

# NF-Y Recruits Both Transcription Activator and Repressor to Modulate Tissue- and Developmental Stage-Specific Expression of Human $\gamma$ -Globin Gene

Xingguo Zhu<sup>1</sup>, Yongchao Wang<sup>1</sup>, Wenhui Pi<sup>1</sup>, Hui Liu<sup>2</sup>, Amittha Wickrema<sup>2</sup>, Dorothy Tuan<sup>1\*</sup>

<sup>1</sup> Department of Biochemistry and Molecular Biology, Medical College of Georgia and College of Graduate Studies, Georgia Health Sciences University, Augusta, Georgia, United States of America, <sup>2</sup> Department of Medicine, University of Chicago, Chicago, Illinois, United States of America

## Abstract

The human embryonic, fetal and adult  $\beta$ -like globin genes provide a paradigm for tissue- and developmental stage-specific gene regulation. The fetal  $\gamma$ -globin gene is expressed in fetal erythroid cells but is repressed in adult erythroid cells. The molecular mechanism underlying this transcriptional switch during erythroid development is not completely understood. Here, we used a combination of in vitro and in vivo assays to dissect the molecular assemblies of the active and the repressed proximal  $\gamma$ -globin promoter complexes in K562 human erythroleukemia cell line and primary human fetal and adult erythroid cells. We found that the proximal  $\gamma$ -globin promoter complex is assembled by a developmentally regulated, general transcription activator NF-Y bound strongly at the tandem CCAAT motifs near the TATA box. NF-Y recruits to neighboring DNA motifs the developmentally regulated, erythroid transcription activator GATA-2 and general repressor BCL11A, which in turn recruit erythroid repressor GATA-1 and general repressor COUP-TFII to form respectively the NF-Y/GATA-2 transcription activator hub and the BCL11A/COUP-TFII/GATA-1 transcription repressor hub. Both the activator and the repressor hubs are present in both the active and the repressed  $\gamma$ -globin promoter complexes in fetal and adult erythroid cells. Through changes in their levels and respective interactions with the co-activators and co-repressors during erythroid development, the activator and the repressor hubs modulate erythroid- and developmental stage-specific transcription of  $\gamma$ -globin gene.

**Citation:** Zhu X, Wang Y, Pi W, Liu H, Wickrema A, et al. (2012) NF-Y Recruits Both Transcription Activator and Repressor to Modulate Tissue- and Developmental Stage-Specific Expression of Human  $\gamma$ -Globin Gene. PLoS ONE 7(10): e47175. doi:10.1371/journal.pone.0047175

**Editor:** Ramani Ramchandran, Medical College of Wisconsin, United States of America

**Received:** July 18, 2012; **Accepted:** September 10, 2012; **Published:** October 10, 2012

**Copyright:** © 2012 Zhu et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by the National Institutes of Health (HL 89519 and MD003383 to D.T., CA98550 to A.W.), and the Giving Tree Foundation to A.W. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: dtuanlo@georgiahealth.edu

## Introduction

The human  $\beta$ -like globin genes, consisting of embryonic  $\epsilon$ -, fetal  $\gamma$ - and  $A\gamma$ - and adult  $\delta$ - and  $\beta$ -globin genes, are expressed in erythroid cells and undergo an ordered developmental switching program:  $\epsilon$ -globin gene is expressed in early embryos until  $\sim 7$  weeks in gestation, when it is switched off and  $\gamma$ -globin genes are switched on; at the time of birth,  $\gamma$ -globin genes are switched off and  $\beta$ -globin gene is switched on. The mechanisms of transcriptional activation and silencing of human  $\gamma$ -globin genes have been under intensive investigation because of the clinical importance of  $\gamma$ -globin gene re-activation in the adult erythroid cells of sickle cell disease and  $\beta$ -thalassemia patients in ameliorating the symptoms of the diseases. Many pharmacological compounds have been found to re-activate fetal  $\gamma$ -globin gene in adult erythroid cells [1]. However, the mechanism by which these compounds re-activate  $\gamma$ -globin gene is not fully understood, as the molecular mechanism of  $\gamma$ -globin gene activation and subsequent inactivation during erythroid development has not been clearly established.

Large-scale sequence analysis of human promoters finds that over 60% of annotated human promoters contain the CCAAT motifs, which are located frequently in the proximal promoter regions, 40–120 bases upstream of the TATA box [2,3]. The CCAAT motifs are present in the proximal promoter regions of all globin genes

(GenBank U01317, NG000006 and X14061). The CCAAT motif binds transcription factor NF-Y and plays a critical role in transcriptional activation and developmental switching of  $\gamma$ -globin genes [4–6]. NF-Y is a ubiquitously expressed protein complex composed of NF-YA, -YB and -YC subunits; YB and YC first associate through their histone fold domains to form a dimer, which then recruits YA—the regulatory subunit; the trimeric NF-Y through a DNA binding domain in YA binds to the CCAAT motif with specificity and affinity among the highest for DNA-binding transcription factors [3,7,8]. Among globin promoters, the proximal  $\gamma$ -globin promoter uniquely contains two tandem CCAAT motifs; in addition, it contains DNA motifs that bind erythroid transcriptional activators: GATA-2, as shown in this study, and the CP2/NF-E4 complex [9] and also general transcription repressors: BCL11A [10,11] and COUP-TFII, an orphan nuclear receptor [12–15] (Fig. 1A). Base mutations in the DNA motifs in the proximal  $\gamma$ -globin promoter cause hereditary persistence of fetal hemoglobin (HPFH) [16–20], indicating the functional importance of these DNA motifs and the transcription factors recruited by them in regulating  $\gamma$ -globin gene expression. However, how these DNA-binding transcription factors interact to assemble the active  $\gamma$ -globin proximal promoter complex in fetal erythroid cells and the inactive promoter complex in adult erythroid cells is largely unknown.



efficiencies as described [25,26]. For transduction protocol of GFP-lentiviruses into primary erythroid cells, FACS analysis and sorting of the transduced GFP-fluorescent cells, see Methods S1.

#### RNA Isolation, RT-PCR, Western Blots, Transfection and EMSA

RNA isolation, RT-PCR, Western blots, transfection and EMSA were carried out as described [25–27]. Human fetal brain RNA was from Biochain (Cat. # R124414).

#### ChIP and Re-ChIP

ChIP and Re-ChIP were carried out as described [25,28]. In ChIP,  $10^6$  cells were used for each antibody pull-down and the pulled down chromatin was quantified by PCR.

#### In vivo and in vitro co-IP, EMSA Probes, PCR Primers and Antibodies

See Methods S1.

### Results

#### Expression Profiles of Human Globin Genes and Transcription Factors that Bind to the Proximal $\gamma$ -globin Promoter in Primary Human Fetal and Adult Erythroid Cells

To study regulation of tissue- and developmental stage transcription of  $\gamma$ -globin gene, we investigated the proximal  $\gamma$ -globin promoter complex in human fetal liver erythroid cells, in which the  $\gamma$ -globin promoter is active, and adult erythroid cells cultured from CD34+ hematopoietic stem/progenitor cells for up to 14 days, during which, the progenitor cells undergo terminal differentiation to become mature, adult erythroid cells [24], in which the  $\gamma$ -globin gene is silenced. To molecularly characterize the cell systems, we determined the expression profiles in these primary cells of globin genes and transcription activators NF-Y, GATA-2 and CP2/NF-E4 and repressors COUP-TFII, GATA-1 and BCL11A that bind to the proximal  $\gamma$ -globin promoter (Fig. 1A).

The transcription profiles of globin mRNAs in fetal liver and adult erythroid cells demonstrated that these primary erythroid cells recapitulated the developmental globin gene switching program: In fetal liver erythroid cells,  $\gamma$ -globin genes were expressed at a high level but  $\beta$ -globin gene at a very low level, at ~1% that of  $\gamma$ -globin genes; in contrast, in adult erythroid cells cultured for 14 days from CD34+ cells,  $\gamma$ -globin genes were suppressed and expressed at ~1% that of the actively expressed  $\beta$ -globin gene (Fig. 1B,  $\gamma$ - and  $\beta$ -globin gene panels). In parallel with the decline of  $\gamma$ -globin gene expression during fetal to adult erythroid development, expression of transcription activators NF-YA, the regulatory and limiting subunit of NF-Y [3], GATA-2 and NF-E4 also declined (Fig. 1B and 1C). However, expression of repressors COUP-TFII, GATA-1 and BCL11A remained relatively stable or increased during fetal to adult erythroid development (Fig. 1C). The expression profile of BCL11A protein with differently-sized isoforms was complex: In fetal erythroid cells, only the median-sized isoforms were expressed at low levels; in Day 14 adult erythroid cells, the long-, median- and short-isoforms were all more abundantly expressed (Fig. 1B).

The human erythroleukemia K562 cells expressed  $\gamma$ - but not  $\beta$ -globin gene (Fig. 1B, top panels), like fetal erythroid cells, and the expression profiles of the transcription factors were also similar to those of fetal erythroid cells (Fig. 1B and 1C). Therefore, K562 cells were used in place of fetal erythroid cells for subsequent

transfection experiments due to the availability of large numbers of easily transfectable cells.

#### CCAAT and GATA Motifs Activate and GGCCGG Motif Represses $\gamma$ -globin Promoter Activity in Erythroid Cells Regardless of their Developmental Stages

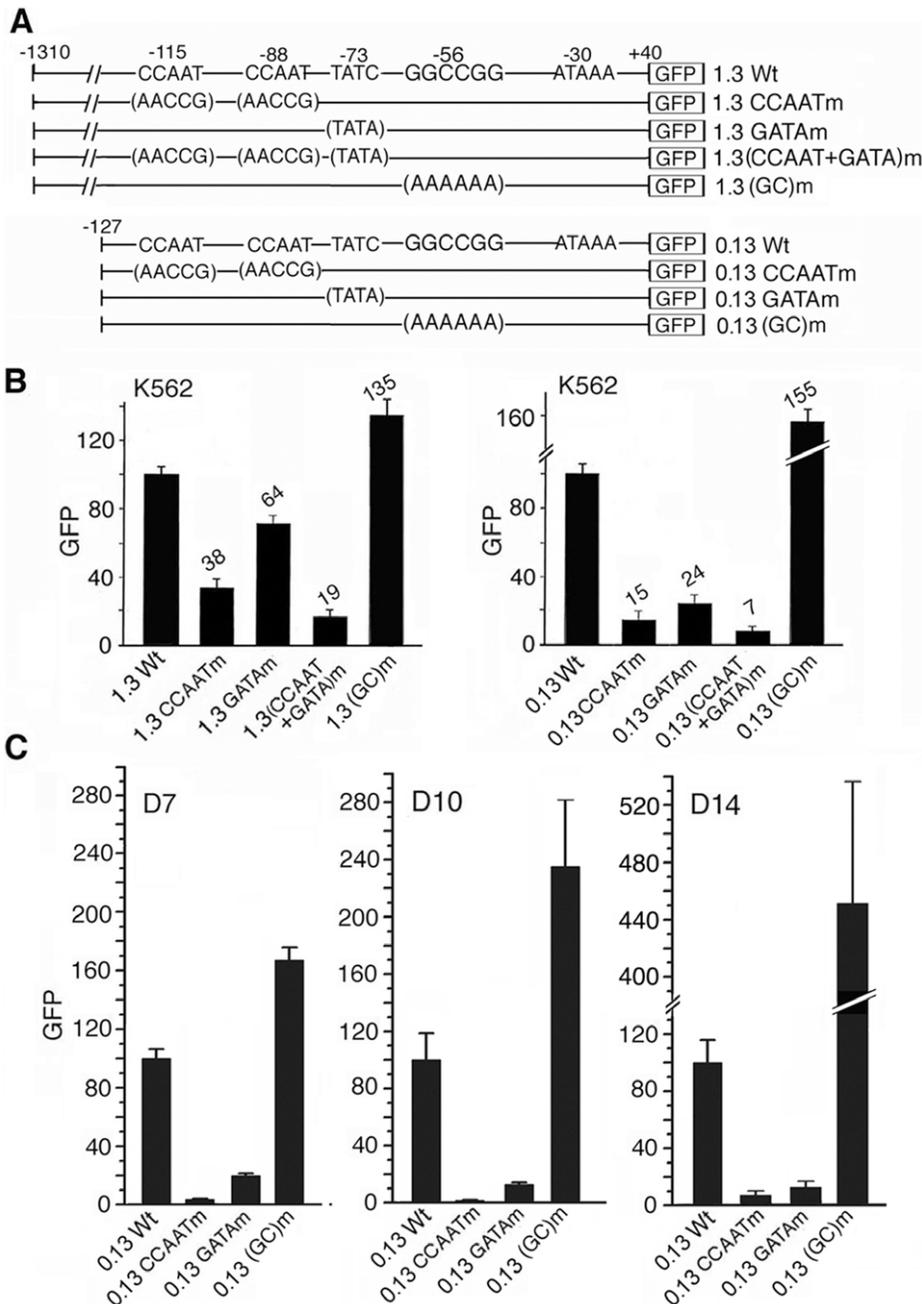
To functionally dissect the proximal  $\gamma$ -globin promoter, we focused on three DNA motifs near the TATA box: the tandem CCAAT motifs with flanking bases, which bind both activator NF-Y and repressor COUP-TFII, the -73 GATA site, whose functional significance was unknown until this study and the -56 GGCCGG motif, which binds repressor BCL11A [10,11] (Fig. 1A). To determine the contribution of these motifs to promoter activity by transfection assays, we generated two sets of GFP reporter plasmids containing either the short 0.13 kb proximal promoter or the long 1.3 kb  $\gamma$ -globin promoter spanning the distal as well as the proximal promoter regions (Fig. 2A). In these wildtype (Wt) plasmids, mutations were introduced into the tandem CCAAT motifs to eliminate the binding site for NF-Y, which also eliminated the binding site for COUP-TFII, as the two binding sites overlapped at the CCAAT motif (Fig. 1A), or into the GATA or the GGCCGG motifs to eliminate binding of GATA factors or BCL11A (Fig. 2A). The plasmids and recombinant lentiviruses were transfected/transduced into K562 and primary adult erythroid cells, respectively.

The results showed that mutations of the CCAAT or the GATA motif significantly reduced the activities of the  $\gamma$ -globin promoter in both K562 and D7–14 primary adult erythroid cells (Fig. 2B and 2C). Thus, these two motifs bound transcription activators in both K562 and adult erythroid cells. Since the CCAAT mutation eliminated the binding of both activator NF-Y and repressor COUP-TFII, the drastically reduced activity of the mutant CCAAT promoter indicated that the CCAAT motif bound predominantly activator NF-Y over repressor COUP-TFII in both K562 and adult erythroid cells.

In contrast, base mutations in the GGCCGG motif drastically increased  $\gamma$ -globin promoter activity in both K562 and Day 7–14 adult erythroid cells (Fig. 2B and 2C), indicating that this motif was a repressive site in both K562 and adult erythroid cells. Since K562 cells, like fetal erythroid cells, expressed only BCL11A-M isoforms, these isoforms thus also bound to the GGCCGG motif and repressed  $\gamma$ -globin promoter activity. The results indicated that CCAAT and GATA motifs bound transcription activators and GGCCGG motif bound transcription repressors in erythroid cells regardless of their developmental stages.

#### NF-Y Bound at the Tandem CCAAT Motifs Recruits and Stabilizes Binding of GATA-2 and BCL11A to the Neighboring GATA and GGCCGG Motifs, which Inherently cannot Bind or Bind Weakly the Respective Transcription Factors

As it had not been shown whether NF-Y and BCL11A indeed bound to the CCAAT and GGCCGG motifs and which GATA factor(s) bound to the GATA motif in primary human erythroid cells, we carried out electrophoretic mobility shift assays (EMSA) with nuclear extracts from human fetal liver, K562 cells and Day 14 adult human erythroid cells, using as probe the proximal  $\gamma$ -globin promoter of 94 bases from the CCAAT motifs to the CP2/NF-E4 site (Fig. 1A). With nuclear extracts from fetal liver and K562 cells, the probe generated three bands (Fig. 3 A and 3B, lanes 1). The top band was generated by NF-Y binding to the CCAAT motifs: the competitor spanning the tandem CCAAT motifs abolished most of the top band and NF-YA antibody



**Figure 2. Function of the tandem CCAAT, GATA and GGCCGG motifs in regulating proximal  $\gamma$ -globin promoter activity, determined in GFP reporter plasmids by transfection/transduction assays. A.** Plasmid maps of Wt and mutant long 1.3 kb and short, proximal 0.13 kb  $\gamma$ -globin promoter coupled to GFP gene. Mutated bases in the motifs were shown in parentheses. **B.** Transfection into K562 cells of GFP plasmids containing the long and the short promoters, left and right panels respectively. Numbers on top of the bars: fluorescence levels of GFP expressed from the mutant plasmids compared to those from Wt plasmids set at 100. Values were averages of two independent transfection experiments. **C.** Transduction into adult erythroid cells of recombinant lentiviruses containing the Wt and mutant proximal  $\gamma$ -globin promoters coupled to GFP gene. Cells were transduced on day 3 and GFP fluorescence was determined on day 7, 10 and 14 of culture, as shown in the 3 panels respectively. The GFP level in control cells transduced with the wildtype 0.13 kb  $\gamma$ -globin promoter-GFP lentivirus was set at 100. Values were averages of two independent transduction experiments.

doi:10.1371/journal.pone.0047175.g002

completely super-shifted this band (Fig. 3A and 3B, lanes 3 and 7); however, antibodies to C/EBP $\gamma$  or C/EBP $\delta$ , which could potentially bind to the CAAT motif, did not supershift the band (not shown). A small portion of this band contained BCL11A, since BCL11A antibody supershifted ~30% of the NF-Y band (Fig. 3A and 3B, lanes 11; Fig. S1A). The GGCCGG motif serving

as competitor, however, did not appreciably diminish the intensity of the NF-Y/BCL11A band (Fig. 3A and 3B, lanes 6), indicating that the GGCCGG motif by itself, even at 100x molar excess than the probe, was unable to bind significant amount of BCL11A to diminish the intensity of the NF-Y/BCL11A band. The GGCCGG motif did not bind Sp1, since Sp1 antibody did not

supershift the NF-Y/BCL11A band (Fig. 3A&B, lane 10). Thus, NF-Y bound at the tandem CCAAT motifs appeared to recruit and stabilize binding of BCL11A to the GGCCGG site. On the other hand, the proximal  $\gamma$ -globin promoter with mutated GGCCGG motif bound much less BCL11A than did the Wt  $\gamma$ -promoter probe (Fig. S1B), indicating cooperation of the GGCCGG motif with the CCAAT motif in recruiting BCL11A to the proximal promoter. Since the NF-YA antibody completely super-shifted the NF-Y/BCL band (Fig. 3A and 3B, lanes 3 and 7), NF-Y and BCL11A appeared to co-exist in the same protein complex.

As NF-Y and COUP-TFII bound to the cognate sites that overlap at the distal CCAAT motif (Fig. 1A), the relative strengths of their binding were determined by competition EMSAs using as competitors the distal CCAAT motif (dCCAAT), which contained the COUP-TFII binding site, and the proximal CCAAT motif (pCCAAT), which did not contain the COUP-TFII site. The results showed that competitors dCCAAT and d+p CCAAT spanning both the COUP-TFII and NF-Y sites were 10% more efficient competitors than pCCAAT and E1CCAAT, spanning only the NF-Y site (Compare the NF-Y/COUP-TFII band intensities of lanes 3–6, Fig. 3D). Quantification of the NF-Y/COUP-TFII band intensities showed that the distal CCAAT motif was occupied predominantly by NF-Y, since COUP-TFII binding constituted only  $\sim$ 10% of the NF-Y band (Fig. S1C).

The faster migrating band below the NF-Y band was generated by GATA-1 and -2, since the (GATA)7 competitor that binds GATA-1 and -2 [29] abolished this band, and both GATA-1 and -2 antibodies supershifted this band (Fig. 3A and 3B, lanes 5, 8 and 9), whereas GATA-3 antibody did not do so (not shown). Again, the oligonucleotide competitor spanning the  $-73$  GATA site failed to diminish the intensity of the GATA band (Fig. 3A and 3B, lanes 4), and when used as a probe it did not bind the GATA factors to generate a GATA band (Fig. S1D). Thus, the  $-73$  GATA site in the absence of the tandem CCAAT motifs was unable to bind the GATA factors. However, the proximal  $\gamma$ -globin promoter with mutated GATA motif bound much less GATA-2/-1 than did the Wt  $\gamma$ -promoter probe (Fig. S1D), indicating cooperation of the GATA motif with the CCAAT motif in recruiting GATA-2/-1 to the proximal promoter.

It was curious to note that although GATA-2 binding to the  $-73$  GATA site required NF-Y bound at the neighboring CCAAT motif, suggesting that the GATA-2 EMSA band should also contain NF-Y bound at the neighboring CCAAT motif, yet the GATA-2 EMSA band did not appear to contain NF-Y, since the NF-Y antibody did not abolish or super-shift the GATA band (see Fig. 3A and 3B, lanes 7–9). One possible explanation for this apparent paradox was that the DNA/protein interaction detected in EMSA was not in a static state but in a dynamic equilibrium of protein/DNA and protein/protein interactions, as discussed previously [25]. In this network of dynamic association and dissociation reactions, NF-Y was first recruited to the CCAAT motif in the  $\gamma$ -globin promoter probe; the NF-Y protein/CCAAT DNA complex then recruited GATA-2 to the adjacent weak  $-73$  GATA site to form a stable GATA-2 protein/GATA DNA complex; NF-Y was subsequently dissociated from the neighboring CCAAT site to bind to other un-occupied CCAAT sites in the free promoter probe, since the free probe was present in a large molecular excess over the transcription factor molecules in the EMSA reaction. Thus, the excess free probe was able to drive the dynamic EMSA reactions to an equilibrium, where the occupied probes contained predominantly only one bound protein per probe to generate either the GATA-2/-1 or the NF-Y band observed in the EMSA gels.

In D14 adult erythroid cells, the proximal promoter probe generated only an NF-Y band, since the tandem CCAAT motifs as competitors abolished the band and NF-Y antibody completely super-shifted the band (Fig. 3E, lanes 3 and 7). Approximately 40% of the NF-Y band contained BCL11A, since BCL11A antibody supershifted  $\sim$ 40% of the band (Fig. 3E, lanes 7 and 11; Fig. S1A). Even though D14 adult erythroid cells expressed abundant GATA-1 protein (Fig. 1C), GATA-1 was not recruited by NF-Y to the  $-73$  GATA site to generate a GATA band (Fig. 3E, lanes 1, 5, 8 and 9), indicating that in the absence of a detectable level of GATA-2 in D14 cells (Fig. 1C), GATA-1 was not recruited to the  $-73$  GATA site. Thus, GATA-1 recruitment appeared to require GATA-2 bound at the  $-73$  GATA site. Indeed, addition to D14 nuclear extract of myc-tagged GATA-2, but not myc-tagged GATA-1, generated a clear GATA band (Fig. 3G, lanes 2 and 6). As this GATA band was completely abolished by either GATA-2 or -1 antibody (Fig. 3G, lanes 3&4 and 7&8), GATA-2 and -1 appeared to form a hetero-dimer at the  $-73$  GATA site.

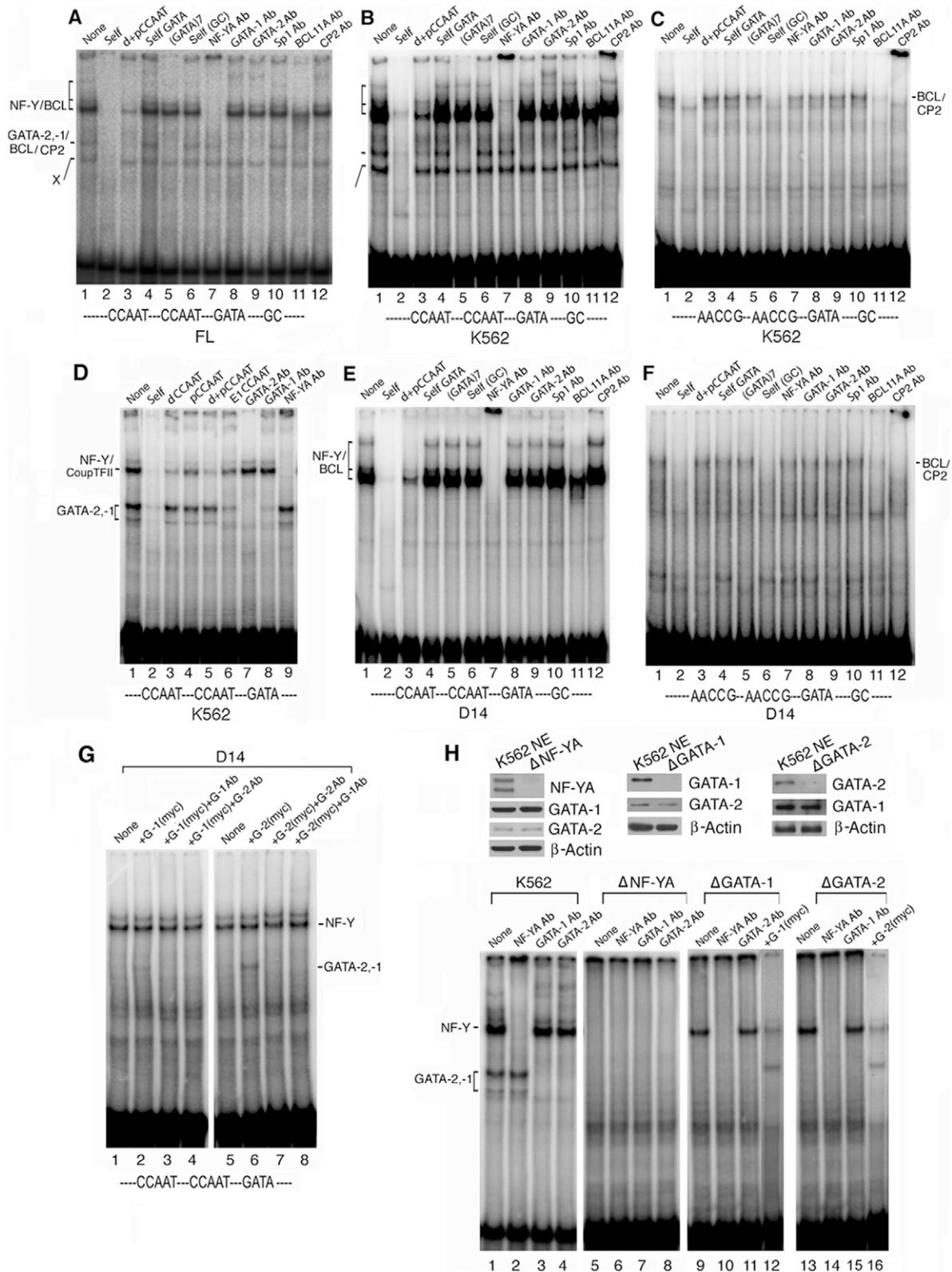
With both K562 and D14 nuclear extracts, the mutant proximal promoter probe containing mutated CCAAT motifs generated no NF-Y band as anticipated, but also no GATA band and only two faint bands that were super-shifted by BCL11A and CP2 antibodies (Fig. 3C and 3F, lanes 3, 7–11; Fig. S1A). This again demonstrated that in the absence of NF-Y bound at the CCAAT motifs, the mutant promoter containing normal GATA and GGCCGG motifs was unable to efficiently bind GATA-2/-1 and BCL11A in both K562 and D14 erythroid cells.

To further investigate the role of NF-Y in recruiting GATA-2 and/or -1, we carried out ID-EMSA with K562 nuclear extract from which NF-Y, GATA-2 or -1 was immuno-depleted (Fig. 3H, top). In ID-EMSA with the NF-YA depleted ( $\Delta$ NF-YA) nuclear extract, the proximal  $\gamma$ -globin promoter probe generated neither the NF-Y band nor the GATA bands (Fig. 3H, lane 5), although GATA-2 and -1 proteins were present in the  $\Delta$ NF-YA extract (Fig. 3H, top left panel). This result again confirmed that recruitment of GATA-2/-1 to the  $-73$  GATA site required NF-Y bound at the neighboring CCAAT motifs. In contrast, the  $\Delta$ GATA-1 and -2 nuclear extracts generated the NF-Y band with the  $\gamma$ -globin promoter probe (Fig. 3H, lanes 9 and 13), indicating that NF-Y binding to the promoter probe did not require presence of either GATA-2 or -1. Addition of myc-tagged GATA-1 or -2 to these nuclear extracts regenerated the GATA band (Fig. 3G, lanes 12 and 16).

Together, the EMSA results indicated a hierarchical binding order of the transcription factors to the proximal promoter: NF-Y bound strongly at the CCAAT motifs formed a stable anchor that subsequently recruited GATA-2 and BCL11A to their cognate sites; the bound GATA-2 in turn recruited GATA-1. The results also indicated that the proximal  $\gamma$ -globin promoter bound both transcription activator NF-Y and repressor BCL11A in both fetal liver and adult erythroid cells, regardless of whether the  $\gamma$ -globin promoter was active or repressed in the respective erythroid cells.

### NF-Y and GATA-2 Activate and GATA-1 and BCL11A Repress $\gamma$ -globin Promoter Activity

To determine the functional roles of NF-Y, GATA-2 and -1 and BCL11A on  $\gamma$ -globin promoter activity, we over-expressed or knocked-down by siRNA the respective transcription factors and measured the resultant effects on transcription of the endogenous  $\gamma$ -globin mRNA in K562 and in primary adult erythroid cells. The transcription factors were either over-expressed or siRNA knocked down with a co-expressed GFP selectable marker gene in a lentiviral vector for NF-YA, GATA-1 and -2 or as a GFP fusion



**Figure 3. Identification of transcription factors that bind to the proximal  $\gamma$ -globin promoter as determined by EMSA. A, B & E.** Wt proximal  $\gamma$ -globin promoter probe with nuclear extracts from human fetal liver, K562 cells and D14 adult erythroid cells, respectively. Top labels: Competitors and antibodies used in EMSAs. Self, self CCAAT, self GATA, (GATA)7 and self GC were respectively 100x molar excess of unlabeled probe, distal -115 and proximal -88 CCAAT motifs, the -73 GATA motif, seven tandem GATA motifs (29), and GC-rich bases spanning GGCCGG motif. Margin labels: bands generated by the transcription factors as marked. **C & F.** Mutant CCAAT proximal promoter probe with K562 and D14 nuclear extracts; all other labels were the same as in **A**. **D.** Binding affinity of COUP-TFII to its cognate site overlapping the distal CCAAT motif in the proximal

$\gamma$ -globin promoter: Competition EMSAs of proximal  $\gamma$ -globin promoter probe spanning the tandem CCAAT motifs to the GATA site (base positions: -120 to -62 in Fig. 1A) with K562 nuclear extract. Competitors d, p and d+p CCAAT: distal and proximal CCAAT motif and distal-proximal CCAAT motifs in the  $\gamma$ -globin promoter; E1CCAAT: CCAAT motif in the ERV-9 LTR enhancer upstream of the  $\beta$ -globin gene locus (25). For sequences of the competitors, see Methods S1. **G.** Binding affinities of GATA-1 and -2 to the proximal  $\gamma$ -globin promoter: EMSA of Wt proximal promoter probe with D14 nuclear extract. Lanes 2 & 6, 3 & 7, 4 & 8: myc-tagged GATA-1, G-1(myc), and myc-tagged GATA-2, G-2(myc), added respectively to D14 nuclear extract alone, and with GATA-1 or -2 antibody. **H.** Binding of GATA-1 and -2 to the proximal  $\gamma$ -globin promoter required presence of NF-Y determined by EMSA with immuno-depleted K562 nuclear extract (ID-EMSA). Upper panel: Western blots of K562 nuclear extracts immuno-depleted with antibody to NF-YA, GATA-1 and -2,  $\Delta$ NF-YA,  $\Delta$  GATA-1 and  $\Delta$ GATA-2 respectively. Lower panel: EMSA of Wt proximal promoter probe spanning the tandem CCAAT motifs to the -73 GATA site with wildtype,  $\Delta$ NF-YA,  $\Delta$  GATA-1 and  $\Delta$ GATA-2 K562 nuclear extracts respectively. doi:10.1371/journal.pone.0047175.g003

gene in a plasmid vector for BCL11A. The transfected K562 cells and the transduced Day 14 adult erythroid cells were sorted by FACS (Fig. 4A). The sorted fluorescent cells were then used for Western blot to assess the degree of over-expression or knockdown of the respective transcription factors (Fig. 4B) and qRT-PCR to determine the subsequent effects on transcription of the endogenous  $\gamma$ -globin mRNA. The protein and RNA analyses showed that over-expression of NF-YA, either the long NF-YAL or the short NF-YAS isoform (see Fig. 1C), and of GATA-2 activated while over-expression of GATA-1 suppressed transcription of  $\gamma$ -globin mRNA in both K562 and D14 adult erythroid cells (Fig. 4C and 4D; NF-YAS results not shown). Thus, NF-Y and GATA-2 were activators while GATA-1 was a repressor of  $\gamma$ -globin promoter activity in both K562 and D14 cells independent of the developmental stages of the cells. In addition, over-expression of each of the four BCL11A isoforms—XL, L, M and S—in K562 cells all suppressed transcription of  $\gamma$ -globin mRNA, with the XL isoform being the most effective repressor (Fig. 4C), in agreement with earlier findings [10,30]. Consistent with the over-expression results, knockdown of NF-Y and GATA-2 conversely suppressed whereas knockdown of GATA-1 conversely activated transcription of  $\gamma$ -globin mRNA (Fig. 4C). Thus, NF-Y and GATA-2 were activators and GATA-1 and BCL11A were repressors of  $\gamma$ -globin promoter activity.

### Over-expression of Transcription Activator NF-Y or GATA-2 or Repressor BCL11A or GATA-1 Changes the *in vivo* Assembly of the Proximal $\gamma$ -globin Promoter Complex

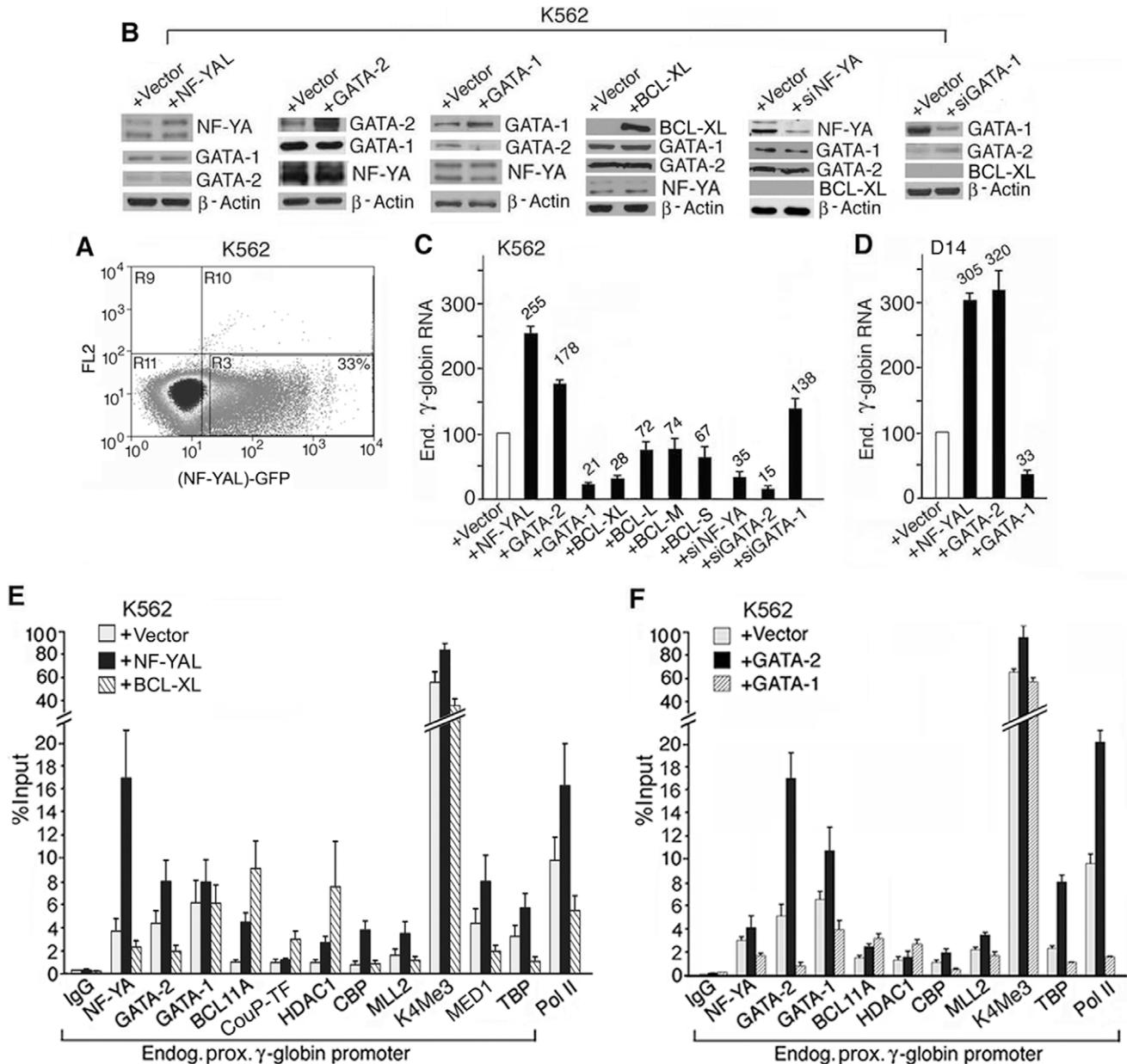
To determine whether changes in transcription of  $\gamma$ -globin mRNA induced by the over-expressed transcription activators or repressors were correlated with corresponding changes in the *in vivo* assembly of the proximal  $\gamma$ -globin promoter complex, we carried out ChIP assays on the sorted fluorescent cells over-expressing each of the transcription factors. ChIP results showed that over-expression of NF-YA, the regulatory and limiting subunit of NF-Y [3], increased occupancy of NF-Y on the proximal  $\gamma$ -globin promoter (Fig. 4E). As a result, GATA-2 occupancy also increased (Fig. 4E), due likely to increased recruitment by NF-Y of GATA-2 to the -73 GATA site, even though the level of GATA-2 was not increased by over-expression of NF-Y (Fig. 4B, first panel). Associated with the increased occupancies of activators NF-Y and GATA-2, occupancies increased also for the following co-activators (Fig. 4E): CBP with histone acetyltransferase activity [21] that can bind to GATA-2 [31], MLL2 with histone methyltransferase activity [22], which could be recruited to the globin gene locus by the transcription activators [32] to methylate lysine 4 in histone 3 and generate the H3K4me3 chromatin mark associated with actively transcribed gene locus, and Mediator 1, a member of the Mediator complex that has been shown to interact with GATA-2 in mouse erythroid and thyroid cells [33,34] and may mediate signal transmission between the activators and the basal transcription machinery. Occupancies of TBP and Pol II that could be recruited by NF-Y, GATA-2 and Mediator 1 to the

TATA box [3,35] also increased at the proximal  $\gamma$ -globin promoter (Fig. 4E). Thus, over-expression of NF-Y increased occupancy of NF-Y, which in turn increased occupancies of GATA-2, the co-activators and TBP and Pol II, resulting in increased transcription of  $\gamma$ -globin mRNA (Fig. 4C).

To determine whether co-activators MLL2 and Mediator 1 indeed regulated transcription of  $\gamma$ -globin gene, the expression of MLL2 or Mediator 1 was knocked down by the respective siRNAs in K562 cells. Transcription of  $\gamma$ -globin mRNA was reduced by 60–80% as a result of the knockdown of either MLL2 or Mediator 1 (Fig. S2), indicating the critical importance of MLL2 and MED 1 in transcriptional regulation of  $\gamma$ -globin promoter activity in human erythroid cells.

Over-expression of NF-YA increased occupancies also of repressor BCL11A and co-repressor HDAC1 at the proximal promoter (Fig. 4E), due likely to recruitment of BCL11A by the increased occupancy of NF-Y, since NF-Y bound at the tandem CCAAT motifs was able to interact with and recruit BCL11A to the neighboring GGCCGG motif, as shown by EMSA (Fig. 3A, 3B and 3E); increased occupancy of BCL11A in turn recruited more HDAC1 due likely to BCL11A/HDAC1 interaction [30]. Since BCL11A-M, the major isoform detectable in K562 cells (Fig. 1C), was a weak transcriptional repressor (Fig. 4C), presence of the weak repressor complex BCL11A-M/HDAC1 in the proximal promoter complex did not overcome the activity of the induced activator complex NF-Y/GATA-2, which activated transcription of  $\gamma$ -globin mRNA at a higher level as a result of NF-Y over-expression (Fig. 4C). Consistent with these findings, knockdown of NF-YA in K562 cells decreased occupancy of not only NF-Y but also GATA-2; however, NF-Y knockdown caused not a decrease in occupancy of BCL11A as anticipated from NF-Y/BCL11A interaction but a slight increase in BCL11A occupancy (Fig. S3A). This was likely due to strong interaction of BCL11A with COUP-TFII [11], whose occupancy at the proximal  $\gamma$ -globin promoter increased as a result of NF-Y knockdown (Fig. S3A). Since COUP-TFII binding site overlapped with the NF-Y binding site in the proximal  $\gamma$ -globin promoter (Fig. 1A), COUP-TFII was able to bind competitively to the proximal promoter at a higher level as a result of the decrease in occupancy of NF-Y. Thus, NF-Y knockdown decreased occupancies of activators NF-Y and GATA-2 and co-activator MLL2 but increased occupancies of repressors COUP-TFII and BCL11A and co-repressor HDAC1 at the proximal promoter (Fig. S3A), leading to transcriptional suppression of  $\gamma$ -globin mRNA (Fig. 4C).

On the other hand, over-expression of strong repressor BCL11A-XL in K562 cells not only increased occupancies of repressors BCL11A and COUP-TFII and co-repressor HDAC1 but also reduced occupancies of activators NF-Y and GATA-2 and co-activators CBP, MLL2 and MED1 and TBP and Pol II (Fig. 4E). The combination of the strong repressor complex BCL-XL/COUP-TFII/HDAC1 with the weakened activators/co-activators complex thus drastically reduced transcription of  $\gamma$ -globin mRNA due to BCL-XL over-expression (Fig. 4C).



**Figure 4. Over-expression of NF-YA, GATA-2, -1 or BCL11A activates or represses  $\gamma$ -globin promoter activity and induces corresponding changes in the *in vivo* assembly of proximal  $\gamma$ -globin promoter complex.** **A.** Fluorescent K562 cells over-expressing NF-YA long isoform (YAL) with co-expressed GFP sorted by FACS. Lower right quadrant: sorted GFP-fluorescent cells comprising 33% of total cell population used for analyses in panels **B–E**. FACS sorting of K562 and D14 cells over-expressing the other three transcription factors (not shown). **B.** Western blots of proteins isolated from K562 cells over-expressing NF-YAL, GATA-2, -1 or BCL-XL and K562 cells in which NF-YA or GATA-1 was knocked down by the respective siRNA, siNF-YA or siGATA-1. +Vector: control K562 cells transduced or transfected by the empty vector. **C & D.** Effects of over-expression/knockdown of NF-YAL, GATA-2, -1 or various isoforms of BCL11A on the transcription level of endogenous  $\gamma$ -globin mRNA in K562 and D14 adult erythroid cells determined by real-time RT-PCR. The level of  $\gamma$ -globin mRNA in K562 and D14 cells transduced/transfected with the vector was set at 100. Levels of  $\gamma$ -globin mRNA in test samples were averages of two independently transduced/transfected cells. Effects of over-expressing NF-YA short isoform, NF-YAS, on transcription of  $\gamma$ -globin mRNA in K562 and D14 cells were similar to the effects of over-expressing NF-YAL (not shown). **E & F.** ChIP assays of the endogenous  $\gamma$ -globin promoter in K562 cells over-expressing NF-YAL, BCL-XL, GATA-2 or -1. The NF-YA and BCL11A antibodies used in the ChIP assays recognized all isoforms of the respective transcription factor (Fig. 1C) and should pull down chromatin associated with all the isoforms of NF-YA and BCL11A. ChIP values were averages of two independent pull-down assays. doi:10.1371/journal.pone.0047175.g004

Over-expression of activator GATA-2 and repressor GATA-1 induced respective changes in occupancies of the transcription activators, repressors, co-activators and co-repressors (Fig. 4F) similar to the respective changes induced by over-expression of NF-Y and BCL11A (Fig. 4E). Notably, over-expression of GATA-

2 increased occupancy not only of GATA-2 but also of GATA-1 on the proximal  $\gamma$ -globin promoter (Fig. 4F), even though GATA-1 expression remained the same in cells over-expressing GATA-2 as in control cells (Fig. 4B, 2<sup>nd</sup> panel from left). Thus, over-expression of GATA-2 resulted in higher occupancy of not only

GATA-2 but also GATA-1. Paradoxically, over-expression of GATA-1 decreased occupancy of GATA-1 as well as of GATA-2 (Fig. 4F). Since GATA-2 expression was suppressed by GATA-1 over-expression (Fig. 4B, 3<sup>rd</sup> panel from left), the decreased cellular level of GATA-2 reduced occupancy of GATA-2, which led in turn to decreased recruitment of GATA-1, despite GATA-1 over-expression. Thus, GATA-1 recruitment to the proximal  $\gamma$ -globin promoter depended on GATA-2 occupancy, in agreement with earlier EMSA results (Fig. 3E and 3G).

It could be argued that the changes in chromatin occupancies due to over-expression of GATA-2 and -1 might not have occurred at the proximal  $\gamma$ -globin promoter as was interpreted above but occurred at the multiple GATA sites in the distal promoter immediately upstream of the proximal promoter, since the *in vivo* ChIP assays might not be able to differentiate between the distal and proximal GATA sites separated by short distances of  $\sim$ 100 DNA bases. To address this question, we carried out ChIP assays of transfected plasmid containing no distal promoter and only the 0.13 kb proximal  $\gamma$ -globin promoter linked to the GFP reporter gene. The results showed that over-expression of GATA-1 or -2 caused similar changes in the occupancies of transcription factors and co-factors on the transfected 0.13 kb proximal promoter as on the endogenous  $\gamma$ -globin proximal promoter (compare Fig. S3B and Fig. 4F). These results supported the original interpretation that changes in the levels of transcription factors caused changes in the assembly of the proximal  $\gamma$ -globin promoter complex.

In summary, the ChIP assays showed that changes in the levels of transcription activators or repressors induced distinct changes in the proximal  $\gamma$ -globin promoter complex in correlation with the induced activation or repression of  $\gamma$ -globin promoter activity: Over-expression of activators induced higher occupancies of the activators and co-activators but also of the repressor and co-repressor, which were apparently weak and did not prevent the activators/co-activators from activating transcription of  $\gamma$ -globin mRNA. On the other hand, over-expression of repressors not only increased occupancies of the repressors and co-repressor but also reduced occupancies of activators and co-activators; thus, the combination of a strong repressor complex with a weakened activator complex repressed transcription of  $\gamma$ -globin mRNA. Both activators and repressors were present in both the activated and the repressed  $\gamma$ -globin promoter complex, due at least in part to interaction between the activator and the repressor: NF-Y with BCL11A and GATA-2 with GATA-1, as was indicated also by EMSA results (Fig. 3A, 3B and 3E). Thus,  $\gamma$ -globin promoter activity appeared to be determined by the relative abundance/strengths of the activator/co-activator complex vs. the repressor/co-repressor complex.

### Molecular Assemblies of the Active and the Repressed Proximal $\gamma$ -globin Promoter Complexes in Human Fetal and Adult Erythroid Cells

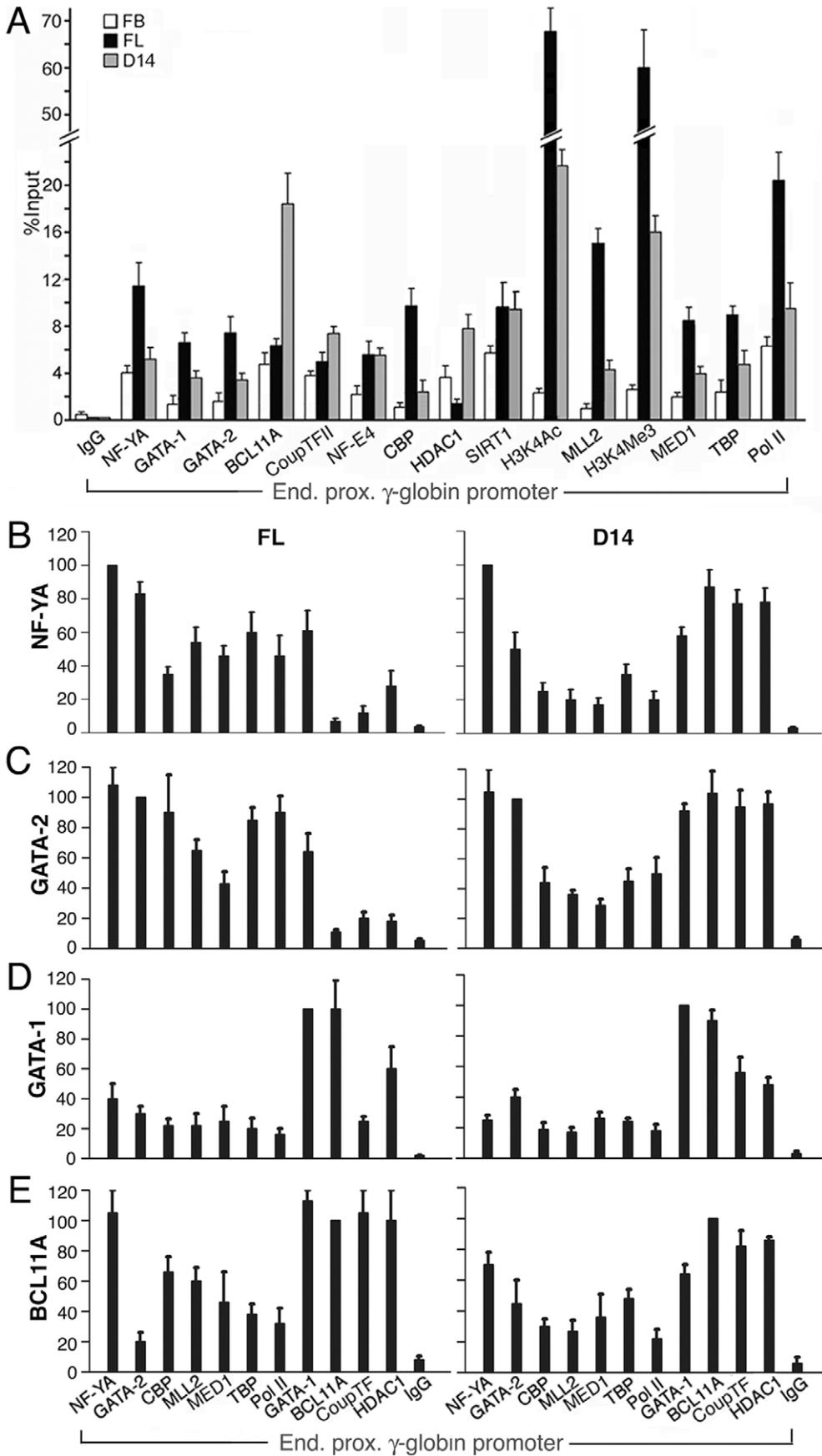
To verify the molecular assembly of the activated and repressed proximal  $\gamma$ -globin promoter complexes obtained in K562 cells due to over-expression of transcription activators NF-Y and GATA-2 or repressors BCL11A and GATA-1 (Fig. 4), we examined by ChIP the *in vivo* assembly of the active and the repressed proximal  $\gamma$ -globin promoter complexes in primary human fetal and adult erythroid cells, in which the transcription activators and repressors were developmentally regulated (see Fig. 1B and 1C). In fetal liver erythroid cells, as compared to human fetal brain cells in which the globin gene locus is transcriptionally inactive, the active, proximal  $\gamma$ -promoter bound activators NF-Y, GATA-2 and NF-E4 and also repressors BCL11A and COUP-TFII (Fig. 5A). The

activators recruited co-activators CBP and MLL2 to generate high levels of active chromatin marks H3K4Ac and H3K4Me3, and Mediator 1, TBP and Pol II (Fig. 5A) to actively transcribe  $\gamma$ -globin gene. However, the repressors BCL11A and COUP TFII appeared unable to recruit a significant level of HDAC1 in fetal erythroid cells (Fig. 5A).

In contrast, in D14 adult erythroid cells, which expressed lower levels of activators NF-Y and GATA-2 but higher levels of BCL11A and GATA-1 as compared to fetal erythroid cells (Fig. 1B and 1C), the repressed  $\gamma$ -globin promoter bound much lower levels of NF-Y and GATA-2 and -1 and co-activators CBP and MLL2 and TBP and Pol II but much higher levels of BCL11A and HDAC1 (Fig. 5A). Even though GATA-1 was expressed at a higher level in adult D14 erythroid cells than in fetal erythroid cells, GATA-1 occupancy was lower in D14 cells than in fetal erythroid cells (Fig. 5A), due again to dependence of GATA-1 recruitment on GATA-2 occupancy, which was low since GATA-2 protein was expressed at an undetectable level in adult D14 erythroid cells (Fig. 1C). It was curious that even though GATA-2 expression was undetectable in D14 erythroid cells by Western blot, GATA-2 was detected by *in vivo* ChIP assay to be present in the repressed proximal  $\gamma$ -globin promoter complex, in opposition to the findings of the *in vitro* EMSA that the proximal  $\gamma$ -globin promoter probe did not bind GATA-2 or -1 in the D14 nuclear extract (Fig. 3E). An explanation for the different *in vivo* and *in vitro* results could be that even though GATA-2 was expressed at a very low level in adult D14 erythroid cells, the expressed GATA-2 protein existed *in vivo* in the nucleus in discreet nuclear structures [36] with relatively high local concentrations of GATA-2 that were detectable by the ChIP assays. However, when the nuclear structure was disrupted to make the protein/nuclear extracts for Western blot and EMSA, the homogenized GATA-2 concentration in the bulk extracts could be too low to be detectable by either of the two *in vitro* methods. In both fetal and adult erythroid cells, ChIP assays showed that the levels of NF-E4 and SIRT1—an HDAC that associates with BCL11A [37]—did not change significantly in the proximal promoter complexes (Fig. 5A). These proteins appeared not to be involved in the developmental repression of the  $\gamma$ -globin promoter in D14 adult erythroid cells and were not further investigated.

In summary, ChIP assays in primary fetal and adult erythroid cells demonstrated that the molecular assemblies of the proximal promoter complexes, as measured by the relative occupancies, therefore abundance/strengths, of the activators/co-activators vs. the repressors/co-repressor, correlated with the developmental expression profiles of the transcription factors and with  $\gamma$ -globin promoter activities in fetal and adult erythroid cells.

To examine the *in vivo* interactions of NF-YA, GATA-2, and -1 and BCL11A among one another and with other proteins in the active and the repressed  $\gamma$ -globin proximal promoter complexes, we next carried out re-ChIP assays of the chromatin initially pulled down by the antibody to each of these four factors from fetal and adult erythroid cells in the first ChIP. In fetal erythroid cells, the re-ChIP results showed that activators NF-YA and GATA-2 associated with each other and with the co-activators at higher levels than with repressors BCL11A and Coup-TFII and co-repressor HDAC1 (Fig. 5B & 5C, left panels), whereas the repressors BCL11A and Coup-TFII associated with each other and with co-repressor HDAC1 at higher levels than with the activators and the co-activators (Fig. 5D & 5E, left panels), except for BCL11A, which also associated at a high level with NF-YA (Fig. 5E, left panel, BCL11A and NF-YA lanes), and for GATA-1, which associated at high levels with both activators NF-YA and



**Figure 5. Molecular assemblies of the active and repressed proximal  $\gamma$ -globin promoter complexes in human fetal and adult erythroid cells.** **A.** ChIP assays of the proximal  $\gamma$ -globin promoter complexes in fetal brain, fetal liver and adult D14 erythroid cells, FB, FL and D14, respectively. Values were averages of two independent pull-down assays. **B–E.** Re-ChIP assays of fetal liver and D14 adult erythroid cells, left and right panels respectively: The antibodies to the respective transcription factors used in first ChIP was shown on the Y-axis. The amount of  $\gamma$ -globin promoter pulled down by antibodies in the 1st ChIP was set at 100 to serve as the reference for comparing the amount of promoter pulled down in the 2nd ChIP by antibodies to proteins shown on the X-axis. Values were averages of two independent re-ChIP assays. doi:10.1371/journal.pone.0047175.g005

GATA-2 and repressor BCL11A (Fig. 5B–E, left panels, GATA-1 lanes).

In re-ChIP of the repressed  $\gamma$ -globin promoter complex in Day 14 adult erythroid cells, the activators appeared to associate at lower levels with the co-activators but at higher levels with the repressors and the co-repressor (Fig. 5B & 5C, right panels), whereas the repressors BCL11A and CoupTFII still associated with each other and with the co-repressor at higher levels than with the activators and the co-activators (Fig. 5D and 5E, right panels), again except for BCL11A, which associated at a high level with NF-YA (Fig. 5E, right panel, BCL11A and NF-YA lanes), and for GATA-1, which associated at high levels with both activators NF-YA and GATA-2 and repressor BCL11A (Fig. 5B–E, right panels, GATA-1 lanes).

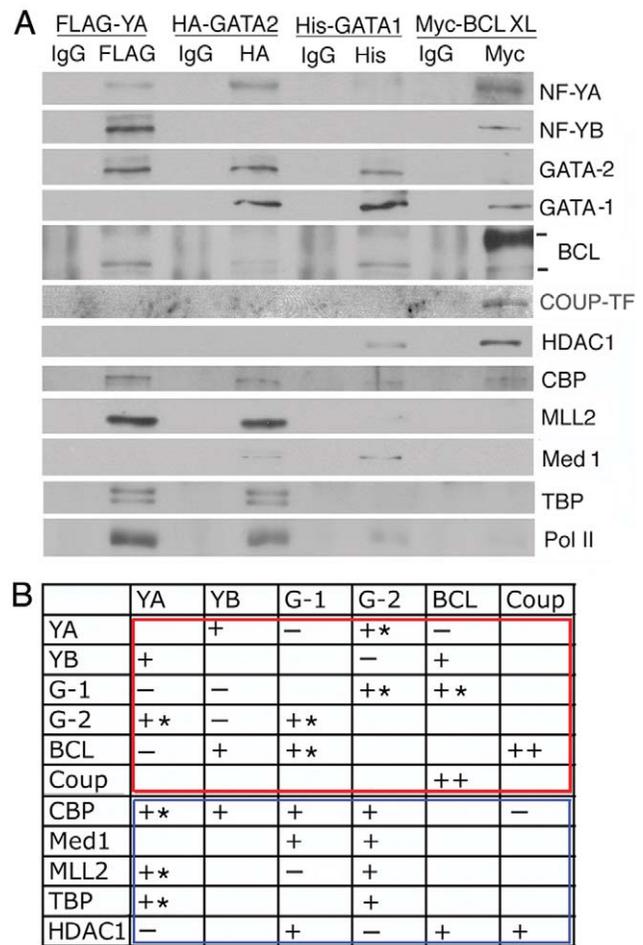
In summary, ChIP assays showed that in the active  $\gamma$ -globin promoter complex of fetal erythroid cells, occupancies of transcription activators NF-YA, GATA-2 and co-activators CBP, MLL2, Med 1, TBP and pol II were higher than those of transcription repressors BCL11A and CoupTFII and co-repressor HDAC1; however, in the repressed  $\gamma$ -globin promoter complex of D14 adult erythroid cells, occupancies of transcription activators and co-activators were lower than those of transcription repressors and co-repressor (Fig. 5A). Re-ChIP assays confirmed that in the active promoter complex, the activators associated at higher levels with co-activators than with repressors and co-repressors (Fig. 5B & 5C, left panels); in the repressed  $\gamma$ -globin promoter complex, the activators associated at higher levels with the repressors and co-repressor than with the co-activators (Fig. 5B & 5C, right panels). In contrast, the repressors and co-repressor associated preferentially with each other in both the active and the repressed promoter complexes (Fig. 5D and 5E, right and left panels). Through cross interactions of the activators with the repressors, NF-Y with BCL11A and GATA-2 with GATA-1, both the activators/co-activators and the repressors/co-repressor were present in both the active and the repressed  $\gamma$ -globin promoter complexes in fetal and adult erythroid cells (Fig. 5A and 5B).

### Protein Interaction Network in the Assembly of the Activator and the Repressor Hubs

Since ChIP and re-ChIP assays showed association of the transcription factors and co-factors with  $\gamma$ -globin promoter DNA but did not reveal whether the component proteins interacted with one another, we next carried out *in vivo* and *in vitro* co-immunoprecipitation (co-IP) to investigate the protein interaction network in the proximal promoter complex. For *in vivo* co-IP, the four key transcription factors, NF-YA, GATA-2 and -1 and BCL-XL were expressed with different tags from plasmids separately transfected into K562 cells. The proteins associated with each of the tagged transcription factors were pulled down by the specific antibodies for each of the tags. Western blots of pulled-down proteins showed particular protein association patterns: Activators NF-YA and GATA-2 associated with each other and with co-activators CBP and MLL2 and TBP and pol II but not with co-repressor HDAC1; in addition, activators NF-YA and GATA-2 cross-associated with repressors BCL11A and GATA-1 respectively (Fig. 6A, 1<sup>st</sup> and 2<sup>nd</sup> columns from left). In contrast,

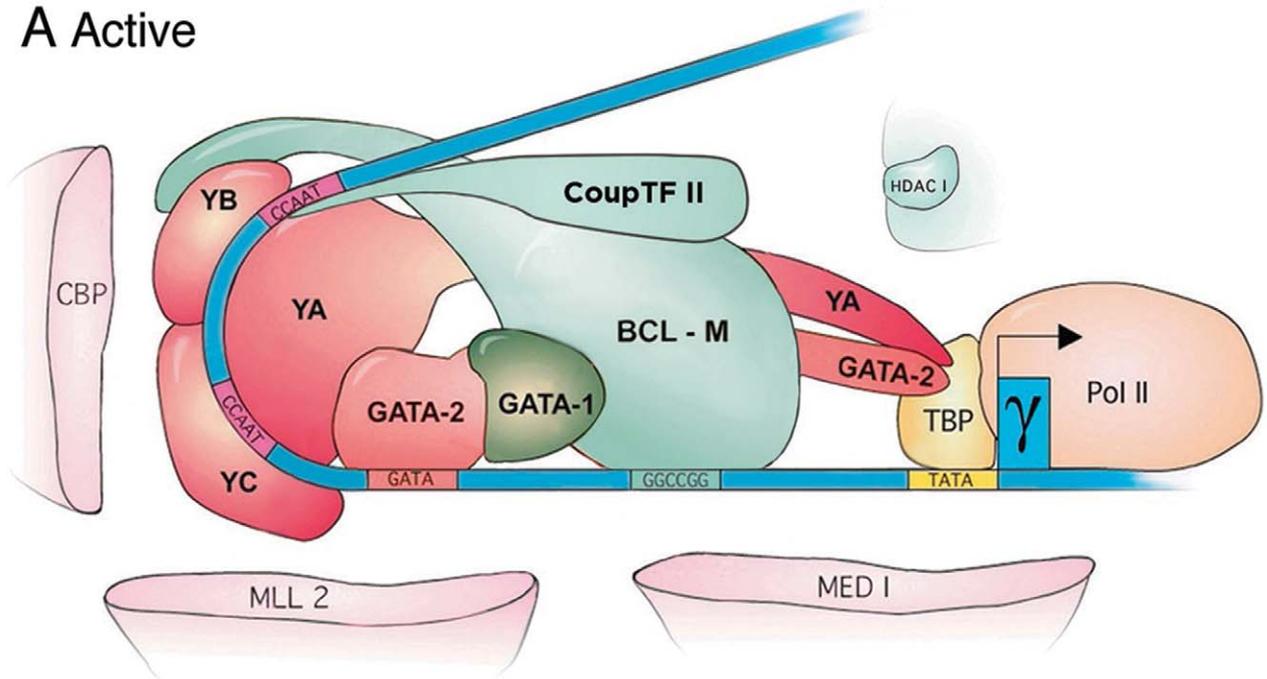
repressor BCL-XL associated only with co-repressor HDAC1 but not with the co-activators, and repressor COUP-TFII associated only with BCL11A (Fig. 6A, 4<sup>th</sup> column).

Whether these *in vivo* associations between two target proteins were direct interactions between the two proteins or indirect association through common third protein partners were assessed by *in vitro* co-IP of individual pairs of differently tagged proteins expressed in and isolated from K562 and/or 293 cells. The *in vitro*

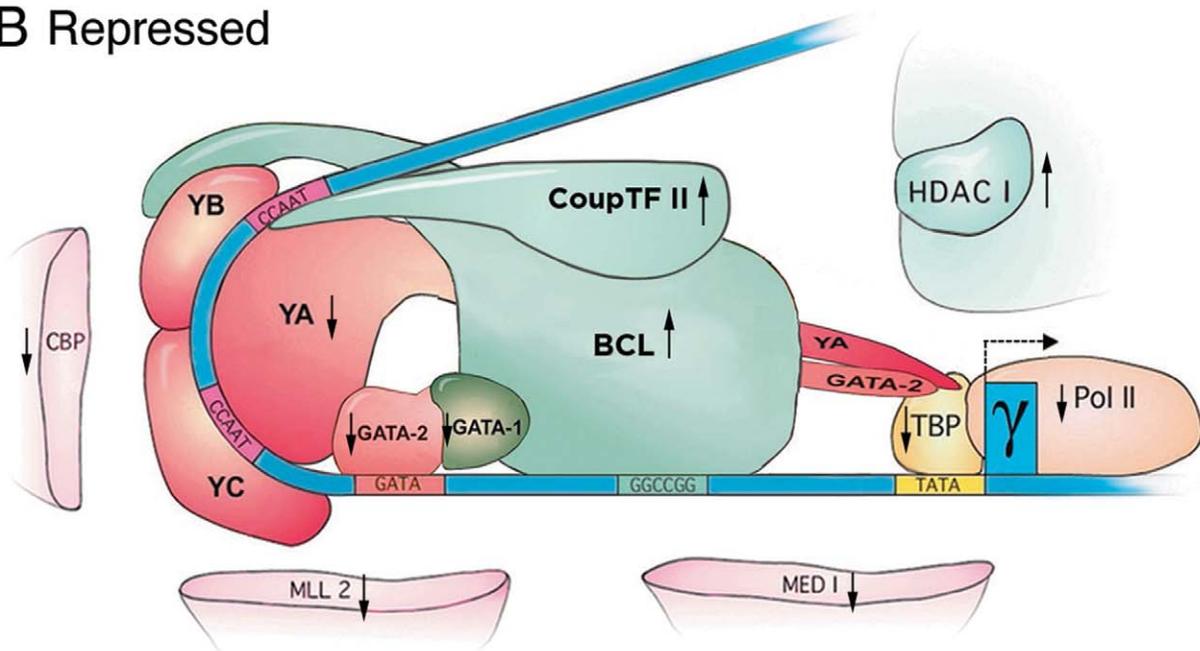


**Figure 6. Protein interaction network in the proximal  $\gamma$ -globin promoter complex.** **A.** Western blots of proteins co-immunoprecipitated with tagged NF-YA, GATA-2 or -1 or BCL11A-XL expressed from plasmids transfected into K562 cells. **B.** Pair-wise *in vitro* interactions between purified transcription factors and co-factors. Pair-wise interactions within the red frame: *in vitro* interactions between transcription factors. Pair-wise interactions within the blue frame: *in vitro* interactions between transcription factors and co-factors. ++, +, –: strong, average or no interaction between the proteins. Interactions marked with \*: the proteins in the pair-wise interaction were isolated from both K562 and HEK293 transfected with the expression plasmids for the target proteins. Otherwise, the proteins were isolated from transfected HEK293 cells. doi:10.1371/journal.pone.0047175.g006

## A Active



## B Repressed



**Figure 7. Models of molecular assemblies of the active and repressed proximal  $\gamma$ -globin promoter complexes. A.** Active complex. Blue ribbon: Proximal  $\gamma$ -globin promoter DNA containing transcription activator-binding motifs, CCAAT and GATA (red bars) and the repressor-binding motif GGCCGG (green bar); blue rectangle with angled arrow:  $\gamma$ -globin gene and transcriptional direction of  $\gamma$ -globin mRNA. Red colored proteins: transcription activators; pink colored proteins: co-activators. Dark green and green colored proteins: transcription repressors; BCL-M: BCL11A median-isoforms expressed predominantly in human fetal and K562 erythroid cells; light green colored protein: the HDAC1 co-repressor complex. NF-Y bound at each of the tandem CCAAT motifs bends the DNA by  $\sim 70^\circ$  (3) to form the pocket for assembly of the proximal promoter complex. **B.** Repressed complex. Larger sizes of BCL, COUPTFII and HDAC1 complex and smaller sizes of YA, GATA-2 and -1 and the co-activators represent higher and lower levels of the respective proteins—additionally indicated also by the upward or downward arrows—in the repressed promoter complex as compared to the levels of these proteins in the active complex in **A**; BCL: all sized-isoforms expressed in D14 erythroid cells (Fig. 1C); dotted angled arrow: lower transcription level of  $\gamma$ -globin mRNA from the repressed promoter. doi:10.1371/journal.pone.0047175.g007

co-IPs confirmed that most of the protein pairs that associated *in vivo* also associated *in vitro* (Fig. 6B), except that BCL11A, which associated *in vivo* with both NF-YA and -YB (Fig. 6A), associated *in vitro* only with NF-YB (Fig. 6B).

Together, the co-IPs revealed specific protein interactions: (1). Transcription activators NF-Y and GATA-2 interacted with each other to form an apparent activator hub. (2). Repressor BCL11A interacted with COUP-TFII with strong affinity and with GATA-1 (Fig. 6B) to form an apparent repressor hub. (3). The activator and the repressor hubs cross-associated through interactions between activator NF-Y and repressor BCL11A, and activator GATA-2 and repressor GATA-1. (4). The activator hub interacted with co-activators CBP, MLL2 and Mediator 1, and also TBP and pol II in the basal transcription machinery but not with co-repressor HDAC1, while the repressor hub interacted with co-repressor HDAC1 but not the co-activators. Incorporating these and earlier results (Fig. 2,3,4,5,6), the molecular assemblies of the active and the repressed proximal  $\gamma$ -globin promoter complexes were depicted in Fig. 7.

## Discussion

In this study, we showed the pivotal role of NF-Y in assembling the active and the repressed proximal  $\gamma$ -globin promoter complexes (Fig. 7): Transcription activator NF-Y stably bound at the CCAAT motifs with affinity reported to be among the highest for DNA binding proteins [3,7,8] recruited and stabilized binding of activator GATA-2 to the neighboring GATA motif to form the activator hub, NF-Y/GATA-2; NF-Y also recruited and stabilized binding to the neighboring GGCCGG motif of repressor BCL11A, which in turn interacted with COUP-TFII bound at its cognate site overlapping the distal CCAAT motif and with GATA-1 recruited by GATA-2 to form the repressor hub, BCL11A/COUP-TFII/GATA-1. The interaction between BCL11A and COUP-TFII in the repressor hub was consistent with earlier reports that BCL11A was originally identified and cloned by its interaction with COUP-TFII and was initially named C<sub>T</sub>IP1, COUP-TFII Interacting Protein 1 [38] and that it was a transcription repressor with cognate DNA binding motif GGCCGG [11]. The association between BCL11A and GATA-1 has also been previously reported [30]. Due apparently to cross-interaction of NF-Y with BCL11A and GATA-2 with GATA-1, both the activator and the repressor hubs were present in both the active and the repressed  $\gamma$ -globin promoter complexes. However, through their relative levels in the proximal promoter complex and their respective interactions with the co-activators and co-repressor, the activator and the repressor hubs together modulated activation or repression of  $\gamma$ -globin promoter activity during erythroid development.

Our findings shed light on the underlying mechanisms of tissue- and developmental stage-specific transcription of  $\gamma$ -globin gene. In fetal erythroid cells, which expressed at high levels the developmentally regulated, ubiquitous activator NF-Y and erythroid activator GATA-2 (Fig. 1C), NF-Y interacted with GATA-2 to assemble a strong activator hub that recruited co-activators CBP and MLL2 to remodel promoter chromatin, and Mediator 1, TBP and pol II to activate transcription of  $\gamma$ -globin mRNA from the proximal promoter; the fetal erythroid cells also expressed at very low levels repressor BCL11A-M isoforms (Fig. 1C), which assembled apparently a weak repressor hub BCL-M/COUP-TFII/GATA-1 that interacted only weakly with the HDAC repressor complex. The strong activator hub predominated over the weak repressor hub to confer erythroid- and fetal stage-specific transcription of  $\gamma$ -globin mRNA (Fig. 1B). In D14 adult erythroid

cells, which expressed lower levels of activators NF-Y and no detectable GATA-2 but much higher levels of all repressor BCL11A isoforms (Fig. 1C), the strong repressor hub predominated over the weakened activator hub to repress transcription of  $\gamma$ -globin mRNA in adult erythroid cells (Fig. 1B). In non-erythroid cells such as brain cells, which expressed a high level of NF-Y and very low to non-detectable levels of GATA-2 (Fig. 1B), but high levels of BCL11A [11] and COUP-TFII [39], the very weak activator hub combined with an apparently strong repressor hub further repressed transcription of  $\gamma$ -globin mRNA in non-erythroid cells (Fig. 1B).

Our finding that BCL11A bound to the GGCCGG motif in the proximal  $\gamma$ -globin promoter was consistent with earlier studies [10,11]. However, it has been reported that BCL11A does not bind to the  $\gamma$ -globin promoter but binds instead to the locus control region (LCR) far upstream of  $\gamma$ -globin genes and the intergenic region 3' of A $\gamma$  globin gene, thus repressing  $\gamma$ -globin gene transcription indirectly by long-range interactions [30,40]. The difference in findings may be due in part to sensitivity differences between the techniques of ChIP-chip using microarrays as the quantification tool [40,41] and the ChIP-qPCR using real-time PCR as the quantification tool employed in this report. In a recent study on global gene expression profiles in human erythroid cells, a side-by side comparison between microarrays and qPCR as quantification tools shows that qPCR is more sensitive in detecting changes in expression levels of genes, including  $\beta$ -globin gene, throughout erythroid development [42]. The human  $\beta$ -globin gene locus contains four GGCCGG binding sites for BCL11A, located respectively near the HS3 site in the LCR, in the proximal promoters of G $\gamma$ - and A $\gamma$ -globin genes and 3' of the A $\gamma$ -globin gene (GenBank U01317). BCL11A bound to the strong LCR HS3 site at a sufficiently high level detectable by both ChIP-chip and ChIP-qPCR, while it bound to the relatively weak  $\gamma$ -globin promoter site at a lower level detectable only by ChIP-qPCR (Fig. S4). Thus, BCL11A may repress  $\gamma$ -globin gene transcription both by binding to the LCR through an indirect, long-range mechanism and by directly binding to the proximal  $\gamma$ -globin promoter.

The presence of both an activator and a repressor hub in both the active and the repressed proximal  $\gamma$ -globin promoter complex may underlie the propensity of  $\gamma$ -globin promoter to respond to pharmacological compounds that activate the repressed  $\gamma$ -globin gene in adult erythroid cells [1]. The molecular models of the proximal  $\gamma$ -globin promoter complexes (Fig. 7) suggest that these pharmacological compounds may enhance expression of activators, NF-Y and GATA-2, in the activator hub and/or suppress expression of repressors BCL11A, COUP-TFII and GATA-1 in the repressor hub. These possibilities are currently under investigation.

## Supporting Information

**Figure S1** The -115 distal CCAAT motif overlapping the COUP-TFII binding site, -73 GATA and -56 GGCCGG did not bind or bound weakly to COUP-TFII, GATA-2 and BCL11A, respectively, as determined by EMSA (related to Fig. 3). **A.** Percentage contribution of binding by BCL11A to the NF-Y/BCL11A EMSA band.  $\gamma$ P wt and  $\gamma$ P(CCAAT)m: Wt and mutant CCAAT proximal  $\gamma$ -globin promoter probes. The quantified NF-Y/BCL11A bands were those in Fig. 3A, 3B, 3C, 3D and 3F. The intensities of the control NF-Y/BCL11A band without supershifts by the antibodies were set at 100. **B.** The GGCCGG motif bound BCL11A weakly. Left and middle panels: The short probe spanning -56 GGCCGG motif (same sequence as self (GC

competitor in Fig. 3) bound BCL11A very weakly in both K562 and D14 nuclear extracts. Right panel:  $\gamma$ P(GC)m (proximal  $\gamma$ -globin promoter with GGCCGG mutated to AAAAAA) bound little BCL11A, since BCL 11A antibodies only slightly decreased the intensity of the NF-Y/BCL band (lanes 8 & 9), indicating requirement of GGCCGG motif to cooperate with CCAAT motif in recruiting and binding of BCL11A. BCL11A Ab-1 and -2: antibodies from Novus NB-100–259 and Abcam Ab19487 respectively. Other designations: same as in Fig. 3. C. COUP-TFII binding to the –115 CCAAT motif overlapping the COUP-TFII site comprised ~10% of the NF-Y/COUP-TFII EMSA band, as indicated by quantification of the competition bands in Fig. 3D. Competitors dCCAAT and d+p CCAAT spanning both the COUP-TFII and NF-Y sites were 10% more efficient competitors than pCCAAT and E1CCAAT, spanning only the NF-Y site. D. Left panel: The –73 GATA motif by itself did not bind GATA-2/–1. EMSA of short probes spanning –175, a GATA site upstream of the –73 GATA site that bound GATA factors, and the –73 GATA motifs of equal length that did not bind the GATA factors (Compare lanes 1 and 2; for probe sequences, see Methods S1). Right panel:  $\gamma$ P(GATA)m, proximal  $\gamma$ -globin promoter with mutated GATA motif, bound GATA-2/–1 much less than the wildtype  $\gamma$ P (Compare lanes 3 and 4), indicating cooperativity between GATA and CCAAT motifs in recruiting/binding of GATA-2/–1. (TIF)

**Figure S2** Effect of Mediator 1 and MLL2 knockdown on mRNA levels of  $\gamma$ -globin and select transcription factors in K562 cells. **A.** Knockdown by siRNA targeting MLL2 as well as MLL1 and Menin 1 in the MLL1/2 hCOMPASS-like complex and PTIP in the MLL3/4 hCOMPASS-like complex [22] and Mediator 1 in the Mediator complex [23]. The RNA level of each of the co-factors in K562 cells transfected by the control plasmid producing scrambled siRNA (Si Ctl) was set at 100 to serve as the reference for percentage knockdown of the co-factors by the specific siRNAs. The RNA levels were determined by RT-PCR. **B.** Effects of knockdown of MLL2 and MED1 on  $\gamma$ -globin mRNA level. **C.** Effects of knockdown of MLL2 and MED1 on mRNA levels of select transcription factors and co-factors in the proximal  $\gamma$ -globin promoter complex. The results showed that reduction in transcription of  $\gamma$ -globin gene did not appear to be the secondary effects of MLL2 or MED1 knockdown, which first reduced transcription of activators NF-Y and GATA-2 and/or increased transcription of repressors BCL11A and GATA-1, since NF-Y and GATA-2 levels did not significantly change and the levels of BCL11A and GATA-1 actually decreased as a result of MLL2 and MED1 knockdown (Fig. S2C). (TIF)

**Figure S3** Effects of NF-YA knockdown and over-expression of GATA-2 and -1 on molecular assembly of the proximal  $\gamma$ -globin promoter complex in the K562 endogenous genome and in transfected GFP reporter plasmids. **A.** Effects of NF-YA knockdown on assembly of the endogenous proximal  $\gamma$ -globin promoter complex: NF-YA knockdown decreased occupancy of NF-Y, which in turn decreased occupancies of GATA-2 and co-activator MLL2; however, NF-Y knockdown increased occupancy

of COUP-TFII, which could competitively bind to its cognate site overlapping the NF-Y binding site at a higher level due to decreased occupancy of NF-Y. On the other hand, occupancy of BCL11A did not correspondingly decrease with a decrease in NF-Y occupancy (Fig. S3A), as anticipated from interaction/association of BCL11A with NF-Y, but increased as a result of the decrease in NF-Y occupancy. This was apparently because BCL11A interacted not only with NF-Y but also strongly with COUP-TFII (11). Thus, an increase in COUP-TFII occupancy increased the recruitment and occupancy of BCL11A. **B.** Over-expression of GATA-2 and -1 and CCAAT mutation (to abolish NF-Y binding) on assembly of the  $\gamma$ -globin promoter complex in plasmids transiently transfected into K562 cells. 0.13Wt-GFP and 0.13CCAATm-GFP: designations same as in Fig. 2A; Vector: pCRFP1 plasmid containing RFP selectable marker gene; GATA-2 and GATA-1: Expression plasmids containing GATA-1 or -1 cloned into the pCRFP1 vector plasmid. Inset: K562 cells doubly transfected with (GATA-2)-RFP and 0.13 Wt-GFP were sorted by FACS. Sorted cells expressing both RFP and GFP, comprising ~12% of total cell population, were used for ChIP assays. K562 cells transfected with (GATA-1)-RFP and 0.13 Wt-GFP were similarly sorted by FACS. ChIP results showed that the effects of GATA-2 and -1 over-expression on assembly of the proximal  $\gamma$ -globin promoter complex in transfected plasmids and in the K562 endogenous genome were similar (Compare Fig. S3B with Fig. 4F). In addition, mutation of CCAAT motif (to AACCG, see Fig. 2A) to abolish occupancy of NF-Y greatly diminished occupancies of GATA-2, -1 as well as of BCL11A (Fig. S3A), even though their cognate GATA and GGCCGG binding sites were not mutated. The AACCG mutation eliminated binding of not only NF-Y but also COUP TFII (Fig. 1A) and would abolish occupancy of COUP-TFII at the proximal  $\gamma$ -globin promoter; in the absence of the protein interaction partners COUP-TFII and NF-Y, BCL11A binding to the GGCCGG motif therefore decreased in the mutant proximal  $\gamma$ -globin promoter. (TIF)

**Figure S4** Relative *in vivo* binding of BCL11A to the LCR HS3 site and the proximal  $\gamma$ -globin promoter in K562 cells. Values were averages of duplicate pull-downs with the BCL antibody from Novus and Abcam, AB-1 and -2 respectively. (TIF)

**Methods S1**  
(DOC)

## Acknowledgments

We thank Drs. S. Tsai for hCOUP-TFII cDNA plasmid, S. Jane for NF-E4 antibody, R. Slany for MLL2 cDNA plasmid and S. Huang for assistance with the construction and packaging of recombinant lentiviruses and we thank N. Kline and R. Smith for the electronic artwork in Fig. 7.

## Author Contributions

Conceived and designed the experiments: XZ DT AW. Performed the experiments: XZ YW WP. Analyzed the data: XZ YW DT AW. Contributed reagents/materials/analysis tools: HL AW. Wrote the paper: XZ DT.

## References

- Mabaera R, West R, Conine S, Macari E, Boyd C, et al. (2008) A cell stress signaling model of fetal hemoglobin induction. *Exp. Hematology* 36: 1057–1072.
- Suzuki Y, Tsunoda T, Sese J, Taira H, Mizushima-Sugano J, et al. (2001) Identification and characterization of the potential promoter regions of 1031 kinds of human genes. *Genome Res.* 11: 677–684.
- Montovani R (1999) The molecular biology of the CCAAT-binding factor NF-Y. *Gene* 239: 16–27.
- Liberati C, Ronchi A, Lievens P, Ottolenghi S, Mantovani R (1998) NF-Y organizes the  $\gamma$ -globin CCAAT boxes region. *J Biol Chem* 273: 16880–16889.
- Duan Z, Stamatoyannopoulos G, Li Q (2001) Role of NF-Y in *in vivo* regulation of the  $\gamma$ -globin gene. *Mol Cell Biol* 21: 3083–3095.

6. Li Q, Fang X, Han H, Stamatoyannopoulos G (2004) The minimal promoter plays a major role in silencing the galago  $\gamma$ -globin gene in adult erythropoiesis. *Proc Natl Acad Sci USA* 101: 8096–8101.
7. Kim CG, Sheffrey M (1990) Physical characterization of CCAAT transcription factor  $\alpha$ -CPI. *J Biol Chem* 272: 13362–13369.
8. Bi W, Wu L, Coustry F, Crombrughe B, Maity S (1997) DNA binding specificity of the CCAAT-binding factor CBF/NF-Y. *J Biol Chem* 272: 26562–26572.
9. Zhou W, Clouston D, Wang X, Cerruti L, Cunningham J, et al. (2000) Induction of human fetal globin gene expression by a novel erythroid factor, NF-E4. *Mol & Cell Biol* 20: 7662–7672.
10. Chen Z, Luo H, Steinberg M, Chui D (2009) BCL11A represses HBG transcription in K562 cells. *Blood Cells Mo. Dis* Doi: 10.1016/j.bcmd.2008.12.003.
11. Avram D, Fields A, Senawong T, Topark-Ngarm A, Leid M (2002) COUP-TF (chicken ovalbumin upstream promoter transcription factor)-interacting protein 1 (CTIP1) is a sequence-specific DNA binding protein. *Biochem J* 368: 555–563.
12. Filipe A, Li Q, Deveaux S, Godin I, Roméo PH, et al. (1999) Regulation of embryonic/fetal globin genes by nuclear hormone receptors: a novel perspective on hemoglobin switching. *EMBO J* 18: 687–697.
13. Pereira F, Tsai MJ, Tsai S (2000) COUP-TF orphan nuclear receptors in development and differentiation. *Cell & Mol Life Sciences* 57: 1388–1398.
14. Liberati C, Cera MR, Secco P, Santoro C, Mantovani R, et al. (2001) Cooperation and competition between the binding of COUP-TFII and NF-Y on human epsilon- and gamma-globin gene promoters. *J Biol Chem* 276: 41700–41709.
15. Aerbajinai W, Zhu J, Kumjaek C, Chin K, Rodgers G (2009) SCF induces  $\gamma$ -globin gene expression by regulating downstream transcription factor COUP-TFII. *Blood* 114: 187–194.
16. Collins FS, Metherall JE, Yamakawa M, Pan J, Weissman SM, et al. (1985) A point mutation in the A gamma-globin gene promoter in Greek hereditary persistence of fetal haemoglobin. *Nature* 313: 325–326.
17. Martin DI, Tsai SF, Orkin SH (1989) Increased gamma-globin expression in a nondeletion HPHH mediated by an erythroid-specific DNA-binding factor. *Nature* 338: 435–438.
18. Berry M, Grosveld F, Dillon N (1992) A single point mutation is the cause of the Greek form of hereditary persistence of fetal haemoglobin. *Nature* 358: 499–502.
19. Li Q, Duan ZJ, Stamatoyannopoulos G (2001) Analysis of the mechanism of action of non-deletion hereditary persistence of fetal hemoglobin mutants in transgenic mice. *EMBO J* 20: 157–164.
20. Liu LR, Du ZW, Zhao HL, Liu XL, Huang XD, et al. (2004) T to C substitution at -175 or -173 of the gamma-globin promoter affects GATA-1 and Oct-1 binding in vitro differently but can independently reproduce the hereditary persistence of fetal hemoglobin phenotype in transgenic mice. *J Biol Chem* 280: 7452–7459.
21. Goodman RH, Smolik S (2000) CBP/p300 in cell growth, transformation, and development. *Genes & Dev* 14: 1553–1577.
22. Eissenberg C, Shilatifard A (2010) Histone H3 lysine 4 (H3K4) methylation in development and differentiation. *Dev Biol* 339: 240–249.
23. Malik S, Roeder RG (2010) The metazoan mediator co-activator complex as an integrative hub for transcriptional regulation. *Nat Rev Genetics* 11: 761–772.
24. Kang JA, Zhou Y, Weis TL, Liu H, Ulaszek J, et al. (2008) Osteopontin regulates actin cytoskeleton and contributes to cell proliferation in primary erythroblasts. *J Biol Chem* 283: 6997–7006.
25. Yu X, Zhu X, Pi W, Ling J, Ko L, et al. (2005) The LTR of ERV-9 human endogenous retrovirus binds to NF-Y in the assembly of an active LTR enhancer complex NF-Y/MZF1/GATA-2. *J Biol Chem* 280: 35184–35194.
26. Ling J, Pi W, Bollag R, Zeng S, Keskintepe M, et al. (2002) The solitary LTRs of ERV-9 endogenous retrovirus are conserved during primate evolution and possess enhancer activities in embryonic and hematopoietic cells. *J Virology* 76: 2410–2423.
27. Zhu X, Ling J, Zhang L, Pi W, Wu M, et al. (2007) A facilitated tracking and transcription mechanism of long-range enhancer function. *Nucleic Acids Res* 35: 5532–5544.
28. Pi W, Zhu X, Wu M, Wang Y, Fulzele S, et al. (2010) Long-range function of an intergenic retrotransposon. *Proc Natl Acad Sci USA* 107: 12992–12997.
29. Ramchandran R, Bengra C, Whitney B, Lancois K, Tuan D. (2000) A (GATA)7 motif located in the 5' boundary area of the human  $\beta$  globin locus control region exhibits silencer activity in erythroid cells. *Am J of Hematology* 65: 14–24.
30. Sankaran VG, Menne TF, Xu J, Akie TE, Lettre G, et al. (2008) Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. *Science* 322: 1839–1842.
31. Blobel GA (2002) CBP and p300: versatile coregulators with important roles in hematopoietic gene expression. *J Leukoc Biol* 71: 545–556.
32. Demers C, Chaturvedi CP, Ranish JA, Juban G, Lai P, et al. (2007) Activator-mediated recruitment of the MLL2 methyltransferase complex to the beta-globin locus. *Mol Cell* 27: 573–584.
33. Gordon DF, Tucker EA, Tundwal K, Hall H, Wood WM, et al. (2006) MED220/thyroid receptor-associated protein 220 functions as a transcriptional coactivator with Pit-1 and GATA-2 on the thyrotropin-beta promoter in thyrotropes. *Mol Endocrinol* 20: 1073–1089.
34. Stumpf M, Waskow C, Krötschel M, van Essen D, Rodriguez P, et al. (2006) The mediator complex functions as a coactivator for GATA-1 in erythropoiesis via subunit Med1/TRAP220. *Proc Natl Acad Sci USA* 103: 18504–18509.
35. Kabe Y, Yamada J, Uga H, Yamaguchi Y, Wada T, et al. (2005) NF-Y is essential for the recruitment of RNA polymerase II and inducible transcription of several CCAAT box-containing genes. *Mol Cell Biol* 25: 512–522.
36. Elefanty A, Antoniou M, Custodio N, Carmo-Fonseca M, Grosveld F (1996) GATA transcription factors associate with a novel class of nuclear bodies in erythroblasts and megakaryocytes. *EMBO J* 15: 319–333.
37. Senawong T, Peterson VJ, Leid M (2005) BCL11A-dependent recruitment of SIRT1 to a promoter template in mammalian cells results in histone deacetylation and transcriptional repression. *Arch Biochem Biophys* 434: 316–325.
38. Avram D, Fields A, Pretty K, Nevriy D, Ishmael J, et al. (2000) Isolation of a novel family of C2H2 zinc finger proteins implicated in transcription repression mediated by COUP-TF orphan nuclear receptors. *J Biol Chem* 275: 10315–10322.
39. Kim BJ, Takamoto N, Yan J, Tsai SY, Tsai MJ (2009) Chicken Ovalbumin Upstream Promoter-Transcription Factor II (COUP-TFII) regulates growth and patterning of the postnatal mouse cerebellum. *Dev Biol* 326: 378–391.
40. Xu J, Sankaran VG, Ni M, Menne TF, Puram RV, et al. (2010) Transcriptional silencing of gamma-globin by BCL11A involves long-range interactions and cooperation with SOX6. *Genes Dev* 24: 783–798.
41. Jawaid K, Wahlberg K, Thein S, Best S (2010) Binding patterns of BCL11A in the globin and GATA1 loci and characterization of the BCL11A fetal hemoglobin locus. *Blood Cells Molecules and Diseases* 45: 140–146.
42. Merryweather-Clarke A, Atzberger A, Soneji S, Gray N, Clark K, et al. (2011) Global gene expression analysis of human erythroid progenitors. *Blood* 117: e96–e108.