Implication of GATA1 and GATA 1 short in ribosomal protein gene regulation in erythroid cells

Master Thesis

Alexandra Vatikioti

Supervised by:

Dr. John Strouboulis

Acknowledgements

I would first like to thank my thesis advisor Dr. John Strouboulis for the continuous support of my study and research, for his patience, motivation and knowledge. The door to Dr. Strouboulis office was always open whenever I ran into a trouble spot or had a question about my research or writing. His guidance helped me in all the time of research and writing of this thesis.

My acknowledgements to my committee Prof. Charalampos Pontikoglou and Prof. George Garinis for their useful suggestions and comments.

I would like to thank Marina Ioannou in particular for the valuable help and support throughout the thesis. Our collaboration has been excellent and has played a role of great importance in the smooth completion of my thesis. Many thanks to my best friend in the lab, Dimitra Vlahocosta, who has made this journey a great fun. Also thanks to the rest members of the lab, for creating a great environment, Elena, John-Mario and Grigoris.

Finally, i would like to thank my loved ones, who have supported me throughout entire process, both by keeping me harmonious and helping me putting pieces together. Many thanks to my beloved parents, Efi and Kiriakos, and my sister, Maria-Grazia, for their love and support. Without them nothing would have come true. I would like to thank Thomas, who was beside me the past couple of years, for his patience and emotional support. Last but not least, I would like to thank Vivian Mara for her support from the first day to the last of my master. Her help and guidance were more than valuable and determinant during moments that important decisions should be taken.

Many thanks!

Περίληψη:

Η Αναιμία Diamond Blackfan (DBA) ανήκει στα σύνδρομα κληρονομούμενης μυελικής ανεπάρκειας. Η DBA είναι μια σπάνια ριβοσωμοπάθεια που χαρακτηρίζεται από μακροκυτταρική αναιμία, δικτυοευθροπενία, απουσία ή υποπλασία των πρόδρομων μορφών της ερυθράς σειράς. Στο 60-70% των περιπτώσεων ανιχνεύονται μεταλλάξεις σε γονίδια που κωδικοποιούν ριβοσωμικές πρωτεΐνες. Πρόσφατη ερευνητική εργασία έχει δείξει ότι μεταλλαγές σε ριβοσωμικά γονίδια μπορούν να οδηγήσουν στην μη επαρκή μετάφραση της full- length ισομορφής του GATA-1 που αποτελεί βασικό μεταγραφικό παράγοντα κατά την διαφοροποίηση των ερυθρών κυττάρων. Αυτό έχει σαν αποτέλεσμα την έκφραση της ισομορφής του GATA-1 (GATA-1 short) από την οποία λείπει το αμινοτελικό άκρο και η οποία αδυνατεί να υποστηρίξει το πρόγραμμα της τελικής ερυθροειδικής διαφοροποίησης. Πρόσφατη δουλειά έχει ακόμη δείξει ότι ο GATA-1 προσδένεται σε πολλά ριβοσωμικά γονίδια στα ερυθροειδικά κύτταρα στον άνθρωπο και το ποντίκι συμπεριλαμβανομένων εκείνων που εμφανίζουν μεταλλαγές στην DBA. Λαμβάνοντας υπ όψη τα παραπάνω προτείνεται ένα μοντέλο ανατροφοδότησης στα πλαίσια του οποίου ο GATA-1 προσδένεται στα ριβοσωμικά γονίδια και ρυθμίζει την έκφραση τους ώστε να ενεργοποιήσει ή και να διατηρήσει την δική του έκφραση στα ερυθροειδικά κύτταρα. Η συγκεκριμένη μεταπτυχιακή εργασία επιχειρεί στα πλαίσια της παραπάνω υπόθεσης να παρέχει αποτελέσματα που να αποδεικνύουν ότι ο GATA-1 πέραν του γεγονότος ότι προσδένεται στα ριβοσωμικά γονίδια που εμπλέκονται στην DBA, ρυθμίζει και την έκφραση τους.

Για το σκοπό αυτό στη συγκεκριμένη εργασία επιλέχθησαν να μελετηθούν δυο ριβοσωμικά γονίδια (RPL11 και RPS19) τα οποία εμπλέκονται στην DBA. Είναι γνωστό από ChiP δεδομένα ότι ο GATA-1 προσδένεται κοντά στην περιοχή του εκκινητή και στα δύο γονίδια. Μάλιστα στο RPS19 υπάρχουν δύο θέσεις πρόσδεσης ενώ στο RPL11 μια θέση πρόσδεσης για τον GATA-1. Με τη χρήση κατάλληλων εργαλείων βιοπληροφορικής ανάλυσης εντοπίσθηκαν οι πιθανές αλληλουχίες πρόσδεσης του GATA-1 για καθένα από τα δύο γονίδια. Έπειτα ελέγχθηκε εάν ο GATA-1 προσδένεται in vitro στις αλληλουχίες αυτές και αν ρυθμίζει την έκφραση των δυο ριβοσωμικών γονιδίων.

Σύμφωνα με τα αποτελέσματα ο GATA-1 προσδένεται *in vitro* στα RPL11 και RPS19, ενώ η έκφραση τους φαίνεται να μειώνεται παρουσία του GATA-1. Η ύπαρξη όμως υψηλού μη ειδικού σήματος στο δείγμα ελέγχου στα πειράματα μέτρησης ενεργότητας της λουσιφεράσης υποδεικνύει την ανάγκη για επανάληψη του πειράματος σε ένα άλλο σύστημα που θα επιτρέπει την λήψη πιο ξεκάθαρων αποτελεσμάτων. Τα πειράματα αυτά αποτελούν ενώ πρώτο βήμα για να διασαφηνισθεί εάν ο GATA-1 ρυθμίζει την έκφραση των ριβοσωμικών γονιδίων και να διαλευκανθεί ο μοριακός μηχανισμός μέσα στα πλαίσια του οποίου λαμβάνει χώρα αυτή η ρύθμιση.

Table of Contents

1. Introduction

1.1. Ribosomes

1.1.1. Ribosomal structure and origin

Ribosomes are small organelles with a major role in the adequate synthesis of all the proteins of a cell. The term was first proposed by Richard Roberts[1] while the structure was originally observed and described by Palade[2] using electron microscopy. They were characterized as small particles or organelles (100-150 Å) of spherical shape and high density.

Their main structure consists of ribonucleoprotein complexes which are made up of 60% of ribosomal RNAs (rRNAs) and 40% ribosomal proteins (RPs). In their active form, the ribosomes are organized into two subunits: the small and the large subunit which are conserved but also unique between the different domains of life. Based on this, the ribosome of bacteria and archea is known as the 70S ribosome, with the 30S small subunit and the 50S large subunit. The small 30S subunit contains 21 different proteins (S1 to S21) and one RNA molecule, the 16S. The large 50S subunit contains 34 different proteins (L1 to L34) and two RNA molecules, 23S and 5S. In the prokaryotic ribosome all components, except for L7 and L12 proteins, are single copy proteins. L7 and L12 form a tetramer and constitute the acetylated and methylated form of the same protein. In the eukaryotes the 80S ribosome consists of a 40S small subunit and a large 60S. The small ribosomal subunit contains 18S rRNA molecule and 33 ribosomal proteins. The large ribosomal subunit, with the peptidyl transferase center that catalyzes peptide bond formation, is formed by three rRNA molecules, 28S rRNA, 5S rRNA and 5.8S rRNA and approximately 49 ribosomal proteins.

Despite the differences between prokaryotic and eukaryotic ribosomes, the ribosomal "core", which includes the main functional center of the ribosomes, is conserved in all organisms [3,4]. In past years, structural information on the small and large subunits has increased considerably. More specifically, Steitz, Ramakrishnan and Yonath, elucidated the structure of the prokaryotic ribosome, resulting in their awarding of the Nobel Prize in Chemistry in 2009. Following this important scientific discovery, the structure of the eukaryotic and mitochondrial ribosomes was also characterized [4,5]. These studies have led to the clarification and understanding of the functional structure of the ribosomes and the way in which the translation machine is organized and operates.

1.1.2. Ribosomal biogenesis and regulation

Ribosome biogenesis is an extremely complex process that requires large amounts of energy. In eukaryotic cells ribosomal biogenesis requires the coordinated activity of all three polymerases and the coordinated participation of many factors in the assembly of its structure. The process begins in the nucleolus where the 18S, 28S and 5.8S rRNAs are transcribed by RNA polymerase I as a polycistronic transcript. The 5S rRNA gene is located in another region of the genome and transcribed by RNA polymerase III outside the nucleolus, in the nucleoplasm.

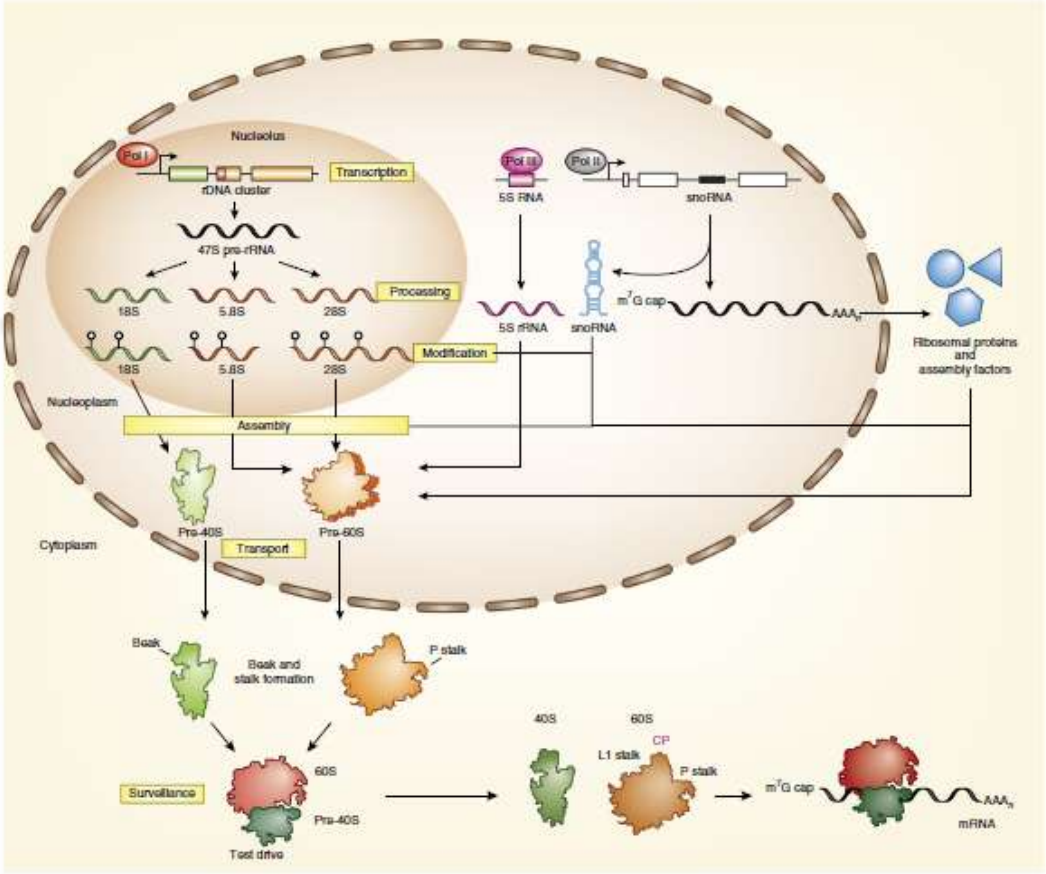

Figure 1 Ribosome biogenesis in eukaryotes. Ribosome biogenesis involves six important steps (yellow frames): (i) transcription of components (rRNAs, ribosomal protein coding (RPs) and assembly factors (AFs) and snoRNAs, (ii) processing (cleavage of pre-rRNAs), (iii) modification of pre-RNA, RPs and AFs, iv) assembly, v) transport (nuclear insertion of RPs and AFs, extraction of pre-ribosomes in the cytoplasm); and vi) quality control. Three of the four rRNAs are transcribed in the nucleus from Pol I as a long 47S precursor (47S pre-rRNA), which is then processed and modified to give the 18S, 5.8S and 28S rRNAs which are assembled into the pre-40S (green) and pre-60S (orange) ribosomal subunits. The 5S rRNA (pink) is transcribed by Pol III in the nucleoplasm and is assembled into 60S subunits, forming the central protuberance (CP). 80 RPs, more than 250 AFs and 200 snoRNAs are transcribedby Pol II. The proteins are synthesized in the cytoplasm and reintroduced into the nucleus for assembly[6].

Genes encoding 80 RPs and over 250 aggregation factors (AFs) are translated from RNA polymerase II, translated into the cytoplasm and enter the nucleus to participate in the assembly of the ribosomal subunits[6]. Upon entry of these proteins into the nucleus, the pre-rRNAs undergo complex modification processes. Such processes include folding, the binding of rRNAs to ribosomal proteins and the removal of spacer sequences. The pre-ribosomes are then released from the nucleus into the nucleoplasm and the process is completed by their export into the cytoplasm.

The ribosomal subunits in their precursor form are transferred from the nucleus to the cytoplasm through the nuclear pores, undergo some final modifications and are then organized into a functional ribosome. A wide arrange of assembly factors are involved in this process. These factors have ATPase, GTPase, helicase, kinase or nuclease activity, and play an important role in the various stages of biogenesis from the nucleolus to the cytoplasm. Although the "core" of the ribosomal biogenesis process in eukaryotes is conserved, there are some differences compared to single cell eykaryotes such as yeast. These differences appear to be due to additional factors identified in higher eukaryotes and involved in the process, thus adding additional steps or paths to the biogenesis process [7,8] .

Due to the high energetic cost that is required for the ribosomal biogenesis, strong regulatory steps exist that adjust according to the cellular needs and the cellular micro-environmental conditions. Specifically, as shown in bacteria, there is a linear relationship between growth rate and ribosome cellular concentrations [9]. The regulation of transcription of rRNAs plays a very important role in the process of biogenesis in prokaryotic cells. The genes of ribosomal proteins are organized into operons that include a protein which acts as a translational repressor targeting the polycistronic mRNA. While the rRNA synthesis continues, suppressor proteins are incorporated into ribosomes and the operons continue to express. When the composition of rRNA decreases, the repressor concentration increases and operons are suppressed[9]. In eukaryotic organisms the corresponding mechanism differs radically. The mechanisms that control the transcription of rRNAs as well as the translation of ribosomal proteins appear to be completely independent and differ in comparison to those occurring in prokaryotes: the genes of ribosomal proteins are scattered throughout chromosomes, as opposed to prokaryotes in which they are organized into operons. Furthermore, in addition to RNA Pol I, three other factors are required: Upstream binding factor (UBF), SL-1 (Selective complex), TIF-A. However, the transcription-translation? coordination mechanism is not yet fully understood[9].

1.2. Ribosomal proteins

The ribosomal proteins are an integral part of the ribosome with an important role. In both subunits there exist approximately 80 different proteins, which are encoded by mRNAs synthesized in the nucleus by RNA polymerase II. The mRNAs are transported to the cytoplasm where they are translated and the proteins produced are transported to the nucleus and eventually to the nucleolus via the nuclear pores, where they are linked to the

newly synthesized rRNAs. Mammalian ribosomal proteins are 164 aa in length with a molecular weight of 18.5KDa. Most proteins are rich in lysine and arginine, with an average isoelectric point of 11.05. They also have specific structural features. In addition to the nuclear localization signal, some of them present with repeated motifs of 3-8 aa. Additional features include zinc finger structures, leucine zippers[10] and KH motifs to which RNA binds.

1.2.1. Biogenesis of ribosomal proteins in eukaryotes

Although the mechanism of coordinated regulation in prokaryotes seems simple, the mechanism in higher eukaryotes such as mammals is significantly more complex and much less understood. Whereas they are coordinately expressed, RP genes in higher eukaryotes are scattered throughout the genome without forming any operon structure [11]. The presence of specific expression patterns under different conditions was examined through the identification of motifs (transcription factor binding sites) but no pattern was found to be present in all the genes exhibiting a common expression profile. Following the sequencing of the human genome, new attempts have been made to identify common motifs involved in regulating RPs in humans, but no common pattern was found [12]. By contrast, ChIP-chip experiments in *Saccharomyces cerevisiae* have shown that Rap1 [13,14] and Fhl1[14] bind to the promoters of almost all of the RP genes. Specifically, the upstream region of the vast majority of the RP genes in *S. cerevisiae* contains one or two Rap1 binding sites. However, Rap1 is responsible for the transcription of many other genes, so it cannot explain the coordinated regulation of the RP genes. It is now known that the factors Fhl1 and Ifh1 play a key role in regulating RP gene transcription [15–19]. When the transcription of the RP genes becomes activated, Rap1 binds to the promoter and recruits Fhl1, which in turn recruits Ifh1[17–19]. When transcription of RP genes is suppressed, only Rap1 and Fhl1 are recruited to the RP gene promoters [18–20]. There are three possible models for determining how activation and suppression are achieved. According to the first, Crf1, which is a cofactor of Fhl1, competes with Ifh1 for binding to Fhl1 at the promoters of the RP genes and thereby transcription is activated or suppressed [21]. Hmo1, according to the second model, can also determine the activation or suppression as it is necessary both for assembling complexes of transcription factors containing Fhl1 and Ifh1 at the promoters of RP and for the maximum transcription of rRNA genes from Poll [15]. The third possible model concerns the CURI complex, which is likely to contribute to the switch as long as Ifh1 is absent from the promoters of the PR genes and becomes part of the CURI complex when the transcription is repressed [18]. Other proteins such as Sfp1 also appear to be involved in the transcriptional control of the RP genes [22,23].

Li *et al*., demonstrated that the regulatory DNA motifs found in yeast are also present in higher eukaryotes. The existence of different common motifs was identified, with at least one of them occurring in all RP genes of the examined species[11]. These motifs are found to be conserved among nearby species (human, mouse, rat), but not in phylogenetically distant species. For mammals, the most common pattern is found in the intronic region of the ribosomal genes, mostly in the first intron, whereas the patterns found in other species are

located 1000 bp upstream of the transcription initiation sites of the RP genes. This motif has its two ends conserved and appears to play an important role in the RPL30 promoter function, which has also been verified experimentally [11,24].

With regards to translational control, it is known that ribosomal mRNAs present with certain structural features at their 5' end which play a very important role in the regulation of translation. In particular, ribosomal mRNAs have a C residue at the cap site, followed by a continuous (non-interrupted) sequence of 4 to 15 pyrimidines (chapter) consisting of a similar ratio of C and U residues (5' TOP motif) and a CG rich region immediately after the 5'TOP motif. The 5'TOP motif, as well as its adjacent sequence, exhibit high conservation in specific members of the mammalian group [25]. The translational regulation of the 5' TOP ribosomal transcripts depends on the mTOR signaling pathway, growth factors, the presence of nutrients and oxygen as well as mitogenic activity [25].

Finally, another mechanism of regulation is the balanced degradation of unassembled ribosomal proteins in the ribosome, which are free in the nucleoplasm[26].

1.2.2. Role of ribosomal proteins

Ribosomal proteins are necessary for various stages of ribosome biogenesis as well as for the translation process. They take part in the processing of rRNAs, assembling proteins, replicating RNA, transporting precursor subunits, stabilizing subunit structures, and interacting with other factors required for ribosomal biogenesis or translation[27,28].

In addition to ribosomal biosynthesis, ribosomal proteins play an important role in extraribosomal functions of both eukaryotic and prokaryotic organisms. They have also been associated with diseases such as cancer. These functions include transcription, translation, processing of RNA, DNA repair[29]. They are also involved in cell growth and proliferation, apoptosis, and as for cancer they are related to the proliferation, metastasis and regulation of tumor suppressor genes. They have also been associated with Turner syndrome, Noonan syndrome and ribosomal diseases, collectively known as ribosomopathies[30]. An example is the S13 protein that prevents splicing of its own mRNA[31]. Another example is L13 for which it has been found that after treatment of U937 cells with IFNg a kinase cascade is activated, resulting in L13 removing from the ribosome and becoming part of the GAIT complex that binds to the ceruloplasmin mRNA and does not allow its translation [32].

1.2.3. Ribosomal protein S19 (RPS19)

The S19 ribosomal protein belongs to the family of S19E ribosomal proteins and is located in the cytoplasm. It participates in the formation of the small ribosomal subunit. The RPS19 gene is located on chromosome 19 in humans. The protein consists of 126 amino acids, has a molecular weight of 16kDa and an isoelectric point of 10.3.

Regarding its function it has been shown that its decreased expression in human erythroleukemic cells may lead to the inappropriate maturation of the 40S ribosomal subunit [33]. Mutations in the RPS19 gene were the first identified to be involved in Diamond Blackfan Anemia (DBA; OMIM 105650) [34].

Since then, 72% of DBA patients have been found to carry heterozygous pathogenic mutations in RP genes [35]. It is the most commonly mutated ribosomal protein found in ~ 25% of patients. Specifically, more than 120 mutations of the gene have been identified in patients presenting with the disease [36,37]. These mutations are believed to affect the stability and /or functionality of the protein. *RPS19* $^{-/-}$ homozygous mice are embryonic lethal. On the other hand, heterozygous mice are either normal or show mild macrocytic anemia depending on the RPS19 mutation. As it has been recently shown, depletion of RPS19 causes a reduction of rRNA synthesis. Specifically, a reduction of 47S rRNA was observed which indicates that the inhibition of rRNA synthesis could be involved in the molecular etiology of the DBA phenotype[38].

1.2.4. Ribosomal protein L11 (RPL11)

The L11 ribosomal protein participates in the formation of the large ribosome subunit. The gene is located on chromosome 1 in humans. The RPL11 protein consists of 178 aa with a molecular weight of 20 kDa.

In yeast the L11 ribosomal protein forms a complex with Rpl5 and 5S rRNA which join pre-ribosomes at an early step in the assembly pathway[39]. In human cells, RPL11 has a similar role in the ribosome biogenesis and further suppresses c-Myc transcriptional activity and plays a feedback regulatory role in c-Myc level coordination and ribosomal biogenesis activity [40].

The ribosomal L11 protein interacts with MDM2 and inhibits MDM2-mediated p53 ubiquitination and degradation in response to ribosomal stress. This leads to a drastic accumulation of ubiquitinated and native MDM2. This effect is dependent on the ubiquitin ligase activity of MDM2, but not p53, and requires the central MDM2 binding domain (residues 51-108) of L11[41]. It has also been found that induction of p53 by RPL11 in human cell lines is a general response to inhibition of biosynthesis of 40S and 60S ribosomal subunits. Inhibition of 40S biogenesis leads to an increase in free RPL11 due to the depression of the polypyrimidine sequence at the 5-prime TOP transcription site resulting in increased translation of the RPL11 mRNA. Therefore, it is considered that the decrease in 60S assembly allows RPL11 to inhibit MDM2 and stabilize p53. The increase in translation of 5'-TOP mRNA is observed despite the continuous biosynthesis of 60S and the overall decrease in translation[42].

RPL11 mutations are involved in Diamond Blackfan Anemia. They were first identified in 2008 by Gazda *et al*. in 6.5% of the DBA patients tested [43]. Since then, more than 25 mutations have been identified in the gene in DBA patients [36,44,45].

RPS19 MUTATION MAP

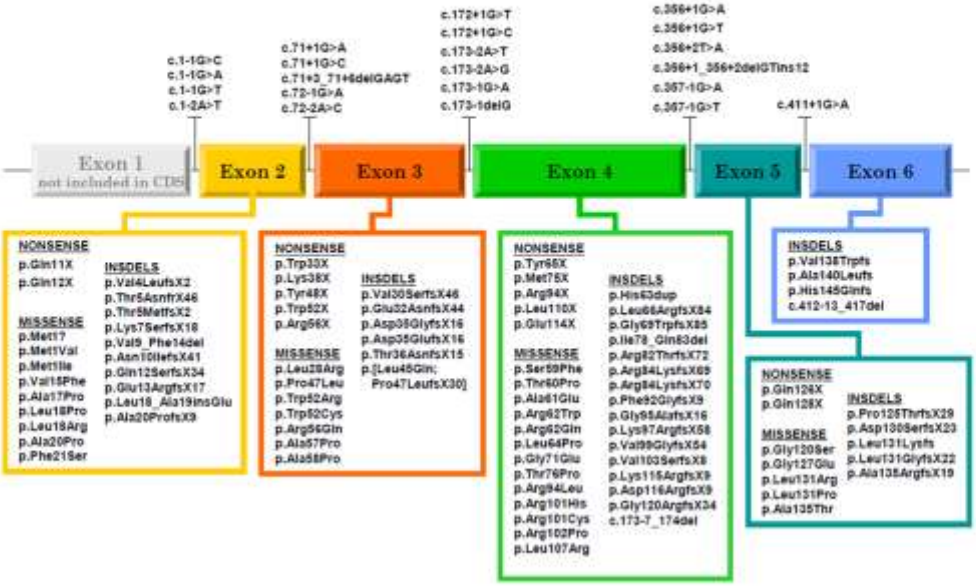

Figure 2 RPS19 mutation map from Leiden Open Variation Database (LOVD; http://www.dbagenes.unito.it/home.php?selet_db=RPS19)

RPL11 MUTATION MAP

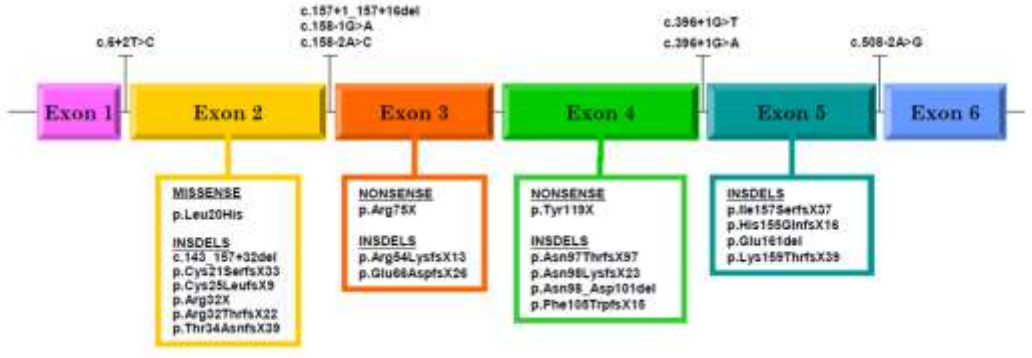

Figure 3 RPL11 mutation map from Leiden Open Variation Database (LOVD; http://www.dbagenes.unito.it/home.php?select_db=RPL11)

1.3. Progression of Normal hematopoiesis and erythropoiesis

1.3.1. Hematopoiesis

Hematopoiesis is a complex process for producing all blood cells starting from the hematopoietic stem cell (HSC). In man, it begins in the first weeks of gestation in the yolk sac. From the 6th week until the 6th- 7th month of fetal life, the main hematopoietic organs are the spleen and the liver and continue to produce blood cells up to two weeks after birth. In childhood and adulthood bone marrow is the only source of blood cells[46].

Hematopoiesis begins with the pluripotent stem cell which, through asymmetric division. In this process a pluripotent stem cell is created which replaces the maternal cell, and also creates a multipotent stem cell. The latter can be self-renewing, but it also has the ability to differentiate, leading to the creation of separate lineages. The hematopoietic stem cell is quite rare (1: 20 million nucleated cells) and is located in the bone marrow in a specific microenvironment, the specialized osteoblastic or vascular "niches"[47].

Cell differentiation proceeds from the stem cell through committed hematopoietic progenitors which are restricted in their developmental potential. The stem cell has the ability to self-regenerate so that the cellularity of the marrow stays stable in a normal healthy steady state. The selection of the cell lineage to which the differentiation will occur depends on both stochastic events and signals that the cells receive from their microenvironment. These include transcription factors that interact with each other to amplify the transcriptional program of a lineage and suppress another, as well as growth factors that regulate the proliferation and differentiation of hematopoietic progenitor cells and the function of mature blood cells[48]. There is considerable reinforcement in the system as a stem cell is capable of producing 10^6 mature blood cells after 20 cell divisions.

1.3.2. Erythropoiesis

All blood cells are divided into three cell types: red cells, white cells, platelets. The most populous are the red cells that are specialized to carry oxygen from the lungs to tissues and carbon dioxide in the opposite direction. They have a lifetime of four months unlike the smaller in size platelet cells that are involved in hemostasis and are in circulation for only ten days.

Humans produce 10^{12} new red blood cells per day through the complex and well-regulated erythropoiesis process of erythropoiesis. Human erythropoiesis is a dynamic multi-stage complex process involving the differentiation of early erythroid progenitors into enucleated erythrocytes. The process begins with the differentiation of HSCs into the multipotent common myeloid progenitor (CMP) that gives rise to multipotent megakaryocytic/erythroid progenitors (MEP).

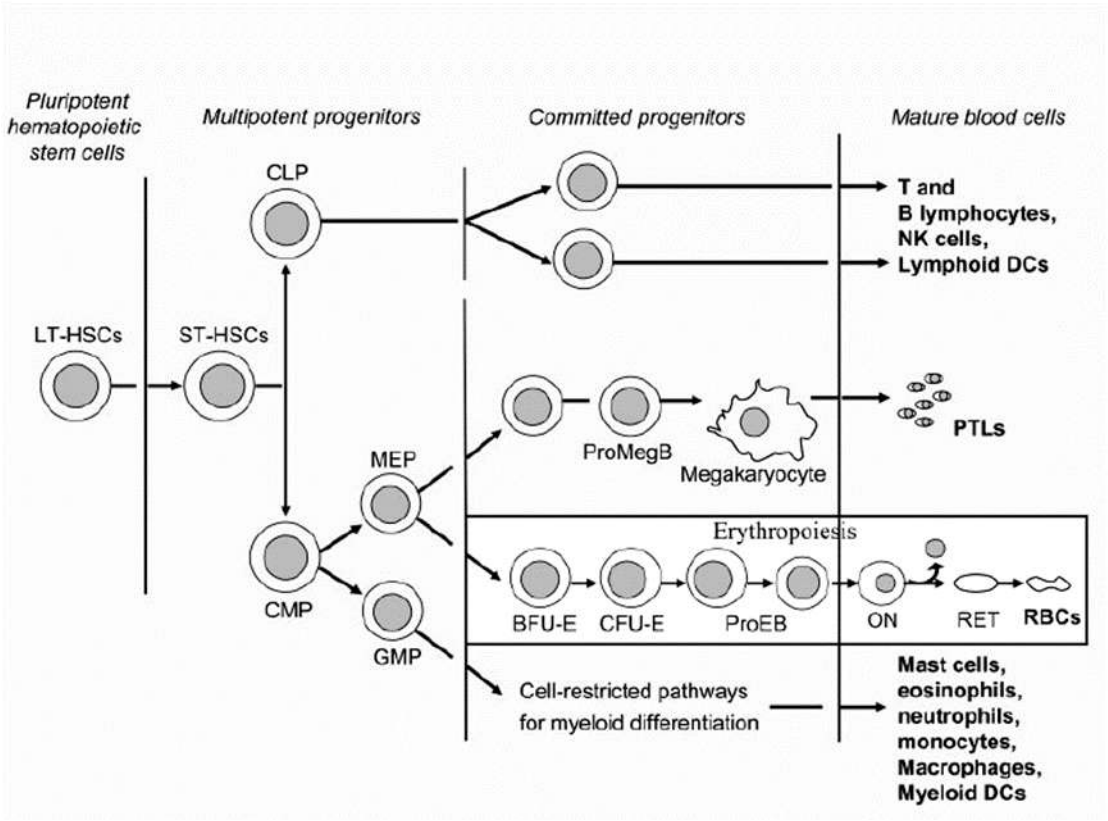

Figure 4 Schematic representation of the erythropoiesis process[49]

MEPs differentiate into the committed progenitors of the erythroid lineage (BFU-E and CFU-E), intermediate forms of proerythroblasts (ProEB) and then into orthochromatic normoblasts. During terminal differentiation, wherein proerythroblasts differentiate into mature red blood cells, specific diffrentiaiton stages with distinct morphologies can be distinguished, which include, in this order, proerythroblasts, basophilic erythroblasts, polychromatic erythroblasts and orthochromatic erythroblasts. Finally, enucleation of orthochromatic erythroblasts leads to the birth of reticulocytes which then mature into enucleated red blood cells (RBCs) once released in the circulation (Figure 4). In erythropoiesis, cell division is coupled to differentiation as each daughter cell is morphologically and functionally different from the maternal cell from which it emerged. During the maturation of the erythroid lineage there is a gradual decrease in cell size, RNA content, DNA synthesis in the polychromatic stage and increase of the cell content in hemoglobin.

Erythropoiesis is regulated by erythropoietin (EPO), which activates erythropoiesis by increasing the number of progenitor cells that are committed in the erythroid lineage. The EPO receptor first appears at low abundance in early BFU-E, it increases in CFU-Es and decreases in later stages. EPO receptors are stimulated by erythropoietin and thus the cells proliferate and differentiate. The GATA-2 transcription factor is involved in the onset of erythroid differentiation at the stage of pluripotent stem cells. Other transcription factors involved in the process are GATA-1 and FOG-which are activated by erythropoietin receptor

stimulation and are important in enhancing expression of antiapoptotic genes. Furthermore, TAL1 / SCL, Erythroid Kruppel-Like Factor, Gfi-1b and BCL11A transcription factors play an important role in the process [49].

As the hematopoietic process evolves from the early stages of maturation of erythroid cells, the cells gradually lose the ability to proliferate and they turn into mature, enucleated red blood cells. Mature red blood cells are biconcave discs without mitochondria and other organelles but are filled with hemoglobin capable of binding and dispensing O_2.

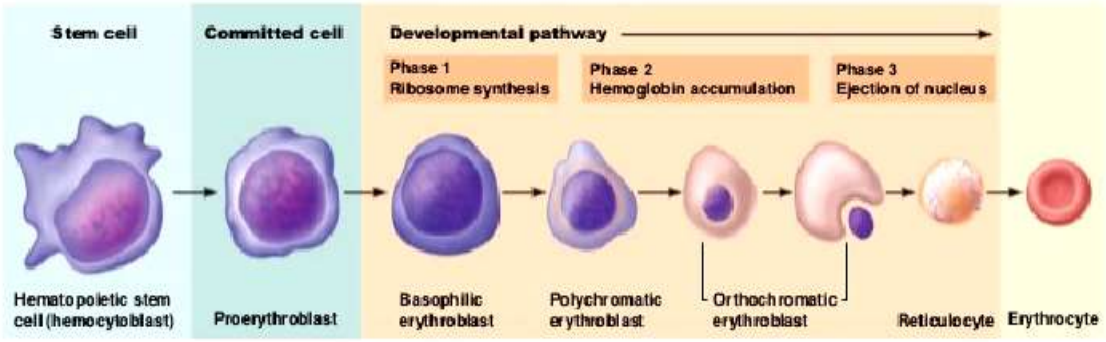

Figure 5 Stages of terminal erythroid differentiation.

1.3.3. Gata-1 in terminal erythroid differentiation

The transcription factor GATA-1 is an essential component in erythropoiesis. It is characterized by the presence of two zing-finger domains and binds to the (A/T) GATA (A/G) consensus DNA binding motif. GATA-1 is expressed in erythroid cells, megakaryocytes, mast cells, eosinophils, and dendritic cells in hematopoiesis. Proper expression of GATA-1 is essential for normal hematopoiesis. The GATA-1 gene knockout in mice results in embryonic lethality at E10.5-E11.5 due to severe anemia[50]. Also, GATA-1 null cells (primitive or definitive) are unable to differentiate beyond the proerythroblastic stage due to apoptotic cell death. GATA-1 plays an instructive role in erythroid differentiation. In particular, the ectopic expression of GATA-1 in myeloid or lymphoid progenitor cells is able to re-program the differentiation and lead it towards the erythroid and megakaryocytic lineages.

The basic functions of GATA-1 in erythroid differentiation are achieved through well-organized regulation, which involves the activation and suppression of various target genes involved in all aspects of the maturation of erythroid cells. Thus, in addition to activating genes associated with the red blood transcription program, GATA-1 is a critical repressor of early hematopoietic progenitor multipotentiality and genes belonging to transcriptional programs of other hematopoietic lineages. For example, GATA-1 suppresses expression of GATA-2 or of the transcription factor of the myeloid lineage, PU.1. Interestingly, in erythroid cells, GATA-1 appears to repress alternative transcriptional programs in other hematopoietic lineages such as eosinophils. GATA-1 also promotes the survival of erythroid cells by up-regulating anti-apoptotic genes such as Bcl-xL.

The ability of GATA-1 to exert a regulatory role in erythropoiesis lays in its ability to interact with other proteins. GATA-1 has been reported to interact with many transcription factors, such as FOG-1, EKLF, TAL-1 / SCL, PU.1 and co-factors such as CBP / p300, Brg1, MeCP1 / NuRD, and others. It has been shown that GATA-1 forms two different FOG-1 complexes with and without the repressive MeCP1 / NuRD complex. GATA-1 also forms discrete complexes with the TAL-1 / SCL or Gfi-1b hematopoietic transcription factors and with the ACF / WCRF chromatin remodeling complex. These complexes probably regulate different target gene groups. For example, the GATA-1 / FOG-1 / MeCP1 complex suppresses early hematopoietic and alternative lineage transcription programs, whereas the GATA-1 / Gfi-1b complex suppresses cell division by suppressing mitogenic genes such as myc. In addition, GATA-1 / FOG-1 and the GATA-1 / TAL-1 / SCL complexes up-regulate the erythroid transcription program.

Also, GATA-1 protein levels play an important role in the regulatory function of the protein. Protein levels decrease in the last stages of erythroid differentiation, with overexpression of GATA-1 inhibiting terminal erythroid differentiation in vivo. According to the existing model, mature erythroid cells stimulate immature cells to degrade GATA-1 in order to complete the final steps of maturation. This process functions as a homeostatic mechanism for maintaining red blood cell production[49].

1.4. Human disorders of ribosome dysfunction and implications of the GATA-1 transcription factor

1.4.1. Ribosomopathies

Changes in ribosomal components or in numerous cellular products with effects on ribosomal structure and function can cause a heterogeneous class of diseases known as ribosomopathies. The family of these diseases first appeared with the discovery of Diamond Blackfan Anemia (DBA). In 1999 it became known that the RPS19 gene is mutated in DBA patients [34]. It is now known that the Rps19 gene is mutated in 25% of the cases whereas at least 60-70% of patients with Diamond-Blackfan anemia carry mutations in ribosomal proteins in general.

Ribosomopathies comprise all disorders in which genetic abnormalities cause impaired biogenesis and function of the ribosome, resulting in the emergence of specific clinical phenotypes[51]. Although all ribosomopathies include ribosomal dysfunction, these diseases differ significantly in the mechanism, clinical presentation and treatment options. It would be expected that disorders related to the biogenesis of ribosomes would have ubiquitous and fatal impacts for the organism. However, ribosomal protein mutations appear to have tissue-specific effects[52]. This finding provides a new perspective of the role of ribosomes in normal development and disease.

A typical example is the haploinsufficiency of the RPSA ribosomal protein which is detected in patients suffering from isolated congenital asplenia and in whom the spleen does not develop normally[53]. These patients, in addition to being susceptible to bacterial infections due to a lack of the spleen, do not show any other pathological symptom. This is surprising because RPSA is a component of the small subunit and is expressed in all tissues as well as its mutated form[54].

A possible explanation for the above phenomenon suggests that the tissues expressing the symptoms divide very quickly, making them particularly prone to mutations that reduce the functionality and number of ribosomes (e.g. bone marrow). Another explanation could be that the composition of the ribosomes differs in the different cell types. This is supported by the fact that the reduction of a particular ribosomal protein can lead to changes in the range a subset of translated mRNAs without affecting overall protein synthesis. Indeed, the amounts of ribosomal mRNAs have been shown to vary in the various developing fetal mouse tissues, but it is not known whether these differences are also reflected in the composition of functional ribosomes[53]. Alternatively, an interpretation that could be attributed is that mutations in the ribosomal genes lead to a reduced number of fully functional ribosomes resulting in them being sufficient for survival, but due to their reduced number there are differences in the translated mRNA which may affect proper cellular growth and differentiation in tissues[54].

Table 1 Known ribosομopathies (41; modified)

Disease	Gene Defect
Diamond Blackfan anemia	RPS19, RPS24, RPS17, RPL35A, RPL5, RPL11, RPS7, RPL36, RPS15, RPS27A
5q-syndrome	RPS14
Shwachman-Diamond syndrome	SBDS
X-linked dyskeratosis congenita	DKC1
Cartilage hair hypoplasia	RMRP
Treacher Collins syndrome	TCOF1

In some cases, mutations may also disrupt ribosome assembly by causing nucleolar stress response in some cells resulting in an increase in p53 protein synthesis leading to apoptosis. For example, deactivation of p53 in a mouse model for Diamond-Blackfan Anemia (Rps19 haploinsufficiency), led to the rescue of certain red cell disorders [54].

The clinical features of ribosomal infections may include bone marrow deficiency, developmental abnormalities and an increased risk of developing cancer. Ribosomal insufficiency can cause a wide range of symptoms, the severity of which may vary dramatically between patients with the same diagnosis [55].

1.4.2. Diamond-Blackfan Anemia (DBA)

DBA was first described by Josephs (Josephs 1936). Later it was characterized by Diamond and Blackfan as a congenital hypoplastic anemia. This disorder is characterized by anemia, macrocytosis, reticulopenia and selective reduction or absence of erythroid precursor cells in an otherwise normocellular bone marrow[56]. The majority of patients are diagnosed in their first year of life with paleness and lethargy being the most common symptoms. There is often a family history of the disease and in the majority of cases, it has an autosomal dominant character. Other features of the disease include elevated levels of red blood cell adenosine deaminase or presence of fetal membrane antigen "i" and a set of physical abnormalities seen in 40% to 62% of patients, ranging from low stature, thumb malformations to heart defects.

Regarding current treatments, corticosteroids were the first effective treatment for DBA and remain the treatment of choice today. Activation of the glucocorticoid receptor promotes erythroid proliferation [57]. Leucine, which stimulates translation through the mTOR pathway, has been shown to improve hematopoiesis in some animal models for DBA[58]. Transplantation of stem cell hematopoietic stem cells is the only definitive treatment for anemia, but still has complications for which there is an effort to overcome [59].

DBA was the first ribosomal disease to be studied in detail. Until today, a list of 16 ribosomal genes involved in DBA has been identified, but the molecular mechanisms due to which pathogenicity is manifested are not elucidated[60]. It is known that all mutations detected in RP genes in DBA are heterozygous. More than 200 mutations in genes related to DBA have been recorded. This results in the disease presenting great genetic heterogeneity and a very variable phenotype [36]. However, some studies have found correlations between the genotype and the phenotype. In particular, mutations in RPL5 have been found to be associated with a higher incidence of natural abnormalities (lip-palate slit), while patients with RPL11 mutations have been found with more thumb abnormalities[43,61]. Another important observation is that the decrease in expression of ribosomal proteins results in a generally reduced translation in both hematopoietic and non-hematopoietic cells[62,63].

Regarding the role of p53 specifically in DBA, it has been shown that in RNAi Rps19 knockdown human CD34+ hematopoietic progenitor cells, which successfully represent the defective erythropoiesis model in DBA[64,65], there is an accumulation of p53. The restoration of RPS19 expression and the inactivation of p53 rescue of the phenotype[66]. A study in CD34+ cells from DBA patients showed that in the case of Rps19 mutation there was a decrease in proliferation and an increase in cell cycle arrest, with no significant difference in differentiation and apoptosis.

In the case of Rpl11, delayed differentiation and increased apoptosis were observed. This suggests that possibly the effect of p53 activation varies depending on the mutated ribosomal gene[67]. It should be noted that there are indications for mechanisms in DBA that are independent of p53. Specifically, the inactivation of p53 in RPS19-deficient zebrafish, led to rescue of morphological abnormalities but not of erythropoiesis problems[68].

1.4.3. Gata1 in DBA

Although approximately 60-70% of DBA patients have been identified to carry heterozygous pathogenic mutations in RP genes[35], there is still a percentage of cases for which the molecular etiology is unknown. In an attempt to elucidate the genetic basis of DBA cases that were not characterized by a specific RP mutation, Gazda *et al*. performed exome sequencing for two siblings and discovered a mutation in the transcription factor GATA1, which is the master regulator of erythropoiesis[69]. This mutation resulted in the replacement of the last nucleotide of the second exon and thus the exclusive expression of the GATA1 short isoform lacking the 83aa N-terminal transactivation domain. Subsequently, other research groups identified similar patient cases [70]. Sankaran *et al*. found that ribosomal protein haploinsufficiency could reduce full length GATA1 mRNA translation, without affecting translation of GATA1s [71]. Particularly, in primary hematopoietic cells from patients with RPS19 mutations, a decrease in GATA1-target gene sets as well as of GATA1 was observed, despite the fact that its mRNA levels were normal. It is also important that the ectopic expression of GATA1 improved the erythroid differentiation indicating that its abnormal expression plays an important role in the DBA phenotype. Overall, mutations in RP or GATA1 in DBA patients result in GATA1 short isoform expression which fails to fully support erythropoiesis[72,73]. Importantly, the Strouboulis lab has shown that GATA 1 appears to bind to ribosomal genes involved in DBA, while exhibiting greater binding for the *RPS19* gene[74]. Considering all the above, it may be concluded that there is a molecular link between the ribosomal genes and the transcription factor GATA1, the mechanism of which is not yet known.

1.5. Aim of this thesis

The aim of this thesis is to investigate the molecular relationship between the transcription factor GATA-1 and ribosomal protein genes involved in the pathogenesis of Diamond Blackfan Anemia (DBA). Recent work in the laboratory of Molecular Hematopoiesis at IMBB has shown that GATA-1 binds to many ribosomal protein genes, including many mutated in DBA[75]. This raises the question of whether GATA-1 regulates the expression of ribosomal protein genes and, possibly through a potential feedback loop model, GATA-1 activates the expression of RP genes to enhance or to maintain its own translation in erythroid cells. In order to investigate this possibility, the **aim of this thesis** is to provide preliminary evidence for GATA-1 regulating the expression of ribosomal protein genes in erythroid cells. This aim translates into the following specific objectives: A) Firstly, to demonstrate GATA in vitro binding to RP genes through EMSA assays and the use of appropriate probes carrying the possible GATA-1 binding-RP sequences as well as probes bearing point mutations for this specific GATA-1 binding sites, B) secondly, to show GATA-1 transcriptional regulatory activity by using transcriptional reporter assays wherby RPS19 and RPL11 gene sequences bound in vivo by GATA-1 will be cloned upstream of a minimal promoter linked to a luciferase reporter gene and transfected together with GATA-1 in a

human cell line (HEK293T and HeLa cells). These experiments are a first step in determining whether GATA-1 regulates the transcription of ribosomal genes implicated in DBA.

2. Materials and methods

2.1. Plasmid constructs

The 290bp fragment of the *human RPS19* gene, which is bound by GATA-1 in vivo (Papadopoulos, Karkoulia *et al*., 2013) and contains two possible GATA-1 binding motifs, and the 207bp fragment of the *human RPL11* gene, also bound by GATA-1 *in vivo* and which contains one possible GATA-1 binding motif, were synthesized by Geneart® (Life Technologies) and cloned into the pMA-T vector (pMA-T_RSP19 and pMA-T_RPL11 constructs). Each fragment was subsequently re-cloned as an *Acc65I/ XhoI* fragment into the pGL3 promoter vector (Promega) to be used in luciferase assays.

Complementary phosphorylated oligonucleotides containing the mutated RPL11 GATA-1 binding motif (67 bp) were annealed (Table 1), purified and subsequently cloned as an *Acc65I/NcoI* fragment() into the pMA-T_RPL11 plasmid. The RPL11 207bp fragment containing the mutated GATA-1 binding motif was then re-cloned as an *Acc65I/XhoI* fragment into the pGL3 promoter vector.

In addition to the above, bioGATA1pcDNA3.1(+) plasmid constructed previously in the lab for expressing GATA-1, as well as pRL-**TK** Vector (Promega) were used for the luciferase experiment.

N-terminal 3xFLAGTEV Avi GATA1 pBUDNeo hBirA GFP plasmid constructed previously in the lab was used for the Electrophoretic Mobility Assay (EMSA).

2.2. Site-directed mutagenesis PCR

Mutations to GATA-1 binding motifs were introduced for the RPS19 fragment which had been cloned into the pGL3 promoter vector (pGL3 prom_RPS19 construct) by PCR mediated site-directed mutagenesis, as described previously[76]. Primers were designed using the web-based program PrimeX[http://www.bioinformatics.org/primerx/index.htm]. Initially, two constructs were created, each with one of the two GATA-1 binding motifs mutated. The mutations were confirmed by sequencing. Then one of the two constructs carrying a mutation in one of the two binding motifs was used as a template for PCR mediated site-directed mutagenesis in the second GATA-1 binding motif, thus generating a construct with both GATA-1 binding motifs mutated. Eventually 3 constructs were created: the first carried a mutation in the first GATA-1 binding motif, the second carried a mutation in the second GATA-1 binding motif and the third had both the GATA-1 binding motifs mutated. Primer sequences are listed in Table 2.

Table 1 Primers used for the site-mutagenesis PCR of the GATA-1 binding motifs for the RPS19 fragment which had been cloned into the pGL3 promoter vector. Mutation sites are highlighted with red color.

Primers for mutagenesis PCR	Sequence
RPS19 1st site	Forward: 5' TTTGTCCCGCCCTTACGTTCTCCCCTCCTCCA 3' Reverse: 5' TGGAGGAGGGGAGAACGTAAGGGCGGGACAAA 3'
RPS19 2nd site	Forward: 5' GAAGGACGGAAGATGCTAGCCACATTTCTTC 3' Reverse: 5' GAAGAAATGTGGCTAGCATCTTCCGTCCTTC 3'

Table 2 Complementary oligonucleotides containing the mutated RPL11 GATA-1 binding motif, Mutated sites are highlighted with red color.

Oligonucleotides used for the cloning of RPL11 mutated fragment	**Forward:** **5′CGCTCTCCATCATGGCGGTGAGTAGCTGGGACCTGGATTTGCTTTCCTTTACGCGTCGCCATC 3′**
	Reverse: 5′CATGGATGGCGACGCGTAAAGGAAAGCAAATCCAGGTCCCAGCTACTCACCGCCATGATGGAGAGCG 3′

2.3. Cell culture

HEK 293T and HeLa cells were grown in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% FBS and 1% penicillin/streptomycin. The maintenance of the cells was done in a cell culture incubator at standard temperature 37 °C, with continuous supply of 5% v/v CO_2 and environment saturated with moisture and density which allow logarithmic growth. Dilutions and renewal of the medium were performed according to the developmental characteristics of each cell line.

2.4. Transient-transfection and luciferase reporter assay

HeLa cells were transfected using jetPEI reagent (PolyPlus-transfection) according to the manufacturer's instructions. Cells were plated one day before transfection in a 24-well plate (0.5×10^5 cells /well) and transfected with 1000ng total DNA. In particular, 0.1 μg of each of pGL3_RPL11, pGL3_RPS19 reporter vectors and their mutants was contrasfected with 10ng pRL-**TK** Vector and 300ng of bioGATApcDNA3.1(+) plasmid. pBluescript II KS +/- (Stratagene) was used to maintain a constant amount of transfected DNA in each well. pGL3 empty vector was used as control. Luciferase assays were performed using the Dual-Luciferase Reporter Assay System (Promega) according to the supplier's instructions. Luciferase activities were measured at 48 hours post-transfection. Luciferase assays were carried out using two technical replicates per transfection for each condition.

HEK 293T cells were transfected using the calcium phosphate transfection method. Briefly, the solution of the plasmid DNA mixed with $CaCl_2$ (final concentration was 250 mM $CaCl_2$) was mixed with a 10-fold excess volume of HEPES-phosphate buffer (280 mM NaCl, 50 mM HEPES, 1.5 mM Na_2HPO_4, pH 7.06), vortexed and incubated at room temperature for 10 minutes prior to its dropwise addition to the cultures of eukaryotic cells. After 24 hours, the supernatant was replaced with complete growth medium. Cells were plated one day before transfection in a 24-well plate (0.5×10^5 cells /well) and transfected with 610 ng total DNA. pBluescript II KS +/- (Stratagene) was also used to maintain a constant amount of transfected DNA. The amount of DNA for each construct used, was the same as for HeLa cells. Luciferase assays were done using the Dual-Luciferase Reporter Assay System (Promega) according to the supplier's instructions. Luciferase activities were measured 48 hours post-transfection.

2.5. Nuclear extracts

The nuclear extracts used for the bandshift experiment were obtained from transiently transfected HEK 293T cells. The cells were plated in 10cm plates one day before transfection and cultured to reach ~60% confluency on the day of transfection. HEK 239T cells were

transfected with 15µg of expression vector 3xFLAGTEV Avi GATA1 pBUDNeo hBirA GFP using the calcium phosphate transfection method, as described above. Nuclear extracts were prepared using the High Salt method, as previously described[77].

2.6. Electromobility shift assays (EMSAs)

10 µg of high salt nuclear extract was used for each reaction with 30 ng of unlabelled EMSA probes. Each EMSA reaction was performed in the presence of 1x binding buffer (50mM KCl, 5mM $MgCl_2$, 20mM Tris-HCl pH 7.8, 10µM $ZnSO_4$, 0.5mM EDTA, 5% glycerol, BSA 1mg/ml, 0.4µg/µl poly (dI-dC), DTT and Proteins inhibitors). Each reaction was incubated at 15°C for 15 minutes. After incubation, samples were loaded on a 14% non-denaturing polyacrylamide gel prepared in 0.5x TBE buffer (45 mM Tris, 45 mM boric acid and 2 mM EDTA) and run for 2.5-3 hours in 0.5x TBE buffer. The gel was then transferred to nitrocellulose membrane (GE Healthcare Life Sciences) following the same procedure as for a Western blot. The membrane was blocked using 5% dry milk in 1X TBS, 0.1% Tween® and incubated with anti GATA-1 (N6) (Santa Cruz Biotechnology, sc-265) antibody. Supershift assays were done using anti GATA-1 antibody which was added to the EMSA samples after the binding incubation was followed by an additional 30 minute incubation on ice.

2.7. Western blotting

Western immunoblot analysis was performed using 40-50µg of protein extract, as previously described[78]. As blocking solution 5% dry milk in 1X TBS, 0.1% Tween®, was used. Primary antibody used was anti GATA-1 N6 (1:5000) (Santa Cruz Biotechnology, sc-265), and secondary anti Rat (1:10000) (Santa Cruz Biotechnology, sc-2006). Western blots were developed by chemiluminescence (ECL prime, GE Healthcare).

3. Results

3.1. Binding of GATA-1 to the human RPS19 and RPL11 proximal promoter region in vitro.

As it is known from ChiPseq data GATA-1 binds to the proximal promoter region of several human Ribosomal Protein (RP) genes including RPL11 and RPS19 (Figure 6). We initially examined the sequence under the GATA-1 binding peaks in the human RPL11 and RPS19 promoter sequences for consensus GATA-1 binding sites (WGATAR) using standard TF binding prediction tools. Within the DNA sequence upstream of RPS19 transcription start site there are two consensus GATA-1 binding sites. There is also one GATA-1 binding site within the promoter of RPL11. We first tested whether GATA-1 could bind to the RPS19 and RPL11 proximal promoter region by EMSA using nuclear extracts from HEK 293T cells that had been transfected with a GATA-1 expression vector.

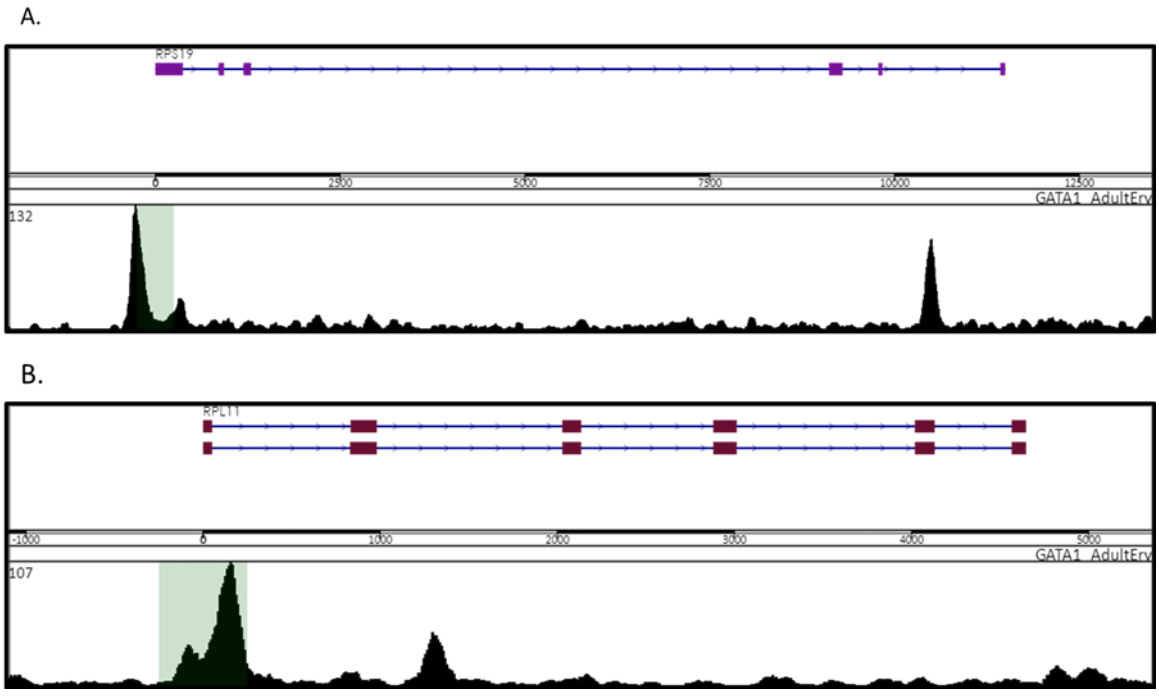

Figure 6 ChiP seq data from human erythroid cell lines showing GATA-1 occupancies of RPS19 (A) and RPL11 (B) genes.

Table 3 Probes used for EMSA assay. Mutations for each probe are highlighted with red color.

EMSA probes	Sequence
RPS19 for the 1st GATA-1 binding site	**TCCCGCCCTTATCTTCTCCCCT**
RPS19 for the 2nd GATA-1 binding site	**ACGGAAGATGATAGCCACATTT**
RPS19 mutated for the 1st GATA-1 binding site	**TCCCGCCCTTACGTTCTCCCCT**
RPS19 mutated for the 2nd GATA-1 binding site	**ACGGAAGATGCTAGCCACATTT**
RPL11	**TTTCCTTTATCCGTCGCCATC**
RPL11 mutated	**TTTCCTTTACGCGTCGCCATC**

Based on the predicted GATA-1 binding motifs in the RPS19 and RPL11 promoter regions, described above, six different EMSA probes were used to test GATA-1 binding *in vitro* (Table 3). Binding of GATA-1 to the possible binding site in probe RPL11 P was indeed confirmed as the protein complex detected by EMSA, was supershifted by the GATA-1 N6 antibody, or abolished by use of a mutated probe (RPL11 Pm) (Figure 7). Also binding of GATA-1 at both possible binding sites in probes RPS19 P1 and RPS19 P2 were confirmed as these protein complexes detected by EMSA were also supershifted by GATA-1 N6 antibody, or abolished by use of mutated probes (RPS19 P1m, RPS19 P2m)(Figure 7). From these data we conclude that GATA-1 binds to the RPS19 and RPL11 proximal promoter region *in vitro*.

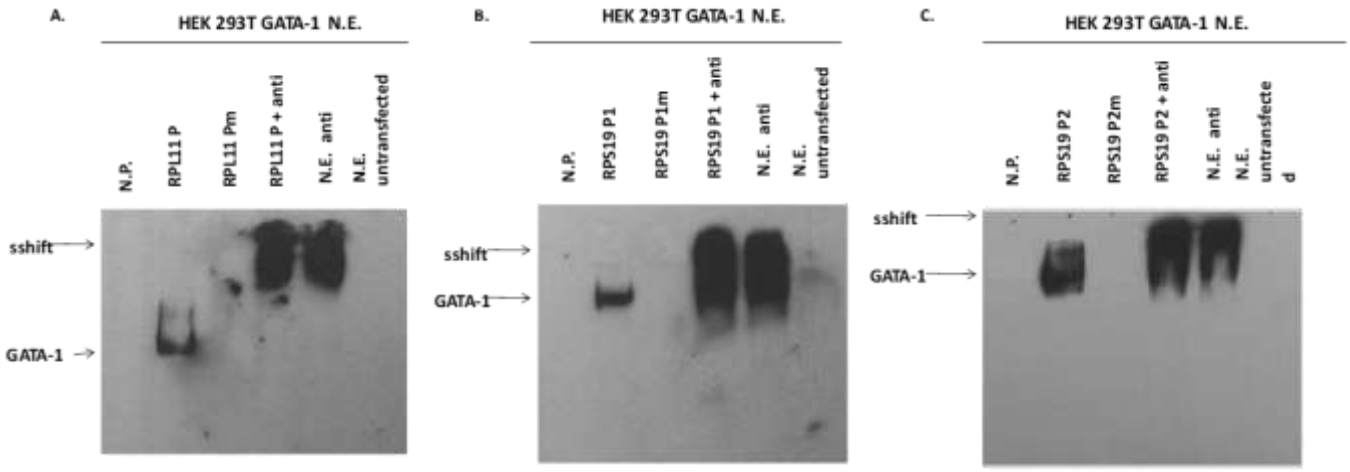

Figure 7 EMSA assays to show GATA-1 binding to the proximal promoter region of the human RPL11 (A) and RPS19 (B, C) genes. N.P.: no probe; RPL11 P: RPL11 probe; RPL11 Pm: RPL11 probe with GATA-1 binding site mutated; RPS19 P1: RPS19 probe 1; RPS19 P2: RPS19 probe 2; RPS19 P1m: RPS19 probe with the first GATA-1 binding site mutated; RPS19 P1m: RPS19 probe with the second GATA-1 binding site mutated; RPL11 P + anti: anti-GATA-1 supershifted reaction; RPS19 P1+ anti: anti-GATA-1 supershifted reaction; N.E. anti: Nuclear Extract and GATA-1 antibody..

3.2. Gata-1 regulates RP's transcription

To determine whether GATA-1 is able to regulate the transcriptional activity mediated by each of RPL11 or RPS19 fragments, we carried out luciferase assays to study the effect of GATA-1 on reporter activity of two plasmids consisting of each RP fragment which includes the GATA-1 binding sites, cloned upstream of a minimal promoter driving luciferase expression. Each reporter plasmid was transfected alone or co-transfected along with a GATA-1 expression plasmid into HeLa or HEK 293T cells which do not express endogenous GATA-1. As shown in Figure 78 in HEK 293T cells, reporter plasmids bearing the RPL11 and RPS19 promoter sequences were capable of leading to variable upregulation of luciferase activity, compared to the minimal promoter control. By contrast, co-transfecting GATA-1 mediated suppression of RP-driven luciferase activity to a similar extent in both RP reporter plasmids.

As shown in Figure 9, In HeLa cells cloning of the Rpl11 and Rps19 fragments upstream of the minimal promoter results in transcriptional upregulation compared to the minimal promoter control (Fig. 9). In the presence of GATA-1, luciferase activity decreases significantly for both RP reporter plasmids (Fig. 9). In the same experiment reporter plasmids consisting of each RP fragment with mutated GATA-1 binding sites were also included. Luciferase expression in reporter plasmids containing the mutated sites was upregulated in the absence of co-transfected GATA-1, when compared to the minimal promoter control (Fig. 9). However when co-transfected with GATA-1, luciferase expression from the RP reporters with the mutated sites presented with marked repression in expression levles, as also seen with the wild type RP promoter reporter plasmids (Fig. 9). Interestingly, besides the fact that these mutations abolish the GATA-1 binding in EMSA assay (Figure 7), transcriptional activity of plasmids containing the non-mutated fragments compared to those containing the mutated RP fragments presents a similar pattern in absence and presence of GATA-1 (Figure 9). This similarity indicates that this pattern is probably unrelated to the GATA-1 binding and regulatory activity.

Although a similar pattern is observed in multiple experiments and in different cell lines, a definite conclusion cannot be drawn as to whether GATA-1 suppresses or activates expression of the RP genes. The reason is that a high background signal in luciferase activity is seen with the empty pGL3 vector which decreases in the presence of GATA-1, in a manner similar to that of luciferase reporter constructs (Figure 8Figure 9). This suggests that in the context of the pGL3 reporter plasmid, the GATA-1 effect on transcription is non-specific. This is also supported by the fact that in constructs carrying mutations in GATA-1 binding motifs which, as shown by EMSA experiments abolish GATA-1 binding luciferase activity is similar to that of non-mutated constructs (Figure 8Figure 9). As a result, after using transcription assays utilizing a basal promoter luciferase reporter plasmid, there was no clear indication for GATA-1 regulating the RPL11 and RPS19 genes.

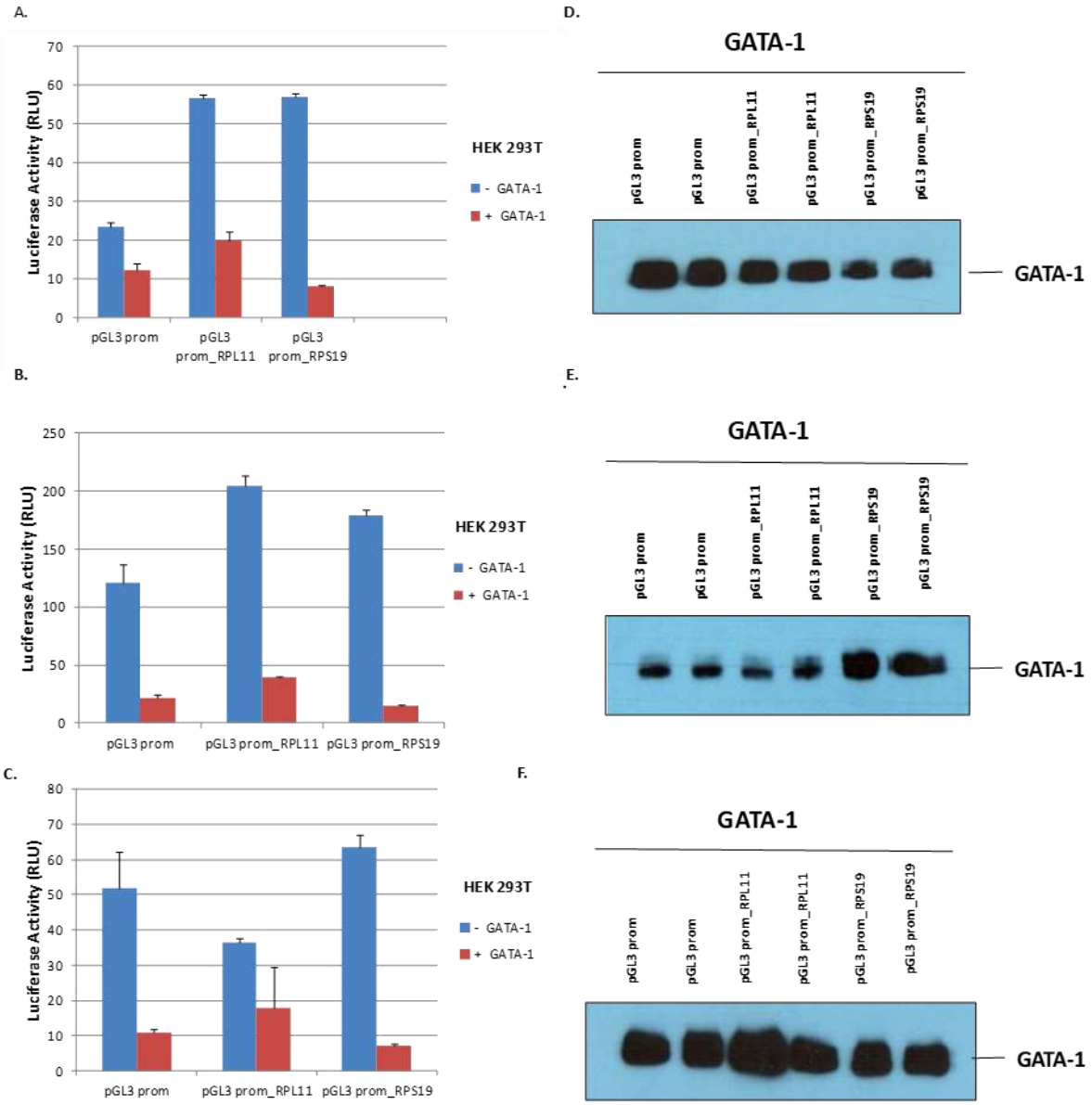

Figure 8 Three independent luciferase assay experiments with two technical replicates each in HEK 293T are shown. (A)(B((C) luciferase reporter assays were carried out in HEK 293T cells transfected with 100 ng of the luciferase reporter with (+) and without (−) the indicated amounts of bioGATApcDNA3.1(+) plasmid expressing GATA-1. 48hours after transfection cell extracts were prepared and analyzed for luciferase activity as described in Materials and Methods. The firefly luciferase enzyme activity was normalized to the Renilla luciferase enzyme activity. GATA-1-stimulated reporter activity is expressed relative to that of the reporter constructs in the absence of bioGATApcDNA3.1(+).(C)(D)(E) Western blot confirming the expression of GATA-1 in the samples co-transfected together with the GATA-1 expression plasmid. Error bars indicate the standard deviations of duplicate assays. Similar results were obtained in at least three experiments

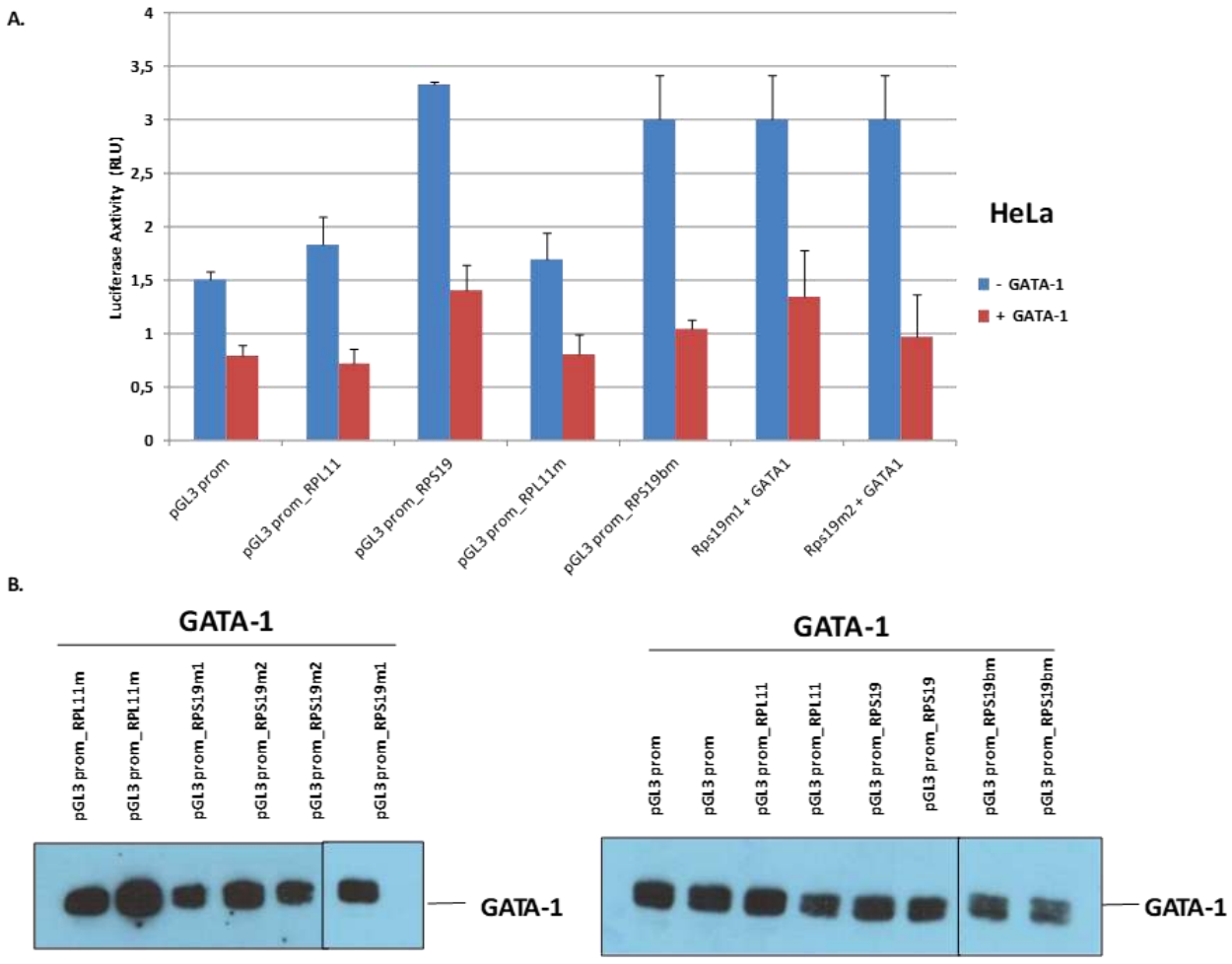

Figure 9 Luciferase assay experiments with two technical replicates in HeLa cells.(A) Luciferase reporter assay was carried out in HeLa cells following the same procedure as for HEK 293T cells. In this experiment constructs containing mutated the GATA-1 binding motifs for RPL1 and RPS19, were also used. (B) Western blot confirms the expression of GATA-1 in the samples co-transfected together with the GATA-1 expression plasmid. Error bars indicate the standard deviations of duplicate assays.

4. Discussion

Ribosomal biogenesis is a universal function that must remain closely coordinated to meet the needs of protein synthesis in each cell type. Despite the widespread expression and necessity of ribosomal proteins in all tissues, the equilibrium in their levels is pathophysiologically manifested by tissue specialization in humans. The correlation between ribosome gene haploinsufficiency and erythropoiesis suggests the necessity of balanced ribosomal biogenesis in the final stages of differentiation and the existence of regulatory mechanisms. In particular haploinsufficiency in specific RP genes that disrupt balanced ribosome biogenesis, result in specific hematopoietic disorders such as DBA. DBA is characterized by a very specific phenotype of erythroid hypoplasia in bone marrow, underlining the need for balanced RP production in final erythroid differentiation[51,79]. Recent data has linked the necessary erythropoietic transcription factor GATA-1 to DBA[79,80]. First, rare mutations in the GATA-1 gene leading to the production of a shorter, N-terminal truncated GATA-1 protein (GATA-1s), were identified in DBA patients with no detectable mutations of the RP genes[69,70]. In addition, it has recently been shown that RP gene haploinsufficiency results in an ineffective translation of the long isoform of GATA-1 into pre-erythroblasts [71], which in turn results in increased apoptosis and diminished erythroid maturation. These findings link the haploinsufficiency of the RP genes to defects in GATA-1 function in the regulation of erythropoiesis[72,73]. However, the possibility that GATA-1 can directly regulate the expression of RP genes in erythroid cells has not been systematically investigated. In fact, transcriptional regulation of the RP genes has been poorly described.

Recent work in the lab showed for the first time that many RP genes, including many involved in DBA, are bound by the GATA-1 transcription factor in erythroid cells in both mouse and human[74]. Taking into account all the above, the purpose of the present study is to provide evidence as to whether the transcription factor GATA-1 is involved in the transcriptional regulation of ribosomal genes, giving priority to those involved in DBA.

For this purpose and knowing from *in silico* analysis of chip data as well as qPCR that GATA-1 binds to the promoters of ribosomal genes involved in DBA[75], two genes were chosen to be studied: RPS19, the first gene which was identified to be involved in DBA and exhibits the greatest binding by GATA-1 and RPL11 which is also bound by GATA-1 and implicated in DBA.

The initial aim is a follow-up to the earlier work to prove that GATA-1, in addition to binding to the ribosomal genes, regulates directly their transcription in erythroid cells. In a first step and in order to investigate the possible regulatory role of GATA-1, transcriptional reporter assays were performed, in which the segments of the RPS19 and RPL11 genes which are bound *in vivo* by GATA-1, were cloned upstream of a minimal promoter linked to a luciferase reporter gene and transfected together with GATA-1 in non-erythroid cell lines of human origin (HEK 293T and HeLa cells). As shown in figures 8 and 9, in the presence of GATA-1, in all cases, the activity of luciferase for both genes decreases markedly indicating a suppressing role for GATA-1 in the regulation of ribosomal genes. It is necessary, however, to note that the high background activity signal of the empty pGL3 prom vector which is also reduced in the presence of GATA-1 does not allow us to fully trust the results of reporter assays. This highlights the need to repeat these experiments using another reporter system that will not display a high background and will allow clear conclusions to be drawn. Furthermore, the experiment is recommended to be repeated in K562 cells or other cellular models for human erythroid differentiation.

The ChiP seq data for GATA-1 binding to MEL cells and immature mouse erythrocytes shows two distinct peaks in the RPS19 gene promoter region and a peak for the RPL11 gene respectively (human). Using bioinformatics analysis, two possible GATA-1 binding motifs were identified in the RPS19 gene promoter and one in the RPL11 gene promoter. GATA-1 as shown has been involved in the formation of discrete protein complexes *in vitro*, as evidenced by the use of suitable probes carrying mutations in the respective binding motifs and the use of specific antibodies in the supershift reactions.

In the future and if the existence of any regulatory role of GATA-1 in the expression of ribosomal genes is established, we will employ CRISPR / Cas9 to mutate GATA-1 binding sites in RP genes in BEL-A (Bristol Erythroid Line Adult) cells by directly testing the potential involvement of GATA-1 in RP gene regulation in erythroid cells *in vivo*. BEL-A cells are the first human immortalized adult erythroid line capable of fully recapitulating normal erythropoiesis, enucleating to generate mature reticulocytes. Characterization of BEL-A cells revealed no differences functionally or at the molecular level to *in vitro*-cultured adult reticulocytes[81].

The hypothesis of the regulation of ribosomal genes also supports the finding of GATA-1 mutations leading to the production of the smallest form of the protein[69]. The reduction in expression of GATA-1 long isoform leads to impaired red differentiation [71]. The reduced GATA-1 levels or production of the short protein isoform are likely to lead to inadequate transcriptional regulation of ribosomal protein genes leading to imbalance in their levels and disease manifestation. For this reason, in future experiments it is necessary to include the investigation of the role of GATA-1s in the regulation of ribosomal genes. This thesis attempts to contribute to the elucidation of the molecular link between the final erythroid differentiation and biogenesis of ribosomes, while at the same time it will shed light on the mechanism of general transcriptional regulation of ribosomal genes. Finally, it is also of great importance to clarify the molecular mechanism leading to the DBA with a future goal of developing therapeutic methods that will relieve sufferers.

Bibliography

1. Roberts, R. B., E. 'Introduction' in Microsomal Particles and Protein Synthesis. *New York Pergamon Press. Inc.*

2. Palade, G. E. A small particulate component of the cytoplasm. *The Journal of biophysical and biochemical cytology* **1,** 59–68 (1955).

3. Doudna, J. A. & Rath, V. L. Structure and function of the eukaryotic ribosome: The next frontier. *Cell* **109,** 153–156 (2002).

4. Wilson, D. N. & Cate, J. H. D. The structure and function of the eukaryotic ribosome. *Cold Spring Harb. Perspect. Biol.* **4,** 5 (2012).

5. Steitz, T. a. A structural understanding of the dynamic ribosome machine. *Nat. Rev. Mol. Cell Biol.* **9,** 242–253 (2008).

6. Lafontaine, D. L. J. Noncoding RNAs in eukaryotic ribosome biogenesis and function. *Nat. Struct. & Mol. Biol.* **22,** 11–19 (2015).

7. Thomson, E., Ferreira-Cerca, S. & Hurt, E. Eukaryotic ribosome biogenesis at a glance. *J. Cell Sci.* **126,** 4815–21 (2013).

8. Tafforeau, L. *et al.* The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of pre-rRNA processing factors. *Mol. Cell* **51,** 539–551 (2013).

9. Nomura, M. Regulation of Ribosome Biosynthesis in Escherichia coli and Saccharomyces cerevisiae : Diversity and Common GUEST COMMENTARY Regulation of Ribosome Biosynthesis in Escherichia coli and Saccharomyces cerevisiae : Diversity and Common Principles. **181,** 6857–6864 (1999).

10. Wool, I. G., Chan, Y.-L. & Glück, A. Structure and evolution of mammalian ribosomal proteins. *Biochem. Cell Biol.* **73,** 933–947 (1995).

11. Li, X., Zhong, S. & Wong, W. H. Reliable prediction of transcription factor binding sites by phylogenetic verification. *Proc Natl Acad Sci U S A* **102,** 16945–16950 (2005).

12. Yoshihama, M. *et al.* The Human Ribosomal Protein Genes : Sequencing and Comparative Analysis of 73 Genes The Human Ribosomal Protein Genes : Sequencing and Comparative Analysis of 73 Genes. 379–390 (2002). doi:10.1101/gr.214202

13. Brown, P. O. *et al.* Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association. *Nat. Genet.* **28,** 327–334 (2001).

14. Tong, I. L. *et al.* Transcriptional Regulatory Networks in Saccharomyces cerevisiae. *Science (80-.).* **298,** 799–804 (2002).

15. Hall, D. B., Wade, J. T. & Struhl, K. An HMG Protein , Hmo1 , Associates with Promoters of Many Ribosomal Protein Genes and throughout the rRNA Gene Locus in Saccharomyces cerevisiae An HMG Protein , Hmo1 , Associates with Promoters of Many Ribosomal Protein Genes and throughout the rRNA Gene. **26,** 3672–3679 (2006).

16. Bachand, F., Baisden, C. & Mason, P. Letters To Nature. *Nature* **432,** 1–4 (2004).

17. Wade, J. T., Hall, D. B. & Struhl, K. The transcription factor Ifh1 is a key regulator of yeast ribosomal protein genes. *Nature* **432,** 1054–1058 (2004).

18. Rudra, D., Mallick, J., Zhao, Y. & Warner, J. R. Potential interface between ribosomal protein production and pre-rRNA processing. *Mol. Cell. Biol.* **27,** 4815–24 (2007).

19. Rudra, D., Zhao, Y. & Warner, J. R. Central role of Ifh1p-Fhl1p interaction in the synthesis of yeast ribosomal proteins. *EMBO J.* **24,** 533–42 (2005).

20. Zhao, Y. *et al.* Fine-structure analysis of ribosomal protein gene transcription. *Mol Cell Biol* **26,** 4853–4862 (2006).

21. Hu, H. & Li, X. Transcriptional regulation in eukaryotic ribosomal protein genes. *Genomics* **90,** 421–423 (2007).

22. Marion, R. M. *et al.* Sfp1 is a stress- and nutrient-sensitive regulator of ribosomal protein gene expression. *Proc. Natl. Acad. Sci.* **101,** 14315–14322 (2004).

23. Jorgensen, P. *et al.* A dynamic transcriptional network communicates growth potential to ribosome synthesis and critical cell size. *Genes Dev.* **18,** 2491–2505 (2004).

24. Hariharan, N., Kelley, D. E. & Perry, R. P. Equipotent mouse ribosomal protein promoters have a similar architecture that includes internal sequence elements. *Genes Dev.* **3,** 1789–1800 (1989).

25. Meyuhas, O. & Kahan, T. The race to decipher the top secrets of TOP mRNAs. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1849,** 801–811 (2015).

26. Lam, Y. W., Lamond, A. I., Mann, M. & Andersen, J. S. Analysis of Nucleolar Protein Dynamics Reveals the Nuclear Degradation of Ribosomal Proteins. *Curr. Biol.* **17,** 749–760 (2007).

27. Blaha, G. & Ivanov, P. in *Protein Synthesis and Ribosome Structure* 53–84 (2004).

28. Nierhaus, K. H. *Ribosome Assembly. . Protein Synthesis and Ribosome Structure.* **3,** (2004).

29. Wool, I. G. Extraribosomal functions of ribosomal proteins. *Trends Biochem. Sci.* **21,** 164–165 (1996).

30. Warner, J. R. & McIntosh, K. B. How Common Are Extraribosomal Functions of Ribosomal Proteins? *Mol. Cell* **34,** 3–11 (2009).

31. Malygin, A. A., Parakhnevitch, N. M., Ivanov, A. V., Eperon, I. C. & Karpova, G. G. Human ribosomal protein S13 regulates expression of its own gene at the splicing step by a feedback mechanism. *Nucleic Acids Res.* **35,** 6414–6423 (2007).

32. Mazumder, B. *et al.* Regulated release of L13a from the 60S ribosomal subunit as a mechanism of transcript-specific translational control. *Cell* **115,** 187–198 (2003).

33. Flygare, J. *et al.* Human RPS19, the gene mutated in Diamond-Blackfan anemia, encodes a ribosomal protein required for the maturation of 40S ribosomal subunits. *Blood* **109,** 980–986 (2007).

34. Draptchinskaia, N. *et al.* The gene encoding ribosomal protein S19 is mutated in Diamond-Blackfan anaemia. *Nat. Genet.* **21,** 169–175 (1999).

35. Mirabello, L. *et al.* Novel and known ribosomal causes of Diamond-Blackfan anaemia identified through comprehensive genomic characterisation. *J. Med. Genet.* **54,** 417–425 (2017).

36. Boria, I. *et al.* The ribosomal basis of diamond-blackfan anemia: mutation and database update. *Hum. Mutat.* **31,** 1269–1279 (2010).

37. Campagnoli, M. F. *et al.* RPS19 mutations in patients with Diamond-Blackfan anemia. *Hum. Mutat.* **29,** 911–920 (2008).

38. Juli, G. *et al.* Depletion of ribosomal protein S19 causes a reduction of rRNA synthesis. *Sci. Rep.* **6,** 35026 (2016).

39. Zhang, J. *et al.* Assembly factors Rpf2 and Rrs1 recruit 5S rRNA and ribosomal proteins rpL5 and rpL11 into nascent ribosomes. *Genes Dev.* **21,** 2580–2592 (2007).

40. Dai, M. S., Sears, R. & Lu, H. Feedback regulation of c-Myc by ribosomal protein L11. *Cell Cycle* **6,** 2735–2741 (2007).

41. Dai, M. S. *et al.* Regulation of the MDM2-p53 pathway by ribosomal protein L11 involves a post-ubiquitination mechanism. *J. Biol. Chem.* **281,** 24304–24313 (2006).

42. Fumagalli, S. *et al.* Absence of nucleolar disruption after impairment of 40S ribosome biogenesis reveals an rpL11-translation-dependent mechanism of p53 induction. *Nat. Cell Biol.* **11,** 501–508 (2009).

43. Gazda, H. T. *et al.* Ribosomal Protein L5 and L11 Mutations Are Associated with Cleft Palate and Abnormal Thumbs in Diamond-Blackfan Anemia Patients. *Am. J. Hum. Genet.* **83,** 769–780 (2008).

44. Cmejla, R. *et al.* Identification of mutations in the ribosomal protein L5 (RPL5) and ribosomal protein L11 (RPL11) genes in Czech patients with diamond-blackfan anemia. *Hum. Mutat.* **30,** 321–327 (2009).

45. Quarello, P. *et al.* Diamond-blackfan anemia: Genotype-phenotype correlations in Italian patients with RPL5 and RPL11 mutations. *Haematologica* **95,** 206–213 (2010).

46. A. Victor Hoffbrand, P. A. H. M. *Hoffbrand's Essential Haematology, 7th Edition*. *Wiley-Blackwell* (2016). doi:978-1-118-40867-4

47. Scadden, D. T. The stem-cell niche as an entity of action. *Nature* **441,** 1075–9 (2006).

48. Heyworth, C. M., Dexter, T. M., Kan, O. & Whetton, A. D. The role of hemopoietic growth factors in self-renewal and differentiation of IL-3-dependent multipotential stem cells. *Growth Factors* **2,** 197–211 (1990).

49. Tsiftsoglou, A. S., Vizirianakis, I. S. & Strouboulis, J. Erythropoiesis: Model systems, molecular regulators, and developmental programs. *IUBMB Life* **61,** 800–830 (2009).

50. Fujiwara, Y., Browne, C. P., Cunniff, K., Goff, S. C. & Orkin, S. H. Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. *Proc. Natl. Acad. Sci.* **93,** 12355–12358 (1996).

51. Narla, A. & Ebert, B. L. Ribosomopathies: Human disorders of ribosome dysfunction. *Blood* **115,** 3196–3205 (2010).

52. Yelick, P. C. & Trainor, P. A. Ribosomopathies: Global process, tissue specific defects. *Rare Dis. (Austin, Tex.)* **3,** e1025185 (2015).

53. Bolze, A. *et al.* Ribosomal Protein SA Haploinsufficiency in Humans with Isolated Congenital Asplenia. *Science (80-.).* **340,** 976–978 (2013).

54. McCann, K. & Baserga, S. Mysterious ribosomopathies. *Science (80-.).* **341,** 849–850 (2013).

55. Nakhoul, H. *et al.* Ribosomopathies: Mechanisms of Disease. *Clin. Med. Insights Blood Disord.* **7,** 7–16 (2014).

56. Lipton, J. M. & Ellis, S. R. Diamond-Blackfan Anemia: Diagnosis, Treatment, and Molecular Pathogenesis. *Hematol. Oncol. Clin. North Am.* **23,** 261–282 (2009).

57. von Lindern, M. *et al.* The glucocorticoid receptor cooperates with the erythropoietin receptor and c-Kit to enhance and sustain proliferation of erythroid progenitors in vitro. *Blood* **94,** 550–559 (1999).

58. Jaako, P. *et al.* Dietary L-leucine improves the anemia in a mouse model for Diamond-Blackfan anemia. *Blood* **120,** 2225–2228 (2012).

59. Narla, A., Vlachos, A. & Nathan, D. G. Diamond Blackfan anemia treatment: Past, present, and future. *Semin. Hematol.* **48,** 117–123 (2011).

60. Gazda, H. T. *et al.* Frameshift mutation in p53 regulator RPL26 is associated with multiple physical abnormalities and a specific pre-ribosomal RNA processing defect in diamond-blackfan anemia. *Hum. Mutat.* **33,** 1037–1044 (2012).

61. Pospisilova, D. *et al.* The Czech National Diamond-Blackfan Anemia Registry: Clinical data and ribosomal protein mutations update. *Blood Cells, Mol. Dis.* **48,** 209–218 (2012).

62. Cmejlova, J. *et al.* Translational efficiency in patients with Diamond-Blackfan anemia. *Haematologica* **91,** 1456–1464 (2006).

63. Avondo, F. *et al.* Fibroblasts from patients with Diamond-Blackfan anaemia show abnormal expression of genes involved in protein synthesis, amino acid metabolism and cancer. *BMC Genomics* **10,** 442 (2009).

64. Flygare, J. *et al.* Deficiency of ribosomal protein S19 in CD34+ cells generated by siRNA blocks erythroid development and mimics defects seen in Diamond-Blackfan anemia. *Blood* **105,** 4627–4634 (2005).

65. Ebert, B. L. *et al.* An RNA interference model of RPS19 deficiency in Diamond-Blackfan anemia recapitulates defective hematopoiesis and rescue by dexamethasone: Identification of dexamethasone-responsive genes by microarray. *Blood* **105,** 4620–4626 (2005).

66. Dutt, S. *et al.* Haploinsufficiency for ribosomal protein genes causes selective activation of p53 in human erythroid progenitor cells. **117,** 2567–2576 (2012).

67. Moniz, H. *et al.* Primary hematopoietic cells from DBA patients with mutations in

RPL11 and RPS19 genes exhibit distinct erythroid phenotype in vitro. *Cell Death Dis.* **3,** e356-11 (2012).

68. Torihara, H. *et al.* Erythropoiesis failure due to RPS19 deficiency is independent of an activated Tp53 response in a zebrafish model of Diamond-Blackfan anaemia. *Br. J. Haematol.* **152,** 648–654 (2011).

69. Sankaran, V. G. *et al.* Exome Sequencing Identifies GATA1 Mutations Resulting in Diamond- - - Blackfan Anemia. *J. Clin. Invest.* **122,** 1–5 (2012).

70. Parrella, S. *et al.* Loss of GATA-1 full length as a cause of Diamond-Blackfan anemia phenotype. *Pediatr. Blood Cancer* **61,** 1319–1321 (2014).

71. Ludwig, L. S. *et al.* Altered translation of GATA1 in Diamond-Blackfan anemia. *Nat. Med.* **20,** 1–9 (2014).

72. Chlon, T. M., Mcnulty, M., Goldenson, B., Rosinski, A. & Crispino, J. D. Global transcriptome and chromatin occupancy analysis reveal the short isoform of GATA1 is deficient for erythroid specification and gene expression. **100,** 575–584 (2015).

73. Byrska-bishop, M. *et al.* Pluripotent stem cells reveal erythroid-specific activities of the GATA1 N-terminus. **125,** 993–1005 (2015).

74. Amanatiadou, E. P., Papadopoulos, G. L., Strouboulis, J. & Vizirianakis, I. S. GATA1 and PU.1 Bind to Ribosomal Protein Genes in Erythroid Cells: Implications for Ribosomopathies. *PLoS One* **10,** e0140077 (2015).

75. Amanatiadou, E. P., Papadopoulos, G. L. & Strouboulis, J. GATA1 and PU . 1 Bind to Ribosomal Protein Genes in Erythroid Cells : Implications for Ribosomopathies. 1–18 (2015). doi:10.1371/journal.pone.0140077

76. Zheng, L., Baumann, U. & Reymond, J.-L. An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.* **32,** e115 (2004).

77. Andrews, N. C. & Faller, D. V. A rapid micrqpreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. *Nucleic Acids Res.* **19,** 2499 (1991).

78. Mahmood, T. & Yang, P. C. Western blot: Technique, theory, and trouble shooting. *N. Am. J. Med. Sci.* **4,** 429–434 (2012).

79. Weiss, M. J., Mason, P. J. & Bessler, M. What ' s in a name ? *J. Clin. Invest.* **122,** 2346–2349 (2012).

80. Boultwood, J. & Pellagatti, A. Reduced translation of GATA1 in Diamond-Blackfan anemia. *Nat. Med.* **20,** 703–704 (2014).

81. Trakarnsanga, K. *et al.* An immortalized adult human erythroid line facilitates sustainable and scalable generation of functional red cells. *Nat. Commun.* **8,** (2017).