeast and mammalian two-hybrid studies (which often result in false positives) (Lee, Chinpaisal et al. 1998; Franco, Farooqui et al. 2001; Yang, Wang et al. 2002; Nakajima, Fujino et al. 2004). Furthermore, to confirm these interactions most of the authors used COS-1 cells, transiently or stably transfected with either His- (Yang, Wang et al. 2002), GST- (Lee, Chinpaisal et al. 1998; Franco, Farooqui et al. 2001) or Flag-tagged (Nakajima, Fujino et al. 2004) forcibly expressed TR2 or TR4 together with those co-factors, followed by immunoprecipitation of nuclear extracts. Examining abundantly expressed proteins in heterologous cell types, forcibly expressed to high concentrations will allow the detection of low affinity interactions that may not be visible or take place at normal concentrations. Such expression studies are no longer generally believed to provide sufficient evidence for relevant in vivo interactions; most current studies demand evidence that such interactions exist through the analysis of cells that express the proteins at their native abundance. Additionally, in most of the cited experiments the stringencies of IPs were much milder than those used by us, again favoring the detection of low affinity interactions. In contrast, we were able to confirm the interactions we identified by mass spectrometry using not only stably transfected cell lines (where we showed that biotin-tagging allows more stringent conditions to be used), but also non-transfected MEL nuclear extracts. In some experiments we also used benzonase digestion to demonstrate that the interactions observed and reported here are not DNA mediated.

A possible explanation for the lack of finding in our mass spectrometry results of previously reported TR2/TR4 interacting proteins might be that the exclusion of the DNA affinity column from our purification steps resulted in the loss of some genuine DRED components during the experimental procedure. Perhaps our method picked up more abundant proteins forming a complex with TR2/TR4 which are not tightly involved in globin gene regulation and failed to detect less abundant proteins that are involved. This may be the reason why we could not see any clear binding by any of these factors to the ε- or γ-globin gene promoters (though we tested primers which mapped only in the vicinity of DR1 elements). It is possible that the pool of free TR2/TR4, observed in the low molecular weight fractions, is actively exchanged between several complexes and that more purification steps are needed in order to identify which of the complexes is the "globin switching" entity. In order to extend our isolation technique to enrich for DR1 binding complexes, we could possibly apply additional steps prior to the streptavidin pull-down step.

On the basis of the mass spectrometry analysis and immunoprecipitation data and assuming we did not miss the relevant complex(es), we hypothesized that TR2/TR4 can form two distinct

transcription factor complexes. Complex 1 consists of TR2/TR4 together with Dnmt1, HDAC3, and TIF1β, whereas complex 2 consists of TR2/TR4 with (at least a subset of) NuRD components (Mi2β, RbAp46/48, Mta1, HDAC1), and possibly also Lsd1 and Dnmt1. Finding which of the complexes confers the DRED repression function would be an immediate goal, followed by the question whether the DRED complex confers only a repressive function (as proposed by (Tanabe, Katsuoka et al. 2002; Tanabe, McPhee et al. 2007))?

The question about function of the DRED complex has been addressed in transgenic and knockout experiments in mice. The results from experiments with TR2/TR4 null mice and Tg^{dnTR4} mutant mice (dominant negative TR4 mutant that does not bind to DNA) suggest that TR2/TR4 may act in a stage–selective manner: as an ε–globin repressor in primitive and definitive erythropoiesis and as a γ-globin repressor only in definitive erythropoiesis (Tanabe, McPhee et al. 2007).

Interestingly, forced expression of TR4 (Tg^{TR4}) in mice resulted in εy–globin reduction in yolk sac, with no difference in βh1 expression, and also in induction of adult β-globin gene expression. Forced expression of TR2 did not cause any changes in the β-type globin genes in the yolk sac (Tanabe, McPhee et al. 2007). These observations are particularly interesting in light of the ChIP results showing binding of TR4 to the β-globin promoters (tested on both the mouse and human genes). We were able to confirm this binding *in vitro* by band shift assays. It is possible that TR4 (in co-operation with TR2 or other nuclear receptors or as a homodimer) may confer an activating function in globin gene regulation as has been previously shown for other genes regulated by TR4 (Hwang, Burbach et al. 1998; Yan, Karam et al. 1998; Koritschoner, Madruga et al. 2001; Lee, Lee et al. 2002; Shyr, Hu et al. 2002; Yang, Wang et al. 2002; Wang, Lee et al. 2006). It is interesting that recent publications reported a link between the androgen and estrogen receptors and the Lsd1 histone demethylase in gene activation (Metzger, Wissmann et al. 2005; Wissmann, Yin et al. 2007; Perillo, Ombra et al. 2008). This could potentially offer an explanation for our observations of TR4 binding *in vivo* to the active β-globin promoter. However we have as yet failed to detect Lsd1 binding to the β globin promoter by ChIP (in MEL cells), probably because the anti-Lsd1 antibodies did not perform well in our ChIP experiments. Lsd1 has also been reported to interact with BHC80 which resembles TIF1β (Moosmann, Georgiev et al. 1996) in that it is a PHD domain-containing protein (Shi, Matson et al. 2005). BHC80 inhibits co-REST/Lsd1-mediated demethylation in *in vitro* experiments (Shi, Matson et al. 2005). In light of this report it would be interesting to test whether there is a functional interaction between TIF1β and Lsd1 in regulating globin gene expression i.e. to test whether TIF1β would negate Lsd1 induced repressive chromatin formation during ε and γ globin silencing.

Additional gain-of-function experiments were designed in order to address the TR2/TR4 function in erythropoiesis (Tanabe, McPhee et al. 2007). These experiments showed that forced expression of TR2 and TR4 resulted in severe repression of εy globin gene in fetal livers and induction of βh1 expression (up to 8-fold over the levels in the wild type littermate livers), which may be the result of lack of εy competition for the LCR. Adult β-globin accumulation was unaffected (Tanabe, McPhee et al. 2007). Furthermore, when the adult spleens from TR2/TR4 overexpressing mice were tested, βh1 levels were higher compared to wild type litter mates while εy or β- globin expression levels were unaffected (Tanabe, McPhee et al. 2007). It has been suggested that in primitive erythroid cells of the yolk sac TR2/TR4 bind only to DR1 sites in the ε-globin gene promoter resulting in silencing. At this stage, TR2/TR4 would have no effect on γ-globin expression. The ε gene–selective binding was proposed to be a result of the greater number of DR1 elements in the ε–globin gene promoter and a greater affinity of TR2/TR binding (Tanabe, McPhee et al. 2007). To explain the induction of γ (and βh1) globin gene expression in transgenic mice with forced expression of TR2/TR4, Tanabe et al (Tanabe, McPhee et al. 2007) suggest two models (Figure 1).

The first model (Figure 1A) suggest that over-expression of TR2/TR4 causes sequestration

and dilution of potential co-repressors (which are present in limited abundance) thus preventing them from binding to globin promoters. That would suggest that the levels of Dnmt1, and possibly of Lsd1 or NuRD complex components in erythroid cells are limiting factors for the silencing of the γ (or βh1) globin genes, which seems unlikely. It is of course possible that there is another, yet undiscovered TR2/TR4 co-repressor that was not detected in our analysis due to its low levels, transient interactions or technical aspects of our purification procedure.

The second model (Figure 1B) suggests that the action of TR2/TR4 would be dependent on co-activators, co-repressors, ligands or post-translational modifications (Tanabe, McPhee et al. 2007). Hence, similarly to the GATA-1 situation (see (Rodriguez, Bonte et al. 2005) and Chapter 2), TR2 and TR4 can form distinct activating and repressive complexes (Figure 1D) depending on the context (and possibly on the post translational modifications or ligand binding). Whether the presumptive complex with Lsd1 plays an activating function while the interaction with NuRD and Dnmt1 would confer a repressive function (Figure 1D) remains to be tested. Such an analysis would require sequential IPs or pull-downs with two different tags to dissect the distinct protein complexes, followed by extensive ChIP experiments showing the binding of all the different complexes to different promoters during different stages of erythropoiesis. In addition different modifications of the factors may be present requiring additional sophisticated mass spectrometry experiments to determine which of the factors carry modifications and whether these modifications are dynamic during development and/or differentiation.

In addition, we suggest that TR2/TR4 may have a secondary effect on globin gene regulation by affecting the expression of other transcription factors such as GATA-1 (Figure 1C). It was recently demonstrated that forced expression of TR2/TR4 in mice repressed GATA-1 in primitive and definitive erythroid cells by binding to the (-3.7kb) GATA-1 hematopoietic enhancer (G1HE) (Tanabe, Shen et al. 2007). We were also able to show the binding of TR2/TR4 to the (-3.7kb) G1HE element in ChIP experiments (after (Tanabe, Shen et al. 2007)). We could hypothesize that TR2/TR4 could bind promoters of other transcription factors genes. GATA-1 binds to several gene promoters including the β- and γ-globin gene promoters (reviewed in (Stamatoyannopoulos 1991)), hence it is possible that some of the effects on globin gene expression in the transgenic TR2/TR4 mice are due to loss of GATA-1 binding (or possibly due to loss of binding of other TR2/TR4-regulated transcription factors) (Figure 1D).

The models described above are rather complex, and do not consider competition between the genes of the β-globin locus for the activating function of the LCR, a mechanism that has been discussed extensively in the literature (Choi and Engel 1988; Enver, Raich et al. 1990; Hanscombe, Whyatt et al. 1991; Dillon, Trimborn et al. 1997; Grosveld, de Boer et al. 1998). It is possible that silencing of the ε-globin gene by forced expression of TR2/TR4 alone is sufficient for activation of the γ-globin to the levels observed by Tanabe et al (Tanabe, McPhee et al. 2007) (Fig. 1E). That would also suggest that the major role of TR2/TR4 is stage specific ε-globin silencing and that those factors play a minor role in γ globin regulation or even that they do not have any function in γ globin gene regulation.

The enrichments we observe in ChIP experiments for the ε and γ globin promoters are unexpectedly low in light of the band shift assays results previously reported for DRED (Tanabe, Katsuoka et al. 2002; Tanabe, McPhee et al. 2007). However, we can not exclude the possibility

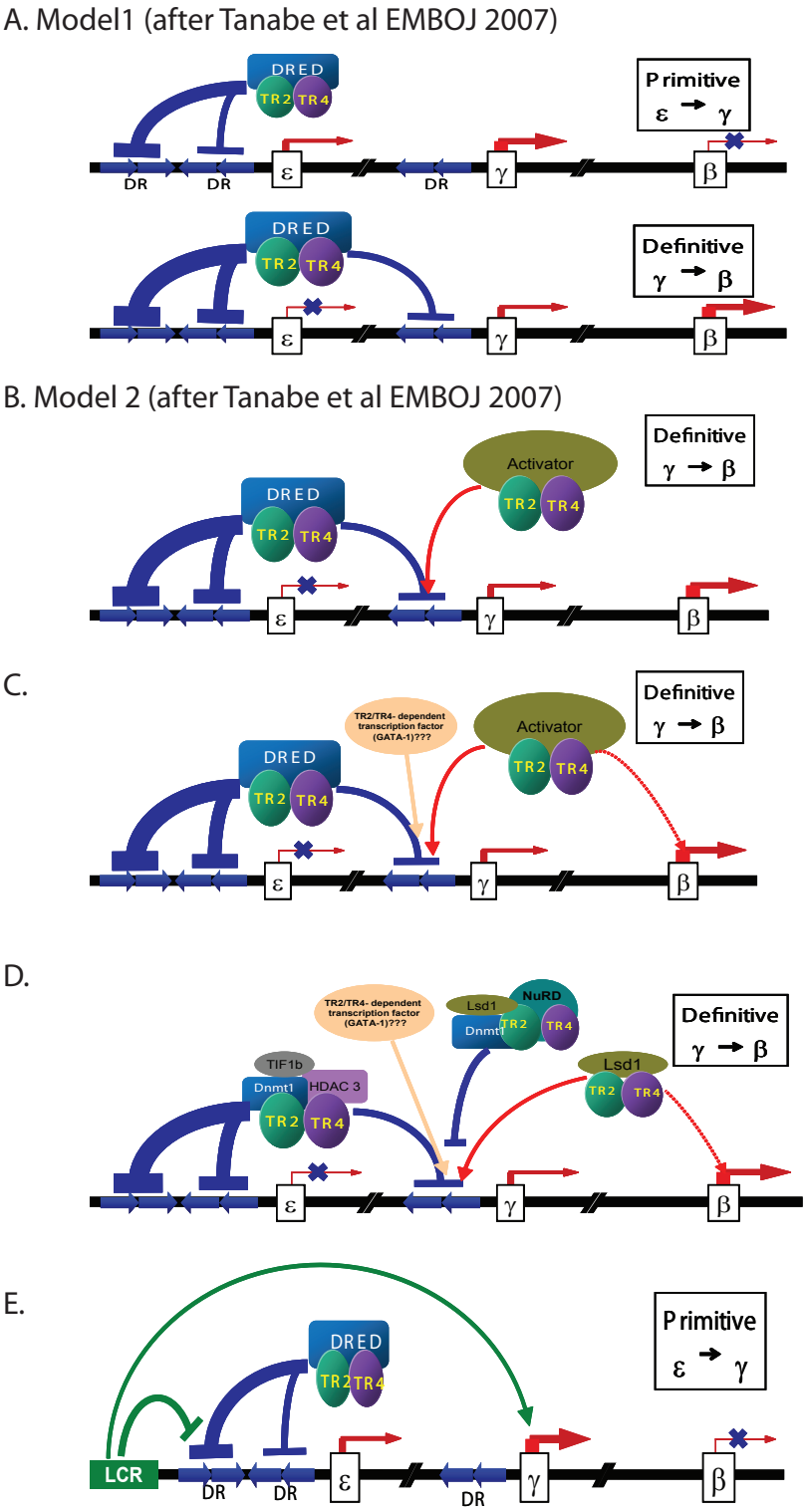


Figure 1. A model of the role of TR2/TR4 complexes in stage specific human globin genes regulation (after (Tanabe, McPhee et al. 2007)). **A**) In primitive erythroid cells TR2/TR4 and co-repressors (DRED) repress ε globin transcription, but have no effect on the γ globin expression. In definitive erythroid cells the activity (abundance?) of DRED increases and allows a gradual repression of the γ-globin gene expression (TR2/TR4 have lower affinity for DR1 in the γ-globin promoter than for DR1 in ε-globin promoter). **B**) Model 2 suggests that TR2/TR4 by interacting with co-factors other than DRED confer an activator function on γ-globin in definitive erythroid cells. **C**) Model 3 suggests that in definitive erythroid cells an activating complex formed by TR2/TR4 acts on the γ and later on the β-globin gene promoters. **D**) It is suggested that TR2/TR4 together with Dnmt1, TIF1β, HDAC3 and possibly other un-known co-factors plays a repressive role in ε globin gene regulation. A complex of TR2/TR4, Lsd1, Dnmt1 and NuRD could possibly function in γ-globin silencing in definitive erythroid cells. In addition, we suggest that TR2/TR4 together with Lsd1 in a distinct complex may bind and possibly activate the adult β-globin. Moreover, GATA-1, the expression of which expression is regulated by TR2/TR4 may bind and regulate γ-globin gene expression. **E**) It is also possible that upon forced TR2/TR4 expression the ε-globin gene is silenced and no longer competes for the LCR. As a result, γ-globin expression is increased.

that despite their high affinity for DR1 elements as measured by *in vitro* band shift assays, the action of TR2/TR4 on globin gene expression (that was shown by loss of function and gain of function studies on transgenic mice) is indirect and it is an effect of other TR2/TR4 sensitive transcription factors binding to the ε and γ (or εy and βh1) globin gene promoters. It may also be possible that the binding of TR2/TR4 requires unidentified but rate limiting co-factors or ligands. We suggest that for further evidence of TR2/TR4 action on globin gene promoters as well as on the essential hematopoietic factors, ChIP-on-chip or ChIP sequencing strategies could be applied.

- **TR2/TR4 interacts with Dnmt1**

Importantly, the results presented in this thesis demonstrate that TR2/TR4 interact with Dnmt1 (Chapter 4). This is a potentially important finding, since DNA methylation has long been associated with transcriptionally silent chromatin. Whether DNA methylation induces transcriptional silencing *per se*, or functions as a stabilizer of silencing remains enigmatic (Rice, Hormaeche et al. 2007). In addition, Dnmt1 was recently reported by an *in vitro* study to bind to an HPFH mutation at a position -3.7 kb from of the γ-globin promoter (Olave, Doneanu et al. 2007).

It was recently demonstrated that the initial activation of embryonic/fetal genes is a result of promoter demethylation as opposed to *de novo* methylation in adults. The differentiation of HSCs (hematopoietic stem cells) derived from either baboon fetal liver or adult bone marrow into mature erythroblasts was accompanied by a progressive decrease in γ-globin promoter methylation and concomitant activation in both fetal liver and adult bone marrow. The difference in γ globin gene demethylation was correlated with the difference of γ globin expression at different developmental stages (Singh, Lavelle et al. 2007). These results suggest the existence of a DNA demethylase activity to counterbalance the repression mediated by Dnmts (Rice, Hormaeche et al. 2007; Singh, Lavelle et al. 2007).

It is therefore essential to assess the functional relevance of the Dnmt1 and TR2/TR4 interaction in erythropoiesis. A Dnmt1 conditional knock-out in the erythroid lineage, which is currently under analysis in the Engel lab, will be a key experiment in this line of investigation.

Recent studies show that 5-azacytidine treatment can induce fetal hemoglobin (HbF) not only by affecting general mechanisms such as DNA methylation but also by a more focused effect (Gore, Baylin et al. 2006; Mabaera, Greene et al. 2007). In studies which demonstrated an enhanced clinical response rate of AML patients that was associated with demethylation of a region in the p15 promoter and acetylation of histones H3 and H4 (Gore, Baylin et al. 2006), the induction of histone acetylation in response to 5-azacytidine was observed before the HDAC inhibitor was administered (Gore, Baylin et al. 2006). In light of these results it would be interesting to test HDAC levels as well as histone acetylation in our 5-azaC experiments in MEL and fetal liver cells. One could hypothesize that the decreasing levels of the factor responsible for recruiting HDAC to the specific site of the promoter may result in decreased HDAC activity on this promoter and possibly further histone acetylation and gene activation.

In addition, G9A mediated methylation in euchromatin is a mark for HP1 and Dnmt1 recruitment, as it was shown that 5-aza-2'-deoxycytidine treatment decreased DNA and/or histone H3K9 methylation levels on the DSC3 and MASPIN tumor suppressor gene promoters, by decreasing levels of Dnmt1 and G9A HMTase levels (Wozniak, Klimecki et al. 2007). This is analogous to the effects observed by us (Chapter 4) and others (Ghoshal, Datta et al. 2005). Therefore we postulate that the effects of 5-azacytidine on TR2/TR4 and Dnmt1 could be extended to include testing the levels of Lsd1 and HDAC proteins, in combination with changes in histone modifications of globin genes.

- **Future directions**

In this thesis we show that using affinity tagging provides a powerful tool for isolating

transcription factor complexes for studying protein interactions and also binding to target genes. Further testing and development of new tags, that can be efficiently used in protein as well as chromatin immunoprecipitations is essential for further investigation of transcription factor interactomes in erythroid (and other) cell types.

We show that erythroid transcription factors can play different functions depending on the interacting partners. A well documented example is GATA-1 (Rodriguez, Bonte et al. 2005), however we cannot exclude the possibility that TR2/TR4 can also function as both repressors and activators depending on the context.

As mentioned above, further continuation of the investigations into TR2/TR4 interacting partners will require further IPs and sequential IPs in order to verify our two complexes hypothesis (Chapter 4). Biotin-tagged Lsd1 (H. Braun, unpublished) and biotin-tagged Dnmt1 (K. Kolodziej, unpublished) MEL cell lines could provide additional useful tools in this analysis. Additional ChIP verification of the binding of the components of particular complexes to globin genes and the G1HE element should be carried out in parallel. We should stress the immediate requirement for *in vivo* verification of the *in vitro* band shift data for TR2/TR4 binding to DR1 elements in the ε and γ globin gene promoters. In addition, the Dnmt1 erythroid-specific conditional knock-out and TIF1β conditional knock-out (both are currently under way in the Engel lab) are going to provide important evidence for the function of these proteins (also in the context of their interaction with TR2/TR4) in globin gene regulation.

Our mass spectrometry results also suggest a possible link between TR2/TR4 and cell cycle proteins. The analysis of TR2/TR4 behavior during different cell stages (by means of GFP tagging or immunofluorescence) in MEL or cultured fetal liver cells, could initiate this (possibly interesting and informative) line of investigation.

Furthermore, for both the TR2/TR4 and GATA-1 projects we envisage developing ChIP-sequencing approaches across the different stages of erythropoiesis. These experiments could be important for designing new therapeutic agents for the treatment of hemoglobinopathies.

References

Cantor, A. B. and S. H. Orkin (2002). "Transcriptional regulation of erythropoiesis: an affair involving multiple partners." Oncogene **21**(21): 3368-76.

Choi, O. R. and J. D. Engel (1988). "Developmental regulation of beta-globin gene switching." Cell **55**(1): 17-26.

Dillon, N., T. Trimborn, et al. (1997). "The effect of distance on long-range chromatin interactions." Mol Cell **1**(1): 131-9.

Enver, T., N. Raich, et al. (1990). "Developmental regulation of human fetal-to-adult globin gene switching in transgenic mice." Nature **344**(6264): 309-13.

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

Franco, P. J., M. Farooqui, et al. (2001). "The orphan nuclear receptor TR2 interacts directly with both class I and class II histone deacetylases." Mol Endocrinol **15**(8): 1318-28.

Ghoshal, K., J. Datta, et al. (2005). 5-Aza-Deoxycytidine Induces Selective Degradation of DNA Methyltransferase 1 by a Proteasomal Pathway That Requires the KEN Box, Bromo-Adjacent Homology Domain, and Nuclear Localization Signal. **25:** 4727-4741.

Gore, S. D., S. Baylin, et al. (2006). "Combined DNA methyltransferase and histone deacetylase inhibition in the treatment of myeloid neoplasms." Cancer Res **66**(12): 6361-9.

Grosveld, F., E. de Boer, et al. (1998). "The dynamics of globin gene expression and position effects." Novartis Found Symp **214**: 67-79; discussion 79-86.

Hanscombe, O., D. Whyatt, et al. (1991). "Importance of globin gene order for correct developmental expression." Genes Dev **5**(8): 1387-94.

Hwang, S. B., J. P. Burbach, et al. (1998). "TR4 orphan receptor crosstalks to chicken ovalbumin upstream protein-transcription factor and thyroid hormone receptor to induce the transcriptional activity of the human immunodeficiency virus type 1 long-terminal repeat." Endocrine **8**(2): 169-75.

Koritschoner, N. P., J. Madruga, et al. (2001). The Nuclear Orphan Receptor TR4 Promotes Proliferation of Myeloid Progenitor Cells. **12:** 563-572.

Lee, C.-H., C. Chinpaisal, et al. (1998). Cloning and Characterization of Mouse RIP140, a Corepressor for Nuclear Orphan Receptor TR2. **18:** 6745-6755.

Lee, Y. F., H. J. Lee, et al. (2002). "Recent advances in the TR2 and TR4 orphan receptors of the nuclear receptor superfamily." J Steroid Biochem Mol Biol **81**(4-5): 291-308.

Mabaera, R., M. R. Greene, et al. (2007). Neither DNA hypomethylation nor changes in the kinetics of erythroid differentiation explain 5-azacytidine's ability to induce human fetal hemoglobin**:** blood-2007-06-093948.

Metzger, E., M. Wissmann, et al. (2005). "LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription." Nature **437**(7057): 436-9.

Moosmann, P., O. Georgiev, et al. (1996). "Transcriptional repression by RING finger protein TIF1 beta that interacts with the KRAB repressor domain of KOX1." Nucleic Acids Res **24**(24): 4859-67.

Nakajima, T., S. Fujino, et al. (2004). TIP27: a novel repressor of the nuclear orphan receptor TAK1/TR4. **32:** 4194-4204.

Olave, I. A., C. Doneanu, et al. (2007). "Purification and identification of proteins that bind to the hereditary persistence of fetal hemoglobin -198 mutation in the gamma-globin gene promoter." J Biol Chem **282**(2): 853-62.

Perillo, B., M. N. Ombra, et al. (2008). "DNA oxidation as triggered by H3K9me2 demethylation drives estrogen-induced gene expression." Science **319**(5860): 202-6.

Rice, K. L., I. Hormaeche, et al. (2007). "Epigenetic regulation of normal and malignant hematopoiesis." Oncogene **26**(47): 6697-714.

Rodriguez, P. (2006). Characterization of Hematopoietic Transcription Factor Complexes in Erythroid Cells. Cell Biology. Rotterdam, Erasmus University Rotterdam**:** 156.

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

Shi, Y. J., C. Matson, et al. (2005). "Regulation of LSD1 histone demethylase activity by its associated factors." Mol Cell **19**(6): 857-64.

Shyr, C.-R., Y.-C. Hu, et al. (2002). Modulation of Estrogen Receptor-mediated Transactivation by Orphan Receptor TR4 in MCF-7 Cells. **277:** 14622-14628.

Singh, M., D. Lavelle, et al. (2007). "The gamma-globin gene promoter progressively demethylates as the hematopoietic stem progenitor cells differentiate along the erythroid lineage in baboon fetal liver and adult bone marrow." Exp Hematol **35**(1): 48-55.

Stamatoyannopoulos, G. (1991). "Human hemoglobin switching." Science **252**(5004): 383.

Tanabe, O., F. Katsuoka, et al. (2002). "An embryonic/fetal beta-type globin gene repressor contains a nuclear receptor TR2/TR4 heterodimer." Embo J **21**(13): 3434-42.

Tanabe, O., D. McPhee, et al. (2007). "Embryonic and fetal beta-globin gene repression by the orphan nuclear receptors, TR2 and TR4." Embo J **26**(9): 2295-306.

Tanabe, O., Y. Shen, et al. (2007). "The TR2 and TR4 orphan nuclear receptors repress Gata1 transcription." Genes Dev **21**(21): 2832-44.

Wang, C.-P., Y.-F. Lee, et al. (2006). "Transactivation of the proximal promoter of human oxytocin gene by TR4 orphan receptor." Biochemical and Biophysical Research Communications **351**(1): 204-208.

Wissmann, M., N. Yin, et al. (2007). "Cooperative demethylation by JMJD2C and LSD1 promotes androgen receptor-dependent gene expression." Nat Cell Biol **9**(3): 347-53.

Wozniak, R. J., W. T. Klimecki, et al. (2007). "5-Aza-2'-deoxycytidine-mediated reductions in G9A histone methyltransferase and histone H3 K9 di-methylation levels are linked to tumor suppressor gene reactivation." Oncogene **26**(1): 77-90.

Yan, Z. H., W. G. Karam, et al. (1998). "Regulation of peroxisome proliferator-activated receptor alpha-induced transactivation by the nuclear orphan receptor TAK1/TR4." J Biol Chem **273**(18): 10948-57.

Yang, Y., X. Wang, et al. (2002). Identification of a novel testicular orphan receptor-4 (TR4) associated protein (TRA16) as repressor for the selective suppression of TR4-mediated transactivation**:** M207116200.

Zhang, Y. and M. L. Dufau (2003). "Dual mechanisms of regulation of transcription of luteinizing hormone receptor gene by nuclear orphan receptors and histone deacetylase complexes." J Steroid Biochem Mol Biol **85**(2-5): 401-14.

Summary

Most proteins mediate their function within macromolecular networks. Studies in model organisms suggest that such complex protein networks have topological and dynamic properties that reflect biological phenomena. Thus, an understanding of biological mechanisms and disease processes demands a 'systems' approach that goes beyond one-at-a time studies of single components to more global analyses of the structure, function and dynamics of the networks in which proteins function (Cusick, Klitgord et al. 2005).

In Chapter 2 of this thesis we present optimal conditions for the efficient application of in vivo biotinylation tagging for studying protein-protein interactions of transcription factors in erythroid cells (de Boer, Rodriguez et al. 2003). This chapter concentrates mostly on the technical aspects of the biotin-tagged transcription factor complex purification and analysis by direct binding to streptavidin beads using GATA-1 complexes from mouse erythroleukemic (MEL) cells as an example.

Chapter 3 presents a more detailed description of the approach we undertook in order to map the network of interactions mediated by GATA-1, an essential transcription factor for erythroid development (Rodriguez, Bonte et al. 2005). Our data indicate that GATA-1 can function as a repressor and activator of its target genes depending on target gene context and the partners that is interacting with. Thus, GATA-1 interacts with the large repressive MeCP1 chromatin remodeling/ histone deacetylase complex, with the Snf2h/ACF1 chromatin remodeling complex and with hematopoietic transcription factors Gfi1b, FOG-1 and the Ldb1/TAL1 complex. Significantly, these GATA-1 protein interactions are distinct. In addition, FOG-1 mediates GATA-1 complexes with the MeCP1 complex. Our data suggest that the GATA-1 complexes with FOG-1/MeCP1 and with Gfi1b confer repressive function potentially on distinct subsets of genes. Other GATA-1 interactions; with FOG-1 (alone) and the Ldb1/TAL1 complex are involved in the transcriptional activation of erythroid genes. The interaction of GATA-1 with Snf2h/ACF1 complex can have either repressive or activating function, however we do not have enough evidence at this stage to clarify this.

In Chapter 4 we show the application of in vivo biotinylation tagging to resolve protein complexes involved in globin gene developmental regulation. TR2 and TR4 nuclear orphan receptors have been previously reported to be part of the DRED complex (Tanabe, Katsuoka et al. 2002), which is implicated in ε and γ globin gene silencing (Tanabe, McPhee et al. 2007). Our data demonstrate that TR2 and TR4 interact with Dnmt1, NuRD components, HDAC3, TIF1β and Lsd1 in MEL cells. We hypothesize that TR2/TR4 may form at least two distinct complexes, one involving Dnmt1, TIF1β and HDAC3 and the other involving Dnmt1, Lsd1 and subset of NuRD components. Identifying an interaction between TR2/TR4 and Dnmt1 is particularly interesting, since methylation has been long been implicated in globin gene regulation. In addition, we could show a specific decrease of Dnmt1 and TR2/TR4 protein levels upon treatment of erythroid cells with the potent DNA methyltransferases inhibitor 5-azacytidine, as has been previously described for Dnmt1(Ghoshal, Datta et al. 2005). Finally, we suggest that TR2/TR4 by binding to a novel DR4 element in the β-globin promoter may be involved in β-globin gene activation in the adult stage.

Chapter 5 shows the application of affinity tagging, with a focus on biotin-tagging, in chromatin immunoprecipitation (ChIP) experiments. We tested several conditions in order to optimize ChIP assays so as to improve specific enrichments and decrease non-specific background. In addition we highlight the requirement of testing additional tags as we show that, next to biotin-tagging, V5 tagging can also be efficiently applied in ChIP assays. These observations have important implications since the study of transcription factor networks and their target genes will require using several, possibly equally efficient, affinity tags.

In summary, this thesis demonstrated the utility of affinity tags, and specifically of vivo biotinylation tagging for the isolation transcription factors complexes in erythroid cells and in identifying in vivo bound transcription factor gene targets. We believe that this approach offers a uniform and powerful platform for the further in-depth studies of protein and gene target networks in erythropoiesis, and effectively, any other cellular system.

Samenvatting

De meeste eiwitten verrichten hun functie binnen een complex netwerk van verschillende macromoleculen. Onderzoek in modelorganismen suggereert dat complexe eiwitnetwerken topologische en dynamische eigenschappen hebben die biologische fenomenen weergeven. Om biologische processen en ziektebeelden die daarmee samenhangen te kunnen begrijpen is een "systems" aanpak noodzakelijk dat verder gaat dan de studie van enkele onderdelen maar een meer globale analyse is van de structuur, functie en dynamiek van het netwerk waarin eiwitten functioneren .(Cusick, Klitgord et al. 2005)

In Hoofdstuk 2 van dit proefschrift tonen we de optimale condities voor het gebruik van " in vivo biotinylation tagging" wat een zeer efficiënte methode is om eiwit-eiwit interacties in erythroide cellen te bestuderen.(de Boer, Rodriguez et al. 2003)
Dit hoofdstuk gaat hoofdzakelijk over technische aanpak van de analyse en zuivering van de complexen van transcriptiefactoren met een biotinetag door directe binding aan streptavidin bollen. Hierbij zijn de GATA-1 complexen van muis erythroleukemie (MEL) cellen als voorbeeld gebruikt

Hoofdstuk 3 bevat een gedetailleerde beschrijving van manier waarop we het netwerk van interacties met GATA-1, een transcriptiefactor die essentieel is voor erythroide ontwikkeling, in kaart gebracht hebben (Rodriguez, Bonte et al 2005) Onze data laten zien dat GATA-1 een functie als activator of repressor van zijn targetgenen kan hebben, dit hangt af van de context van het targetgen en de partners waarmee GATA-1 samenwerkt. Zo heeft GATA-1 een interactie met het grote repressieve MeCP1 chromatine hermodulering/histon deacetylase complex , met het Snf2h/ACF1 chromatine hermodulering complex , met de hematopoëtische transcriptiefactoren Gfi1b, FOG-1 en het Ldb1/TAL-1 complex. Dit zijn duidelijk afzonderlijke eiwitinteracties.
Bovendien helpt FOG-1 de samenwerking van GATA-1 complexen met het MeCP1 complex. Onze data laten zien dat GATA-1 complexen met FOG-1/MeCP1 en met Gfi1b
hun onderdrukkende rol op een bepaalde groep genen uitoefenen. Andere GATA-1 interacties; met FOG-1 alleen en met het Ldb1/TAL1 complex zijn betrokken bij de activering van de transcriptie van erythroide genen . De interactie van GATA-1 met het Snf2h/ACF1 complex kan een activerende of repressieve functie hebben hoewel we op dit moment nog niet genoeg data hebben om dit te verklaren.

In Hoofstuk 4 laten we de toepassing zien van in vivo biotinylering "tagging"om eiwitcomplexen te ontrafelen die betrokken zijn bij de regulatie van de ontwikkeling van globine genen. TR2 en TR4 zijn nucleaire "orphan receptors"en al eerder is getoond dat zij deel uitmaken van het DRED complex (Tanabe, Katsuoka et al 2007) wat betrokken is bij het "silencen"van het epsilon en gamma globine gen (Tanabe, McPhee et al 2007)
Onze resultaten laten zien dat TR2 en TR4 een interactie hebben met Dnmt1, onderdelen van NuRD, HDAC3, TIF1 beta en Lsd1 in MEL cellen. We veronderstellen dat TR2/TR4 tenminste twee afzonderlijke complexen kan vormen, de een met Dnmt!, TIF1 beta en HDAC3, de ander met Dnmt1, Lsd1 en een aantal onderdelen van NuRD. Het vaststellen van een interactie van TR2/TR4 met Dnmt1 in het bijzonder is interessant omdat methylatie al lang een rol schijnt te spelen bij de regulatie van globine genen. Bovendien tonen we een specifieke afname van de hoeveelheid Dnmt1 en TR2/TR4 na behandeling van erythroide cellen met de krachtige DNA methyltransferase remmer 5-azacytide, zoals al eerder beschreven is voor Dnmt1(Ghoshal, Datta et al 2005) Tenslotte suggereren we dat TR2/TR4 door binding aan een nieuw DR4 element in de beta globine promoter betrokken kan zijn bij activering van het beta globin gen in het volwassen stadium.

Hoofstuk 5 bevat de toepassing van affiniteit"tagging"in het bijzonder biotine"tagging", bij chromatine immunoprecipitatie (ChIP) proeven.

We hebben verschillende condities getest om de ChIP proeven te optimaliseren om de specifieke verrijking te verbeteren en de niet-specifieke achtergrond te verlagen. Daarbij leggen we de nadruk op de noodzaak van het testen van verschillende “tags”want zoals we laten zien kan, behalve biotine”tagging” , ook V5 “tagging” efficiënt in ChIP proeven gebruikt worden. Dit is een belangrijke waarneming omdat voor de studie van transcriptiefactor netwerken en hun targetgenen het nodig zal zijn om verschillende, even doelmatige affiniteit”tags”te gebruiken.

Samengevat demonstreert dit proefschrift het nut van affiniteit”tags”, in het bijzonder dat van in vivo biotinylatie “tagging”voor de isolatie van transcriptiefactor complexen in erythroide cellen en de identificatie van in vivo gebonden gen-targets.

We geloven dat deze aanpak een breed en krachtig platvorm is voor de verdere bestudering van eiwitten en gen target netwerken in erythropoese en effectief in elk ander celsysteem.

Vertaling: Ernie de Boer

Curriculum Vitae

Name: Katarzyna Ewa Kołodziej

Born: 14 April 1977 in Wrocław, Poland

1992- 1996 XIV LO Wrocław, Poland

Scientific research experience:

1996 – 1999 Bachelor in Biotechnology University of Wroclaw, Poland
Comparison of antioxidant enzymes activity in non-small cell lung cancer and healthy tissue. (Prof. A. Dżugaj)

1999 to 2001 Master in Biotechnology University of Wroclaw, Poland
Activity of antioxidant enzymes in medullary thyroid carcinoma cell line. (Prof. A. Dżugaj)

2000- 2001 Socrates/Erasmus studentship
Department of Plant Physiology and Biolchemistry
University of Antwerp (UIA), Belgium
Cell cycle kinases inhibitors in BY2 cells. (Dr. L. Roef)

2001- 2002 Assistant in research (Prof. H. van Onckelen)
Department of Plant Physiology and Biochemistry
University of Antwerp (UIA), Belgium

2003- 2007 PhD student (Dr. J. Strouboulis, Prof. F. Grosveld)
Department of Cell Biology, ErasmusMC

Feb-May 2007 EMBO short term fellowship (Prof. J.D. Engel)
Department of Cellular and Developmental Biology
University of Michigan

January 2008 - Post-doc (Dr. B. van Steensel)
Department of Molecular Biology NKI-AVL

Publications:

1. Rodriguez P., Bonte E., Krijgsveld J., **Kolodziej K.E**., Guyot B., Heck AJ., Vyas P., de Boer E., Grosveld F., Strouboulis J. (2005) GATA-1 forms distinct activating and repressive complexes in erythroid cells. *EMBO J*. 24(13) 2354-66

2. Grosveld F., Rodriguez P., Meier N., Krpic S., Poufarzad F., Papadopoulos P., **Kolodziej K.E.**, Patrinos G., Hostert A. and Strouboulis J. (2005) Isolation and characterization of hematopoietic transcription factor complexes by in vivo biotinyation tagging and mass spectrometry. *Ann. N.Y.Acad.Sci*.,

3. Rodriguez P., Braun H., **Kolodziej K.E**., de Boer E., Bonte E., Grosveld F., Philipsen S., Strouboulis J. Isolation of transcription factor complexes by in vivo biotinylation tagging and direct binding to straptavidin beads. (2005) *Gene Mapping, Discovery and Expression (Methods and Protocols). Humana Press*

Acknowledgments

I would like to start with thanking Professor Frank Grosveld for giving me the opportunity to learn and obtain a PhD in the Department of Cell Biology. The stay in your lab was a real eye-opener on many aspects of science. Thank you for generating an open and creative atmosphere and for constructive discussions, especially during the very important, final stage of my PhD studies.

I especially want to acknowledge to John Strouboulis for his help and support throughout the course of my PhD. I immensely enjoyed all the time we spent "sightseeing" the bars and restaurants of Rotterdam (and the Red Hawk in Ann Arbor!).

I would like to thank Professor Doug Engel for a fruitful collaboration, and for accepting the invitation to be part of my defense committee. I am very grateful for your support, and that you allowed me to stay and learn in your lab in Ann Arbor, which was a fantastic experience. Here I would like to thank members of Engel lab for their hospitality. In particular I want to express my gratitude to Dr Osamu Tanabe for his help in TR2/TR4 project.

I would like to thank the members of the committee, Sjaak Philipsen, Dies Meijer, Marieke von Lindern, Peter Verrijzer and Raymond Poot for reading my thesis. Special thanks to Sjaak Philipsen and Dies Meijer for their help with the final corrections to the manuscript.

Thank you Raymond, for a great continuation of the fantastic "706- atmosphere", and for exchanging many good scientific and many more, even better, non-scientific ideas.

I would like to express my gratitude to Jeroen Demmers and Karel Bezstarosti for all the mass spec analyses.

Thank you Marike and Jasperina for all the practical help. Without you, any PhD student, including me, would be completely lost.

Big, big THANK YOU to all the past and present lab 706 members: Ernie, Edgar, Patrick, Marianne, Vesna, Debbie, Tiago, Javier, Amy, Jennifer, Aga, Alireza, Erik, Tim, Boet , Anouk and also people from the neighboring labs (702,708, 710). Thank you for great atmosphere.

I would particularly like to thank Patrick, whom I shadowed for the first three months of my PhD and who shared with me the bench, desk, experience (especially in western blotting and pull-downs) and sense of humor. It was never a boring time with you, merci!

Edgar- thank you very much for all your help with biochemistry issues and for the crash course in cloning!

Thank you Marianne for all the help with cloning (Edgar's crash course covered only the basics, Marianne you are a PRO in cloning!) and the great time we had sharing the bench. Also thank you for helping with the drinks on the 26th.

Debbie and Tiago, thank you for all the great times we had (especially when going for "just one drink"), for always extremely inspiring scientific discussions in the office, and also for very interesting gossip-exchange sessions. Thank you for helping me with organizing this event.

My acknowledgements of my paranimfs I will start with Ernie, who bravely accepted yet another request to be a paranimf and to stand in front of the committee by my side. Thank you for all the super-useful comments and advice for experiments (and most important controls!). I also enjoyed our girls-night outs and outings to the North Sea Jazz festival. You have been a great friend to me!

Thank you Katrin, for all the great evenings we spent together and fantastic holidays (when are we going back to Venice?). Special thanks for your help with editing this text and thank you for coming to be there for me on this big day. Here I would also like also to thank Colin, for great exchange of many wild and abstract ideas and for always being ready to "sample" some wine or beer with me.

Dorotko, bardzo ci serdecznie dziękuję za wszystkie rady i porady, za wysłuchiwanie moich utyskiwań, za pomoc w mikroskopowaniu i składaniu tej książki. Chcę również podziękować

Jarkowi za nieocenioną pomoc przy prezprowadzkach! Buziaczki dla Zosi.

Miyata, thank you for introducing me to the North Sea festival! And for great discussions in the office!

Gusia and Arnie, thank you for the fun evenings in Sijf and Sapporo! Gusiu, dziękuję za wszystkie przepisy na mazurki, pysznego murzynka i za wszystkie pogaduchy przy kawce, winku i nad sushi.

Polish team: Magda, Monika, Gosia, Marcin, and Andrzej, dziękuję za bardzo fajniaste wypady na piwko i plotki w coffee corner.

I would like to express my gratitude to Bas van Steensel for taking me on board in his lab and for creating a great opportunity for me to develop further scientifically. Thanks to all the members of my new lab: Maartje, Joke, Wendy, Inez, Daan, Aleksey, Ulrich, Jop, Ludo, and Guillaume, for your hospitality and the really friendly and open atmosphere. I am looking forward to the great scientific adventure we are about to start together.

I would like to thank the Verhoeven family: Hendrik, Simonne, Ann, Karl, Els, Wim Chris, Jolien, Naomi, Marlijn, Steffi, Charlotte, Matthias and Jonas – Bedankt voor de gastvrijheid en voor het delen van die bijzondere momenten: verjaardagen, feestdagen en de vele andere familiegebeurtenissen - I geniet er altijd van.

Lastly I would like to express my deep gratitude to Luc. Without your support, help and optimism I could not have achieved any of it. For all those years, thank you!

Kochani rodzice, bardzo dziękuję za wsparcie przez te wszystkie lata. Wiem, że mogę na was liczyc w każdej sytuacji. Bez Was nie byłoby mnie tu dzisiaj.

There are not enough pages here to accommodate all the names of the people who helped me during my time in ErasmusMC. If I missed some names, forgive me.

To you all, THANK YOU FOR A GREAT TIME IN ROTTERDAM

Kasia