The EMBO Journal vol.7 no.10 pp.3099-3107, 1988

The −117 mutation in Greek HPFH affects the binding of three nuclear factors to the CCAAT region of the γ-globin gene

Giulio Superti-Furga[1,2,3], Alcide Barberis[1], Gotthold Schaffner[1,2] and Meinrad Busslinger[1,2,3]

[1]Institut für Molekularbiologie II, Universität Zurich, Hönggerberg, CH-8093 Zürich, Switzerland, [2]Institut für Molekulare Pathologie, Dr Bohr-Gasse 7, A-1030 Wien, Austria and [3]Genentech Inc., Department of Molecular Biology, 460 Point San Bruno Boulevard, South San Francisco, CA 94080, USA

Communicated by M.Birnstiel

The Greek form of hereditary persistence of fetal hemoglobin (HPFH) is associated with a point mutation immediately upstream of the distal of the two CCAAT elements of the Aγ-globin gene. Three proteins present in nuclear extracts of erythroleukemia cells bind to this CCAAT region and contact the nucleotide mutated in Greek HPFH. The ubiquitous CCAAT-binding factor CP1 interacts preferentially with the proximal CCAAT sequence. An erythroid cell-specific factor, referred to as NF-E, binds with a higher affinity to the distal CCAAT region and interacts only with sequences flanking the CCAAT motif. The third protein is the vertebrate homologue of the sea urchin CCAAT displacement protein and recognizes sequences in both CCAAT elements and their flanking sequences. While the point mutation in Greek HPFH slightly strengthens the binding of CP1 and the CCAAT displacement protein, the same base change strongly reduces the binding of NF-E to the distal CCAAT region, suggesting a possible role of NF-E in the repression of γ-globin genes in adult erythroid cells.
Key words: Greek HPFH/Aγ-globin gene/erythroleukemia

Introduction

The human β-like globin genes are a model system for the study of gene regulation, as they are expressed in a developmental and tissue-specific manner. The ε-globin gene is active during the first weeks of gestation in erythroid cells of the embryonic yolk sac. Later the Aγ- and Gγ-globin genes are expressed in fetal liver cells and their expression is gradually reduced to a very low level after birth. Finally, the adult β- and δ-globin genes are expressed at a high level only after birth, when the site of erythropoiesis shifts to the bone marrow (reviewed by Collins and Weissman, 1984; Karlsson and Nienhuis, 1985).

Hereditary persistence of fetal hemoglobin (HPFH) is a clinically benign condition manifested by the continued expression of one or both γ-globin genes in the adult. Different molecular lesions are associated with this phenotype and can be categorized as either large deletions 3′ of the γ-globin genes or point mutations in the γ-globin gene promoters (reviewed by Stamatoyannopoulos and Nienhuis, 1987). The Greek type of nondeletion HPFH is characterized by continued expression of the Aγ-globin gene and is associated with a single G to A mutation 2 bp upstream of the distal of the two CCAAT elements in the Aγ-globin gene promoter (Collins *et al.*, 1985; Gelinas *et al.*, 1985; see Figure 1). Screening genomic DNA of many individuals by oligonucleotide hybridization demonstrated complete concordance of this point mutation with Greek HPFH, further suggesting that it is the cause of the Greek HPFH phenotype (Waber *et al.*, 1986). Two hypotheses that are not mutually exclusive have been put forward to explain the effect of this mutation. This base change either increases the binding of a positive transcription factor to the distal CCAAT region or, even more likely, it interferes with the interaction of a transcriptional repressor that is required to inactivate γ-globin genes in adult erythroid cells (Collins *et al.*, 1985; Gelinas *et al.*, 1985). To date little is known about negative regulation of cellular genes in higher eukaryotes. As suggested by the genetic evidence of Greek HPFH, the duplicated CCAAT region of the γ-globin gene may provide a good model system for the study of proteins involved in negative regulation of transcription.

The CCAAT sequence is an essential promoter element of many mRNA-coding genes (Benoist *et al.*, 1980; Efstratiadis *et al.*, 1980). Several distinct nuclear factors have been characterized that interact with different subsets of promoter elements containing a CCAAT sequence (Jones *et al.*, 1985; Graves *et al.*, 1986; Cohen *et al.*, 1986; Dorn *et al.*, 1987; Lichsteiner *et al.*, 1987; Raymondjean *et al.*, 1988; Chodosh *et al.*, 1988a). The CCAAT-binding transcription factor (CTF) from HeLa cells consists of a family of polypeptides of 52−66 kd, stimulates the transcription of several genes in a reconstituted *in vitro* system and appears to be identical with NF-I, a cellular co-factor of adenovirus replication (Jones *et al.*, 1985, 1987; Morgan *et al.*, 1987). The CCAAT-binding protein (CBP) from rat liver was originally identified as a heat stable factor binding to the CCAAT box of the herpes simplex virus (HSV) thymidine kinase (tk) gene and the Moloney murine sarcoma virus (MSV), but was also shown to interact with enhancer core motifs (Graves *et al.*, 1986; Johnson *et al.*, 1987). The factors NF-Y (Dorn *et al.*, 1987) and CP1 (Chodosh *et al.*, 1988a), which are most likely identical proteins, interact with yet another set of CCAAT sequences as exemplified by the CCAAT motif of the MHC Eα gene and the adenovirus major late promoter respectively. The proteins CP1 and CP2 have recently been shown to consist of two heterologous subunits (Chodosh *et al.*, 1988a) that, in the case of CP1, are functionally homologous to the yeast activators HAP2 and HAP3 (Olesen *et al.*, 1987; Chodosh *et al.*, 1988b). While all these CCAAT-binding factors are thought to be transcriptional activators, we have recently identified a potential repressor in extracts of sea urchin embryos. This CCAAT displacement protein binds with high affinity to sequences overlapping the CCAAT element of the sperm H2B gene and thus prevents the interaction of the CCAAT-

binding factor in extracts of sea urchin embryos, where the sperm H2B gene is not expressed (Barberis *et al.*, 1987).

Here we show that three nuclear factors bind to overlapping sequences in the duplicated CCAAT region of the human γ-globin gene. All three proteins contact the G residue mutated in Greek HPFH, but their DNA interaction is differently affected by this mutation. The binding of the CCAAT-binding factor CP1 and of the vertebrate CCAAT displacement protein is slightly increased, while the same point mutation strongly reduces the binding of a novel erythroid-cell-specific factor suggesting it to be the postulated repressor of γ-globin genes into adult erythroid cells.

Results

Three nuclear proteins interact with the CCAAT region of the γ-globin gene

The structure of the γ-globin gene promoter is schematically shown in Figure 1. It contains at least four elements that have previously been shown to be important regulatory sequences of other genes (see Maniatis *et al.*, 1987). Particularly interesting is the duplicated CCAAT region where a single point mutation was identified 2 bp upstream of the distal CCAAT sequence in individuals with Greek HPFH (Collins *et al.*, 1985; Gelinas *et al.*, 1985). This mutation occurs in a TGA trinucleotide within a sequence that is repeated three times at intervals of 13 and 14 bp in the CCAAT region (Figure 1). We have previously noted that TGA motifs with a similar spacing constitute part of the interaction site for the sea urchin CCAAT displacement protein in the sperm H2B promoter (Barberis *et al.*, 1987). To test whether this protein also interacts with the γ-globin promoter, we have compared the binding of nuclear proteins to the CCAAT region of the sperm H2B and the γ-globin genes by the electrophoretic mobility shift assay (Fried and Crothers, 1981; Garner and Rezvin, 1981). An end-labeled *Pvu*II−*Fok*I DNA fragment of the sperm H2B promoter (Figure 2A, lane 1) and a *Hpa*II−*Bal*I DNA fragment of the γ-globin promoter (lane 2) were separately incubated in a nuclear extract from sea urchin blastula embryos prior to analysis of the protein−DNA complexes on a low ionic strength polyacrylamide gel. Both promoter sequences gave rise to the same two complexes corresponding to the CCAAT displacement protein (CDP) and to the CCAAT-binding factor (CBF). Furthermore, an excess of a CDP-specific oligonucleotide derived from the sperm H2B promoter (lane 4, see also Barberis *et al.*, 1987), but not of CCAAT sequence (lane 3), specifically competed for the binding of the CCAAT displacement protein to the γ-globin promoter. This demonstrates that the CCAAT region of the γ-globin gene contains a recognition site for the sea urchin CCAAT displacement protein.

We next analyzed different mammalian cell extracts for nuclear proteins binding to the γ-globin *Hpa*II−*Bal*I DNA fragment (Figure 2B). For this experiment we used the human B cell line BJA-B (lane 6) and HeLa cells (lane 7), which do not express the endogenous γ-globin genes, the human embryonic/fetal erythroid cell line K562 (lane 9), which transcribes the endogenous γ-globin genes and mouse erythroleukemia (MEL) cells (lane 8), which have been extensively used in the past as a model system to study adult globin gene regulation. Three different protein−DNA complexes could be identified in these extracts. The protein in complex B1 exhibits an identical electrophoretic mobility

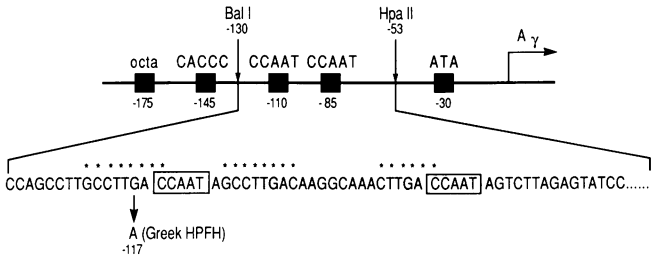

Fig. 1. Promoter structure of the Aγ-globin gene. Nucleotides −69 to −130 relative to the transcription initiation site of the Aγ-globin gene are shown (Slightom *et al.*, 1980). The point mutation present at position −117 in Greek HPFH is indicated by an arrow and the repeated sequence by asterisks (Collins *et al.*, 1985; Gelinas *et al.*, 1985). Octa refers to the octamer sequence ATTTGCAT.

to the sea urchin CCAAT displacement protein (lane 5) and therefore appears to be the vertebrate homologue of the sea urchin protein. We refer to this protein as the vertebrate CCAAT displacement protein (CDP) only in analogy to the sea urchin protein, since we have no evidence so far that the vertebrate protein is able to displace *in vitro* the CCAAT-binding factor from the γ-globin gene promoter. This nuclear factor is ubiquitous, as it is not only present in all four mammalian extracts shown, but also in many other cell types analyzed (data not shown). The protein in complex B2 is also found in all extracts tested and corresponds to a CCAAT-binding factor (see below). By contrast, the protein in complex B3 was so far detected only in nuclear extracts of the erythroid MEL and K562 cells and is therefore referred to as NF-E. This erythroid-specific factor is clearly more abundant in the MEL extract than in the K562 extract, while both extracts contain similar amounts of proteins B1 and B2.

The binding specificity of these three proteins for other globin promoters was studied in the MEL extract by the mobility shift experiment of Figure 2C. To this end, part of the β-, ε- and α1-globin promoters (lanes 11−13) were synthesized as oligonucleotides of identical length, where the single CCAAT box of these globin genes was aligned with the distal CCAAT element of the γ-globin promoter. CDP (B1) binds exclusively to the γ-globin promoter (lane 10), while nuclear factors with a mobility similar to protein B2 interact with all four globin promoters, although with different affinities (lanes 11−13). NF-E (B3) binds strongly to the CCAAT region of both the γ- and β-globin genes (lanes 10 and 11). The same or a similar protein interacts also with the ε- and α1-globin promoter generating a complex of slightly faster mobility (lanes 12 and 13) that can, however, be selectively competed with a B3-specific oligonucleotide (data not shown).

We next addressed the question of whether the three proteins recognize different or identical features in the duplicated CCAAT region of the γ-globin gene. For this purpose we have used cloned oligonucleotides containing either just the distal or the proximal CCAAT sequences in the mobility shift analysis of Figure 1D. CDP (B1) binds only to the *Hpa*II−*Bal*I DNA fragment containing the entire duplicated CCAAT region (lane 14), but not to the distal (lane 15) and the proximal (lane 16) CCAAT oligonucleotide or to the CCAAT sequence of the human α1-globin gene (lane 17). This suggests that CDP interacts with an extended region of the γ-globin gene promoter. By contrast, the two proteins B2 and NF-E (B3) are able to bind

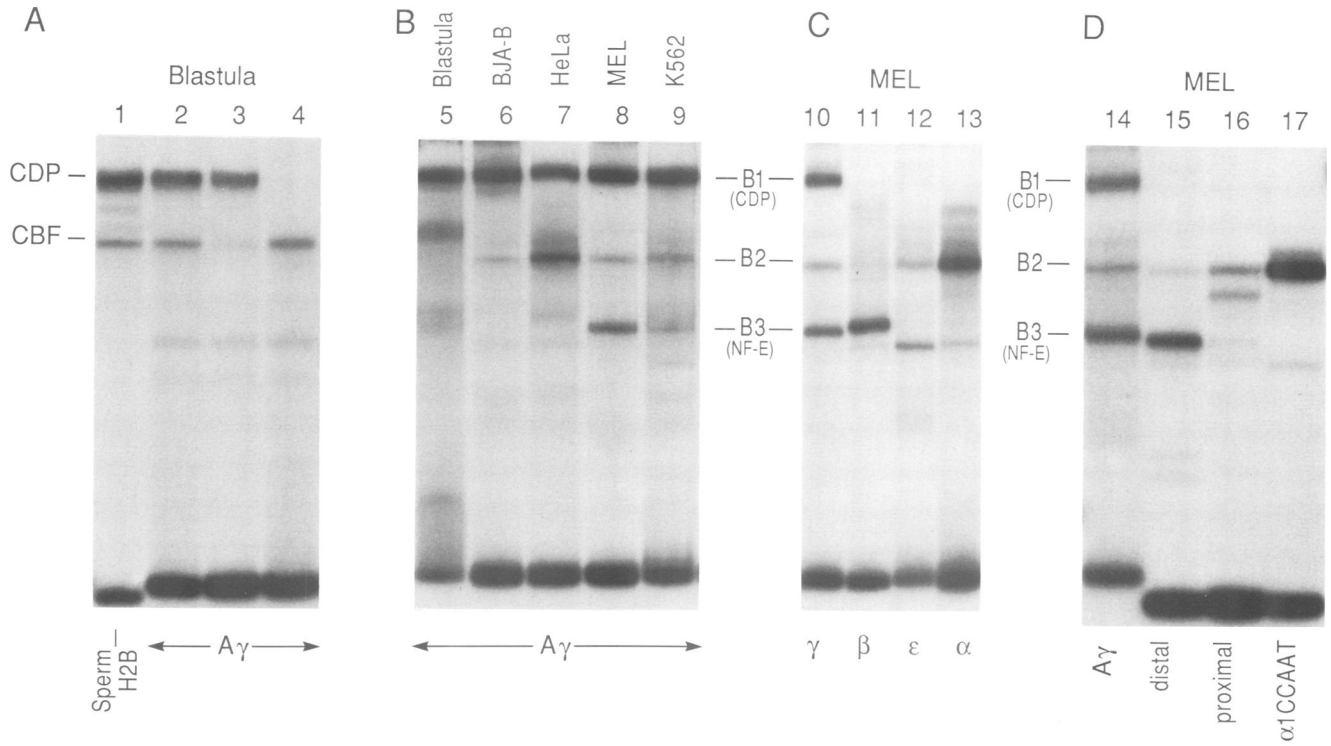

Fig. 2. Electrophoretic mobility shift analysis of nuclear proteins binding to the γ-globin CCAAT region. (**A**) Comparison of nuclear proteins binding to the sperm H2B and γ-globin gene promoters. A 5′ end-labeled *Pvu*II−*Fok*I DNA fragment (**lane 1**) of the sperm H2B-1 gene (Barberis *et al.*, 1987) and a *Hpa*II−*Bal*I DNA fragment (**lane 2**) of the Aγ-globin promoter (Figure 1) were incubated with a nuclear extract of sea urchin blastula embryos followed by electrophoresis through a native 4% polyacrylamide gel, as described in Materials and methods. A double-stranded CCAAT oligonucleotide and a CDP-specific oligonucleotide derived from the promoter of the sperm H2B-1 gene (for sequence see Barberis *et al.*, 1987) were used in a 1000- and 100-fold molar excess for competition in **lanes 3** and **4** respectively. (**B**) Nuclear proteins binding to the γ-globin CCAAT region in vertebrate cell extracts. Nuclear extracts of the human B-cell line BJA-B, of HeLa cells, of the human erythroleukemia cell line K562 and of murine erythroleukemia (MEL) cells were analyzed for proteins binding to the end-labeled *Hpa*II−*Bal*I DNA fragment of the γ-globin promoter. (**C**) Specificity of protein binding for different globin gene promoters. The β-, ε and α-globin oligonucleotides contained the sequences from −15 to −91 upstream of the β-globin gene (**lane 11**), from −24 to −100 of the ε-globin gene (**lane 12**) and from −10 to −86 of the α1-globin gene (**lane 13**) respectively. The γ-globin *Hpa*II−*Bal*I DNA fragment was used as a reference (**lane 10**). (**D**) Sequence requirement for protein binding. Double-stranded oligonucleotides derived from the distal (**lane 15**) and proximal (**lane 16**) CCAAT regions of the γ-globin gene and from the CCAAT region of the α1-globin gene (**lane 17**) were cloned into the *Hin*dIII site of plasmid pSP64 (for oligonucleotide sequences see Table I). These cloned oligonucleotides were released from the polylinker by *Ava*I and *Hin*dIII digestion prior to end labeling and were used together with the γ-globin *Hpa*II−*Bal*I DNA fragment (**lane 14**) for mobility shift analysis in MEL extracts.

to the proximal and distal CCAAT oligonucleotides, although with opposite affinities. Protein B2 interacts preferentially with the proximal CCAAT element (lane 16), while NF-E (B3) has a markedly higher affinity for the distal of the two CCAAT regions (lane 15). The additional complex migrating below protein B2 in lane 16 is an artefact band resulting from protein binding to the polylinker sequences present in the cloned proximal CCAAT fragment (data not shown).

All three proteins recognize different, but overlapping DNA sequences

The contact sites of the three proteins on the γ-globin promoter were determined by the method of methylation interference (Siebenlist *et al.*, 1980). The γ-globin *Hpa*II−*Bal*I restriction fragment was end labeled either in the coding or the noncoding DNA strand, partially methylated at G residues by dimethyl sulfate and then used in mobility shift experiments with MEL extracts (see Materials and methods). Free and bound DNA was eluted from the gel, cleaved by piperidine and analyzed on a denaturing sequencing gel. Methylated G residues that interfere with protein binding are under-represented in the G-ladder of the bound DNA. As shown in Figure 3, CDP (B1) makes multiple contacts along a 48-bp-long DNA sequence of the duplicated CCAAT region. Among these contact sites are all three TGA motifs of the repeat sequence (Figure 1) as well as the two G residues in the complementary strand of the two CCAAT elements. The sea urchin CCAAT displacement protein generates a nearly identical interference pattern on the γ-globin promoter (data not shown) which further identifies protein B1 as the vertebrate homologue of the sea urchin protein. The interference pattern of protein B2 centers over the proximal CCAAT sequence and is typical for a CCAAT-binding factor. The contact sites of NF-E (B3) flank the distal CCAAT sequence on both sides and include the central and distal of the three TGA motifs. Interestingly, most of these contact sites are clustered in the downstream sequence GACAAGG (−104 to −98) which may therefore constitute the core recognition sequence of NF-E. This interference analysis confirms the results of the mobility shift experiment of Figure 2D in that CDP (B1) interacts with an extended sequence of the duplicated CCAAT region, while the CCAAT-binding factor (B2) and NF-E (B3) preferentially recognize sequences in the proximal and distal CCAAT regions respectively. It is interesting to note that CDP and NF-E both contact the G residue −117 which has been mutated in Greek HPFH.

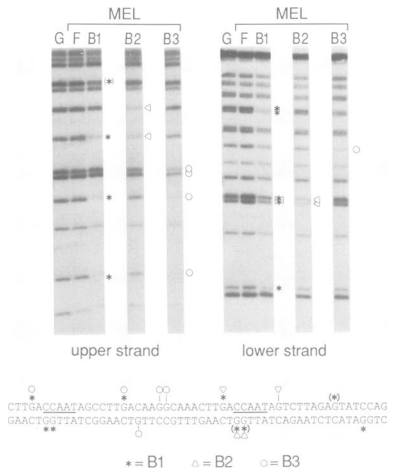

Fig. 3. Methylation interference analysis of proteins binding to the γ-globin CCAAT region. The *Hpa*II–*Bal*I DNA fragment of the γ-globin gene was cloned into the *Hin*dII site of plasmid pSP64, 5′ end labeled in the coding or nocoding strand at the *Bam*HI or *Hin*dIII site and then released from the polylinker by digestion with *Hin*dIII or *Bam*HI respectively. These DNA probes were used for methylation interference analysis of complex B1, B2, B3 and of unbound DNA (F) in MEL extracts as described in Materials and methods. Methylated G residues that interfere with protein binding are indicated to the left of the autoradiograph as well as above and below the relevant sequence of the γ-globin gene promoter (Slightom *et al.*, 1980). Brackets indicate sites of weak methylation interference. All lanes are from the same autoradiographic exposure.

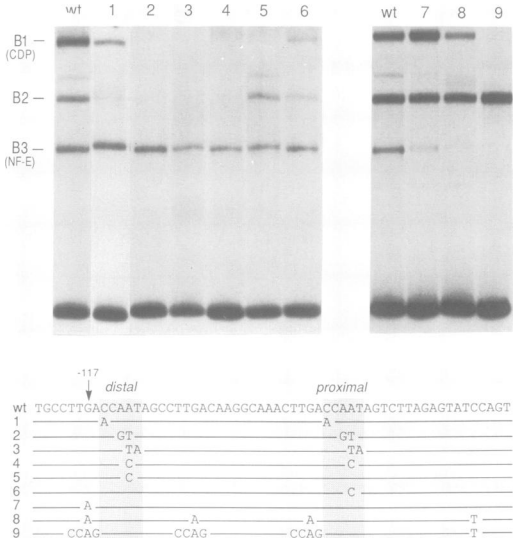

Fig. 4. Sequence requirement of proteins binding to the γ-globin CCAAT region. Oligonucleotides containing the point mutations indicated in the lower part of the figure were synthesized from positions −53 (*Hpa*II) to −130 (*Bal*I) of the γ-globin gene promoter. Complementary strands were annealed and used for electrophoretic mobility shift analysis in MEL extracts.

The methylation interference analysis clearly demonstrated that the three proteins interact with overlapping sequences of the γ-globin CCAAT region. We were therefore interested to see whether mutagenesis of this promoter region would allow us to discriminate between the binding requirements of these three proteins. In a first series of mutations we introduced base changes in the CCAAT sequences and tested the mutant oligonucleotides for protein binding in the mobility shift assay (Figure 4). A C to A transversion in the first position of both CCAAT sequences strongly reduces the interaction of the CCAAT-binding protein (B2) and moderately affects binding of CDP (B1; lane 1). In contrast, single and double mutantions in the 3′ half of both CCAAT sequences not only interfere with the interactions of the CCAAT-binding factor (B2), but completely eliminate CDP binding (lanes 2–4). Even a single point mutation in just the distal CCAAT box (lane 5) completely abrogates CDP binding, while a point mutation in the proximal CCAAT box still allows the interaction of CDP, although with a markedly reduced efficiency. We conclude therefore that the distal CCAAT sequence is more important than the proximal CCAAT element for CDP binding. In agreement with this, methylation of the two G residues in the complementary strand of the distal, but not of the proximal CCAAT box interferes strongly with the interaction of CDP (Figure 3). Interestingly, NF-E (B3) is little affected by these CCAAT mutations indicating that the mutated positions in the distal CCAAT box do not critically contribute to NF-E binding.

In a second series of mutations we changed the sequence context around the CCAAT motifs. The Greek HPFH mutation (G to A) at position −117 has opposite effects on the proteins CDP (B1) and NF-E (B3) both of which contact this nucleotide position (lane 7). As shown by densitometric scanning of the autoradiographic signals, binding of NF-E (B3) was reproducibly decreased by a factor of 8–10, while the affinity of CDP (B1) was 2-fold increased for the mutated sequence. The contrasting effects of the −117 mutation clearly do not result from competition of NF-E and CDP for DNA binding, as a mutant oligonucleotide of just the distal CCAAT region that cannot be recognized by CDP reveals an identical decrease of NF-E binding (see below). Four contact sites of the protein CDP (including the three TGA motifs) and two sites of NF-E have been mutated in oligonucleotide 8. These point mutations reduce the binding of CDP 3-fold and the interaction of NF-E 20-fold with respect to the wild-type sequence. Both these proteins are unable to bind to oligonucleotide 9, where the tetranucleotide CCAG normally flanking the CCAAT element of the α1-globin gene replaces the three TTGA motifs of the γ-globin promoter. We conclude from this that two intact CCAAT boxes alone are not sufficient for the interaction of CDP and that their sequence context also critically contributes to CDP binding. NF-E is strongly influenced by all the mutations in the flanking sequences, as is expected if it does not recognize the CCAAT motif *per se*. By contrast, the DNA interaction of the CCAAT-binding protein (B2) is not affected by the point mutations in oligonucleotides 7 and 8 and is even 2-fold increased by the more extensive mutation in oligonucleotide 9.

Identification of the CCAAT-binding protein B2 as CP1

Different types of CCAAT-binding proteins are known to date, which can be distinguished from each other by their specific interaction with different subsets of eukaryotic CCAAT sequences (see Introduction). In Figure 5 we have compared the ability of various unlabeled oligonucleotides to compete for complex formation with the labeled proximal CCAAT sequence of the γ-globin gene. Protein B2 clearly differs from the CCAAT-binding transcription factor CTF/NF-I, as the high-affinity NF-I-binding site of the

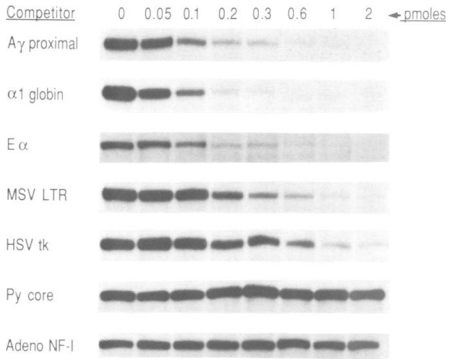

Fig. 5. Identification of the CCAAT-binding factor as protein CP1. The end-labeled oligonucleotide containing the proximal CCAAT region of the γ-globin gene was used for mobility shift analysis with MEL extracts in the presence of increasing amounts of competitor DNA. The following oligonucleotides were used as competitors: the homologous CCAAT sequence of the γ-globin gene, the CCAAT elements of the α1-globin gene, of the class II major histocompatibility Eα gene, of MSV LTR, of the HSV tk gene and the polyoma enhancer core motif (Py core) as well as the NF-I binding site of the adenovirus replication origin. The sequences of all these oligonucleotides are shown in Table I. Only the relevant part of the autoradiograph corresponding to the DNA complex of the CCAAT-binding factor B2 is shown for each competition experiment.

adenovirus replication origin (Jones *et al.*, 1987) does not compete at all for B2 binding (Figure 5). In support of this, purified CTF/NF-I (a kind gift of R.Tjian) does not interact with the γ-globin CCAAT region (data not shown). The CCAAT-binding protein CBP/EBP is a heat-stabile protein and interacts with CCAAT sequences as well as with viral enhancer core sequences (Graves *et al.*, 1985; Johnson *et al.*, 1987). Our B2-binding activity also differs from this protein, as it is not heat stable (data not shown) and does not compete with the polyoma enhancer core motif (Figure 5). The CCAAT sequences of the human α1-globin gene and the class II major histocompatibility Eα genes compete, however, as effectively as the homologous γ-globin CCAAT site, while the coresponding elements of the MSV LTR and the HSV tk gene compete less efficiently for protein binding. This competition pattern is characteristic of the CCAAT-binding factors CP1 (Chodosh *et al.*, 1988a) and NF-Y (Dorn *et al.*, 1987) which are almost certainly identical proteins. We conclude therefore, that protein B2 corresponds most likely to the CCAAT-binding factor CP1/NF-Y.

Simultaneous interaction of CP1 and NF-E with the γ-globin CCAAT region

Our DNA-binding studies suggest the possibility that the proteins CP1 and NF-E may interact simultaneously with the duplicated γ-globin CCAAT region, as these proteins have opposite affinities for the proximal and distal CCAAT region. Intermediate bands between complex B1 and B2 are indeed observed in the mobility shift experiments of Figures 2 and 4. We have identified the nature of these intermediate bands (referred to as Y and Z) by the competition experiment of Figure 6, where the protein to DNA ratio was increased in favor of these complexes. Competition with an excess of the proximal CCAAT oligonucleotide prevents the generation of the three complexes Y, Z and B2, while an excess of a B3-specific oligonucleotide (see legend to Figure 7) interferes only with complex Z and B3. We interpret this to mean that complex Y consists of two CP1 molecules,

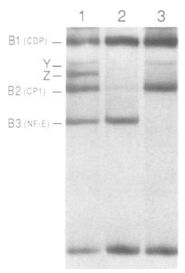

Fig. 6. Simultaneous binding of NF-E and CP1 to the γ-globin CCAAT region. The end-labeled *Hpa*II−*Bal*I fragment of the γ-globin gene promoter was incubated in the MEL extract in the absence of specific competitor DNA (**lane 1**) or in the presence of a 1000-fold excess of the proximal CCAAT oligonucleotide (**lane 2**) or of a B3-specific competitor DNA (**lane 3**). As B3-specific competitor we have used the oligonucleotide C of Figure 7, where the entire CCAAT sequence was replaced by the unrelated pentanucleotide AATCA to prevent the DNA interaction of CP1 and CDP.

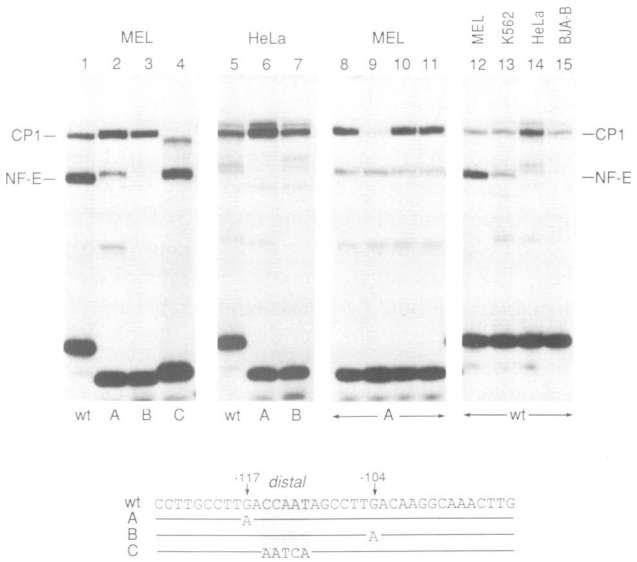

Fig. 7. Mutational analysis of the distal CCAAT region of the γ-globin gene. The mutations introduced into the distal CCAAT region of the γ-globin gene are indicated in the bottom part of this figure. The wild-type (wt) oligonucleotide contains the promoter sequence from the *Hpa*II site (−53) to the *Bal*I site (−130) except for a deletion of the proximal CCAAT region from position −65 to −91. The type of nuclear extracts and of end-labeled oligonucleotides used in this mobility shift analysis is shown above and below the autoradiograph, respectively. In **lanes 9−11** an 300-fold excess (0.3 pmol) of the following competitor DNAs was included in the binding reaction; CCAAT oligonucleotide of the α1-globin gene (**lane 9**), the NF-I binding site of the adenovirus replication origin (**lane 10**) and the enhancer core sequence of the polyoma virus (**lane 11**).

while complex Z results from the simultaneous binding of protein CP1 and NF-E to the proximal and distal CCAAT region respectively. This interpretation of complex Z is supported by the mutational analysis of Figure 4. The formation of this complex is prevented by single point mutations that either eliminate the binding site of CP1 in the proximal CCAAT sequence (Figure 4, lane 6) or that interfere with the binding of NF-E to the distal CCAAT region (Figure 4, lane 9). The interaction of CDP (B1) with the γ-globin CCAAT region is not affected in the competition analysis of Figure 6. We conclude therefore that CDP (B1) can neither be a homodimer nor a heterodimer of protein CP1 and/or NF-E.

Opposite effect of the −117 mutation on the binding of CP1 and NF-E to the distal CCAAT region

The immediate flanking sequences of the distal and proximal CCAAT elements are identical suggesting that the interaction of CP1 with the distal CCAAT region may be similar to that observed with the proximal binding site (Figure 3). This predicts that CP1 contacts the G residue at position −117 in the distal CCAAT region. The Greek HPFH mutation at this position, however, did not affect the overall interaction of CP1 with the duplicated CCAAT region in the experiment of Figure 4B, probably because the preferential interaction of CP1 with the proximal CCAAT element may have obscured any effect on CP1 binding to the distal low affinity site. To test this possibility and to further investigate the sequence requirement for NF-E binding, we have used oligonucleotides containing just the distal CCAAT region for the mutational analysis of Figure 7. The natural G to A transition at position −117 reproducibly increased the binding of CP1 to the distal CCAAT sequence by a factor of 3 in both the MEL and HeLa extracts (lanes 2 and 6). This increase is the direct consequence of an improved affinity of CP1 for the mutated sequence rather than an indirect effect of the reduced binding of NF-E in the MEL extract, as NF-E is absent in the HeLa extract. We conclude therefore that CP1 interacts indeed with the nucleotide that is mutated in Greek HPFH. An identical base change at position −104 resulted in a slight increase of CP1 binding in the MEL extract (lane 3), but had no effect in the HeLa extract (lane 7). Conceivably the −117 point mutation could have created a binding site for a CCAAT-binding factor other than CP1. This is ruled out, however, by the fact that only the high-affinity CP1-binding site of the α1-globin gene (lane 9), but neither the adenovirus NF I binding site (lane 10) nor the polyoma enhancer core sequence (lane 11), competes for protein binding.

The methylation interference and mutational analyses of Figures 3 and 4 have indicated that NF-E does not recognize the CCAAT sequence *per se*, but interacts with flanking sequences in the distal CCAAT region. The entire CCAAT sequence can indeed be replaced by the unrelated pentanucleotide AATCA without affecting the binding of NF-E (Figure 7, lane 4). By contrast, the natural mutation −117 just 2 bp upstream of the distal CCAAT box strongly reduces NF-E binding confirming our previous result (lane 2). The most striking effect is observed, however, with a single point mutation in the downstream interaction site of NF-E at position −104 which completely prevents NF-E binding (lane 3). This further supports the idea that the downstream interaction site constitutes the core recognition sequence of NF-E.

In the last part of Figure 7 we have reinvestigated the cell-type distribution of NF-E using an oligonucleotide of just the distal CCAAT region. NF-E is only abundant in extracts of the two erythroid cell lines MEL (lane 12) and K562 (lane 13) and could not be detected, even on overexposed autoradiographs, in the nonerythroid HeLa (lane 14) and BJA-B (lane 15) cells. The NF-E binding activity is less abundant in the embryonic/fetal erythroid cell line K562 than in MEL cells, while the same amount of CP1 is present in both extracts. The abundance of NF-E appears therefore to inversely correlate with the ability of these erythroid cells to express γ-globin genes in cell transfection experiments (Wright *et al.*, 1984; Kioussis *et al.*, 1985; Anagnou *et al.*, 1986).

Discussion

A great variety of molecular lesions associated with abnormal expression of the γ-globin genes indicates that multiple sequence elements both 5′ and 3′ of these genes are involved in their developmental regulation. One of the key elements controlling this process appears to be the duplicated CCAAT region, since there is a strong correlation between a single point mutation 2 bp upstream of the distal CCAAT sequence of the Aγ-globin gene and the Greek type of nondeletion HPFH (Collins *et al.*, 1985; Gelinas *et al.*, 1985; Waber *et al.*, 1986). This base change could cause the abnormal expression of the Aγ-globin gene in adult erythroid cells either by increasing the binding of a positive transcription factor or be weakening the interaction of a repressor (Collins *et al.*, 1985; Gelinas *et al.*, 1985). In search of such proteins we have identified three nuclear factors that interact with the CCAAT region of the γ-globin gene and that contact the G residue mutated in Greek HPFH.

The CCAAT-binding protein CP1

One of these proteins was identified as the CCAAT-binding factor CP1 by virtue of its contact sites in the proximal CCAAT element of the γ-globin gene and by competition experiments with CCAAT sequences of different genes (Chodosh *et al.*, 1988a). While CP1 has a higher affinity for the proximal CCAAT element, it can also bind the distal element. In fact, the immediate flanking sequences of the two CCAAT boxes are identical. We assume therefore, that CP1 contacts the same residues in both these sequences and that differences in its binding affinity are determined by the larger sequence context. The G residue 2 bp upstream of the proximal CCAAT element that interacts with CP1 in methylation interference experiments corresponds to the site of the Greek HPFH mutation in the distal CCAAT region. This could suggest an involvement of CP1 in the abnormal expression of the HPFH allele. Indeed, we have shown that this point mutation increases the binding of CP1 to the distal CCAAT box 2-fold in MEL and HeLa nuclear extracts. In agreement with this, the mutated sequence exhibits a stronger similarity to high-affinity CP1-binding sites (Chodosh *et al.*, 1988a) than the wild-type sequence. The CCAAT-binding protein CP1 is most likely a positive transcription factor of the γ-globin gene in analogy to its function in other genes where deletion of its interaction site from the promoter reduces the transcriptional efficiency (Dorn *et al.*, 1987; Raymondjean *et al.*, 1988; Chodosh *et al.*, 1988a). Transient expression studies in nonerythroid 293 (Rixon and Gelinas, 1988) and HeLa cells (our unpublished data) indicate that the Greek HPFH mutation results in a 1.4- to 2-fold increase of γ-globin gene transcription. This moderate transcriptional effect is most probably caused by the increased binding of CP1 to the mutated CCAAT region. Analogous point mutations 2 bp upstream of the CCAAT sequences of the mouse β-major globin gene (Myers *et al.*, 1986) and the HSV tk gene (Graves *et al.*, 1986) have revealed a similar increase in transcription which may also be a consequence of improved binding of a CCAAT-binding factor to the mutated sequence (Chodosh *et al.*, 1988a).

The CCAAT displacement protein

A second nuclear protein interacts with the entire duplicated CCAAT region of the γ-globin gene. We have identified

this factor as the vertebrate homolog of the sea urchin CCAAT displacement protein (Barberis *et al.*, 1987), since the vertebrate and the sea urchin factor show an identical mobility in gel retardation experiments, interact with the same three TGA motifs in the γ-globin CCAAT region and can be equally well competed by a promoter fragment of the sea urchin sperm histone H2B gene. We refer to the vertebrate protein also as CCAAT displacement protein (CDP) in analogy to the sea urchin protein, although we have no evidence so far that the vertebrate protein is able to displace the CCAAT-binding factor from the γ-globin gene promoter. CDP binds to the CCAAT region of the γ-globin gene, but not to the corresponding promoter fragments of the ε-, β- and α1-globin genes. The most striking difference to these promoter sequences is the duplication of the CCAAT region in the γ-globin genes. This duplication has created a binding site for CDP, which recognizes sequences in the two CCAAT elements and in their flanking sequences. The CCAAT sequences are, however, more important, as a single point mutation in one of the two CCAAT elements severely interfered with protein binding, whereas extensive mutations had to be introduced in the flanking sequences to abolish protein binding. In fact, the Greek HPFH mutation upstream of the distal CCAAT element even resulted in a 2-fold increase in CDP binding. The requirement for two intact CCAAT elements could be interpreted to mean that CDP is a complex of two CCAAT-binding proteins. However, this can be ruled out for several reasons. The CCAAT-binding protein CP1 and the CCAAT displacement protein behave differently in the mutational analysis of Figure 4. Moreover, the complex containing two CP1 molecules (band Y in Figure 6) migrates faster in mobility shift experiments than that of CDP. Both proteins can be biochemically separated from each other by gel filtration chromatography (unpublished data) and, furthermore, react differently with respect to various competitor DNAs. Most notably, long 'unspecific' DNA is efficiently bound by CDP (data not shown) suggesting that this protein has a high affinity for DNA in general besides its highly sequence-specific interaction with the γ-globin CCAAT region. It remains therefore to be seen, whether this protein also binds to the γ-globin promoter *in vivo* in the context of the entire genome. If so, it could prevent the access of positive transcription factors to the duplicated γ-globin CCAAT region in analogy to the postulated function of the sea urchin CCAAT displacement protein (Barberis *et al.*, 1987). The ubiquitous cell-type distribution and the increased binding to the Greek HPFH mutation argue, however, against this protein being a specific repressor of γ-globin genes in adult erythroid cells. Ultimately, functional studies will be required to elucidate the role of this protein in γ-globin gene regulation.

The erythroid cell-specific protein NF-E

The third nuclear protein is an erythroid-cell-specific factor that we refer to as NF-E. This protein binds preferentially to the distal of the two γ-globin CCAAT regions and has therefore affinities opposite to those of protein CP1 for these two elements. Indeed, NF-E and CP1 are able to bind simultaneously to their respective high-affinity sites in the duplicated CCAAT region. Protein NF-E does not recognize the distal CCAAT sequence *per se*, as a substitution of the entire CCAAT sequence by the unrelated pentanucleotide AATCA does not abolish NF-E binding. This protein interacts instead with flanking sequences on both sides of the distal CCAAT element including the downstream core recognition sequence GACAAGG (−104 to −98) and the upstream G residue at −117. Recently, erythroid-cell-specific factors have been shown to bind to the promoter of the chicken α^D-globin gene (Kemper *et al.*, 1987), to region IV in the 3′ enhancer of the chicken adult β-globin gene (Emerson *et al.*, 1987) and to the second intron of the mouse β^{maj}-globin gene (Galson and Housman, 1988). The binding sites of the chicken and murine factor conform in all three cases to the consensus sequence GATA$^{A}_{G}$G which is similar to the NF-E core recognition sequence GACAAGG in the γ-globin CCAAT region. It is therefore possible that NF-E is identical or closely related to these previously described proteins.

The Greek HPFH mutation at position −117 efficiently reduces DNA binding of NF-E by a factor of 8−10. This strongly suggests that NF-E is the postulated repressor of γ-globin genes in adult erythroid cells. It is conceivable that both CCAAT sequences of the γ-globin gene must be occupied by two molecules of the CCAAT-binding protein CP1 to allow optimal interaction between CP1 and activator proteins bound to upstream and downstream sites of the γ-globin gene promoter. NF-E could act as a repressor by disrupting productive interaction between these transcription factors and might achieve this by preventing the binding of CP1 to the distal CCAAT element for which NF-E has a higher affinity than CP1. According to this model the −117 mutation in Greek HPFH would change the competition between these two factors in favor of the activator CP1 by increasing the affinity of CP1 and by simultaneously reducing that of the repressor NF-E for the distal CCAAT element. Interestingly, the G to A mutation at position −104 in the NF-E core recognition sequence has so far not been found associated with nondeletion HPFH, although it completely eliminates NF-E binding *in vitro*. One possible reason could be that this base change does not sufficiently increase CP1 binding.

NF-E also binds strongly to the −70 CCAAT region of the β-globin gene which, in analogy to the rabbit and mouse β-globin genes, is most likely a site of positive transcriptional

Table I. Oligonucleotides used in competition experiments

Aγ distal CCAAT box	5′ GGCAGCCTTGACCAATAGCCTTGACAAGGCA 3′
Aγ proximal CCAAT box	5′ GGCAAACTTGACCAATAGTCTTAGAGTGGCA 3′
α1 CCAAT box	5′ CTCCGCGCCAGCCAATGAGCGCCGCCCCTCC 3′
Eα (MHC II) CCAAT box	5′ AGCACTCAACTTTTAACCAATCAGAAAAATGTTTCAGAC 3′
MSV LTR CCAAT box	5′ CTTATTTGAACTAACCAATCAGTTCGCTTCT 3′
HSV tk CCAAT box	5′ GCGTGTTCGAATTCGCCAATGACAAGACGCT 3′
Adeno NF-I site	5′ TATACCTTATTTTGGATTGAAGCCAATATGATTGC 3′
Py enhancer core	5′ AGAGGGCAGTGTGGTTTTGCAAGAGGAAGCAAA 3′

regulation (Dierks *et al.*, 1983; Charnay *et al.*, 1985). NF-E may therefore have opposite effects on γ- and β-globin genes by repressing the transcription of the γ-globin gene and stimulating that of the β-globin gene in analogy to the yeast DNA-binding protein RAP1 that plays a role in either transcriptional activation or repression depending upon the context of its binding site (Shore and Nasmyth, 1987). This simplistic and speculative model predicts a higher level of NF-E in adult erythroid cells than in fetal cells. Indeed, we have observed more NF-E-binding activity in murine MEL cells than in the human embryonic/fetal cell line K562 (Figure 7) provided that extrapolation from these cell lines to the *in vivo* situation is valid.

All arguments for the existence of a repressor of γ-globin genes are based on the presence of the same −117 mutation in Greek (Collins *et al.*, 1985; Gelinas *et al.*, 1985) and Sardinian HPFH (Ottolenghi *et al.*, 1988). In both cases, population studies have demonstrated complete concordance of the −117 mutation with the HPFH phenotype (Waber *et al.*, 1986; Ottolenghi *et al.*, 1988) and thus strengthened the evidence that this point mutation causes an estimated 20- to 40-fold overexpression of the Aγ-globin gene in adult erythroid cells (cf. Stamatoyannopoulos and Nienhuis, 1987). However, only minor effects, if any, of the −117 mutation have so far been observed in gene transfer experiments with erythyroleukemia cells (Charnay and Henry, 1986; Collins *et al.*, 1986; Stoeckert *et al.*, 1987) or in transgenic mice (G.Kollias, unpublished results). The reasons for this discrepancy are largely unknown at present. Here we have subjected the critical γ-globin CCAAT region to an extensive biochemical and mutational analysis. Some of these mutants may provide, in combination with an improved globin expression system (Grosveld *et al.*, 1897), a new entry point to assess the functional role of the CCAAT region and its DNA-binding proteins in the developmental regulation of γ-globin genes.

Materials and methods

Preparation of nuclear extracts

HeLa, BJA-B and K562 cells and the MEL cell line FBU (Conscience and Meier, 1980) were grown in suspension to 5×10^5 cells/ml. Nuclear extracts were prepared according to Dignam *et al.* (1983) with minor modifications. After $(NH_4)_2SO_4$ precipitation the extract was resuspended in and dialyzed against 20 mM Hepes pH 7.9, 20% glycerol, 20 mM KCl, 2 mM $MgCl_2$, 0.2 mM EDTA. The protein concentration of these extracts was 5−10 mg/ml as determined by the method of Bradford (1976).

Preparation of poly(dI-dC) DNA

We have observed that the average length of ds DNA molecules varies considerably between different batches of commercially available poly(dI-dC) (Pharmacia, Sweden). In order to obtain consistent results for all three proteins (CDP, CP1 and NF-E) in mobility shift experiments, we had to shear poly(dI-dC) by sonication to an average length of 300−500 bp. After phenol extraction and ethanol precipitation the sonicated DNA was dissolved in 10 mM Tris pH 7.5, 1 mM EDTA.

Sequence and annealing of oligonucleotides

The sequences of those oligonucleotides that are not already shown in Figures 4 and 7 are listed in Table I. Only one of the two DNA strands is indicated. All synthetic oligonucleotides were purified as described by Barberis *et al.* (1987). Of the 5′ end-labeled oligonuclotide 1 pmol was annealed with a 5-fold excess of the unlabeled complementary DNA strand in 100 mM NaCl, 50 mM Tris−HCl pH 7.5, 6 mM $MgCl_2$ and 5 mM DTT by heating to 85°C for 2 min followed by slowly lowering the temperature to 4°C. Annealed oligonucleotides were diluted to the appropriate concentration in 50 mM NaCl, 10 mM Tris pH 7.5, 1 mM $MgCl_2$.

Electrophoretic mobility shift assay

Binding reactions were carried out by incubating end-labeled DNA (1 fmol/5000 c.p.m.) with 4−6 μg of nuclear protein and 2 μg of sheared poly(dI-dC) in a buffer containing 10 mM Hepes (pH 7.9), 100−150 mM KCl, 4% Ficoll, 1 mM DTT, 1 mM EDTA. After 20 min on ice, the reaction mixtures were loaded onto a 4% polyacrylamide gel in 0.25 × TBE and electrophoresed at 10 V/cm for 2 h at room temperature. The gel was then soaked in 20% methanol/10% acetic acid, dried and autoradiographed.

Methylation interference analysis

End-labeled DNA was partially methylated by dimethylsulfate according to Maxam and Gilbert (1980), incubated with the nuclear extract and then electrophoresed as described above. Bound and free DNA was eluted from the acrylamide, purified on DEAE 52 cellulose and then cleaved at methylated G residues in 100 μl of 1 M piperidine at 90°C for 30 min prior to analysis on denaturing sequencing gels.

Acknowledgements

We are especially grateful to Dr D.Goeddel for providing the opportunity for M.B. and G.S.-F. to work in his laboratory and to Genentech's oligonucleotide synthesis group for their service. We thank Drs S.Wright, D.Picard and A.Rosenthal for critical reading of the manuscript, Dr F.Grosveld for helpful discussion, Dr R.Tjian for providing purified CTF and C.Andow for graphical work. This work was supported in part by the State of Zürich and the Swiss National Research Foundation (grant 3.438-2.83), by the Institute of Molecular Pathology and by Genentech Inc.

References

Anagnou,N.P., Karlsson,S., Moulton,A.D., Keller,G. and Nienhuis,A.W. (1986) *EMBO J.*, **5**, 121−126.
Barberis,A., Superti-Furga,G. and Busslinger,M. (1987) *Cell*, **50**, 347−359.
Benoist,C., O'Hare,K., Breathnach,R. and Chambon,P. (1980) *Nucleic Acids Res.*, **8**, 127−142.
Bradford,M.M. (1976) *Anal. Biochem.*, **72**, 248−254.
Charnay,P. and Henry,L. (1986) *Eur. J. Biochem.*, **159**, 475−478.
Charnay,P., Mellon,P. and Maniatis,T. (1985) *Mol. Cell. Biol.*, **5**, 1498−1511.
Chodosh,L.A., Baldwin,A.S., Carthew,R.W. and Sharp,P.A. (1988a) *Cell*, **53**, 11−24.
Chodosh,L.A., Olesen,J., Hahn,S., Baldwin,A.S., Guarente,L. and Sharp, P.A. (1988b) *Cell*, **53**, 25−35.
Collins,F.S. and Weissman,S.M. (1984) *Prog. Nucleic Acid Res. Mol. Biol.*, **31**, 315−438.
Collins,F.S., Metherall,J.E., Yamakawa,M., Pan,J., Weissman,S.M. and Forget,B.G. (1985) *Nature*, **313**, 325−326.
Collins,F.S., Bodine,D.M., Lockwood,W.K., Cole,J.L., Mickley,L. and Ley,T. (1986) *Clin. Res.*, **34**, 454a.
Cohen,R.B., Sheffery,M. and Kim,C.G. (1986) *Mol. Cell. Biol.*, **6**, 821−832.
Conscience,J.-F. and Meier,W. (1980) *Exp. Cell Res.*, **125**, 111−119.
Dierks,P., van Ooyen,A., Cochran,M.D., Dobkin,C., Reiser,J. and Weissmann,C. (1983) *Cell*, **32**, 695−706.
Dignam,J.D., Lebovitz,R.M. and Roeder,R.G. (1983) *Nucleic Acids Res.*, **11**, 1475−1489.
Dorn,A., Bollekens,J., Staub,A., Benoist,C. and Mathis,D. (1987) *Cell*, **50**, 863−872.
Efstratiadis,A., Posakony,J.W., Maniatis,T., Lawn,R.M., O'Connell,C., Spritz,R.A., DeRiel,J.K., Forget,B.G., Wiessman,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.
Emerson,B.M., Nickol,J.M., Jackson,D.P. and Felsenfeld,G. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 4786−4790.
Fried,M. and Crothers,D.M. (1981) *Nucleic Acids Res.*, **9**, 6505−6525.
Galson,D.L. and Housman,D.E. (1988) *Mol. Cell Biol.*, **8**, 381−392.
Garner,M.M. and Rezvin,A. (1981) *Nucleic Acids Res.*, **9**, 3047−3060.
Gelinas,R., Endlich,B., Pfeiffer,C., Yagi,M. and Stamatoyannopoulos,G. (1985) *Nature*, **313**, 323−324.
Graves,B.J., Johnson,P.F. and McKnight,S.L. (1986) *Cell*, **44**, 565−576.
Grosveld,F., Blom van Assendelft,G., Greaves,D.R. and Kollias,G. (1987) *Cell*, **51**, 975−985.
Johnson,P.F., Landschulz,W.H., Graves,B.J. and McKnight,S.L. (1987) *Genes Dev.*, **1**, 133−146.
Jones,K.A., Yamamoto,K.R. and Tjian,R. (1985) *Cell*, **42**, 559−572.

Jones,K.A., Kadonaga,J.T., Rosenfeld,P.J., Kelly,T.J. and Tjian,R. (1987) *Cell*, **48**, 79–89.
Karlsson,S. and Nienhuis,A.W. (1985) *Annu. Rev. Biochem.*, **54**, 1071–1108.
Kemper,B., Jackson,P.D. and Felsenfeld,G. (1987) *Mol. Cell. Biol.*, **7**, 2059–2069.
Kioussis,D., Wilson,F., Khazaie,K. and Grosveld,F. (1985) *EMBO J.*, **4**, 927–931.
Lichtsteiner,S., Wuarin,J. and Schibler,U. (1987) *Cell*, **51**, 963–973.
Maniatis,T., Goodbourn,S. and Fischer,J.A. (1987) *Science*, **236**, 1237–1244.
Maxam,A. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499–560.
Morgan,W.D., Williams,G.T., Morimoto,R.I., Greene,J., Kingston,R.E. and Tjian,R. (1987) *Mol. Cell Biol.*, **7**, 1129–1138.
Myers,R.M., Tilly,K. and Maniatis,T. (1986) *Science*, **232**, 613–618.
Olesen,J., Hahn,S. and Guarente,L. (1987) *Cell*, **51**, 953–961.
Ottolenghi,S., Camaschella,C., Comi,P., Giglioni,B., Longinotti,M., Oggiano,L., Dore,F., Sciarratta,G., Ivaldi,G., Saglio,G., Serra,A., Loi,A. and Pirastu,M. (1988) *Hum. Genet.*, **79**, 13–17.
Raymondjean,M., Cereghini,S. and Yaniv,M. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 757–761.
Rixon,M.W. and Gelinas,R.E. (1988) *Mol. Cell Biol.*, **8**, 713–721.
Siebenlist,U., Simpson,R. and Gilbert,W. (1980) *Cell*, **20**, 269–281.
Shore,D. and Nasmyth,K. (1987) *Cell*, **51**, 712–732.
Slightom,J.L., Blechl,A.E. and Smithies,O. (1980) *Cell*, **21**, 627–638.
Stamatoyannopoulos,G. and Nienhuis,A.W. (1987) Hemoglobin switching. In Stamatoyannopoulos,G., Nienhuis,A.W., Leder,P. and Majerus,P.W. (eds), *The Molecular Basis of Blood Diseases*. W.B.Saunders, Philadelphia, pp. 66–105.
Stoeckert,C.J., Metherall,J.E., Yamakawa,M., Weissman,S.M. and Forget,B.G. (1987) *Mol. Cell Biol.*, **7**, 2999–3003.
Waber,P.G., Bender,M.A., Gelinas,R.E., Kattamis,C., Karaklis,A., Sofroniadou,K., Stamatoyannopoulos,G., Collins,F.S., Forget,B.G. and Kazazian,H.H.,Jr (1986) *Blood*, **67**, 551–554.
Wright,S., Rosenthal,A., Flavell,R. and Grosveld,F. (1984) *Cell*, **38**, 265–273.

Received on May 17, 1988; revised on July 1, 1988