

# Role of the duplicated CCAAT box region in $\gamma$ -globin gene regulation and hereditary persistence of fetal haemoglobin

Antonella Ronchi<sup>1</sup>, Meera Berry<sup>2</sup>,  
Selina Raguz<sup>2</sup>, Ali Imam<sup>3</sup>,  
Nikos Yannoutsos<sup>3</sup>, Sergio Ottolenghi<sup>1</sup>,  
Frank Grosveld<sup>2,3,4</sup> and Niall Dillon<sup>2,5</sup>

<sup>1</sup>Università Degli Studi di Milano, Dipartimento di Genetica e di Biologia di Microrganismi, Via Celoria 26, 20133 Milano, Italy,

<sup>2</sup>Laboratory of Gene Structure and Expression, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK and <sup>3</sup>Erasmus University Rotterdam, Faculty of Medicine, Department of Cell Biology and Genetics, PO Box 1738/dr. Molewaterplein 50, 3000 DR Rotterdam, The Netherlands

<sup>5</sup>Present address: MRC Clinical Sciences Centre, Royal Postgraduate Medical School, Hammersmith Hospital, Du Cane Road, London W12 0NN, UK

<sup>4</sup>Corresponding author

**Hereditary persistence of fetal haemoglobin (HPFH) is a clinically important condition in which a change in the developmental specificity of the  $\gamma$ -globin genes results in varying levels of expression of fetal haemoglobin in the adult. The condition is benign and can significantly alleviate the symptoms of thalassaemia or sickle cell anaemia when co-inherited with these disorders. We have examined structure-function relationships in the -117 HPFH  $\gamma$  promoter by analysing the effect of mutating specific promoter elements on the functioning of the wild-type and HPFH promoters. We find that CCAAT box mutants dramatically affect expression from the HPFH promoter in adult blood but have little effect on embryonic/fetal expression from the wild-type promoter. Our results suggest that there are substantial differences in the structure of the wild-type  $\gamma$  promoter expressed early in development and the adult HPFH promoter. Together with previous results, this suggests that  $\gamma$  silencing is a complex multifactorial phenomenon rather than being the result of a simple repressor binding to the promoter. We present a model for  $\gamma$ -globin gene silencing that has significant implications for attempts to reactivate the  $\gamma$  promoters in human adults by pharmacological means.**

**Keywords:** CCAAT box/ $\gamma$ -globin gene/hereditary persistence of fetal haemoglobin/HPFH promoter

## Introduction

Hereditary persistence of fetal haemoglobin (HPFH) is a genetically inherited condition which results in continued expression of the fetal  $\gamma$ -globin genes in adult life (reviewed by Poncz *et al.*, 1989). The adult expression of  $\gamma$ -globin has no adverse effects and, when a HPFH mutation is co-inherited with a  $\beta$ -thalassaemia allele, the presence of  $\gamma$  chains can partially compensate for the absence of  $\beta$ -chains, giving rise to a much milder disease (Stamatoyannopoulos

and Nienhuis, 1987). Adult expression of  $\gamma$  chains can also alleviate the symptoms of sickle cell anaemia (Schechter *et al.*, 1987). This ameliorating effect is thought to occur because the addition of a fetal haemoglobin (HbF) molecule to the polymerizing sickle haemoglobin results in a block to further polymerization (Schechter *et al.*, 1987). Similar beneficial effects would be obtained if adult  $\gamma$  expression could be induced by pharmacological means. Reactivation of the  $\gamma$  genes, if it could be achieved, would provide a realistic alternative to gene therapy for these diseases (Stamatoyannopoulos and Nienhuis, 1992).

The non-deletion HPFHs have been found to be associated with a number of different point mutations in the  $\gamma$ -globin promoter (reviewed by Poncz *et al.*, 1989). A recent study demonstrated that a single point mutation at -117 in the duplicated CCAAT box region of the A $\gamma$ -globin gene promoter associated with Greek HPFH (Collins *et al.*, 1985; Gelinas *et al.*, 1985) was sufficient to produce a HPFH phenotype in transgenic mice (Berry *et al.*, 1992). The levels of  $\gamma$  expression observed in these mice were at least as high as those associated with the same mutation in humans.

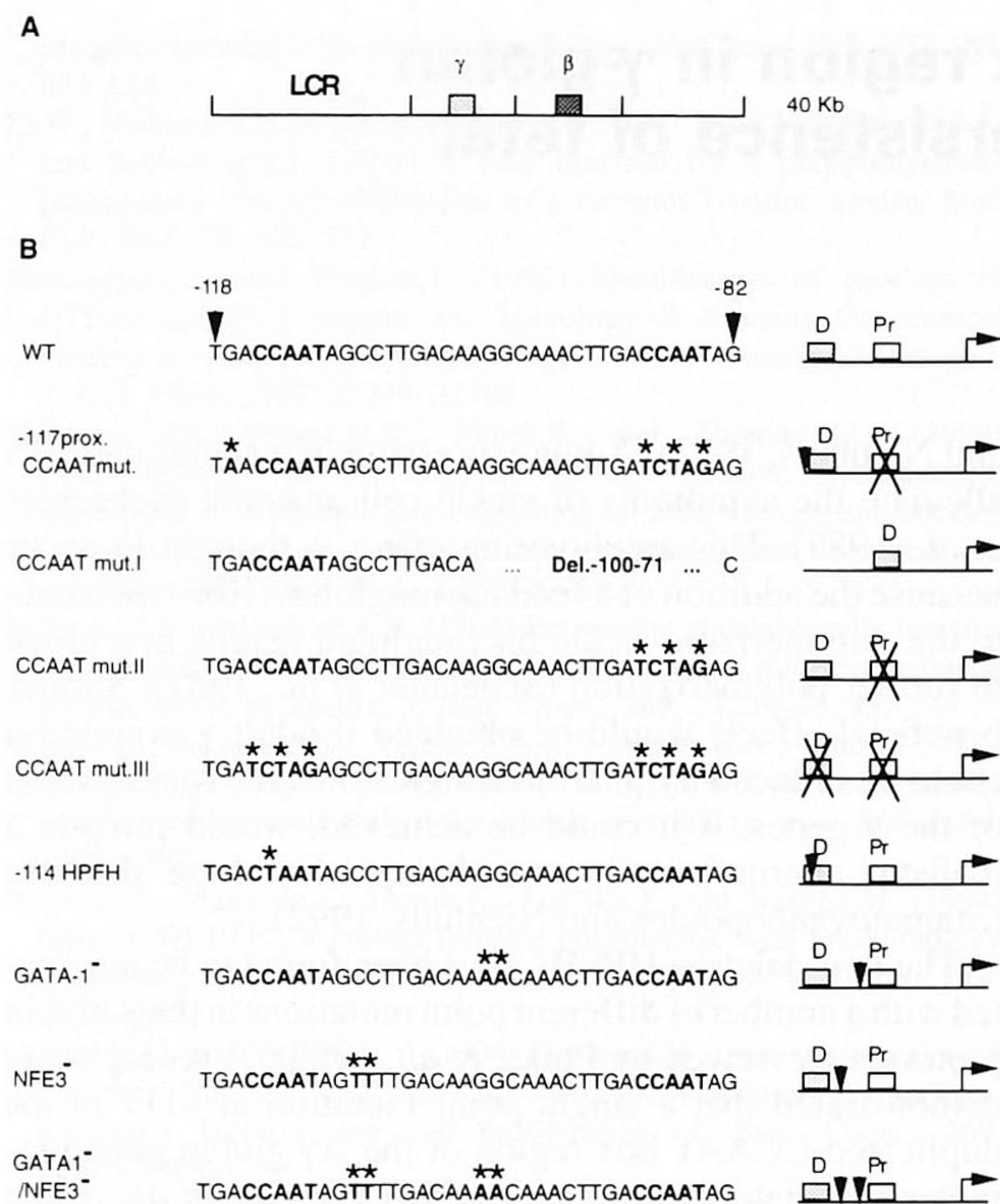
Adult expression from the -117 HPFH promoter represents a change in the developmental specificity of the  $\gamma$  promoter. A key question is whether the fetal-specific promoter uses the same combination of factors as the HPFH promoter which is active in adult erythroid cells. According to this scenario, expression in adult life would result primarily from a lifting of repression. An alternative possibility is that the factor environment is significantly different in adult erythroid cells compared with those of the fetus, with the result that the transcriptionally active HPFH promoter has a quite different structure from the fetal-specific promoter. The location of the duplicated CCAAT box elements suggested that they might be closely involved in the functioning of the -117 HPFH promoter, and differences in factor binding to oligonucleotides spanning the region that contains the -117 mutation have also been observed (Superti-Furga *et al.*, 1988; Mantovani *et al.*, 1989; Berry *et al.*, 1992). We have therefore used mutagenesis to examine the role of these sequences in transcription from the wild-type and HPFH promoter. We find that the proximal CCAAT box is relatively unimportant for embryonic expression but becomes progressively more important as development proceeds. Our results strongly suggest that the HPFH promoter has quite different properties and factor requirements from the wild-type fetal specific promoter. Taken together with studies involving mutagenesis of factor binding sites, they argue against a simple repressor model for  $\gamma$  silencing.

## Results

### **The -117 HPFH phenotype is dependent on the presence of a functional proximal CCAAT box**

The strategy for analysing the effect of mutations on  $\gamma$  expression has been described previously (Berry *et al.*,

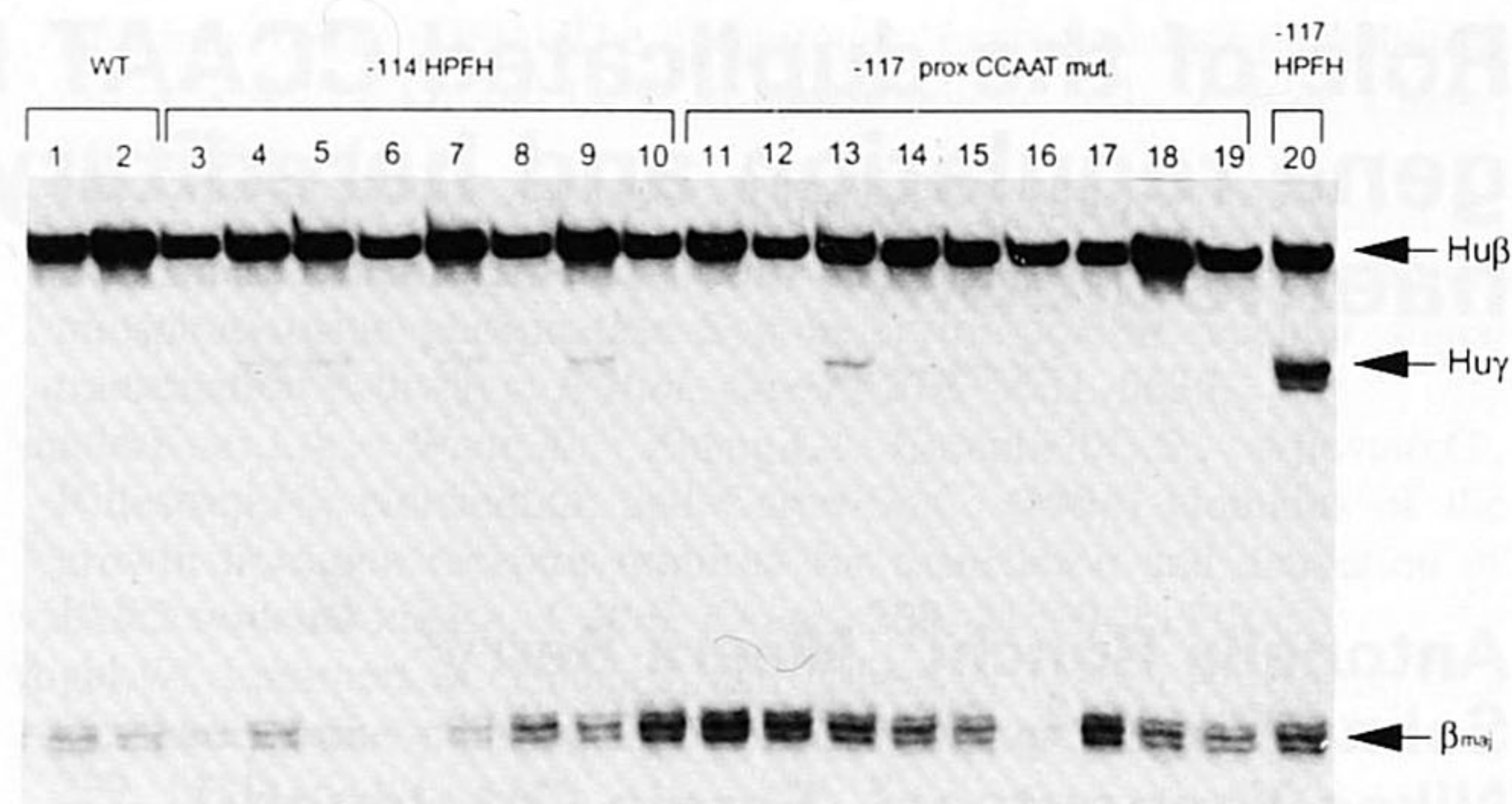




**Fig. 1.** (A) The construct used to analyse the promoter mutants contains the human  $\gamma$  and  $\beta$  genes under the control of the entire locus control region (Berry *et al.*, 1992). (B) Mutation tested in transgenic mice: the CCAAT box regions are shown in bold type. Mutated bases are underlined. -117prox.CCAAT mut.: the HPFH mutation at -117 was combined with a triple mutation destroying the proximal CCAAT box. CCAAT mut.I: the proximal CCAAT box region between nucleotides -100 and -71 was deleted, bringing the distal CCAAT box into the position normally occupied by the proximal. In CCAAT mut.II a triple mutation disrupts the proximal CCAAT box; the same mutation eliminates both proximal and distal CCAAT boxes in CCAAT mut.III. Mutants -98/-99 GG $\rightarrow$ AA and -107/-108 CC $\rightarrow$ TT abolish the binding of GATA-1 and NFE3 respectively, in gel shift experiments. The double mutant GATA1<sup>-</sup>/NFE3<sup>-</sup> fails to bind both factors.

1992). Oligonucleotide mutagenesis was used to introduce mutations into the  $\gamma$  promoter, and transcription was assessed in transgenic mice using a construct that also contained the locus control region (LCR) and a human  $\beta$  gene (Figure 1). The  $\beta$  gene acts as a reference allowing us to carry out precise quantitation of the relative levels of  $\gamma$  expression in adult blood and to measure the ability of the  $\gamma$  gene to compete with  $\beta$  at different stages of development.

The strategy described above was first used to test whether the proximal CCAAT box is required for the high levels of  $\gamma$  expression observed with the -117 HPFH mutation. A construct was made in which this mutation (-117 G $\rightarrow$ A) was combined with three base changes that completely disrupt the proximal CCAAT box (Figure 1). A total of nine founder animals was generated with this construct. Analysis of adult blood RNA from these animals shows that three animals fail to show any detectable  $\gamma$  expression while the remaining six express  $\gamma$  at levels ranging from 2.0 to 13.8% of  $\gamma + \beta$  (Figure 2, lanes 11–19). The mean of these values represents a reduction of >5-fold compared with the mean level of  $\gamma$  expression observed with the -117 mutation alone (Figure 5). We



**Fig. 2.** S1 analysis of RNA from adult blood from mice transgenic for the -114 and -117 prox. CCAAT mut. constructs. Each lane is derived from a separate founder animal. WT indicates RNA derived from animals carrying a wild-type  $\gamma$ -globin gene and is used as a negative control. The -117 HPFH lane derives from Line 61 described by Berry *et al.* (1992) and is used as a positive control. RNA in lanes 3–10 is derived from animals carrying the -114  $\gamma\beta$  gene globin construct. Lanes 11–19 contain RNA from animals carrying the -117 proximal CAAT box mutant  $\gamma\beta$  construct (see Figure 1). S1 analysis was carried out as described by Berry *et al.* (1992). The protected fragments from the human  $\beta$  probe, the human  $\gamma$  probe and the murine  $\beta$  maj probe are indicated as Hu $\beta$ , Hu $\gamma$  and  $\beta$ maj.

conclude, therefore, that the high level of  $\gamma$  expression which is a consequence of the -117 mutation is completely dependent on the presence of the proximal CCAAT box.

#### **Effect of CCAAT box mutations on embryonic expression of the $\gamma$ gene**

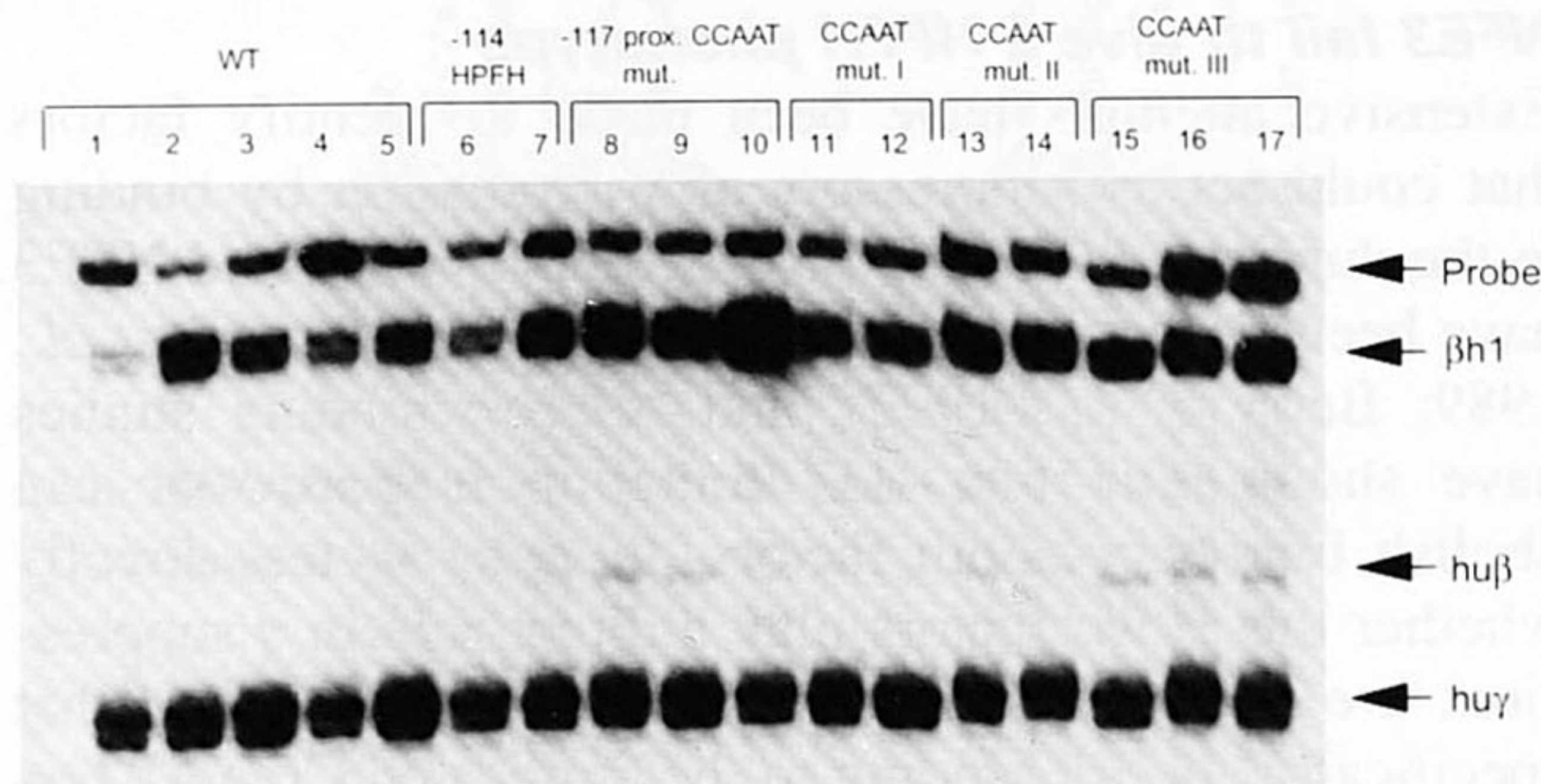
Since it was clear that the proximal CCAAT box is essential for efficient expression of a -117 HPFH  $\gamma$  gene in adult erythroid tissues, we also wanted to know whether it was equally important for expression during the fetal stages when the wild-type gene is normally expressed. In transgenic mice, the wild-type  $\gamma$  genes are expressed in the embryonic yolk sac and the early fetal liver before being silenced in the late fetal liver and adult bone marrow. Expression levels of mutated transgenes at these stages would be expected to give an indication of the role of specific elements in potentiating expression in human fetal liver. We tested three different constructs containing mutations in the CCAAT boxes (Figure 1). In mutant I a deletion was introduced that removed the proximal CCAAT box and placed the distal CCAAT box in the position normally occupied by the proximal CCAAT box. Mutant II contained three base changes in the proximal CCAAT box that completely destroyed the consensus sequence, while in mutant III the same base changes were used to disrupt both proximal and distal CCAAT boxes.

We compared the function of the mutant  $\gamma$  genes with that of the wild-type gene in embryonic yolk sac and fetal liver. This allowed us to assess the relative contribution of the CCAAT box sequences to  $\gamma$  expression at different stages of development.

The construct containing the wild-type gene gives high levels of  $\gamma$  expression in the embryonic yolk sac and no detectable expression of  $\beta$  (Berry *et al.*, 1992 and Figure 3, lanes 1–5). The  $\beta$  gene has been shown to be developmentally permissive and is suppressed in the embryo only when there is transcriptional competition from a linked  $\gamma$  gene (Behringer *et al.*, 1990; Enver *et al.*, 1990; Hanscombe *et al.*, 1991). The competition is highly polar



and depends on the  $\gamma$  gene being located between the  $\beta$  gene and the LCR (Hanscombe *et al.*, 1991). Competition between the genes in the locus is related directly to the level of transcription of the competing genes (N.Dillon, J.Strouboulis and F.Grosveld, in preparation). Therefore, the expectation is that any significant disabling of the  $\gamma$  promoter as a result of the CCAAT box mutations would lead to a decline in  $\gamma$  expression and a corresponding increase in  $\beta$  expression. However, with each of the mutant constructs,  $\gamma$  continued to express at high levels while only a slight increase in  $\beta$  expression was observed (Figure 3, Table I). The highest level of  $\beta$  expression was observed with the double CCAAT box mutant, while only trace



**Fig. 3.** S1 analysis of RNA from yolk sacs obtained from 10.5 day embryos transgenic for wild-type and mutant constructs. The different mutations (Figure 1) are indicated above the lanes. All the embryos were founders, with the exception of the -114 transgenics which were obtained from timed matings of transgenic males to non-transgenic females. S1 analysis was carried out as in Figure 2. The protected fragment of the mouse embryonic  $\beta$ h1 globin probe is indicated as  $\beta$ h1.

levels were observed when the distal CCAAT box was moved into the position normally occupied by the proximal CCAAT box.

Figure 4 shows the analysis of the wild-type and mutant constructs in 13.5 day fetal liver when  $\beta$  and  $\gamma$  are co-expressed. The wild-type  $\gamma$  gene is still expressed at a higher level than  $\beta$  at this stage in this construct. When compared with the wild-type gene, all three mutant constructs show a significant (25–40%) drop in the expression of  $\gamma$  relative to  $\gamma + \beta$  (Figure 4, Table I). However, it should be noted that  $\gamma$  is still expressed at high levels which are comparable with those of  $\beta$  even when both CCAAT boxes have been inactivated.

A comparison of the mean values for all of the CCAAT box mutants is shown in Figure 5. It is clear that inactivating the CCAAT boxes has only a very slight effect on the wild-type promoter early in development, while inactivation of the proximal CCAAT box results in a very substantial reduction in the functioning of the -117 HPFH promoter in the adult.

**Analysis of the -114 HPFH mutation in mice**

Since we are attempting to draw quite detailed conclusions about the regulation of a human promoter using transgenic mice as a functional assay, it is important to know whether the mechanism of  $\gamma$  silencing is the same in mice and humans. A point mutation at -114 has been associated with adult  $\gamma$  expression of 3–5% per copy relative to  $\beta$  in a Georgian family (Oner *et al.*, 1991). This is considerably lower than the level of 10–20% observed in humans carrying the -117 Greek HPFH mutation (Wood, 1993). We reasoned that a comparison of the effect of the two mutations in mice would provide a good indication of

**Table I.** Levels of  $\gamma$  expression obtained from wild-type and mutated constructs in embryonic yolk sac, fetal liver and adult blood

Yolk Sac (Figure 3)			Fetal liver (Figure 4)			Blood of adult animals (Figure 2)		
Construct	Lane	%Hu $\gamma$ /Hu $\gamma$ + $\beta$	Construct	Lane	%Hu $\gamma$ /Hu $\gamma$ + $\beta$	Construct	Lane	%Hu $\gamma$ /Hu $\gamma$ + $\beta$
Wt	1	>99	Wt	1	85.9	Wt	1	<1
	2	>99		2	64.1		2	<1
	3	>99		3	68.5	-114HPFH	3	5.2
	4	>99		3*	86.6		4	7.3
	5	>99		4*	79.4		5	7.8
-114 HPFH	6	>99	-114HPFH	1*	92.5	6	4.0	
	7	>99		2*	92.3	7	3.6	
-117Prox.mut	8	94.1	CCAATmut.I	4	59.8	8	2.1	
		95.3		5	55.9	9	7.4	
		>99		6	48.3	10	<1	
CCAATmut.I	11	>99	CCAATmut.II	7	59.5	-117Prox.mut	11	<1
		>99		8	33.3		12	<1
CCAATmut.II	13	95.7	CCAATmut.III	9	48.3		13	13.8
		97.8		10	59.0	14	2.0	
		92.5		11	56.2	15	9.0	
CCAATmut.III	16	94.3				16	7.6	
		93.6				17	<1	
						18	2.2	
						19	6.3	
						-117HPFH <sup>a</sup>	20	35.6

<sup>a</sup>-117HPFH values taken from ref.7: 30.5, 27.7, 22.2, 22.2, 16.6, 16.6, 33.3, 16.6, <1

Values were obtained by quantitation of the gels shown in Figures 2, 3 and 4 using a PhosphorImager. The levels are expressed as the percentage of  $\gamma$  expression relative to the total expression of  $\gamma$  and human  $\beta$ . The \* in the WT and -114 construct are the lines shown in Figure 2. The value obtained for -117 HPFH line 61 (Figure 2, lane 12) was used as a reference to correct the values for the other -117 HPFH lines reported by Berry *et al.* (1992), allowing them to be compared directly with the constructs used in this study.



whether silencing of the  $\gamma$  genes in humans and mice occurs through similar mechanisms.

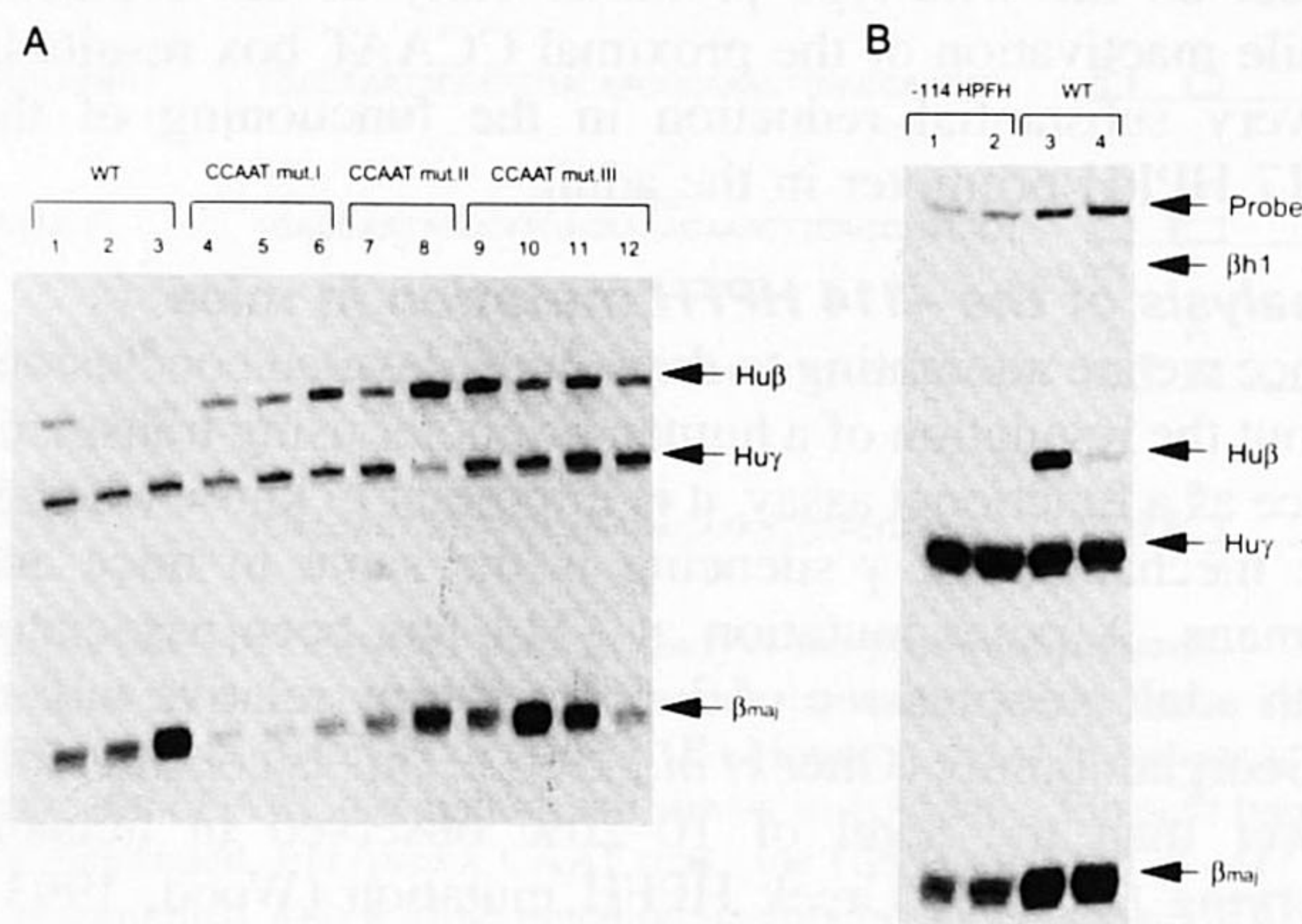
A construct containing the -114 mutation was used to generate a total of eight adult founder animals (Figure 1). S1 mapping analysis was used to measure the level of  $\gamma$ -globin mRNA compared with human  $\beta$ -globin mRNA. RNA from two animals carrying the wild-type gene in a similar construct was analysed for comparison. In all of the animals analysed, the -114 mutation resulted in a significant increase in  $\gamma$  expression (Figure 2, Table I). The level of  $\gamma$  expression ranged from <1 to 7.8% of human  $\beta$ . The fluctuation in levels between different founders is similar to that observed with the -117 HPFH and is presumably due to position effects (Berry *et al.*, 1992). The mean value of 4.7% is very similar to that observed in humans, and suggests a high degree of

homology between the mechanism of  $\gamma$  silencing in humans and transgenic mice.

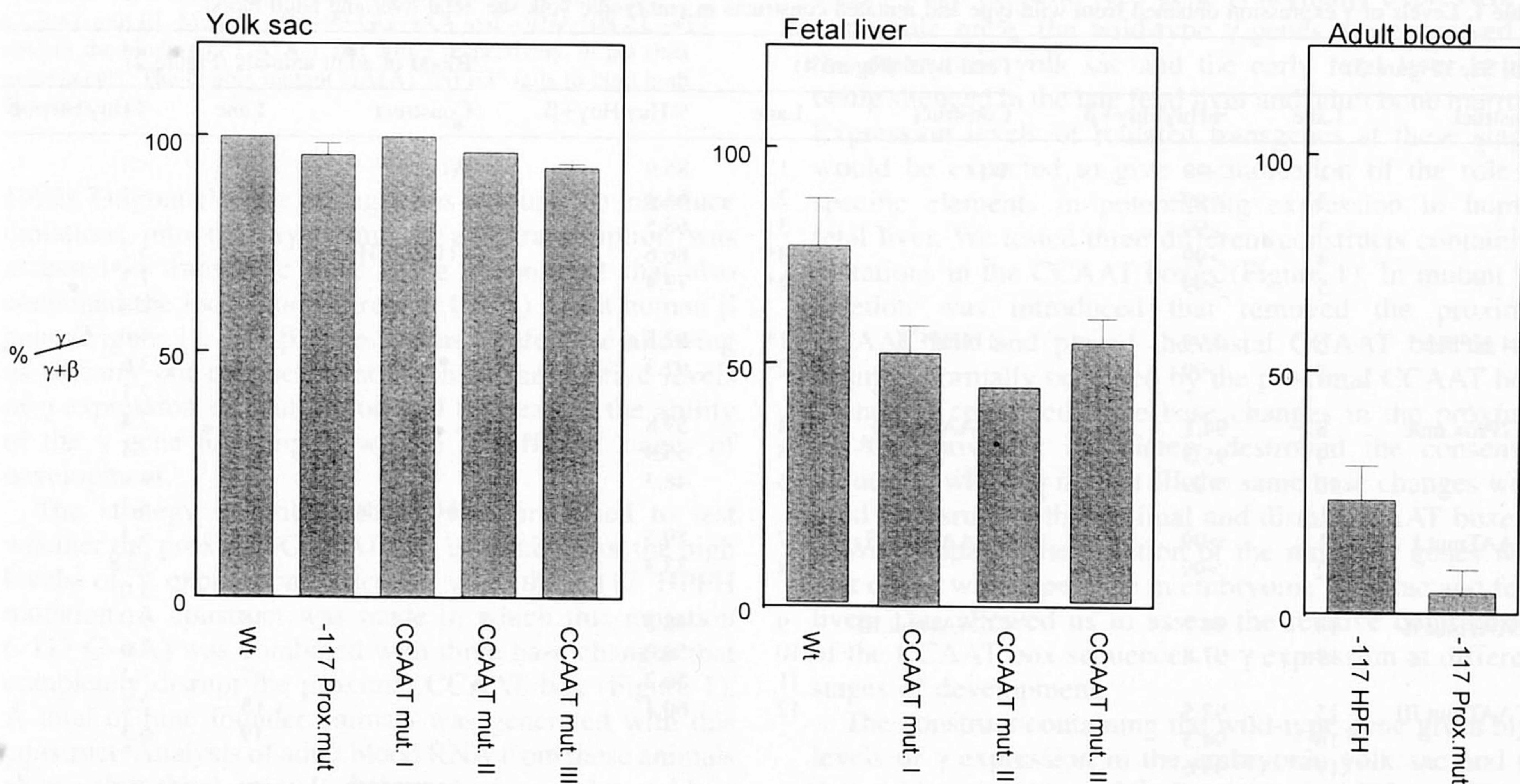
We also carried out timed matings and analysed the effect of the -114 mutation on  $\gamma$  and  $\beta$  expression in 13.5 day fetal liver. We find that the relative expression of the mutant  $\gamma$  gene is much higher than that of the wild-type gene at the same stage and that  $\beta$  expression is partially suppressed (Figure 4, right panel). Increased competition of  $\beta$  was also observed in 13.5 day fetal liver with a  $\gamma$  transgene containing the -117 HPFH mutation (Berry *et al.*, 1992).

#### Mutations that abolish binding of GATA-1 and NFE3 fail to give a HPFH phenotype

Extensive attempts have been made to identify factors that could act as suppressors of  $\gamma$  expression by binding to the duplicated CCAAT box region. GATA-1 and NFE3 have been shown to bind to this region (Mantovani *et al.*, 1989; Berry *et al.*, 1992), and *in vitro* binding studies have shown that the -117 mutation in particular can abolish binding of both factors. In order to test directly whether these two factors play a direct role in  $\gamma$  suppression, we generated mutations in the contact sites that specifically abolish binding to the duplicated CCAAT box sequences (Figures 1 and 6). A double mutation at -103/-104 prevents binding of GATA-1 without affecting binding of the other factors (Figure 6). Band-shift analysis showed that a double mutation at positions -107/-108 completely abolishes binding of NFE3 without affecting binding of GATA-1, CP1 or NFE6 (Figure 6). Constructs containing these mutations were used to generate transgenic mice, and blood from adult transgenic founders was analysed for  $\gamma$  expression. Neither mutation resulted in

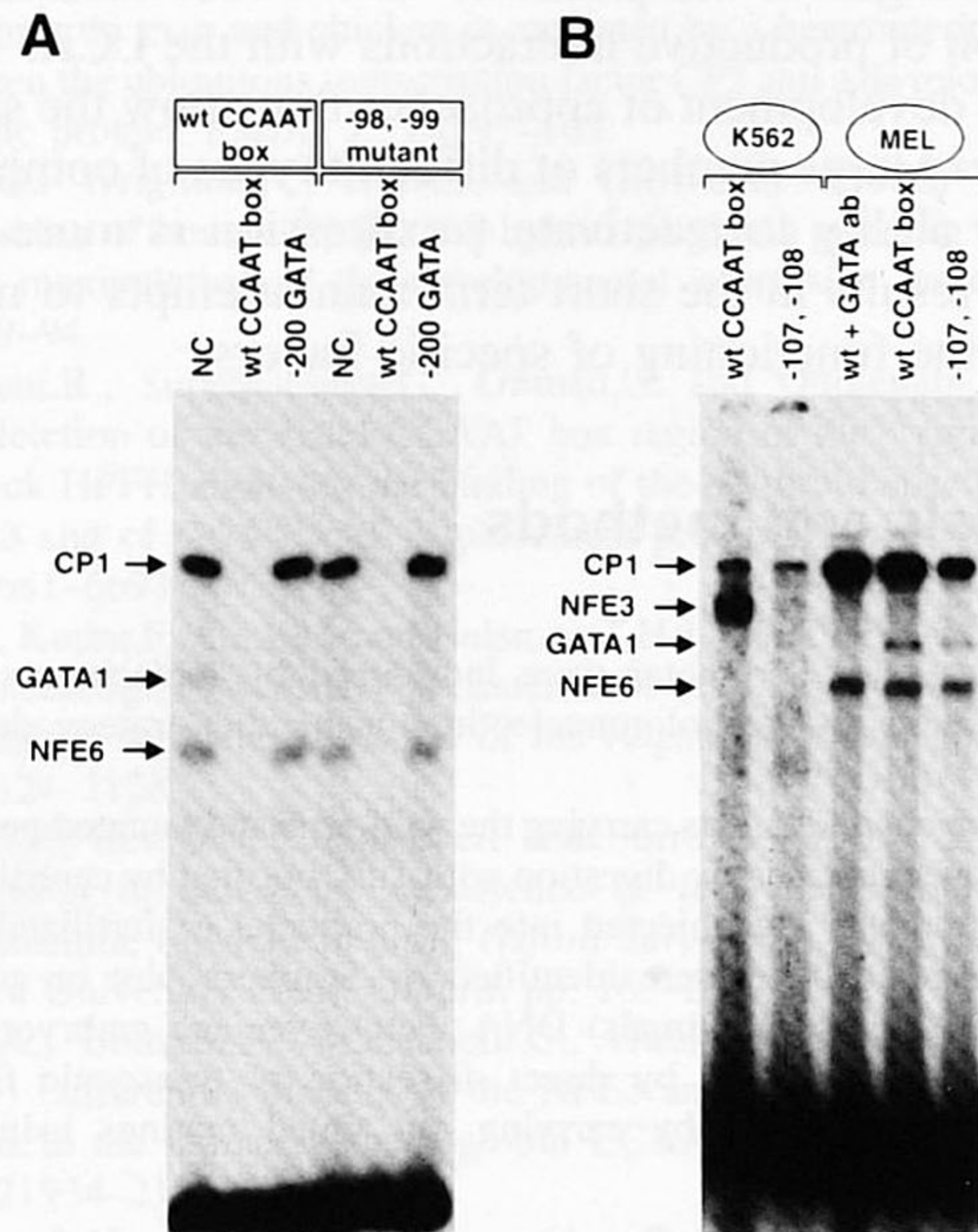


**Fig. 4.** S1 analysis of RNA from liver obtained from 13.5 day founder fetuses transgenic for wild-type and mutant constructs. The different mutant constructs are indicated above the lanes.



**Fig. 5.** Mean values for the percentage  $\gamma$  expression relative to total human  $\gamma + \beta$ . Error bars represent standard deviations. Because position effects result in occasional extreme values, trimmed means were calculated by omitting the lowest and highest five percentile (1.645 standard deviations, Fischer and Yates, 1963). The value for the -117 HPFH was obtained by averaging the values reported by Berry *et al.* (1992) after first correcting them using the value obtained for HPFH 61 in this analysis. In this study, we have quantitated specific activities of the probes by PhosphorImager analysis following gel electrophoresis. This more stringent quantitation has resulted in a lower value for the -117 HPFH mutation than that reported by Berry *et al.* (1992).



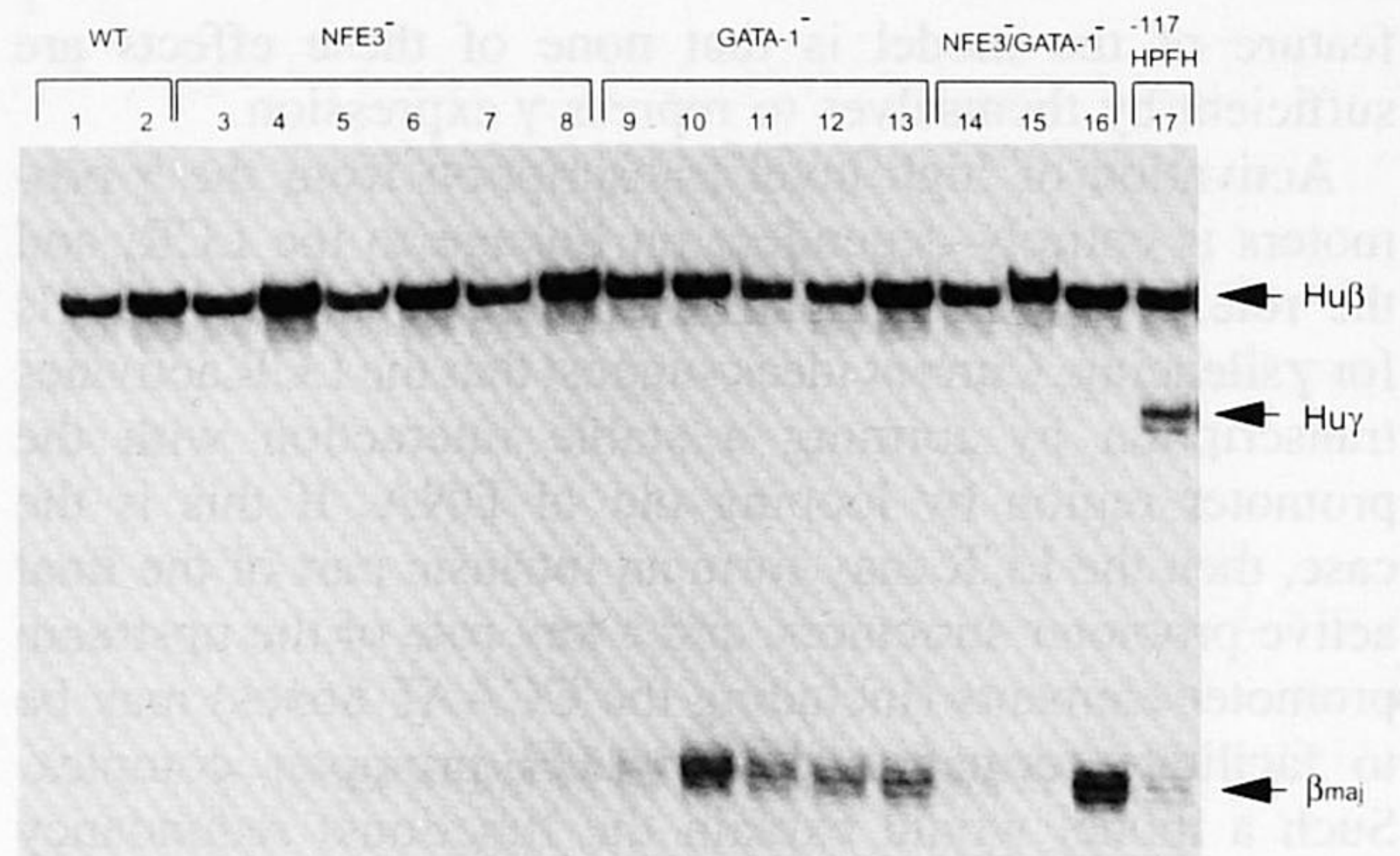


**Fig. 6.** Top: bases contacted by GATA-1 and NFE3. Depurination (Wall *et al.*, 1988) was used to identify the GATA-1 contact points. Filled circles indicate strong contact points; open circles correspond to weak protections (Whyatt *et al.*, 1993; M. Berry *et al.* unpublished). Data on NFE3 were obtained by DMS interference. Black rectangles indicate bases contacted by NFE3 (Ronchi *et al.*, 1995). Arrows indicate the bases that were mutated to abolish GATA-1 and NFE3 binding. **(A)** Bandshift competition analysis of wild-type (WT) CCAAT box and -98, -99 mutant oligonucleotides was performed using nuclear extract from the mouse adult erythrocyte (MEL) cell line C88. The WT oligo was used as a probe in the first three lanes and the -98, -99 mutant oligo as a probe in the last three lanes. Competitions were performed with a 50× molar excess of each of the competitors over probe. Competitor oligonucleotides or no competitors are indicated above the lanes. **(B)** Bandshift analysis of wild-type (WT) CCAAT box and mutant oligonucleotides was performed with a purified nuclear extract from the human erythrocyte cell line K562 (lanes 1 and 2) and a crude nuclear extract from mouse erythrocyte (MEL) cells. The probes are WT CCAAT box oligo, and the -107, -108 oligo; +GATA ab indicates addition of GATA antibody to the shift. The low amount of NFE3 present in MEL extracts is not visible without purification.

any increase in adult  $\gamma$  expression (Figure 7). A construct containing both mutations together, and therefore mimicking the -117 mutation in terms of factor binding *in vitro*, also failed to give a HPFH phenotype. Our data do not therefore support a direct role for either of these factors as suppressors of  $\gamma$  transcription through binding to the CCAAT box region.

## Discussion

There are a number of features that render the mechanism of  $\gamma$  silencing enigmatic. The fact that single point



**Fig. 7.** S1 analysis of adult blood from founder animals transgenic for the GATA-1<sup>-</sup> and NFE3<sup>-</sup> mutants described in Figure 1.

mutations can give rise to very high levels of  $\gamma$  expression in adult life would seem to indicate that all of the factors required for  $\gamma$  expression are present in the adult erythrocyte cell, making the binding of a suppressor factor or factors the most likely mechanism for silencing of the  $\gamma$  genes. However, while three different mutations in the CCAAT box region are associated with HPFH phenotypes, mutations in the upstream region of the promoter are also associated with high levels of  $\gamma$  expression in adults (Poncz *et al.*, 1989). Phylogenetic footprinting and functional studies have also suggested that a region around -50 may be important for  $\gamma$ -globin gene regulation (Jane *et al.*, 1995). This suggests that the sequences mediating suppression may be spread across a broad region of the promoter, extending at least from -202 to -50. In addition, a residual level of  $\gamma$  expression is observed in the blood of adult humans (Wood, 1993). Although this is generally <1% of  $\beta$  expression, the total level of globin expression is very high (90% of red cell protein) so that, in absolute terms, there is still a significant amount of  $\gamma$  transcription in the adult. Silencing of the  $\gamma$  genes is also a labile phenomenon which can be readily reversed by the presence of enhancers near the genes (Kollias *et al.*, 1986).

Our results provide further insights into the problem by showing that the HPFH promoter is a quite different entity from the promoter that mediates fetal  $\gamma$  expression. While we do not know which elements are responsible for early  $\gamma$  gene expression, the CCAAT boxes are relatively redundant. In contrast, inactivation of the proximal CCAAT box substantially reduces the level of  $\gamma$  expression observed with the Greek HPFH mutation at later stages. Any model for  $\gamma$  silencing must explain this dependence of the -117 HPFH phenotype on the proximal CCAAT box and must also account for the HPFH mutations between -156 and -202. The simplest explanation would be that the wild-type fetal promoter is a large multicomponent structure in which the different elements are partially redundant, while the -117 HPFH promoter is a much more restricted structure which depends strongly on a single element (the proximal CCAAT box) for its activity. This would imply that repression of the  $\gamma$  promoters occurs through a relatively mild interference with several different segments rather than through a powerful repressor binding to a single sequence element. The interference could be the result of an absence of a positively acting factor or the binding of a weak repressor, but the most important



feature of the model is that none of these effects are sufficient by themselves to repress  $\gamma$  expression.

Activation of high level transcription from the  $\gamma$  promoters is entirely dependent on linkage to the LCR, and the role of the LCR must be considered in any models for  $\gamma$  silencing. Current ideas suggest that the LCR activates transcription by forming a stable interaction with the promoter region by looping out of DNA. If this is the case, then the LCR may form an intrinsic part of the final active promoter structures, and a key role of the upstream promoter elements (including the CCAAT boxes) may be to facilitate formation of the LCR promoter complex. Such a model would explain the functional redundancy of promoter elements in the fetal stage, since the loss of one element would not on its own be sufficient to prevent interaction, although it would be expected to reduce the frequency of productive interaction and the overall level of expression. In the adult stage, mutations that restore the function of a single element such as the CCAAT box would be sufficient to allow interaction with the LCR and activation of high level transcription. Placing enhancers close to the  $\gamma$  genes would also tip the balance in favour of LCR interaction, perhaps by providing an extra target for the initial LCR contact. Overall, this model suggests that in the adult stage, rather than being actively suppressed, the  $\gamma$  genes are finely balanced between being able to interact productively with the LCR and being unable to form such interactions.

The mechanism by which the -117 HPFH mutation activates CCAAT box function remains obscure. On the basis of *in vitro* binding studies, GATA-1 and NFE3 have both been suggested as candidate suppressor factors that would act by binding to the region. However, we have found that mutations that specifically abolish binding of both factors to the CCAAT box region *in vitro*, fail to give a HPFH phenotype in transgenic mice. Although we cannot exclude the possibility that these factors work in conjunction with other factors to modulate CCAAT box function, for the present there is no evidence to suggest that they participate directly in  $\gamma$  silencing. CAAT displacement protein (CDP) has also been suggested as a factor which could affect CCAAT box function (Superti-Furga *et al.*, 1988), and our data showing the importance of the CCAAT box for the -117 HPFH phenotype is consistent with this hypothesis. However, once again, there is little direct evidence to suggest that this factor binds differentially to the CCAAT box region in adult erythroid cells, and further studies will be required to examine this possibility. Another interesting observation is the fact that the -117 and -114 HPFH mutations enhance the ability of the  $A\gamma$  gene to compete with  $\beta$  at 13.5 days, with the result that  $\beta$  expression is almost completely suppressed. This observation further emphasizes that  $\gamma$  silencing is the result of a gradual change in the total factor environment affecting expression of the  $\beta$ -globin locus.

Patients suffering from  $\beta$ -thalassaemia and sickle cell anaemia generally have intact and fully functional  $\gamma$  genes. Pharmacological reactivation of these genes in human adults represents an attractive approach to the treatment of  $\beta$ -thalassaemia and sickle cell anaemia. A level of  $\gamma$  expression similar to that observed in Greek HPFH would be sufficient to provide a very significant alleviation of the symptoms of these conditions. Possible strategies for

achieving this goal could include the identification of a suppressor factor and the design of compounds that would interfere with its action. Our results, together with the total body of data from the HPFH syndromes, suggest that this strategy is unlikely to be successful. It seems likely that  $\gamma$  silencing is mediated by an accumulation of weak effects spread across the promoter, with the CCAAT box region being one participant in this effect. The key to reactivation is likely to be the introduction of quite subtle changes in the promoter that would facilitate the formation of productive interactions with the LCR. Therefore, the development of approaches that allow the screening of very large numbers of different types of compounds for their ability to reactivate  $\gamma$  expression is more likely to yield results in the short term than attempts to modify directly the functioning of specific factors.

## Materials and methods

### Transgenic mice

Mutations in the  $\gamma$  promoter were introduced by replacing wild-type sequences with mutated oligonucleotides using the strategy described by Berry *et al.* (1992).

The 40 kb cosmid inserts carrying the wild-type and mutated promoters (Figure 1) were excised by digestion with *Sall*, purified by centrifugation on NaCl gradients and injected into the pronuclei of fertilized mouse eggs. Transgenic mice were identified by Southern blot on placental (fetuses) or tail (adult animals) DNA. Fetal liver and embryonic yolk sac were obtained either by direct dissection of transgenic founders from foster mothers or by carrying out timed matings using adult founder animals.

### Bandshifts

Bandshift analyses were as described previously (De Boer *et al.*, 1988). Each lane contained 5  $\mu$ g of nuclear extract and 0.5 ng of labelled double-stranded DNA probe. The -200 GATA oligo used for competitions is from the human  $\beta$ -globin gene promoter (-183 to -209).

### RNA and S1 protection analysis

RNA extracted from 10.5 day yolk sacs, fetal livers and adult blood was analysed by S1 nuclease protection assay exactly as described by Berry *et al.* (1992). Quantitation of signals obtained from S1 analysis was carried out using a PhosphorImager.

## Acknowledgements

We are grateful to the following organizations for supporting this work: Fondazione Italiana L.Giambrone per la Guarigione della Talassemia, Progetti Finalizzati Ingegneria Genetica and Biotecnologia e Biostrumentazione CNR, The Medical Research Council UK, The EEC Biotechnology Programme, and The Howard Hughes Medical Foundation. A.R. was the recipient of a short-term EMBO fellowship.

## References

- Behringer, R., Ryan, T.M., Palmiter, R.D., Brinster, R.L. and Townes, T. (1990) Human  $\gamma$  to  $\beta$ -globin switching in transgenic mice. *Genes Dev.*, **4**, 380-389.
- Berry, M., Dillon, N. and Grosveld, F. (1992) A single point mutation is the cause of the Greek form of hereditary persistence of fetal haemoglobin. *Nature*, **358**, 499-452.
- De Boer, E., Antoniou, M., Mignotte, V., Wall, L. and Grosveld, F. (1988) The human  $\beta$ -globin gene promoter; nuclear protein factors and erythroid-specific induction of transcription. *EMBO J.*, **7**, 4203-4212.
- Collins, F., Metherall, J., Yamakawa, M., Pan, J., Weismann, S.M. and Forget, B. (1985) A point mutation in the A gamma-globin gene promoter in Greek hereditary persistence of fetal haemoglobin. *Nature*, **313**, 325-326.
- Enver, T., Raich, N., Ebens, J., Papayannopoulos, T. and Stamatoyannopoulos, G. (1990) Developmental regulation of human



- fetal-to-adult globin gene switching in transgenic mice. *Nature*, **344**, 309–313.
- Fischer, R.A. and Yates, F. (1963) *Statistical Tables for Biological Agricultural and Medical Research*. Oliver and Boyd Ltd, Edinburgh.
- Gelinas, R., Endlich, B., Pfeiffer, C., Yagi, M. and Stamatoyannopoulos, G. (1985) G to A substitution in the distal CCAAT box of the A gamma-globin gene in Greek hereditary persistence of fetal haemoglobin. *Nature*, **313**, 323–325.
- Hanscombe, O., Whyatt, D., Fraser, P., Yannoutsos, N., Greaves, D., Dillon, N. and Grosveld, F. (1991) Importance of globin gene order for correct developmental expression. *Genes Dev.*, **5**, 1387–1394.
- Jane, S.M., Nienhuis, A.W. and Cunningham, J.M. (1995) Hemoglobin switching in man and chicken is mediated by a heteromeric complex between the ubiquitous transcription factor CP2 and a developmentally specific protein. *EMBO J.*, **14**, 97–105.
- Kollias, G., Wrighton, N., Hurst, J. and Grosveld, F. (1986) Regulated expression of human  $\gamma$ ,  $\beta$ -, and hybrid  $\gamma\beta$ -globin genes in transgenic mice; manipulation of the developmental expression patterns. *Cell*, **46**, 89–94.
- Mantovani, R., Superti-Furga, G., Gilman, G. and Ottolenghi, S. (1989) The deletion of the distal CCAAT box region of the  $\gamma$  globin gene in black HPFH abolishes the binding of the erythroid specific protein NF-E3 and of the CCAAT displacement protein. *Nucleic Acids Res.*, **17**, 6681–6691.
- Oner, R., Kutlar, F., Gu, L.H. and Huisman, T.H.J. (1991) The Georgia type of nondeletional hereditary persistence of fetal hemoglobin has a C–T mutation at nucleotide –114 of the A gamma-globin gene. *Blood*, **77**, 1124–1128.
- Poncz, M.P., Henthorn, P., Stoeckert and Surrey, S. (1989) Globin gene expression in hereditary persistence of fetal hemoglobin and  $\delta\beta$  thalassaemia. In McLean (ed.), *Oxford Surveys on Eukaryotic Genes*. Oxford University Press, Oxford, pp. 163–203.
- Ronchi, A., Bottardi, S., Mazuchelli, C., Ottolenghi, S. and Santoro, C. (1995) Differential binding of the NFE3 and CPI/NFY transcription factors to the human  $\gamma$  and  $\epsilon$ -globin CCAAT boxes. *J. Biol. Chem.*, **270**, 21934–21941.
- Schechter, A.N., Noguchi, C.T. and Rodgers, G.P. (1987) In Stamatoyannopoulos, G., Nienhuis, A.W., Leder, P. and Majerus, P.W. (eds), *The Molecular Basis of Blood Diseases*. W.B.Saunders, Philadelphia, PA, pp. 179–218.
- Stamatoyannopoulos, G. and Nienhuis, A.W. (1987) In Stamatoyannopoulos, G., Nienhuis, A.W., Leder, P. and Majerus, P.W. (eds), *The Molecular Basis of Blood Diseases*. W.B.Saunders, Philadelphia, PA, pp. 66–105.
- Stamatoyannopoulos, G. and Nienhuis, A.W. (1992) Therapeutic approaches to hemoglobin switching in treatment of hemoglobinopathies. *Annu. Rev. Med.*, **43**, 497–521.
- Superti-Furga, G., Barberis, G., Schaffner, G. and Busslinger, M. (1988) The –117 mutation in Greek HPFH affects the binding of three nuclear factors to the CCAAT region of the  $\gamma$ -globin gene. *EMBO J.*, **7**, 3099–3107.
- Wall, L., De Boer, E. and Grosveld, F. (1988) The human  $\beta$ -globin 3' enhancer contains multiple binding sites for an erythroid-specific protein. *Genes Dev.*, **2**, 1089–1100.
- Whyatt, D., DeBoer, E. and Grosveld, F. (1993) The two zinc finger-like domains of GATA-1 have different DNA binding specificities. *EMBO J.*, **12**, 4993–5005.
- Wood, W.G. (1993) Increased HbF in adult life. *Baillieres Clin. Haematol.*, **6**, 177–213.

Received on May 23, 1995; revised on September 12, 1995