Distal CCAAT box deletion in the $\Lambda\gamma$ globin gene of two black adolescents with elevated fetal $\Lambda\gamma$ globin

J.G.Gilman*, N.Mishima, X.J.Wen, T.A.Stoming, J.Lobel¹ and T.H.J.Huisman

Department of Cell and Molecular Biology, Medical College of Georgia, Augusta, GA 30912-2100 and ¹Department of Pediatrics, Geisinger Medical Center, Danville, PA 17822, USA

Received July 25, 1988; Revised and Accepted October 18, 1988

ABSTRACT

Point mutations in $\zeta\gamma$ and $\Lambda\gamma$ globin gene promoters are associated with increased production of $\zeta\gamma$ and $\Lambda\gamma$ globin, respectively. To determine whether an upstream promoter mutation could account for elevated $\Lambda\gamma$ in a Black adolescent with $\Lambda\gamma$ - β^+ -HPFH and sickle cell trait, we cloned the 13 kb BglIII fragment containing both γ genes into phage lambda vector EMBL3. For one clone, the $\Lambda\gamma$ upstream promoter showed no hybridization to a 19 bp oligonucleotide whose sequence centered at -117. $\Lambda\gamma$ promoter sequence data for this mutant clone revealed a 13 bp deletion which eliminated the $\Lambda\gamma$ distal CCAAT box. Amplified $\Lambda\gamma$ genomic DNA of this and a similar case showed hybridization to both deletion-mutant and normal oligonucleotide probes. We propose that this 13 bp deletion removes part of the binding site for a repressor protein which is abundant in adult erythroid cells.

INTRODUCTION

Expression of the fetal $\zeta\gamma$ and $\Lambda\gamma$ globin genes is under strict developmental control. In the red cells of the fetus and newborn baby, fetal hemoglobin or Hb F ($\alpha_2\gamma_2$) predominates, while adults have primarily hemoglobin A ($\alpha_2\beta_2$) and Hb F is 0.3-1.2% of total hemoglobin (1).

In the heterogeneous set of genetic variants known as "hereditary persistence of fetal hemoglobin" (HPFH), adult erythroid cells produce γ globin as 2-30% of total β -like globin (1). Point mutations in the upstream region of the γ promoters are associated with increased γ gene expression in several HPFH (2). Some of the mutations occur in or near DNA sequence motifs which bind regulatory proteins in other systems (3). In particular, the -175 T $\rightarrow$ C (4,5,6) and -117 G $\rightarrow$ A (7,8) mutations alter protein binding to the γ -globin gene octamer and distal CCAAT box motifs, respectively (9,10).

We here describe a 13 bp deletion of the $\Lambda\gamma$ gene distal CCAAT box that is associated with increased $\Lambda\gamma$ globin production in two Black adolescents with $\Lambda\gamma$ - β^+ -HPFH and sickle cell trait.

MATERIALS AND METHODS

Blood collection, DNA preparation and restriction enzyme digestion of the DNA were as previously described (11,12).

Haplotype analysis of DNA was for the eight β -globin gene cluster restriction fragment length polymorphisms (RFLP) listed in reference 13.

Composition of hemoglobin (percentages of Hb F, sickle cell hemoglobin or Hb S, and normal major and minor adult hemoglobins Hb A and Hb A₂) was determined by DEAE cellulose chromatography (14). Reverse phase HPLC (15) was used to show the $^G\gamma$ and $^A\gamma$ values as percentages of total γ .

Cloning was into lambda phage vector EMBL3 and subcloning into Bluescript plasmid (Stratagene, San Diego, CA, U.S.A.) as previously described (2,16). DNA sequencing (17) used Sequenase (United States Biochemical Corporation, Cleveland, OH, U.S.A.).

Oligonucleotides were prepared using the Applied Biosystems Model 380B DNA Synthesizer. Nineteen bp oligonucleotides were used as probes: Normal -117: CAGCCTTGCCCTTGACCAAT; Greek HPFH mutant -117: CAGCCTTGCCCTTAACCAAT; 13 bp CCAAT deletion mutant: AGCCTTGCCCTTGACAAGGC; normal -196: GGCCCTTCCCCACACTAT. For hybridization studies, probes were 5' end-labelled with [γ - 32 P]-ATP as described in reference 18.

To distinguish between EMBL3 clones from the $^A\gamma$ - β^+ -HPFH and normal chromosomes, EcoRI-SalI digests of several γ -containing clones were electrophoresed in 0.8% agarose minigels. After blotting to nitrocellulose and baking 2 hr at 80°C, filters were hybridized (11) to either normal or mutant -117 probes overnight at 55°C. They were then washed in 6XSSC at 37°C for 30 min, at 41°C for 30 min, and at 55°C for 2 min.

Polymerase chain reaction DNA amplifications used Taq polymerase (Perkins Elmer Cetus, Norwalk, CT, U.S.A.), following procedures of reference 19. For amplification of the $^A\gamma$ promoter, the primers were: TGAAACTGTGGTCTTTATGAAAATTG (-622 to -597 of $^A\gamma$), and GCGCTCTGGACTAGGAGCTTATTG (complement of +30 to +53 for both $^G\gamma$ and $^A\gamma$). Hybridization procedures for dot blots were as in reference 18. DNA sequencing of amplified DNA used Sequenase and primers that were 5' end-labelled with [γ - 32 P]-ATP.

RESULTS

Two Black adolescents with sickle cell trait, K.B. and L.T., had high (30%) Hb F composed primarily of $\alpha_2^A\gamma_2$. Table I summarizes their hematology and hemoglobin composition data.

The presence of both β^A and β^S ruled out HPFH associated with deletions of the adult β gene (1). RFLP analysis showed that both chromosomes of K.B. lacked polymorphic HindIII sites in $^G\gamma$ and $^A\gamma$ globin genes (12) and the XmnI site 5' of the $^G\gamma$ gene (20). BglII digests and hybridization to γ IVSII probe showed the normal 13 kb band.

The mother of K.B. had sickle cell trait and normal levels of Hb F (<1%). She was homozygous for RFLP haplotype 19 or [-----++] (haplotype nomenclature is from reference 21). K.B. was heterozygous for haplotypes 19 and 4 or [-----++]. The father of K.B. was unavailable, but if one assumes that the high Hb F determinant is linked to the β -globin gene cluster then one can deduce that the HPFH determinant of K.B. is on haplotype 4.

Cloning of K.B. γ Genes and DNA Sequencing

To analyze the molecular basis for the elevated Hb F in this case, we decided to clone and sequence the upstream promoter regions of $^G\gamma$ and $^A\gamma$ globin genes. The lack of polymorphic

TABLE I: HEMATOLOGICAL DATA AND HEMOGLOBIN COMPOSITION FOR TWO BLACK $\alpha\gamma$ - β^+ -HPFH/Hb S

SUBJECT	K.B.	L.T.
Hb (g/dL)	9.3	13.4
PCV (L/L)	0.265	0.375
RBC (10^{12} /L)	3.44	5.04
MCV (fL)	77.0	74.0
MCH (pg)	27.0	26.6
Hb A ₂ (%)	2.1	2.1
Hb S (%)	44.7	47.5
Hb A (%)	21.4	20.3
Hb F (%)	31.8	30.1
γ (%)	20.8; 16.7	7.7; 12.1
α (%)	79.2; 83.3	92.3; 87.9
Sex/Age	F/16 (pregnant)	M/12

markers to distinguish the γ genes of the two chromosomes led to the strategy of screening clones for known up-mutations near -200 and -117.

The BglII 13 kb fragment containing both γ genes was cloned in lambda phage vector EMBL3. Clones were digested with SalI (which cuts at the boundary of phage arms and the insert) and EcoRI (which digests within the insert). This was expected to yield 2.6 and 3.1 kb fragments containing the $\alpha\gamma$ and γ genes, respectively.

Fig. 1: Identification of normal and HPFH EMBL3 clones which were selected to have the 13 kb BglII fragment containing both γ genes. DNA was digested with SalI and EcoRI, which was expected to yield 2.6 kb $\alpha\gamma$ and 3.1 kb γ fragments. Hybridization was with 19-mer oligonucleotide with normal γ sequence centered at -117. See text for further details.

Fig. 2: DNA sequence data for the cloned HPFH mutant A_γ promoter identified by the hybridization data of Fig. 1. See text for details.

Both bands were seen for all clones after agarose gel electrophoresis and blot hybridization to a 19-mer with normal sequence centered at -196 (data not shown). These data suggested that this HPFH did not have the -196, -198 or -202 mutations.

Use of a 19-mer with normal sequence centered at -117 showed both bands in some clones, but only the 3.1 kb G_γ fragment in others. This is illustrated in Fig. 1. The 19-mer with the -117 G→A Greek HPFH mutation (7,8) did not give any hybridization, suggesting that this case had a new mutation near -117 of A_γ .

The clone corresponding to lane 2 of Fig. 1 was used for subcloning and DNA sequencing from -383 to the Cap site. Fig. 2 reveals that the A_γ gene has a 13 bp deletion in the vicinity of the distal CCAAT box. (Normal DNA sequence data are from reference 22.) In addition, both G_γ and A_γ had G instead of C at -369, which is a known polymorphism for A_γ (23). The G_γ gene also had the -309 A→G mutation previously seen in the β^S Benin haplotype (24).

Dot Blot and Sequencing of Amplified DNA

In order to test for this 13 bp deletion in L.T., and to exclude artifacts due to cloning, the A_γ promoter regions of K.B. and L.T. were amplified from -622 to +53. Fig. 3 shows the results of dot blot hybridization to 19-mer oligonucleotides corresponding to normal and mutant (13 bp deletion) sequences.

Fig. 3: Dot blot hybridization of cloned and amplified genomic DNA with normal and $\Delta\gamma$ - β^+ -HPFH 19-mer oligonucleotide probes. Sample 1: Plasmid subclone with the 13 bp deletion (mutant) $\Delta\gamma$ promoter. Sample 2: Amplified $\Delta\gamma$ promoter DNA from $\Delta\gamma$ - β^+ -HPFH/Hb S heterozygote K.B. Sample 3: Amplified $\Delta\gamma$ promoter DNA from $\Delta\gamma$ - β^+ -HPFH/Hb S heterozygote L.T. Sample 4: Amplified $\Delta\gamma$ promoter DNA from a normal control.

Amplified DNA of both cases (samples 2 and 3) showed strong hybridization to mutant and normal probes. This confirms that L.T. and K.B. are heterozygotes for the 13 bp deletion.

Amplified DNA of L.T. was sequenced, and the sequence data (not shown) were consistent with the 13 bp deletion.

DISCUSSION

In non-deletion HPFH, elevated γ globin production *in vivo* has been associated with various point mutations in the upstream promoters of the fetal γ globin genes. Mutations associated with strong effects on γ gene expression occur near the distal CCAAT box: $\Delta\gamma$ -117 G>A (7,8). They also are found in the octamer, -175 T>C in $\epsilon\gamma$ (4,5) and $\Delta\gamma$ (6), and in a cluster further upstream, $\Delta\gamma$ -196 C>T, $\Delta\gamma$ -198 T>C, $\Delta\gamma$ -202 C>T, and $\epsilon\gamma$ -202 C>G (reviewed in reference 2). *In vitro* studies of expression of transfected DNA in K562 cells have confirmed that several of these mutations cause increased γ gene expression (25).

In this report we have described the first instance of a deletion in the γ globin upstream promoter associated with increased *in vivo* Hb F. The net effect of this deletion is to remove one of two duplicated GCCTTGAC sequences (at -122 to -115, and -109 to -102), as well as the CAATA sequence (-114 to -110) between them. Hence, the $\Delta\gamma$ distal CCAAT box (-115 to -111) is eliminated by this deletion, as illustrated in Fig. 2.

One should note that the precise location of the 5' end of this deletion is ambiguous: It could be anywhere between -122 and -114. This suggests that misalignment between the two GCCTTGAC sequences that occur 13 bp apart could account for the 13 bp deletion. "Slipped mispairing" during DNA replication is the likely mechanism (16,26).

Elimination of a CCAAT box would be expected to affect transcription, since point mutations in the sequence GACCAAT have significant effects (27). Furthermore, the $\Delta\gamma$ -117 G>A mutation in the distal GACCAAT sequence is associated with greatly elevated transcription in Greek HPFH (7,8).

GACCAAT binds polypeptides that activate transcription in a variety of promoters (28). Nuclear protein CP1, found in both erythroid and non-erythroid cells, binds more tightly to the proximal than to the distal γ CCAAT box (9). Two specifically

Fig. 4: Model to explain the effect of the 13 bp promoter deletion in activating transcription. See text for details.

erythroid nuclear proteins also bind in this region: NFE-1 binds at the GACAAGG sequence (-104 to -98) and at -117, while NFE-2 binds at the distal CCAAT box (9,10,29).

The -117 mutation decreased binding of both NFE-1 and NFE-2, but increased binding of CP1 (9,10). Preliminary data indicate that the 13 bp HPFH deletion abolishes binding of CP1 and NFE-2, but reduces binding of NFE-1 by about 50% (29).

While NFE-1 may be a repressor, the fact that the HPFH -175 T→C mutation increases NFE-1 binding (10,29) suggests a role for NFE-1 as an erythroid-specific activator. In addition, both the CCAAT and GACAAGG sequences diverged from identical segments of a triplicated region: GCCTTGACcaata (-122 to -110), GCCTTGACAaggc (-109 to -97) and aaCTTGACcaata (-95 to -83). Perhaps an ancestral protein of CP1 and NFE-1 duplicated and diverged in concert with the diverging binding sites. One may therefore hypothesize that the 13 bp HPFH deletion eliminates binding of the repressor NFE-2 to the distal CCAAT box, thus permitting activation by NFE-1 and CP1 binding to the GACAAGG motif and the proximal CCAAT box, respectively.

An alternative hypothesis to account for the activating effect of the 13 bp HPFH deletion is suggested by the observation that a "CCAAT-displacement protein" (CDP) is found in erythroid and non-erythroid cells and binds to all three TGA elements of this triplicated region (9). As illustrated in Fig. 4, we suggest that binding of CDP could interfere with CCAAT-binding factors such as CP1. If adult erythroid cells produced CDP in relatively high concentration, then γ -gene transcription would be repressed. The 13 bp deletion, by removing one of the TGA sites, would decrease the affinity of CDP for this region and permit CP1 to activate transcription by binding to the proximal CCAAT box.

ACKNOWLEDGEMENTS

We thank Ms. J.S. Dansie for assistance in preparing this manuscript. This research was supported in part by National Institutes of Health Research Grants Nos. DK35443 (to J.G.G.), HL05168 and HL15158 (to T.H.J.H.), and HL15966 to the Children's Hospital Research Foundation of Cincinnati and the Cincinnati Comprehensive Sickle Cell Center. This is contribution number 1139 of the Department of Cell and Molecular Biology, Medical College of Georgia.

*To whom correspondence should be addressed

REFERENCES

1. Karlsson, S. and Nienhuis, A.W. (1985) *Ann. Rev. Biochem.* **54**, 1071-1108.
2. Gilman, J.G., Mishima, N., Wen, X.J., Kutlar, F. and Huisman, T.H.J. (1988) *Blood* **72**, 78-81.
3. Gilman, J.G. (1988) Hemoglobin (In press).
4. Surrey, S., Delgrosso, K., Malladi, P. and Schwartz, E. (1988) *Blood* **71**, 807-810.
5. Ottolenghi, S., Nicolis, S., Taramelli, R., Margaretti, N., Mantovani, R., Comi, P., Giglioni, B., Longinotti, M., Dore, F., Oggiano, L., Pistidda, P., Serra, A., Camaschella, C., and Saglio, G. (1988) *Blood* **71**, 815-817.
6. Stoming, T.A., Stoming, G.S., Lanclos, K.D., Fei, Y.S., Altay, C., Kutlar, F. and Huisman, T.H.J. (1988) *Blood* (In press).
7. Collins, F.S., Metherall, J.E., Yamakawa, M., Pan, J., Weissman, S.M. and Forget, B.G. (1985) *Nature* **313**, 325-326.
8. Gelinas, R., Endlich, B., Pfeiffer, C., Yagi, M. and Stamatoyanopoulos, G. (1985) *Nature* **313**, 323-325.
9. Superti-Furga, G., Barberis, A., Schaffner, G. and Busslinger, M. (1988) *Embo. J.* **7**, 3099-3107.
10. Mantovani, R., Margaretti, N., Nicolis, S., Ronchi, A., Giglioni, B. and Ottolenghi, S. (1988) *Nucleic Acids Res.* **16**, 7783-7797.
11. Gilman, J.G., Huisman, T.H.J. and Abels, J. (1984) *Brit. J. Haemat.* **56**, 339-348.
12. Gilman, J.G. and Huisman, T.H.J. (1984) *Blood* **64**, 452-457.
13. Hattori, Y., Kutlar, F., Kutlar, A., McKie, V.C. and Huisman, T.H.J. (1986) *Hemoglobin* **10**, 623-642.
14. Schroeder, W.A. and Huisman, T.H.J. (1980) *The Chromatography of Hemoglobin*, Marcel Dekker, New York.
15. Shelton, J.B., Shelton, J.R. and Schroeder, W.A. (1984) *J. Liq. Chromatogr.* **7**, 1969-1977.
16. Gilman, J.G., Johnson, M.E. and Mishima, N. (1987) *Brit. J. Haemat.* **68**, 455-458.
17. Gilman, J.G., Mishima, N. and Johnson, M.E. (1988) *United States Biochemical Corp. Newsletter "Editorial Comments"* **14**, 17.
18. Diaz-Chico, J.C., Yang, K.G., Yang, K.Y., Efremov, D.G., Stoming, T.A. and Huisman, T.H.J. (1988) *Biochim. Biophys. Acta* **949**, 43-48.
19. Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) *Science* **239**, 487-491.
20. Gilman, J.G. and Huisman, T.H.J. (1985) *Blood*, **66**, 783-787.

21. Antonarakis, S.E., Boehm, C.D., Serjeant, G.R., Theisen, C.E., Dover, G.J. and Kazazian, H.H., Jr. (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81, 853-856.
22. Shen, S-h., Slightom, J.L. and Smithies, O. (1981) *Cell* 26, 191-203.
23. Collins, F.S., Stoeckert, C.J., Jr., Serjeant, G.R., Forget, B.G. and Weissman, S.M. (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81, 4894-4898.
24. Month, S., Delgrosso, K., Orchowski, P., Rappaport, E., Malladi, P., Schwartz, E. and Surrey, S. (1986) *Blood* 68, 76a.
25. Bodine, D.M., Gray, W.G., Mickley, L.A. and Ley, T.J. (1986) *Blood* 68, 71a.
26. Efstratiadis, A., Posakony, J.W., Maniatis, T., Lawn, R.M., O'Connell, C., Spritz, R.A., DeRiel, J.K., Forget, B.G., Weissman, S.M., Slightom, J.L., Blechl, A.E., Smithies, O., Baralle, F.E., Shoulders, C.C. and Proudfoot, N.J. (1980) *Cell* 21, 653-668.
27. Myers, R.M., Tilly, K. and Maniatis, T. (1986) *Science* 232, 613-618.
28. Jones, K.A., Kadonaga, J.T., Rosenfeld, P.J., Kelly, T.J. and Tjian, R. (1987) *Cell* 48, 79-89.
29. Gilman, J.G., Mantovani, R. and Ottolenghi, S. (1988) *Blood* (In press).