Switching genes on and off in haemopoiesis

David Garrick, Marco De Gobbi, Magnus Lynch and Douglas R. Higgs

MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Oxford OX3 9DS, U.K.

Abstract
At present, the molecular mechanisms by which stem cells commit to and differentiate towards specific lineages are poorly characterized, and will need to be better understood before stem cells can be exploited fully in experimental and clinical settings. Transcriptional regulation, the ability to turn genes on and off, lies at the heart of these processes of lineage commitment and specification. We have focused on fully understanding how these decisions are made at a single mammalian gene locus, the α-globin genes, which become up-regulated in a tissue- and developmental-stage specific manner during haemopoiesis. The studies summarized in the present article have revealed that complete regulation of this gene cluster involves not only activating mechanisms in expressing erythroid cells, but also repressing mechanisms, involving the Polycomb complex and histone deacetylases which are present in non-erythroid tissues. Taken together, these observations provide a well-characterized model of how gene expression is fully regulated during the transition from stem cells through lineage commitment and terminal differentiation.

Introduction
If we are to fully realize the enormous potential promised by stem cell biology, there is a fundamental need to understand the mechanisms by which pluripotent or multipotent stem cells commit to and differentiate towards specific lineages. At the molecular level, the commitment of a stem cell to any specific lineage from among multiple potential fates involves up-regulation of the expression of a distinct subset of genes (lineage-affiliated genes), while, at the same time, maintaining transcriptional repression of others (associated with an alternative fate). Thus the tight control of gene expression, turning genes on and off, lies at the heart of lineage commitment and differentiation. Although changes in transcription programmes in differentiating tissues can be monitored by global expression profiling [1–3], the molecular mechanisms by which individual genes become activated or repressed to set up those transcriptional profiles are still poorly understood. We have focused on fully characterizing the molecular mechanisms at work at a single genetic locus to regulate its activity during lineage specification and terminal differentiation.

The α-globin gene cluster, situated close to the telomere on human chromosome 16p, has proved to be a very useful model locus for addressing questions of gene regulation during differentiation. These genes are inactive in pluripotent ES (embryonic stem) cells and multipotent HSCs (haemopoietic stem cells), a self-renewing stem cell population which gives rise to the eight different lineages of the blood cells commit to and differentiate towards specific lineages. Initial efforts were directed towards demarcating the complete α-globin regulatory domain, containing all known erythroid-specific DNaseI-hypersensitive sites and multispecies conserved sequence motifs [5]. When this entire domain was recombineered to replace the mouse α-globin locus, it was found to be appropriately regulated during mouse erythropoiesis, confirming that all required cis-acting regulatory elements are present in this region [6]. ChIP (chromatin immunoprecipitation) and 3C (chromosome conformation capture) experiments have been carried out to document epigenetic changes, the binding of

Key words: gene switching, α-globin, haemopoiesis, stem cell, transcriptional regulation.

Abbreviations used: ChIP, chromatin immunoprecipitation; ES, embryonic stem; EZH2, enhancer of zeste homologue 2; H3ac (etc.), acetylated histone H3 (etc.); H3K4me (etc.), histone H3 methylated at Lys4 (etc.); H3K9me (etc.), histone H3 methylated at Lys9 (etc.); HDAC, histone deacetylase; HSC, haemopoietic stem cell; PRC, Polycomb repressor complex; siRNA, short interfering RNA; SUZ12, suppressor of zeste homologue 2; H3ac (etc.), acetylated histone H3 (etc.); H3K4me (etc.), histone H3 methylated at Lys4 (etc.).
key transcription factors, three-dimensional conformation and ultimate recruitment of the transcription machinery throughout this domain during erythroid differentiation in both mice and humans [7–9]. As summarized previously [10], transcriptional activation of the α-globin genes involves the sequential recruitment of key haemopoietic transcription factors initially to remote regulatory elements within the domain in multipotent progenitors. Following commitment to an erythroid fate, a conformational change results in apposition of these upstream regulatory elements with the genes themselves as part of a chromatin loop, facilitating eventual recruitment of the transcription PIC (pre-initiation complex) and RNA polymerase II to the gene promoters late in terminal erythroid differentiation. This transcriptional programme is associated with widespread epigenetic changes across the domain, in the form of activating histone modifications H3ac (acetylated histone H3), H4ac and H3K4me (histone H3 methylated at Lys4) (Figure 1 and [7,9]).

Keeping the α-globin genes off in non-erythroid cells: an unexpected role for the Polycomb complex

It has become clear that this series of up-regulatory events taking place in erythroid cells is only a part of the complete picture of α-globin regulation. More recently, we have been addressing the possibility that repressive regulatory mechanisms are also required to keep the α-globin genes inactive in non-erythroid cell types. Despite their tightly tissue-restricted expression pattern, these genes lie in a GC-rich gene-dense region of the genome (Figure 1), and the majority of surrounding genes are widely expressed. Thus, although the α-globin genes themselves are transcriptionally silent in most tissues, the region that contains them remains broadly transcriptionally active, early-replicating [11], accessible and localized away from pericentromeric heterochromatin [12] and thus bears the hallmark features of constitutively active chromatin. The promoters of all of the genes in the region, including the α-globin genes, are associated with CpG islands (Figure 1) and remain unmethylated in all tissues. The domain containing the α-globin cluster contrasts starkly with the genomic region surrounding the β-globin genes (also erythroid-specific), which is relatively gene-poor and GC-poor. Except in erythroid cells, the whole region surrounding the β-globin cluster is late-replicating, relatively inaccessible and localizes with pericentromeric heterochromatin (reviewed in [13]).

The location of the tissue-specific α-globin genes within a constitutively active chromatin domain prompted us to investigate whether these genes might be targeted by specific repression events in non-erythroid tissues. We initially surveyed repressive histone modifications (H3K9me3, a marker of constitutive heterochromatin, H4K20me and H3K27me3) across this domain in non-erythroid cells by ChIP. We observed only background levels of H4K20me and H3K9me3 (consistent with the generally euchromatic character of this broad domain). However, we observed that there was strong enrichment for the H3K27me3 modification at the inactive α-globin genes. When analysed at a high resolution by applying immunoprecipitated material to a custom tiled microarray covering a genomic region of 500 kb at 16p13.3 (ChIP-on-chip), we observed that there was a discrete block of enrichment of H3K27me3 which marked all of the genes of the α-globin cluster, with only background enrichment detected for ~80 kb on either side, in the regions containing the widely expressed flanking genes (summarized in Figure 1). The presence of this histone modification correlated well with the inactive state of the α-globin genes as it strongly marked the genes in all non-erythroid cell types tested, and was markedly reduced at or absent from the α-globin genes in primary erythroblasts and erythroid cell lines (in which the α-globin genes are expressed).

Modification of histones by H3K27me3 is associated with gene repression mediated by the PcG (Polycomb group) machinery. The modification is put in place by PRC (Polycomb repressor complex) 2, which includes the core components SUZ12 (suppressor of zeste 12 homologue) and the histone methyltransferase EZH2 (enhancer of zeste homologue 2) [14,15]. It is believed that the H3K27me3 signal deposited by PC2 then targets the marked chromatin for recruitment of a second PcG complex (PRC1) and that together these complexes exert a transcriptionally repressive activity by mechanisms that are still poorly understood [16]. Consistent with the distribution of H3K27me3, we also found, using ChIP and ChIP-on-chip, that two core components of the PRC2 complex (SUZ12 and EZH2) were localized at the inactive α-globin genes in non-erythroid cells and had been cleared from the genes upon transcriptional activation in erythroid cell types. Clearing of PcG from the α-globin genes in erythroid cells was a localized effect, as PcG binding at other targets within the region (DIST1, RGS11 and ARHGDIG) was observed in both erythroid and non-erythroid cells (Figure 1). Depletion of EZH2 or SUZ12 by siRNA (short interfering RNA) resulted in inappropriate expression of the α-globin genes in non-erythroid cell lines, confirming the functional importance of the Polycomb pathway for repressing this locus. In contrast with the α-globin locus, we found that the β-globin genes were not associated with H3K27me3 or the PRC2 complex, and were not induced by siRNA-mediated knockdown of EZH2. Thus, although these two clusters are co-ordinately up-regulated during erythroid differentiation, only the α-globin genes are dependent on the PcG machinery for appropriate repression in non-erythroid cells.

At present, the molecular mechanisms by which targeting by PRC2 brings about downstream transcriptional silencing are poorly understood. It has been demonstrated previously that components of the PRC2 complex can interact with both DNA methyltransferases [17] and HDACs (histone deacetylases) [18], suggesting that either or both genomic DNA methylation and histone deacetylation might be important for Polycomb-mediated gene repression. However, these biochemical interactions are based on
**Figure 1** | Comparison of the human α-globin locus in expressing and non-expressing cells

Schematic representation of the terminal 300 kb of human chromosome 16p, containing the α-globin (HBA) locus. Genes of the α-locus are shown in black, and other genes are in grey. Pseudogenes are in white. Genes shown above the line are transcribed towards the centromere, and those below the line are transcribed towards the telomere (black oval). The region of conserved synteny is indicated by vertical broken lines. CpG islands (black bars) are indicated. Also shown is the epigenetic state of the region in primary human pro-erythroblast cells in which the α-globin genes are highly expressed (A) and non-erythroid lymphocytes [primary T-lymphocytes or EBV (Epstein-Barr)-transformed B-lymphocytes] in which the α-globin genes are repressed (B). DNaseI-hypersensitive sites are shown as black (erythroid specific) or grey (ubiquitous) arrows. Activating (green) and repressing (red) factors (SUZ12 and HDAC1) and histone modifications (H4ac, H3ac, H3K4me2, H3K4me3, H3K27me3) were analysed by ChIP (HDAC1) or ChIP-on-chip (all others), and the results are summarized as coloured bars, with darker bars indicating stronger enrichment. The HDAC1 enrichment was analysed only by real-time PCR at the α-globin genes themselves and so has not been analysed elsewhere throughout the domain. Based on data from [9] and by D. Garrick and D.R. Higgs, unpublished work.
in vitro observations only, and the functional importance of these mechanisms has been tested at few PcG-targeted genes in healthy tissues. We found that the large CpG islands at the α-globin promoters remain in an unmethylated state in all tissues and cell lines studied, regardless of whether the α-genes are expressed, indicating that DNA methylation is not involved in the PcG-mediated repression of these genes. In contrast, we observed that HDACs did play an important role in maintaining the α-globin genes in an inactive state in non-erythroid tissues. Using ChIP, we found that HDAC1 was localized to the α-globin genes in non-expressing (T-lymphocytes), but not in expressing (pro-erythroblasts) primary cells (Figure 1). Moreover, the genes become inappropriately expressed in a range of non-erythroid cells upon treatment with HDAC inhibitors. This expression is accompanied by localized histone hyperacetylation and the formation of hypersensitive sites at the α-gene promoters. A direct link between HDACs and PcG was demonstrated by the fact that treatment of lymphocytes with the HDAC inhibitor TSA (trichostatin A) caused the SUZ12 component of PRC2 to be removed from the inactive α-globin genes. Thus HDACs, but not DNA methylation, are intimately associated with PcG-mediated repression of the α-globin genes.

**Signals for recruitment and clearing of Polycomb complexes to the α-globin genes**

Although it was originally believed that the PcG machinery constituted a robust epigenetic memory (imposed in stem cells) to stably maintain the repressed state of a gene throughout development, more recent genome-wide studies have revealed that the pattern of PcG repression is actually highly dynamic. In ES cells, PcG is particularly targeted to genes whose expression would drive commitment and differentiation towards specific cell lineages, and so which must be repressed to maintain the pluripotent undifferentiated state [19,20]. Lineage commitment is associated with the removal of PcG and activation of lineage-affiliated genes. At the same time, PcG can become recruited to other genes which are not required (or would even repress) the selected lineage. Despite these dynamic changes in PcG targeting during differentiation, at present, very little is known about the signals (both cis-acting DNA sequences and trans-acting factors) which regulate PcG-recruitment and clearance from a given locus.

In order to use the α-globin genes as a model locus to address some of these issues, we have initially investigated the dynamics of PcG recruitment to these genes and clearing during erythroid differentiation. We found that PRC2 is already engaged at the α-globin genes in pluripotent human ES cells (Figure 2). High-resolution analysis using ChIP-on-chip suggests that the PcG complex binds primarily to sequences within the large CpG islands of the α-globin genes. It is of interest that, unlike the human α-globin genes, PRC2 and the H3K27me3 modification are not detected at the orthologous mouse α-globin genes in non-expressing cell types. This was observed directly in heterozygous ‘humanized’ mouse ES cells, bearing an endogenous mouse α-globin cluster on one chromosome 11 and a targeted
human α-globin cluster at the corresponding site on the other allele [6]. We found that the human α-globin genes recruit PRC2, but the mouse α-globin genes (in the same nucleus) do not. The human and mouse clusters are contained within a syntenic block, and the genomic context and structural organization of these orthologous clusters has been highly conserved [5]. However, one of the most prominent differences between them is that the large CpG islands present at the human α-globin promoters have been eroded and are not present at the mouse α-globin genes. The differential targeting of PcG to these two clusters suggests that, in this case, the DNA elements which recruit PcG complexes may be contained within the promoter CpG islands. Early mouse embryos containing the humanized α-globin locus may prove to be a useful tool for studying the dynamics of PcG recruitment and the molecular forces driving it during very early mammalian development.

To investigate clearing of the PcG complex, we have also carried out ChIP experiments to analyse PcG binding and H3K27me3 at the α-globin genes in purified primary cell populations during erythroid commitment and differentiation. Preliminary experiments reveal that the complex is present at the α-globin genes in multipotent HSCs (which can give rise to all blood lineages) and in progressively restricted progenitor populations, but becomes rapidly cleared from the genes soon after commitment to an erythroid fate (summarized in Figure 2). It will now be of interest to correlate the exact timing of clearance of the PcG complex with the dynamics of transcription factor binding across the locus during erythroid differentiation (summarized in [10]). Furthermore, preliminary experiments are now underway using natural and engineered deletions to identify cis-acting DNA sequences throughout the locus which are required for clearing of the PcG complex.

**General conclusions**

Our findings of PcG- and HDAC-mediated repression add an unexpected new dimension to the mechanisms required for complete regulation of the well-characterized α-globin genes. The specific requirement of this repressive machinery at the α-globin, but not the β-globin, genes suggests that PcG may play a generally important role in situations where tissue-specific genes lie in gene-dense regions of open chromatin and must be silenced while neighbouring genes remain fully active. Indeed, this hypothesis is supported by genome-wide bioinformatic analyses, which confirm a bias among PcG targets to tissue-specific genes which have promoter CpG islands and which lie in gene-dense regions of the genome. For such genes, more widespread repression involving the formation of facultative heterochromatin may not be possible. In the case of α-globin, PcG-mediated repression is associated with the activity of HDACs, a functional association which has been previously observed at genes which are aberrantly silenced during malignancy, but which has been demonstrated at few endogenous genes in normal mammalian tissues. Finally, the relative accessibility of the haemopoietic system, together with the availability of the α-globin ‘humanized’ mouse makes the α-globin genes a useful model locus to experimentally address issues relating to PcG recruitment during early development and clearance during lineage commitment and differentiation.

**References**


Received 29 April 2008
doi:10.1042/BST0360613