Fetal hemoglobin in sickle cell anemia

Idowu Akinsheye1*
Abdulrahman Alsultan2*
Nadia Solovieff3*
Duyen Ngo†
Clinton T. Baldwin1
Paola Sebastiani3
David H.K. Chui1
Martin H. Steinberg1,†

1Department of Medicine, Boston University School of Medicine, Boston, MA; 2Department of Pediatrics, College of Medicine, King Saud University, Riyadh, Saudi Arabia; 3Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts

*These authors contributed equally to this work.
†Correspondence: mhsteinb@bu.edu
Abstract

Fetal hemoglobin (HbF) is the major genetic modulator of the hematologic and clinical features of sickle cell disease, an effect mediated by its exclusion from the sickle hemoglobin polymer. Fetal hemoglobin genes are genetically regulated and the level of HbF and its distribution among sickle erythrocytes is highly variable. Some patients with sickle cell disease have exceptionally high levels of HbF that are associated with the Senegal and Saudi-Indian haplotype of the \textit{HBB}-like gene cluster; some patients with different haplotypes can have similarly high HbF. In these patients, high HbF is associated with generally milder but not asymptomatic disease. Studying these individuals might provide additional insights into HbF gene regulation. HbF appears to benefit some complications of disease more than others. This might be related to the premature destruction of erythrocytes that do not contain HbF, even though the total HbF concentration is high. Recent insights into HbF regulation have spurred new efforts to induce high HbF levels in sickle cell disease beyond those achievable with the current limited repertory of HbF inducers.
Fetal Hemoglobin in Sickle Cell Anemia

Appreciating the role of fetal hemoglobin (HbF; $\alpha 2\gamma 2$) in sickle cell disease started more than 60 years ago when Janet Watson confirmed that infants with sickle cell disease had few symptoms and that their deoxygenated erythrocytes took longer to sickle and did not deform as extensively as did their sickle cell trait-carrying mother's cells. She attributed these observations to high HbF levels in infant blood. Sickle hemoglobin (HbS) gelation studies showed that HbF did not interact with HbS; it was also reported that compound heterozygotes for sickle cell trait and hereditary persistence of HbF (HPFH) were clinically normal despite having a very high HbS concentration. (reviewed in 1)

HbF is the most powerful modulator of the clinical and hematologic features of sickle cell anemia (defined as homozygosity for glu6val in the $\beta$-globin gene or $HBB$). To protect against various complications of disease, different concentrations of HbF were postulated to be required, although any increment in HbF had a beneficial effect on mortality.2,3 Higher HbF levels were associated with a reduced rate of acute painful episodes, fewer leg ulcers, less osteonecrosis, less frequent acute chest syndromes and reduced disease severity. However, HbF level had a weak or no clear association with priapism, urine albumin excretion, stroke and silent cerebral infarction, systemic blood pressure and perhaps sickle vasculopathy as estimated by tricuspid regurgitant velocity (reviewed in 4). The failure of HbF to modulate uniformly all complications of sickle cell disease might be related to the pathophysiological events that impact the likelihood of developing these complications. Many epidemiological studies suggested that disease complications most closely linked to sickle vasoocclusion and blood viscosity were robustly related to HbF concentration while complications associated with the intensity of hemolysis were less affected5 (and references therein) although HbF is protective for leg ulcers, one complication closely associated with hyper-hemolysis.6,7 After defining hyper-hemolysis by the highest quartile of serum lactic dehydrogenase and controlling for liver disease by examining only patients with normal serum alanine aminotransferase, HbF
levels were lower in patients with the highest quartile compared with the lowest quartile of hemolysis.⁸

Even when total HbF levels are high, perhaps the intravascular hemolysis of erythrocytes containing little or no HbF leads to sufficient nitric oxide (NO) scavenging by plasma hemoglobin to provoke hemolysis-related complications.⁹ (Fig. 1)

The Saudi-Indian and Senegal haplotypes of the HBB-like globin gene complex, discussed below, are associated with high HbF levels and carriers of these haplotypes can have milder disease.¹⁰-¹⁴ Some patients who have other HbS-associated haplotypes also have very high HbF. Notwithstanding the high HbF levels of all these patients, acute painful episodes and other symptoms of sickle cell disease still occur, perhaps because the heterogeneous cellular distribution of HbF does not equally protect all erythrocytes from polymerization-induced damage. In contrast, individuals who are compound heterozygotes for HbS and gene deletion HPFH (see below), where HbF is more evenly apportioned among erythrocytes, or pancellularly distributed, are clinically asymptomatic with nearly normal hemoglobin levels.

**HbF and the retardation of HbS polymerization**

HbF is composed of 2 α-globin polypeptide chains and 2 γ-globin chains. The γ-globin chains are encoded by 2 nearly identical genes (HBG2 and HBG1) within the β-globin gene-like cluster on chromosome 11p that differ by a glycine or alanine residue at amino acid position γ136. Gγ- and Aγ-globins have similar effects on HbS polymerization.¹⁵ With the rapid decrease in the numbers of circulating fetal erythrocytes, the ratio of Gγ to Aγ-globin falls from 0.7 at birth to 0.4 at age 5 months. This is accompanied by a progressive decline in the number of erythrocytes with measurable HbF, called F-cells. In normal adults, HbF is less than 1% of total hemoglobin and is distributed unevenly among erythrocytes. HbF levels in sickle cell anemia range between 5 and 8%. In African Americans with sickle cell anemia, 2% to 80% of erythrocytes were F-cells compared with 2.8 ± 1.6% in normal African Americans. Sickle cell trait carriers have a mean HbF of 1.4% and 14.1 ± 7.5% F-cells. In sickle cell anemia, F-cells
survive longer than non-F-cells and this depends on the amount of HbF/ F-cell. A high correlation ($R^2 = 0.967$) is present between the number of F-cells and the percent HbF.

The pathophysiology of sickle cell disease is dependent on the polymerization of deoxy sickle hemoglobin (deoxyHbS). Increased levels of HbF retard this process. HbF reduces HbS concentration, but more importantly, both HbF and its mixed hybrid tetramer ($\alpha_2\beta^S\gamma$) cannot enter the deoxyHbS polymer phase. In contrast, the hybrid tetramer containing $\beta^S$ and $\beta^A$ chains has only half the probability of entering the polymer as the HbS molecule hence the special value of HbF compared with other hemoglobins. The anti-polymerization effect of HbF resides primarily in $HBG$ (both $\gamma$-globin genes) residues glycine $\gamma 87$ and aspartic acid $\gamma 80$. By inhibiting the tendency of deoxyHbS to polymerize, sufficient HbF thwarts the cellular damage evoked by HbS polymer.(reviewed in 4)

Engineering recombinant HbF and HbA by adding additional substitutions can enhance the capacity of the molecule to inhibit polymerization, an approach exploited when devising vectors for gene therapy. Conversely, the natural mutant hemoglobin, HbS-Antilles ($HBB$ glu6val; val23ile) has enhanced polymerization tendencies; contrasted with individuals with sickle cell trait, heterozygotes with this variant are symptomatic and homozygotes have severe sickle cell disease.

**Genetic Basis of HbF Regulation**

*Globin Gene Switching*

HbF is the predominant hemoglobin from early gestation until 1 to 2 months postnatally when adult HbA predominates. Although erythroid precursors of normal adults express $HBG$ at a low level stress erythropoiesis is associated with increased HbF. A stochastic model posits that the increase in HbF is due to recruitment of erythroid progenitor cells that prematurely undergo terminal differentiation and are committed to producing $\gamma$-globin. The stress signal transduction model suggests that cytokines like erythropoietin, stem cell factor
and transforming growth factor-β \cite{24} initiate downstream intracellular signaling pathways that activate HBG expression and is the premise on which HbF induction by cytostatic agents was based \cite{25}.

In sickle cell anemia there is a delayed switch from HBG to HBB expression and the replacement of HbF by HbS, and HbF levels remain above normal in most patients. The mechanism accounting for this is unknown but might reflect the slower centripetal regression of red, or hematopoietic marrow to the axial skeleton in the presence of expanded erythropoiesis that is due to sustained hemolysis. The goal of HbF-inducing treatments is to reverse this switch to the largest degree possible. Compound heterozygotes with HbSC disease and HbS-β\textsuperscript+ thalassemia, who usually have lower levels of hemolysis, most often have HbF levels near normal or only slightly increased \cite{26,27}. This could be a result of less intense hemolysis or a lack of those genetic modulators linked to the HbS gene.

The molecular basis of hemoglobin switching, including more detailed descriptions of the role of many of the genes involved in this process has been recently reviewed \cite{28}.

**Haplotypes of the HBB gene-like cluster**

The HbS β-globin gene is found on 4 or 5 common haplotypes reflecting its regions of origin in Africa, the Middle East and the Indian subcontinent \cite{29-31}. Patients with a Bantu haplotype have the lowest HbF and those with a Senegal or Saudi-Indian haplotype, have the highest; individuals with a Benin haplotype have HbF levels that are intermediate \cite{32-34}. Carriers of all haplotypes, have considerable variance in HbF levels suggesting the importance of other quantitative trait loci (QTL) modulating HBG expression.

In carriers of Senegal and Saudi–Indian haplotypes, a C-T polymorphism 158 base pairs (bp) upstream of \textit{HBG2} (rs7482144), the Xmn1 C-T restriction site polymorphism, is associated with high HbF and G\textsubscript{γ}-globin levels \cite{29}. Carriers of the Bantu and Benin haplotypes lack this single nucleotide polymorphism (SNP) \cite{34,35}. Most likely this SNP is in linkage disequilibrium (LD) with functional elements responsible for increased expression of \textit{HBG2}. 
Recent studies that sequenced about 87 kb within the HBB gene-like cluster suggested that a SNP in LD with rs7482144, (rs10128556; p=1.2E-09), was more strongly associated with HbF than rs7482144 (p=3.7E-07), however the effect (beta coefficients) of the 2 SNPs on HbF was very similar and the stronger p-value may be due to the higher minor allele frequency of rs10128556. Also, rs10128556 had an effect on HbF independent of rs7482144 (p = 0.047) and rs7482144 had no effect on HbF independent of rs10128556 (p = 0.78).\(^3\) However, this determination was based on conditional haplotype tests that excluded uninformative haplotypes. The test excluded less than 10% of the data for rs10128556 but nearly 90% for rs7482144 which severely affected its power and is reflected by the non-significant p-value for rs7482144.

The mechanism whereby this region influences HbF is unclear.

Homozygotes for the Saudi-Indian haplotype, which also includes rs7482144, have HbF levels substantially higher than homozygotes for the Senegal haplotype.\(^3\) These observations suggest that other elements, perhaps linked to these haplotypes are differentially effecting HBG2 transcription. While it is likely that the association of a haplotype with the clinical features of sickle cell anemia is mediated by haplotype-related differences in HbF concentration, in nearly all instances the actual functional elements responsible for HbF modulation are unknown.

**QTL modulating HbF**

HbF levels are hereditable.\(^3\) Two quantitative trait loci (QTL) in addition to the HBB-linked regions discussed above, have major influences on HbF expression. These loci are the HBS1L-MYB intergenic region and BCL11A. SNPs in these QTL and in the HBB gene-like cluster explain about one third to one half of HbF variation in sickle cell anemia leaving much of the variance in HbF level unexplained. Rare variants are likely to explain this "missing" heritability but are difficult to detect using genome-wide association studies (GWAS) and will require genomic sequencing.\(^3\)

**HBS1L-MYB intergenic region (HMIP, 6q23)**
Polymorphisms in the HMIP region were associated with F-cell levels and accounted for 19.4% of the F-cell variance in normal Europeans. SNPs were distributed in three LD blocks called the HMIP blocks 1, 2 and 3, however the sequences of HBS1 and MYB and other genes in this region were uninformative regarding HbF regulation. Nevertheless, the expression profile of MYB and HBS1L in adults with non-gene deletion HPFH was downregulated. Overexpression of MYB in K562 cells inhibited HBG expression. Low levels of MYB were associated with reduced cell expansion and accelerated erythroid differentiation, suggesting that variation in the intrinsic levels of MYB might affect HbF by its effect on the cell cycle.

Overexpression of microRNA-15a and -16-1 downregulated MYB in CD34+ erythroid progenitors and increased HbF. Among few individuals found to harbor one of 3 rare missense variants in MYB, their HbF was 7.5% compared with 6.1% in noncarriers. In another study, only HBS1L expression was correlated with elevated HbF levels.

The HBS1L-MYB intergenic polymorphisms are also highly associated with HbF expression among Chinese β-thalassemia heterozygotes. It was recently reported that the most significant functional motif accounting for HMIP modulation of HbF is a 3 bp deletion polymorphism which is in complete LD with the SNP rs9399137 shown by several GWAS to be most significantly associated with HbF in Europeans, Asians, and African Americans. It is located near erythroid specific DNase I hypersensitive site 2 within the HMIP block 2, 42.6 kb upstream of HBS1L and 83.8 kb upstream of MYB. In close proximity to the 3 bp deletion polymorphism, there is binding of 4 erythropoiesis-related transcription factors, TAL1, E47, GATA2, and RUNX1. Furthermore, the short DNA fragment encompassing the 3 bp deletion polymorphism appears to have enhancer-like activity based on in vitro transient transfection experiments.

The HBS1L-MYB intergenic polymorphism is also associated with HbF among sickle cell anemia patients of African descent though much less significantly when compared with Europeans or Chinese due to their much lower minor allele frequencies. It could be that there
are other HMIP variants associated with HbF level among people of African descent that are not tracked well by SNP rs9399137.36,51

**BCL11A (2p16)**

*BCL11A*, a zinc finger protein gene was first associated with lymphoid malignancies in man. In one of the singular successes of GWAS, *BCL11A* polymorphisms were strongly associated with HbF concentrations in normal individuals and several different populations of patients with β thalassemia and with sickle cell anemia. By its effects on HbF concentration, *BCL11A* modified the clinical features of both diseases.44,45,53,54 Binding sites for *BCL11A* have been described in HS3 of the LCR, the Aγ-δ intragenic region by chromosome immunoprecipitation assays and a GGCCGG motif in proximal promoter of the *HBG* and complexes of *BCL11A* with other proteins might mediate the suppressive effects of *BCL11A* on *HBG* expression.

Polymorphism within the 14 kb intron 2 of *BCL11A* correlated with F-cell numbers in healthy Northern Europeans and Chinese with β-thalassemia trait. In Chinese with β thalassemia, Thais with HbE/β-thalassemia patients, Tanzanians and African Americans with sickle cell anemia polymorphisms in *BCL11A* were associated with HbF. Individual variants and haplotypes at this locus account for up to 18% of HbF variance sickle cell anemia.55

**8q**

A QTL at chromosome 8q appeared to interact with the *HBG2*-158 C-T SNP to modulate HbF levels.56-58 This locus did not appear to be associated with HbF in GWAS of patients with sickle cell anemia and β thalassemia but these studies were likely underpowered.44,45,54 Haplotype tagging SNPs were used to probe chromosome 8q in 2 independent sickle cell anemia patient groups. In 1,518 individuals, 3 SNPs in *TOX* (8q12.1) were associated with HbF and 3 additional SNPs in *TOX* showed significant association in 211 individuals used for validation. Joint analysis of all SNPs and covariates confirmed the association of HbF with *TOX*.59 SNPs in *TOX* were also associated with the HbF response to
hydroxyurea treatment in sickle cell anemia.\textsuperscript{60} \textit{TOX} belongs to a conserved high-mobility group box protein family that binds the minor groove of DNA and might modulate gene expression by an effect on chromatin structure and transcriptional regulatory complexes.\textsuperscript{61}

\textbf{SAR1A}

A small guanosine triphosphate (GTP)-binding protein, secretion-associated and RAS-related (SAR1A) protein is inducible by hydroxyurea and might play a pivotal role in induction of \textit{HBG} expression via its role in erythroid maturation. Polymorphisms in the \textit{SAR1A} promoter were associated with differences in HbF levels or the HbF response to hydroxyurea in sickle cell anemia.\textsuperscript{62,63}

\textbf{F-cell production locus}

A putative F-cell production locus was localized between DXS143 and DXS16 within Xp22.3-22.2 and hypothesized to account, in part, for the higher HbF levels in females compared with males, an observation found in both the normal population and in patients with sickle cell anemia.\textsuperscript{64,65} GWAS have yet to find a gene or polymorphism in the F-cell production locus at Xp22, the phenotype of which was associated with HbF in sickle cell anemia; probing this region with haplotype-tagging SNPs did not reveal a strong candidate gene.\textsuperscript{59}

Other genetic loci as modulators of HbF have been proposed but are less well established.\textsuperscript{66-68}

\section*{Sickle Cell Disease and Unusually High HbF}

\textbf{HbF in African Americans}

We define a "HbSF" phenotype as a HbF concentration of at least 10\% in sickle cell anemia patients aged 4 years or more, the time by which HbF levels stabilize.\textsuperscript{42,43} After excluding the known causes of the HbSF phenotype, like point mutations in the \textit{HBG} promoters or large deletions within the \textit{HBB}-like globin gene cluster, the molecular basis of this phenotype
was studied in African Americans. Twenty "high" HbF patients and 30 "low" HbF controls were first studied and the results validated in a replication set with 56 "high" HbF (HbF 20.7±8.2%) patients and 489 "low" HbF (HbF 3.1±1.5%) control patients. The 20 "high" HbF cases were aged 16.3±8.3 years and had a hemoglobin level of 9.0±1.3 g/dL, MCV 87.9±9.0 fl and HbF 17.2±4.8%. The 30 "low" HbF cases were aged 19.3±9.8 years, had a hemoglobin of 8.6±1.4 g/dL, MCV 81.4±11.1 fl and HbF of 5.0±2.5% (Akinsheye I, personal communication, and Akinsheye I, Solovieff, N, Malek, A, et al. Fetal hemoglobin in sickle cell anemia: Molecular characterization of the high fetal hemoglobin phenotype in African American patients. *Blood*. 2010;116(21):858.). *BCL11A* rs766432 and *HMIP* rs9399137 had a higher minor allele frequency in patients with "high" HbF in both data sets and accounted for 20% of HbF variance. The aforementioned 3 bp (TAC) in the *HBS1L-MYB* intergenic region was present in 18% of total chromosomes in the "high" HbF group, compared with a 3% frequency in the "low" HbF controls and was in complete linkage disequilibrium with rs9399137.

A 14.1 kb DNA fragment between *HBG1* and *HBD* was sequenced in 15 "high" and 15 "low" HbF patients. This fragment included the 7.2 kb Corfu deletion that was associated with elevated HbF levels and contains binding sites for BCL11A. Four SNPs had significant frequency differences between "high" and "low" HbF groups (Fig 2A). Three of the SNPs were between positions 49213 and 49994 and 1 was at position 54541. In silico analysis showed that the G-A polymorphism at position 49876 (rs968856) created a C/EBP binding site which is not present in the minor allele. The G-A polymorphism at position 49994 (rs968857) eliminated an AP-1 and NF-E2 binding sites, which are present in the minor alleles. Any functional significance of these SNPs remains to be explored (Fig. 2A).

**HbF in Saudi Arabian sickle cell anemia**

Sickle cell anemia in Saudi Arabia has population concentrations in the Southwestern and Eastern Provinces. Most Eastern Province patients carry the Saudi-Indian β-globin gene-
like cluster haplotype and have very high levels of HbF. Southwestern Province patients have
typical African-derived haplotypes and lower HbF levels, albeit higher than comparable
haplotype groups of African descent.

**Southwestern Province**

The HbS gene in Saudi patients from the Southwestern Province was introduced from
Africa and is present on typical African \textit{HBB} haplotypes. Nevertheless, patients differ from
African Americans phenotypically and have fewer episodes of stroke, priapism, and leg ulcers
and a higher prevalence of splenomegaly.\textsuperscript{14,71-73} This might be related to differences in HbF
levels or co-inheritance of $\alpha$ thalassemia.

To examine genetic modifiers of HbF level in patients from the Southwestern Province, the
3 HbF QTLs were genotyped in 77 patients, aged 17.7±10 years, 69% with sickle cell anemia
and 31% with HbS-$\beta^0$ thalassemia.\textsuperscript{74} The distribution of \textit{HBB} gene cluster haplotypes was;
74% Benin, 22% Bantu, 4% others. \textit{BCL11A} was the sole QTL associated with HbF level but
the study was underpowered to detect associations with small effect sizes and with SNPs
having a low minor allele frequency like those in HMIP and \textit{HBB}. These findings were compared
with 2 studies of African Americans with sickle cell anemia who had a similar distribution of \textit{HBB}
haplotypes. Southwestern Province cases had HbF levels almost twice that of African
Americans (P<0.0001), after adjusting for the \textit{BCL11A} genotype. When the genetic population
structure of the African Americans and Saudi patients was compared using genome-wide SNP
data, African Americans were similar to Yoruban, Mandenka and Bantu Africans while Saudi
patients resembled Arab populations. The commonality of \textit{HBB} haplotypes coupled with the
genetic distance between these populations suggested that additional genetic modifiers have
evolved in this population or that unknown environmental influences were likely to account for
the higher HbF in these Saudi patients.

**Eastern Province**
In the Eastern Province of Saudi Arabia, sickle cell anemia is usually associated with the Saudi-Indian (sometimes called Arab-Indian) HBB-gene cluster haplotype, high levels of HbF and a milder, but not asymptomatic clinical course. As in Southwestern patients, splenomegaly is common and stroke and leg ulcers are rare. The rarity of stroke might be a result of the higher hemoglobin concentration and a high incidence of α thalassemia; little information on the incidence of pulmonary vasculopathy is available. Sickle cell trait carriers or individuals with HbA and the Saudi Indian haplotype did not have high HbF but the cultured erythroblasts of sickle cell trait patients with the Saudi Indian haplotype made increased amounts of HbF suggesting that the kinetics of erythropoiesis played a role in the expression of the high HbF determinant.

HbF levels are higher in Saudi Indian haplotype patients when they are compared with African Americans homozygous for the Senegal haplotype whose mean HbF was 12.5±5.3%, even though both haplotypes share the 5' HBG2 C-T polymorphism (rs7482144). Five Saudi families from the Eastern Province that included 7 patients with sickle cell anemia who had a median HbF of 30.3% (18-41%) were studied. (Alsultan A, personal communication and, Alsultan A, Solovieff, N., Akinsheyeye, et al. Fetal hemoglobin in sickle cell anemia: Molecular characterization of Saudi patients from the Eastern Province. Blood. 2010;116 (21):684.) It was hypothesized that the elements that modify HbF in Eastern Province Saudi patients evolved with, and were in LD with the βS-globin gene. Runs of SNP homozygosity of 160 kb to nearly 2 mb were present within and surrounding the HBB cluster only in individuals with sickle cell anemia (Fig. 2B) and were not present elsewhere in the genome. Limited sequencing of the HBD-HBG1 intergenic region and core regions of HS 2, 3, and 4 of the LCR showed that the cores of HS3 and HS4 were identical to the reference sequences and the core of HS2, contained the 10TA.2CA.2TA.CG.12TA motif associated with the Saudi-Indian haplotype. Polymorphisms were identified in HBD-HBG1 intergenic region and many lead to creation or abolition of transcription factor binding sites and some bind transcription factors presumed to
have regulatory roles in globin gene expression when examined in silico.

Deep sequencing of more than 800 kb of DNA in all three major HbF QTLs has been completed. The results of these studies might uncover variants in Eastern Province sickle cell anemia patients that could be associated with their unusually high HbF level and be candidates for functional and mechanistic studies.

**HbS-HPFH**

Genetic causes of high HbF in normal adults, when not a result of thalassemia, *HBG* duplications or rearrangements within the *HBB* gene-like cluster are called HPFH. HPFH is phenotypically and genetically diverse and has been divided into non-gene deletion and gene deletion types. The former group can be linked or unlinked to the *HBB* gene-like cluster. Known causes of non-deletion HPFH result from mutations in the *HBG* promoters and SNPs in the known QTLs that modulate *HBG* expression. Gene deletion HPFH is caused by large deletions in the *HBB* gene-like cluster. Compound heterozygotes with HbS-HPFH and sickle cell anemia with non-gene deletion HPFH have been reported. Deletion mutants leading to pancellular HbF distribution are associated with a benign phenotype while the heterogeneous cellular distribution of HbF and the lower levels of HbF associated with point mutations and homozygosity for minor alleles in the HbF QTL have the usual disease complication, albeit at what might be a lower rate. A database of reported HPFH mutations is available at [http://globin.bx.psu.edu/hbvar/menu.html.](http://globin.bx.psu.edu/hbvar/menu.html).

**Point mutations in HBG promoters**

Point mutations associated with non-deletional HPFH were found in three regions around positions -114, -175, and -200 in the 5' promoter regions of *HBG2* and *HBG1*. The sequences around position -200 have been shown to be a binding sites for various erythroid transcription factors. Heterozygosity for the T-C SNP at -175 in the promoter regions of *HBG* is associated with HbF levels of 20-40%. This region contains the DNA sequences that serve as binding sites for OCT-1 and GATA-1. The third region involved in non-deletional HPFH is
located at position -114 and -117. The G-A mutation at -117 in the *HBG*1 promoter affects the CCAAT box, a regulatory element of globin and other genes. The C-T SNP at position -114 or a 13 bp deletion of this region (from -102 to -114) interferes with erythroid specific factor NF-E3 and ubiquitous trans-acting factors such as CP-1 and CDP. These SNPs interfere with the binding of erythroid-specific and ubiquitous transcription factors to *HBG* promoters and could result in decreased binding affinity of negative regulatory factors or, alternatively, increased interactions with positive regulatory factors. SNPs in the *HBG* promoters must be very uncommon causes of high HbF in sickle cell anemia; none were present in the selected group of African American patients with high HbF described above.

**Hbs-Gene deletion HPFH**

At least 8 HPFH-causing deletions within the *HBB* gene-like cluster of 10 to more than 80 kb have been described. They are associated with HbF levels of 20-30% that is distributed nearly equally among erythrocytes (pancellular). Compound heterozygotes with Hbs-gene deletion HPFH have mild microcytosis and do not have features of sickle cell disease. In a recent examination of 28 cases of Hbs-HPFH (6 HPFH-1 and 22 HPFH-2), the hemoglobin concentration was 13±1 g/dL, MCV 75±6 fl and HbF 31±2%.(Ngo D, personal communication and Ngo DA AB, Akinsheye I, Steinberg MH, et al. Characterization of the HbF decline in compound heterozygotes for Hbs and deletional hereditary persistence of fetal hemoglobin. *Blood*. 2010;116(21):683.) The absence of sickle vasoocclusive events or hemolytic anemia illustrates the critical importance of a pancellular distribution of HbF.

HPFH-1 and HPFH-2 deletions span more than 80 kb and are the most common types of HPFH found in American and African blacks. Three possible mechanisms have been proposed to explain the increased HbF associated with these large deletions: 1) the deletion of regulatory sequences affecting *HBG* expression; 2) juxtaposition of enhancers of *HBG* normally located downstream of *HBB*; 3) LCR interactions with *HBG* due to deletion of *HBB*.
A form of deletion HPFH has been described in association with Hb Kenya, an abnormal hemoglobin resulting from a ~22.5 kb deletion leading to a fusion product composed of the \(HBG1\) and \(HBB\). HbS-Hb Kenya compound heterozygotes had mild microcytic anemia and an average of 10% HbF in a pancellular distribution. These patients had few, if any, sickle cell-related events. HbF and Hb Kenya make up 30% or more of the total hemoglobin and this might be sufficient to inhibit sickling.\(^{79}\)

**HbF Inducing Agents**

**Hydroxyurea**

The beneficial effects of high HbF in sickle cell anemia and in \(\beta\) thalassemia where HbF can substitute for HbA launched an effort to find drugs capable of increasing HbF levels. The DNA hypomethylating agent, 5-azacytidine was used to induce HbF in anemic baboons, a species whose hemoglobin composition and regulation is nearly identical to man's. Based on the impressive results of these preclinical studies, this agent was used in sickle cell anemia and \(\beta\) thalassemia with promising results, however continued trials were abandoned because of potential carcinogenicity. In these studies it was unclear if HbF induction was due to \(HBG\) hypomethylation or to the cytotoxic effects of 5'azacitadine.(reviewed in \(^1\)) Therefore, trials of hydroxyurea, an S-phase specific agent without primary hypomethylating activity, with a long history of use in myeloproliferative disorders and with tolerable side effects were started. The culmination of this work was the Multicenter Study of Hydroxyurea (MSH), a double-blind, placebo-controlled study of patients with symptomatic sickle cell disease. Patients randomized to hydroxyurea had fewer pain episodes, less acute chest syndrome, and a lower transfusion requirement than placebo-treated cases and this agent rapidly received approval for use in sickle cell anemia in the United States and elsewhere. How therapeutic induction of HbF affects less common complications like priapism, leg ulcers, pulmonary vasculopathy and stroke is not
known The average increment in HbF achieved in the MSH was only 3.6% over the baseline level of 5.1% but other studies using different dosing regimens found higher increases in HbF.\textsuperscript{81-83} Long-term follow-up studies of MSH patients suggested that mortality was reduced in patients who took this drug and that side effects were minimal\textsuperscript{83-86}; other studies confirmed the benefits of hydroxyurea.\textsuperscript{86}

A trial of hydroxyurea in babies has been completed (NCT00006400). Although the trial's primary endpoints of preservation of renal and splenic function were not met during the relatively brief observation period, treated patients had less pain, higher hemoglobin concentrations, increased HbF and reduced leukocyte counts, with minimal short-term toxicity. (Wang, W, 4th Annual Sickle Cell Disease Research and Education Symposium, Feb 14-19, 2010, Hollywood FL and\textsuperscript{87}) The long-term effects of hydroxyurea begun in the neonatal period will require careful follow-up.

**Experimental HbF inducing agents**

Not all patients respond to hydroxyurea and the erythrocytic distribution of HbF in treated patients is heterocellular.\textsuperscript{88} Among responders, the increment in HbF is variable suggesting the need for additional agents capable of inducing HbF and perhaps broadening its cellular distribution. One class of promising agents are histone deacetylase (HDAC) inhibitors whose inhibition is associated with increased expression of \textit{HBG}. Arginine butyrate, a short chain fatty acid with HDAC inhibitory activity used as single agent or with hydroxyurea has been associated with increases in HbF.\textsuperscript{89, 90} However, a pulsed, or intermittent, dosing regimen was necessary to avoid cytotoxicity from butyrate while retaining the targeted promoter activation. Another oral short chain fatty acid derivative, sodium 2,2 dimethylbutyrate showed HbF induction in thalassemia in early phase clinical trials (NCT00842088). Suberoylanilide hydroxamic acid (Vorinostat), an orally available agent approved for treatment of cutaneous T-cell lymphoma is an HDAC inhibitor that induced HbF expression in K562 cells.\textsuperscript{91} A Phase1/2 trial of this agent in sickle cell anemia is presently enrolling patients (NCT01000155).
High throughput screening studies with follow-up of promising candidates have suggested that strong inhibitors of HDAC1 and HDAC2 were associated with substantial increments in both HBG expression and HbF in vitro. BCL11A has been shown to interact with HDAC1 and HDAC2.

Decitabine (5-aza-2'-deoxycytidine), a less toxic and perhaps non-carcinogenic deoxynucleotide is also associated with DNA hypomethylation and has been used as single agent therapy and with hydroxyurea in a small number of patients with sickle cell anemia but Phase 2 or 3 studies have not been done.

The recent studies of BCL11A and HMIP have stimulated a search for new agents that might act by modulating the expression of these genes and their signaling pathways to augment HbF expression in the β hemoglobinopathies.

Conclusion

HbF has beneficial effects in sickle cell anemia. The contrast between asymptomatic individuals HbS-gene deletion HPFH and symptomatic patients with sickle cell anemia with similarly high HbF levels suggests that if it were possible to induce high HbF levels in most sickle erythrocytes, and if this could be done before organ damage occurs, one might expect the disease to be "cured." Presently, this is not possible but a better understanding of how HbF levels are modulated might suggest new therapeutic approaches and combinations of HbF-inducing agents that could allow this goal to be met.
Acknowledgements

Supported by: NIH grants R21 HL080463 (PS); R01 HL68970 (MHS); R01 HL70735 (MHS); R01 DK069646 (DHKC); RC2 HL101211 (MHS).

Authorship

AI and CTB performed experiments, analyzed and interpreted data and wrote the manuscript, AA, NS, DN, PS DHKC and MHS analyzed and interpreted data and wrote the manuscript.

Conflict-of-interest disclosure: The author declare no competing financial interests.
References


Figures

**Figure 1.** Sickle erythrocytes are a mixture of cells with measurable HbF (F-cells) and non F-cells. F-cells are long lived, do not acquire the same increment of HbS-induced damage as non F-cells, are less likely to initiate adhesive events and are associated with protection from sickle vasoocclusion (left arrow). The heterocellular distribution of HbF in sickle cell anemia, even when total HbF concentrations are high at baseline or in response to hydroxyurea, means that some erythrocytes with no HbF or with suboptimal concentrations of HbF are present. Some of these cells hemolyze intravascularly liberating hemoglobin which scavenges nitric oxide and contributes to certain vascular complications of this disease (right arrow). This might account for the failure of high HbF that is heterocellularly distributed to protect against all disease complications.
Figure 2A. Results of sequencing 14.1 kb of chromosome 11 in the region of the Corfu deletion. BCL11A binding sites are present between coordinates 44000, 45000, 53000, and 54000. A PYR site is near 54000; rs10128556 is located just 5’ to 47000. Thirty-eight SNPs were found; SNPs marked in red and detailed in the table had a significantly different distribution in 15 patients with HbF>11% compared with 15 patients with HbF <9%; SNPs in blue had a similar distribution between these groups. Binding sites for BCL11A and PYR are shown in the blue ovals and diamond respectively. Below the coordinates are shown the locations of the β-globin pseudogene (HBBP1) and the Corfu deletion. In the table, minor and major alleles are indicated by black and red respectively. High and Low HbF represents the major allele frequencies in the 2 groups. TF binding sites were determined using TFSEARCH and a minimal threshold score of 85.0. Changes in TF binding sites occur in the major allele in the high HbF patients.
**Figure 2B.** Runs of SNP homozygosity (ROH) on chromosome 11 in Saudi-Indian and Senegal HBB gene cluster haplotype patients. In the remainder of the genome analyzed with the 610 Illumina array, no other substantial runs of SNP homozygosity were present. (personal communication, Alsultan A. (Alsultan A, Akinsheye I, Solovieff, N., et al. Fetal hemoglobin in sickle cell anemia: Molecular characterization of Saudi patients from the Eastern Province. Blood. 2010;116(21):684 (abstract))
Fetal hemoglobin in sickle cell anemia

Idowu Akinsheye, Abdulrahman Alsultan, Nadia Solovieff, Duyen Ngo, Clinton T. Baldwin, Paola Sebastiani, David H.K. Chui and Martin H. Steinberg

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include digital object identifier (DOIs) and date of initial publication.