The Breakpoint of a Large Deletion Causing Hereditary Persistence of Fetal Hemoglobin Occurs Within an Erythroid DNA Domain Remote From the β-Globin Gene Cluster

By Elise A. Feingold and Bernard G. Forget

The DNA juxtaposed to the γ-globin genes as a result of a large deletion associated with hereditary persistence of fetal hemoglobin (HPFH) was studied to define the role it may play in maintaining active expression of these genes in adult erythroid cells. The DNA located immediately 3' to the deletion breakpoint was found to function as an enhancer element in gene transfer experiments and to be specifically hypomethylated in normal erythroid cells of both fetal and adult origin. This DNA also contains a long open reading frame encoding a polypeptide chain 292 amino acids in length. Therefore, in this form of HPFH (HPFH-1), the continued expression of γ-globin genes in adult life may result from the inclusion of these genes within a new chromosomal domain that is potentially transcriptionally active in adult erythroid cells. The 3' breakpoint of another large deletion causing δβ thalassemia rather than HPFH was also identified. This deletion (Spanish δβγ γδβδ) thalassemia is nearly identical in size and location to that of HPFH-1, but extends an additional 8.5 to 9 kb in the 3' direction, and therefore results in loss of the sequences near the 3' breakpoint of HPFH-1. Thus, the presence of these sequences appears to be important for the expression of the HPFH phenotype.

DURING HUMAN development there is a perinatal switch from the synthesis of fetal hemoglobin (Hb F; α2γ2) to that of adult hemoglobin (Hb A, αβδ). Hereditary persistence of fetal hemoglobin (HPFH) consists of a heterogenous group of conditions in which this switch fails to occur or is incomplete, resulting in a high level of Hb F expression in adult life (for reviews, see references 1 and 2). One type of HPFH is associated with large deletions involving the β-globin gene cluster. The precise locations of the breakpoints of the four known HPFH deletions (HPFH-1, -2, -3 and -4) have been established and in all cases, both of the adult globin genes (δ and β) are deleted (references 1 and 2 and references therein) (Fig 1). The deletions causing HPFH-1 and HPFH-2 are of nearly identical size and have recently been shown to encompass approximately 105 kb of DNA.3 The deletions causing HPFH-3 and HPFH-4 are also quite similar to one another in size but they encompass considerably less DNA, their 3' endpoints being located approximately 50 kb more proximal to the site of the β-globin gene.67

Several hypotheses have been generated in an attempt to explain how these deletions result in persistent γ-globin expression in adult life.8-10 These hypotheses have focused on the possible role of the DNA sequences located at the 5' breakpoint of the deletions, in particular sequences deleted in HPFH but not in δβ thalassemia. The δβ thalassemias are a related group of deletion mutants in which the expression of Hb F is considerably lower and, in the heterozygote, is 5% to 7% of normal. Several mechanisms have been proposed to explain this phenomenon. These include protective linkage of other genes, specific transcriptional silencing of the γ-globin genes, or a specific role of the γ-globin gene promoter linked to the gene for chloramphenicol acetyl transferase (CAT) with or without the SV40 enhancer sequences, respectively, were the kind gift of D. Bodine (National Institutes of Health, Bethesda, MD). pRSVCAT, containing the promoter and enhancer of Rous sarcoma virus (RSV) was kindly provided by B. Howard (National Institutes of Health, Bethesda, MD). pRSVL containing the luciferase gene driven by the RSV promoter/enhancer was kindly provided by S. Subramani (UCSD, La Jolla, CA). pγCAT2 was constructed by ligation of the 2.1 kb BamHI γCAT fragment from pγCAT2 into the BamHI site of pBR322. pγCAT3D.2 and pγCAT3D.3 were constructed by ligation of the 2.1 kb BamHI γCAT fragment into the BamHI site of pBR322. pγCAT3D.2 and pγCAT3D.3 were constructed by ligation of the 2.1 kb BamHI γCAT fragment into the BamHI site of pHPFH-3D. pγCAT3D.1 and pγCAT3D.4 were constructed by isolation of the 2.1 kb SalI γCAT fragment from pSVγCAT1, addition of EcoRI linkers to the SalI ends, and ligation into the EcoRI site of pHPFH-3D.

Electroporation of DNA into K562 cells and assay of CAT activity. K562 cells were grown in Improved Minimal Essential Medium (Biofluids, Rockville, MD) with 10% fetal calf serum (Biofluids), 50 mg/L of gentamicin sulfate (GIBCO, Grand Island, NY), and 2.5 mg/L of fungizone (GIBCO). Forty micromicrograms of pRSVCAT or the equivalent molar amount of thetest plasmids plus 20 µg of pRSV gpt17 were introduced into 3 x 10⁶ K562 cells by electroporation.14 The cells were harvested after 48 hours and cell lysates prepared by three cycles of freeze-thawing in 0.1 mL of 0.25 mol/L Tris-HCl, pH 7.8. Five to 10 µL of each cleared cell lysate were used to determine the amount of guanine phosphoribosyltransferase (gpt) activity using the method of Chu and Berg18 in order to standardize for efficiency of gene transfer.14 Each extract...
expressed gpt levels that were significantly higher than those obtained from either extracts of mock-electroporated cells or reaction mixtures alone without added cell lysate. The amounts of the different lysates that were then assayed for CAT activity contained equivalent levels of gpt activity and these amounts were determined separately to exhibit CAT activity in a linear range of the assay. The level of CAT activity was determined by the method of Gorman et al. CAT activity was quantitated by scintillation counting of acetylated and nonacetylated forms of 14C chloramphenicol separated by thin layer chromatography on silica gel sheets and identified by fluorography. The quantitative results were subjected to statistical analysis using Dunnett's multiple comparison test.

Southern blot analyses. Genomic DNAs were prepared from various tissues and tissue culture cell lines by the method of Blin and Stafford. Nucleated cells were isolated from peripheral blood following lysis of erythrocytes. The lymphoblastoid cell line LAZ 149 was derived from an individual homozygous for HPFH-1. Twenty micrograms of genomic DNA were digested with three-to fivefold excess of restriction endonuclease for 5 to 16 hours under conditions recommended by the supplier. In representative experiments probing methylation sites M₁ and M₂, aliquots of the digestion mixtures were removed, added to 1 μg of λ phage DNA and incubated in parallel. These test digests were then analyzed by gel electrophoresis to control for completeness of digestion of the DNA. The DNA was then fractionated by electrophoresis in 0.8% agarose gels and blotted onto nitrocellulose or nylon filters. The nitrocellulose filters were prehybridized for 5 to 6 hours at 67°C in a solution containing 6X SSC, 10 mmol/L EDTA, pH 8.0, 100 mmol/L potassium phosphate, pH 7.0, 5X Denhardt's solution, 0.5% SDS, and 250 μg/ml of denatured salmon sperm DNA. After addition of denatured nick-translated probe, hybridization was carried out in the same solution for 16 to 20 hours at 67°C. The nylon filters were prehybridized and hybridized at 42°C in a solution containing 50% formamide, 5X SSC, 5X Denhardt's solution, 50 mmol/L sodium phosphate, pH 6.5, and 200 μg/ml of denatured salmon sperm DNA. After hybridization, the filters were rinsed two or three times for 15 minutes each at room temperature in 2X SSC, 0.5% SDS, two times for 30 minutes in 2X SSC, 0.5%, SDS at 67°C, then two times for 30 minutes in 0.1X SSC, 0.5% SDS at 67°C. The filters were subjected to autoradiography for 1 to 7 days at -80°C with intensifying screens.

The DNA probes were nick-translated to a specific activity of 10⁷ to 10⁸ cpm/μg using a nick-translation kit from Amersham (Arlington Heights, IL), according to the manufacturer's specifications. Probes 1 and 2 (see Fig 3A), also designated pHPFH-3D and pHPFH-2 respectively, were previously described. Probe 3 (see Fig 3A) was a 0.7 kb EcoRI/XbaI fragment excised from the 5 kb BamHI subclone (pHPF-4) originally derived from clone Λ5. Probe B (see Fig 5) contained the 0.4 kb HindIII/BglII fragment at the 3' end of the insert of clone Λ5.

Nucleotide sequence determination. DNA sequencing was carried out by the dideoxy method of Sanger et al. DNA fragments were subcloned into the M13 mp phage vectors and sequenced using the cloning and sequencing kits from Bethesda Research Laborato-
**RESULTS**

**Enhancer activity located at the 3' breakpoint of the HPFH-1 deletion.** Because the DNA at the 3' breakpoint of HPFH-1 is brought to within the same distance of the γ-globin genes in both HPFH-1 and HPFH-2 (Fig 1), we examined the ability of this DNA to function as an enhancer element, using the assay system of Gorman et al\(^\text{15}\) that is based on the transfer and expression in tissue culture cells of the gene for the bacterial enzyme chloramphenicol acetyl transferase (CAT). We initially used the test plasmid pA\(_{\gamma\text{CAT}}\)\(^2\),\(^3\) in which the CAT gene is linked to the SV40 early promoter region lacking the viral enhancer sequences. Expression of CAT activity is markedly increased in the presence of an added enhancer element. A number of different DNA fragments surrounding the 3' deletion breakpoint of HPFH-1 were subcloned into pA\(_{\gamma\text{CAT}}\)\(^2\),\(^3\) and the resulting plasmids tested for increased CAT expression in various types of tissue culture cells. One fragment, (HPFH-3D) (Fig I), located approximately 1 kb 3' to the HPFH-1 breakpoint was associated with increased CAT expression in both mouse L cells\(^2\)\(^,\)\(^3\) and K562 human erythroleukemia cells, (unpublished observations, March through April, 1985) but not in all four combinations of position and orientation.

To explore the possibility that the level of transcriptional activation may, in part, be determined by the promoter linked to the CAT gene, the HPFH-3D fragment was subcloned in different positions and orientations adjacent to a fusion gene consisting of the CAT gene linked to the γ-globin gene promoter.\(^1\)\(^\text{1}\) These γCAT plasmids were introduced into K562 cells by electroporation together with a RSV\(_{\text{gpt}}\) plasmid as an internal control for transfection efficiency. Cell extracts with equivalent amounts of gpt activity were then assayed for CAT activity. The results of
this analysis are presented in Fig 2. HPFH-3D had the ability to enhance CAT expression 2.7- to 5.3-fold over the control (enhancerless) γCAT plasmid when situated in all four possible combinations of position and orientation with respect to CAT gene transcription. The geometric mean of each set of experimental results was found to be significantly different (P < .05) from the enhancerless control. Separate experiments, using different preparations of the plasmid DNAs and an RSV enhancer-driven luciferase gene (pRSVL) as an internal standard for DNA transfer efficiency, demonstrated that the HPFH-3D region enhanced CAT activity 3.4- to 11.6-fold, depending on the orientation and position of this region with respect to the γ-globin gene promoter (data not shown). Use of the pRSVL plasmid as an internal control provided a more rapid and reproducible method for standardization of transfection efficiency. When the same γCAT plasmids were introduced into HeLa cells by calcium phosphate coprecipitation, the HPFH-3D fragment enhanced CAT activity 1.3- to 3.0-fold (data not shown). Thus, based on these and previous results obtained in L cells, the HPFH-3D fragment can elevate the level of CAT activity in nonerythroid cells, but in a less consistent manner and to a lesser extent than in erythroid cells. Attempts to map the cap site of the γCAT mRNA in transfected K562 cells were not successful, presumably because of the very low levels of the γCAT mRNA in these cells.11

Methylation status of DNA at the 3' deletion breakpoint of HPFH-1. One of the hallmarks of gene expression is the relative hypomethylation of a gene and its flanking DNA in tissues where it is expressed (for review, see reference 23). The DNA at the HPFH-1 breakpoint was examined in a variety of tissues and cell lines to determine if any differences in methylation pattern exist in erythroid versus nonerythroid cells. Southern blot analysis of total cellular DNA was carried out using the methylation-sensitive restriction endonuclease HpaII, an isoschizomer of Mspl.

The Mspl site (M1) located 1.4 kb 3' to the HPFH-1 breakpoint within HPFH-3D, was assayed by digestion with XbaI or BamHI plus HpaII, as illustrated in Fig 3A. M1 was found to be hypomethylated in the DNA from both human adult and fetal erythroid tissues as judged by the predominance of the two smaller bands over the larger methylated bands (Fig 3B, lanes 1-7). Specifically, DNA from the K562 human erythroleukemia cell line (lane 4) was completely unmethylated at M1 and DNA from human fetal liver (lanes 5-7), which consists of approximately 50% nucleated erythroid cells, was roughly 50% unmethylated. DNA from peripheral blood leukocytes of two patients with homozygous β thalassemia that contained 50% to 80% normoblasts (lanes 1 and 2), and from the erythroid-rich bone marrow of a patient with immune hemolytic anemia (lane 3) was also substantially hypomethylated. This same Mspl site was also found to be hypomethylated by Poncz et al in peripheral nucleated red blood cells of an individual doubly heterozygous for HPFH-2 and β' thalassemia. In contrast, the DNA from nonerythroid tissues was either completely methylated or only slightly hypomethylated since the larger methylated fragment was either the only fragment detected or was represented in vast excess over the smaller fragments (lanes 8-13). DNA from peripheral blood leukocytes (lane 8), adult spleen (lane 9), a lymphoblastoid cell line (LAZ 149) from an individual homozygous for HPFH-1 (lane 10), the HL60 myeloid leukemia cell line (lane 11), and placenta (lanes 12 and 13) were either completely or highly methylated.

We also analyzed the nearest Mspl sites located 5' and 3' to M1, designated M2 and M3, respectively, as illustrated in Fig 3A. M2 was examined by digestion of DNA with XbaI + HpaII followed by hybridization of Southern blots with probe 2 (Fig 3C). As shown in Fig 3C, probe 2 hybridized exclusively to the larger methylated fragment in DNA from all of the tissues examined except for placenta (lanes 6-8), where the smaller unmethylated fragment predominated and the larger methylated fragment was barely detectable. Similar results were obtained in the case of the Mspl site situated 3' to M1 (M3). DNA was digested with BamHI + HpaII and the blots hybridized to probe 3 (Fig 3A). Probe 3 hybridized only to the larger methylated fragment in DNA from all the tissues examined, except placenta, which yielded almost exclusively the smaller unmethylated fragment (data not shown). Thus, the Mspl sites located 4.6 kb 5' and 3.5 kb 3' to M1 are completely methylated in DNA from both erythroid and nonerythroid tissues, but hypomethylated in placental DNA, which is known to be generally undermethylated, in an apparently nonspecific fashion.25,26
Fig 4. (A) Comparison of the DNA sequence at the HPFH-1 deletion breakpoint with that from normal DNA in the regions corresponding to the 5' and 3' ends of the deletion. The box indicates the insertion of 5 nucleotides (AAATA) at the breakpoint of HPFH-1. The normal 5' DNA sequence has been previously published.27 The analogous region of DNA from a non-HPFH individual3 was also determined. These sequences and those from normal DNA in the region 5' to the deletion breakpoint27,29 are aligned in Fig 4A. The nucleotide sequence of the HPFH-1 DNA matches perfectly both the sequence of normal DNA 5' to the deletion breakpoint and the sequence of normal DNA 3' to the deletion breakpoint. However, there is an insertion of 5 nucleotides (AAATA) at the deletion breakpoint (boxed in Fig 4A). The origin of these sequences is unknown. Similar insertions of DNA of unknown origin or from remote sites have been reported at the breakpoint of other deletions involving the β-globin gene cluster,3–32 as well as in the α-globin gene cluster.33 Computer analysis of the normal nucleotide sequences in the regions of the 5' and 3' breakpoints failed to reveal any extensive regions of homology to one another (data not shown). In particular, the 5' breakpoint of the HPFH-1 deletion occurs within a repetitive DNA sequence of the Alu family,27 whereas no similar sequence is present in the vicinity of the 3' breakpoint of the deletion.3 Therefore, the deletion appears to have arisen by an illegitimate recombination event.

The nucleotide sequence of the DNA located to the 3' side of the HPFH-1 deletion breakpoint was found to contain a number of long (>90 amino acid) open reading frames. The longest open reading frame (292 amino acids) is shown in Fig 4B. Based on its codon usage (data not shown), this region of DNA has a high likelihood of encoding an authentic protein.
features of the sequence suggest that it may correspond to a central exon of a larger gene. The long open reading frame contains near its 5' end a consensus sequence for an acceptor splice site (underlined in Fig 4B). In addition, 26 nucleotides 5' to this potential splice site is a region identical to the consensus sequence (CTGAC) for the site of lariat formation shown to be an intermediate in RNA splicing.35,36

On the other hand, the long open reading frame could be part of an intronless gene. When compared with known protein sequences in various data banks, the derived amino acid sequence shown in Fig 4B was found to share 24% and 27% sequence identity, respectively, with that of the human \( \beta_2 \)-adrenergic receptor37 and another related receptor.38 Of interest, the genes for the \( \beta_2 \)-adrenergic and related receptors are intronless39-40 and are members of a family of related intronless genes for various G-protein-coupled membrane receptors (for review see reference 41). It is therefore conceivable that the DNA at the 3' breakpoint of HPFH-1 encodes a previously unidentified erythroid-specific membrane receptor. The DNA sequence 5' to the open reading frame does contain two TATA sequences (boxed in Fig 4B) consistent with a promoter element, but no CCAAT sequence. No polyadenylation signals were found in the 226 nucleotides 3' to the long open frame. However, since the length of 3' untranslated sequences is quite variable and can be quite considerable, one cannot definitively exclude the possibility that this region of DNA constitutes an intronless gene.

Because of the possibility of this region of DNA being transcriptionally active, we sought evidence for the presence of mRNA transcripts in human erythroid cells by a number of different techniques (Northern blot, S1 nuclease analysis, and screening of erythroid cDNA libraries), all of which yielded negative results (data not shown). These results indicate that if this region of DNA is in fact transcriptionally active in erythroid cells, the resulting mRNA is of extremely low abundance or expressed during a very restricted period of erythroid cell proliferation or differentiation.

The DNA 3' to the HPFH-1 breakpoint has short stretches of identity with a number of previously described enhancer elements. In particular, it contains four regions that are identical in 7 of 8 bases to the following SV40 core enhancer sequence:

\[
\text{GTGG AAA TTT G}
\]

Two of these are located within the putative enhancer fragment HPFH-3D, one at nucleotide position 1501 and the other on the complementary strand at position 1770; the two others are located 5' to this region at positions 59 and 584, respectively (underlined in Fig 4A and B). There is also a sequence at positions 687 to 705 (dashed underline in Fig 4B) that demonstrates a region of partial identity with a conserved sequence motif found in the enhancer element located 3' to the human \( \beta_2 \)-globin gene (14 of 18 bases) as well as that located 3' to the chicken \( \beta_2 \)-globin gene (13 of 18 bases). Another portion of the HPFH-1 DNA sequence (positions 91 to 132) also shares partial sequence identity (27 of 42 bases) with an overlapping sequence of the 3' enhancer region of the chicken \( \beta_2 \)-globin gene.

**Localization of the 3' deletion breakpoint of the Spanish \( \beta_2 \)-\( \gamma \) thalassemia.** We have mapped the 3' deletion breakpoint of the Spanish \( \beta_2 \)-\( \gamma \) thalassemia because of some of the features that this mutation shares with HPFH-1. The 5' breakpoint of this deletion is located approximately 1 kb 3' to that of HPFH-1. The 3' breakpoint of the deletion has previously been shown to extend at least 7 kb beyond that of HPFH-1, and more recently, using a probe from the Spanish \( \beta_2 \)-thalassemia DNA, the breakpoint was mapped to a point approximately 9 kb from the 3' breakpoint of HPFH-1. We have confirmed this mapping assignment using a probe from DNA located downstream of the HPFH-1 3' breakpoint and have precisely located the breakpoint by DNA sequence analysis.

The previously determined restriction map3 of normal DNA located in the region of the 3' breakpoint of the HPFH-1 deletion is illustrated in Fig 5, with the location of various probes; probes 3E and 3D used in the previous mapping studies; probe 7.6 shown by Saglio et al7 to be absent in DNA from an individual homozygous for the Spanish \( \beta_2 \)-thalassemia; and probe B, the most distal DNA available in the DNA cloned from the HPFH-1 individual. This was analyzed by Southern gel blotting and hybridized to probe B, bands of ~20, 4.5, and 3.3 kb were revealed (data not shown). The location of the 3' sites of these enzymes in cloned DNA, the location of the 3' sites were assigned as illustrated in Fig 5 (line 1). The location of these 3' restriction sites in normal DNA coincides with their location in DNA downstream from the 3' breakpoint of the Spanish \( \beta_2 \)-thalassemia, as illustrated in Fig 5 (line 2). Our results therefore confirm the earlier results of Camaschella et al.17

---

**Fig 5.** Restriction maps of normal and mutant DNAs in the region of 3' deletion breakpoint of HPFH-1. The positions of various probes and DNA fragments are indicated above the map. The restriction sites 3' to the breakpoint of the Spanish (Sp.) \( \beta_2 \)-thalassemia deletion were determined by Saglio et al7 and Camaschella et al.17 BgII: X. XbaI: B. BamHI: E. EcoRI: P. PstI: H. HindIII.
chella et al. that the 3' breakpoint of the Spanish (δβ) thalassemia is located approximately 8.5 to 9.0 kb downstream from that of HPFH-1.

To localize more precisely the 3' breakpoint of the Spanish (δβ) thalassemia deletion, we determined the nucleotide sequence of the DNA between the PsiI and HindIII sites (fragment A, Fig 5), determined by gene mapping to contain the breakpoint. Figure 6 shows the alignment of this DNA sequence with that of the Spanish (δβ) thalassemia and normal inter-δβ-globin gene DNA. There is no discrete point at which the DNA sequence of the Spanish (δβ) thalassemia breakpoint shifts identify from that of inter-γδ-globin gene DNA to that of fragment A. Instead, there is a stretch of sequence where all three sequences are virtually identical (16 of 22 bases) and within which a recombinational event could have occurred. Of interest is the fact that this conserved sequence coincides with the 5' end of a repetitive DNA sequence of the AluI family, the site of the 5' breakpoint of Spanish (δβ) thalassemia in the region 5' to the δ-globin gene. However, the sequence of the downstream DNA (fragment A) does not resemble that of an AluI family repetitive DNA sequence beyond the 22 bases noted above. It is conceivable that the region of the 3' breakpoint of the Spanish (δβ) thalassemia (in normal DNA) contains a truncated AluI family repetitive DNA sequence that served as the focal point for the recombination event leading to the deletion. Recombination between AluI family repetitive DNA sequences is thought to have been the cause of deletional mutagenesis in the globin

DISCUSSION

We have studied several properties of the DNA at the 3' breakpoint of the HPFH-1 deletion with the aim of determining what role, if any, these sequences may play in maintaining activity of γ-globin gene expression in adult erythroid cells of affected individuals. These normally quite remote sequences are brought into the vicinity of the γ-globin genes by the deletion event and may interfere in some way with the normal regulation of γ-globin gene expression during development.

Using the CAT assay, we have found that a 1 kb region of DNA (HPFH-3D), located approximately 1 kb 3' to the HPFH-1 deletion breakpoint, functions as an enhancer element in K562 human erythroleukemia cells. Because enhancer elements are usually found within or adjacent to functional genes, we searched for and obtained additional evidence suggesting the presence of a gene in this region of DNA. By Southern gel blotting analysis, we determined that a HpaII restriction endonuclease site located within HPFH-3D is specifically hypomethylated in both fetal and adult erythroid tissues but remains hypermethylated in nonerythroid tissues including placenta. The identical region of DNA contains a DNase I hypersensitive site in erythroid but not non-erythroid tissues (J.T. Elder and M. Groudine, personal communication, September 1986). Both hypomethylation and the presence of DNase I hypersensitive sites have been correlated with tissue-specific or stage-specific expression of many genes including the globin genes. Finally, nucleotide sequence analysis of this region of DNA has revealed the presence of a number of large open reading frames including one 292 amino acids in length, which has characteristics suggestive of an intronless gene similar to genes that encode members of the family of G-protein-coupled membrane receptors. However, we were unable to demonstrate the presence of stable mRNA transcripts corresponding to this region of DNA in either erythroid or nonerythroid cells. The inability to detect mRNA transcripts may be due to a very low steady state abundance of the mRNA or to selective expression of the putative gene at a specific stage of development, differentiation, or cellular proliferation that cannot be readily assayed.

We propose that the DNA sequences located at the 3' breakpoint of the HPFH-1 deletion may contribute to maintaining the activity of γ-globin gene expression in adult erythroid cells of individuals with HPFH-1 and HPFH-2 by one or both of the following mechanisms. The effect may be simply that of one or more enhancer elements increasing the transcription of the now neighboring γ-globin genes. Alternatively, the effect may be more nonspecific and mediated by the juxtaposition of a potentially active or "open" chromatin domain, evidenced by hypomethylation and DNase I hypersensitivity, that maintains the neighboring γ-globin genes in a transcriptionally active configuration by some type of spreading effect.

There exists a well-characterized and quite relevant precedent for the general mechanism that we propose. Hemoglobin Kenya is a mutant hemoglobin caused by nonhomologous recombination between the γ- and β-globin genes resulting in a "γβ"-fusion gene and globin chain (reviewed in reference 1). The interesting feature of this hemoglobinopathy is that it is in fact a form of HPFH with pancellular expression of the δγ-globin gene in cis to the fusion gene. Initial theories proposed that the HPFH phenotype associated with Hb Kenya might be due to the loss of control elements located in the inter-γβ gene region that is deleted in the recombination event. However, recent observations on the presence of an enhancer element in the 3'-flanking DNA of the human β-globin gene support an alternative theory, as illustrated.

| Normal 5' DNA | GATGTTATCC TTTCGGATTT TAGAGTTT CTGTGCGGGOG GAGCCCGGC GAGCGGAGG CAGAGATCTT |
| (δβ) 3' thal Sp. | GATGTTATCC TTTCGGATTT TAGAGTTT CTGTGCGGGOG GAGCCCGGC GAGCGGAGG CAGAGATCTT |
| Normal 3' DNA | CCGCGCGGG CCGCGCGGG CCGCGCGGG CCGCGCGGG CCGCGCGGG CCGCGCGGG CCGCGCGGG |

Fig 6. Sequence of normal DNA in the region corresponding to the 3' of the deletion causing Spanish (δβ) thalassemia. The normal 3' DNA sequence is contained within fragment A (Fig 5A) in DNA approximately 8.5 to 9.0 kb downstream from the HPFH-1 breakpoint. The normal 5' DNA sequence has been previously published. The arrow over the normal 5' DNA sequence indicates a repetitive DNA sequence of the AluI family.

From www.bloodjournal.org by guest on September 24, 2017. For personal use only.
in Fig 7. It is plausible and likely that in Hb Kenya the 3' β enhancer translocated into the vicinity of the γ-globin genes by the recombination event mediates the continued expression of both the Gγ and fusion Aγβ genes in adult erythroid cells of affected individuals. The HPFH-3D region, as a result of the deletions in HPFH-1 and HPFH-2, is juxtaposed to the γ genes at a further distance than is the 3' β enhancer in Hb Kenya, ie, approximately 15 kb and 10 kb from the Aγγ and Gγ promoters, respectively, versus 6 kb and 2 kb in the case of Hb Kenya (Fig 7). Nevertheless, it is quite conceivable that similar mechanisms influencing γ-globin gene expression are operative in all three of these disorders.

Additional evidence for an important role of the HPFH-3D region in the expression of the HPFH phenotype is provided by the characterization of the deletion causing Spanish (Aββ) thalassemia. In the latter disorder, Hb F synthesis is heterocellular and significantly lower than in HPFH-1 and HPFH-2 even though the deletions are very similar in size and location (Fig 1). The deletion causing Spanish (Aββ) thalassemia has a 5' breakpoint within 1 kb of that of HPFH-1, but extends an additional 8.5 to 9.0 kb in the 3' direction, thereby removing the HPFH-3D region and adjacent DNA. Therefore, the loss of the HPFH-3D region in a deletion otherwise very similar to that associated with HPFH-1 and HPFH-2 results in a phenotype characterized by significantly lower levels of Hb F. On the other hand, another deletion, that causing the Chinese type of (Aγββ) thalassemia, does not appear at first glance to support the proposed model because the 3' breakpoint of the deletion is approximately 18 kb upstream of the HPFH-3D region, thereby translocating it into the vicinity of the γ-globin genes but without a resulting HPFH phenotype. The apparent discrepancy may be explained by the fact that the Chinese (Aγββ) thalassemia is associated with partial deletion of the Aγ γ-globin gene and that two intact γ (or γ-like) genes may be required for the expression of the HPFH phenotype.

REFERENCES

16. Dunnett CW: A multiple comparisons procedure for compar-
ing several treatments with a control. J Am Statist Assoc 50:1096, 1955
42. Weiher H, Konig M, Gruss P: Multiple point mutations affecting the simian virus 40 enhancer. Science 219:626, 1983
The breakpoint of a large deletion causing hereditary persistence of fetal hemoglobin occurs within an erythroid DNA domain remote from the beta-globin gene cluster

EA Feingold and BG Forget