MOLECULAR AND CELLULAR BIOLOGY, Oct. 1993, p. 6290–6303
0270-7306/93/106290-14$02.00/0

Vol. 13, No. 10

Transcriptional Up-Regulation of the Mouse Cytosolic Glutathione Peroxidase Gene in Erythroid Cells Is Due to a Tissue-Specific 3′ Enhancer Containing Functionally Important CACC/GT Motifs and Binding Sites for GATA and Ets Transcription Factors

JIM O'PREY, SCOT RAMSAY, IAN CHAMBERS,† AND PAUL R. HARRISON*

Beatson Institute for Cancer Research, Cancer Research Campaign Beatson Laboratories, Garscube Estate, Switchback Road, Bearsden, Glasgow G61 1BD, Scotland

Received 26 April 1993/Returned for modification 31 May 1993/Accepted 13 July 1993

Nuclear run-on experiments have shown that the high level of expression of the mouse cytosolic glutathione peroxidase mRNA in erythroid cells is due to up-regulation of the gene at the transcriptional level. Studies of the chromatin structure around the cytosolic glutathione peroxidase gene have revealed a series of DNase I hypersensitive sites (DHSS) in the 3′ flanking region of the gene in erythroid and other high-expression tissues that are lacking in low-expression cells, in addition to a DHSS over the promoter region in both high- and low-expression tissues. Functional transfection experiments have demonstrated that one of the 3′ DHSS regions functions as an enhancer in erythroid cells but not in a low-expression epithelial cell line; and site-directed mutagenesis and footprinting experiments reveal that the activity of the erythroid cell-specific enhancer requires a cluster of binding sites for the CACC/GT box factors and the GATA and Ets families of transcription factors.

The glutathione peroxidases (GSHPXs) constitute a family of selenocysteine (SeCys)-containing enzymes whose function is to remove organic peroxides and H_2O_2 (see references 9 and 14 for reviews). To date, four enzymes with different tissue distributions have been characterized: the cytosolic GSHPX (cGSHPX) was the first to be isolated and sequenced, but a different GSHPX present in plasma and a third form of GSHPX that specifically reduces hydroperoxide derivatives of intact phospholipids have been characterized (reviewed in reference 66); very recently, another cellular GSHPX different from cGSHPX has also been characterized (12). As we first showed for cGSHPX (8), the SeCys-containing selenoproteins are unique in that the SeCys is encoded by an in-frame UGA codon; this has now been demonstrated for all four GSHPX types (12, 61, 67) and several other eukaryotic and bacterial SeCys-containing selenoproteins (6, 20, 27, 75, 76). All cells contain low levels of cGSHPX, but high levels are particularly important in tissues in which peroxides are produced at high levels, for example, in erythrocytes as a by-product of the redox reactions involved in oxygen transport or during detoxification processes in liver or kidney. We have shown previously (1, 2, 8) that the high-level expression of cGSHPX in erythroid cells, liver, and kidney in comparison to most other tissues (brain tissue, fibroblasts, and epithelial cells, for example) is due to up-regulation at the level of cGSHPX mRNA. In this report, we analyze further the molecular mechanisms responsible for this tissue-specific regulation: we show that the cGSHPX gene is in an altered chromatin structure in high-expression tissues, as judged by a series of tissue-specific DNase I-hypersensitive sites (DHSS) 3′ to the gene and that up-regulation of transcription in erythroid cells is due to a tissue-specific 3′ enhancer. Detailed analysis of the nature of the nuclear proteins binding to the enhancer and mutagenesis studies have shown that members of the GATA, CACC/GT box, and Ets families of transcription factors are involved in enhancer function.

* Corresponding author.
† Present address: Centre for Genome Research, Edinburgh EH9 3JQ, Scotland.

MATERIALS AND METHODS

Cells and cell culture. Adherent mouse erythroleukemia (MEL) cells (F4-12B2; a gift from W. Ostertag, Hamburg, Germany), mouse epithelial cell line C5 (35), STO mouse fibroblasts, HepG2 liver cells, and HeLa cells were grown in a modification of Eagle minimal essential medium with double concentrations of amino acids and vitamins, 10% fetal calf serum (Tissue Culture Services), and in transient transfection experiments with 100 μg of streptomycin per ml and 37.5 μg of penicillin per ml. The erythroid/megakarocyte HEL cell line (39) and the erythroid/myeloid K562 cell line were grown in RPMI 1640 medium supplemented with 10% fetal calf serum. KMOE cells (47) were grown in a mixture consisting of RPMI 1640 (33%), α minimal essential medium (33%), and Ham's F12 (33%) supplemented with 10% fetal calf serum. FDCP-mix A4 cells were grown as previously described (25).

Nuclear run-on experiments. Briefly, 2×10^7 nuclei in 200 μl of a solution of 30 mM Tris-HCl (pH 8.3), 5 mM $MgCl_2$, 150 mM KCl, 0.05 mM EDTA, 1 mM dithiothreitol, 20% (vol/vol) glycerol, and 0.25 mM (each) ATP, GTP, and CTP were labelled with 3.7 MBq of [α-^{32}P]UTP for 30 min at 30°C, and RNA was extracted by the RNAzol method (10). RNA (4×10^6) cpm from MEL or C5 nuclei was hydridized at 42°C for 72 h to prehydridized filters containing 2 μg of test or control plasmid in a solution of 50% formamide, 5×

SSPE (1× SSPE is 0.18 M NaCl, 10 mM $NaPO_4$, and 1 mM EDTA [pH 7.7]), 5× Denhardt's solution (59), 0.5% sodium dodecyl sulfate (SDS), and 100 μg of denatured herring testis DNA per ml and then washed in 1× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.1% SDS at 65°C.

Mapping of DHSS. Nuclei were prepared (4), and DNase I digestions were performed for 10 min at 0°C in a solution of 60 mM KCl, 15 mM NaCl, 15 mM Tris-HCl (pH 7.5), 5 mM $MgCl_2$, 0.5 mM spermidine 0.15 mM spermine, 1 mM phenylmethylsulfonyl fluoride, and 0.2 mM EGTA [ethylene glycol-bis(β-aminoethyl ether)-*N,N,N',N'*-tetraacetic acid] at a nucleic acid concentration of 0.5 mg/ml (16). Genomic DNA was then isolated by the proteinase K treatment, phenol-chloroform extraction, and RNase treatment (1).

Plasmid constructions. Plasmid pLW2 (17) contains the bacterial chloramphenicol acetyltransferase (CAT) gene under the control of the herpes simplex virus type 2 (HSV-2) immediate-early 5 (IE-5) gene promoter with the transcriptional termination sequences of the HSV-2 IE-5 gene inserted downstream of the CAT gene. Plasmid p22 is a promoterless CAT construct used as the basis for construction of all cGSHPx-CAT recombinants. p22 was derived from pLW2 by insertion of the 900-bp *Bam*HI-*Xba*I fragment containing the CAT gene and transcription termination sequences into plasmid pUC13. Plasmid pGPY289 contains cGSHPx promoter sequences from −289 to +10 bp (with respect to the transcription initiation site) immediately upstream of the CAT gene of p22. pGPY289 was constructed by cloning the 470-bp *Xba*I-*Pst*I fragment of the cGSHPx promoter into pUC12. The *Xba*I-*Hin*fI fragment was isolated from this plasmid, blunt ended, and cloned into the *Sma*I site of p22. Plasmid pGPY158 was derived from pGPY289 by removal of the 5′ 141 bp by *Xma*III-*Sst*I digestion. Plasmid pGPY1.7, containing promoter sequences from −1700 to +10 bp, was constructed by cloning a 1.55-kb *Eco*RI-*Xma*III fragment of promoter sequence into pGPY289. Plasmid pGPY1.7(4.0)+ contains the 3′ 4-kb partial *Eco*RI-*Bam*HI fragment (see Fig. 3), pGPY1.7(1.8)+ contains the 1.8-kb 3′ *Eco*RI fragment, and plasmid pGPY1.7(2.2)+ contains the 3′ 2.2-kb *Eco*RI-*Bam*HI fragment cloned into the polylinker sequence 3′ to the CAT gene in pGPY1.7. pGPY1.7(650)+, pGPY1.7(HS II), and pGPY1.7(HS III) contain 650-, 340-, and 310-bp polymerase chain reaction (PCR) fragments corresponding to regions containing DHSS II and III, DHSS II, and DHSS III, respectively. These fragments were cloned in the *Hin*dIII site of pGPY1.7 in either their natural (+) or opposite (−) orientations.

Plasmid pHSV β-Gal contains the β-galactosidase gene under the control of the HSV IE-5 gene promoter and was used to standardize for transfection efficiency in the functional experiments.

Preparation of PCR fragments. Specific DNA sequences were amplified by the PCR with Perkin-Elmer Cetus AmpliTaq DNA polymerase. To amplify sequences from plasmid DNA, 20 ng of plasmid DNA was mixed with the appropriate oligonucleotide primers (1 μM) in 100 μl of a solution of 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.1% gelatin, 200 μM deoxynucleoside triphosphates, and 2.5 U of AmpliTaq DNA polymerase. Amplification was achieved by 30 cycles of 1 min at 93°C, 1 min at 58°C, and 1.5 min at 75°C in a Perkin-Elmer Cetus thermal cycler. Extension time of the last cycle was 10 min to ensure that all single-stranded DNA had been copied or reannealed to form double strands. The products of the PCR were resolved by agarose gel electrophoresis and purified by Geneclean (BIO 101 Inc.).

Nuclear protein preparation. Approximately 10^{10} cells were washed in phosphate-buffered saline and then twice with TMS (5 mM Tris-HCl [pH 7.5], 2.5 mM $MgCl_2$, 125 mM sucrose). Cells were lysed in 200 ml of TMS plus 0.25% Triton X-100, and nuclei were harvested by centrifugation (1,600 × *g*, 20 min). Nuclei were washed three times in 200 ml of TMS and resuspended in approximately 5 ml of TMS (5 to 10 mg of DNA per ml), and 0.1 volumes of 4 M NaCl was added dropwise with stirring. The solution was centrifuged at 10,000 × *g* for 20 min, and the supernatant was spun at 100,000 × *g* for 60 min. Solid ammonium sulfate was added to 0.35 g/ml and left on ice for 30 min. The precipitate was pelleted at 10,000 × *g* for 30 min and redissolved in 5 ml of E50 buffer (50 mM ammonium sulfate, 20 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid [HEPES; pH 7.9], 5 mM $MgCl_2$, 0.1 mM EDTA, 0.1% [vol/vol] Brij 35, 20% [vol/vol] glycerol, 1 mM dithiothreitol) and dialyzed for 16 h against 1 liter of storage buffer (50 mM NaCl, 20 mM HEPES [pH 7.9], 5 mM $MgCl_2$, 20% [vol/vol] glycerol, 1 mM DTT). The crude protein extract was cleared by centrifugation at 100,000 × *g* for 60 min, and aliquots were stored at −70°C.

A microscale technique for the preparation of DNA-binding proteins (3) was also employed for electrophoretic mobility shift assays (EMSAs). Where possible, microscale and large-scale nuclear protein preparations were compared in EMSAs and found to produce identical results.

All manipulations were carried out at 4°C, and all solutions contained 0.5 mM phenylmethylsulfonyl fluoride, 0.5 mM benzamidine, 10 mM sodium butyrate, 10 mM β-glycerophosphate, 2 mM levamisole, 50 mM sodium orthovanadate (pH 8.0), and 1 μg (each) of leupeptin, aprotinin, bestatin, and pepstatin per ml.

FP analysis. Footprinting probes were prepared by 5′ end labelling of DNA restriction fragments with T4 polynucleotide kinase and [γ-^{32}P]ATP and isolated after secondary restriction (53). Alternatively, probes were generated by PCR with one of the primers first end labelled with [γ-^{32}P]ATP (34). Markers were prepared by chemical sequencing reactions (40). DNase I footprint (FP) protection assays were performed in a volume of 100 μl of storage buffer in the presence of 6 μg of poly(dI-dC) · poly(dI-dC), 2 ng of end-labelled restriction fragment, and up to 80 μl (30 μg) of nuclear protein extract. After partial DNase I digestion, the nucleic acid was purified and resolved by denaturing 6% polyacrylamide gel electrophoresis and autoradiography.

EMSA. Double-stranded oligonucleotides were 5′ end labelled with T4 polynucleotide kinase and [γ-^{32}P]ATP, and 5′ overhangs were filled in with Klenow fragment of DNA polymerase and excess deoxynucleotide triphosphates and purified by electrophoresis on an 8% polyacrylamide gel. Labelled oligonucleotide (0.1 ng) was incubated on ice for 30 min with protein extracts (0 to 10 μg) in the presence of 6 μg of poly(dI-dC) · poly(dI-dC) in a final volume of 20 μl of storage buffer. Samples were electrophoresed in a 5% polyacrylamide gel in 0.2× Tris-borate-EDTA for 2 to 3 h at 4°C at 150 V, and the gel was dried for autoradiography. Oligonucleotide competitors were added at a 200-fold molar excess. The following oligonucleotides were used in this study (all sequences are written 5′ to 3′): Fp1, CCCAAACGTG AACTTCTCCTTTCTGAGTTT; Fp2, TTGGTGGATGAGG AAGTGACTGTGTGCCCC; Fp3, GACTCTAGAAATCTA GATCTTTGCCCACAG; Fp4, TGTTTTGCACAGCTGCA ACAGCCTTGCACA; GATA, CCGGGCAACTGATAAG GATTCCCTG; mGATA, CCGGGCAACTGATGAGGAT TCCCTG; Ets3, ATACTCCAGAAGGAGGAAGCCAGTA GACAA; PU-1, GATCCATAACCTCTGAAAGAGGAACT

TGGTTAGGT; PEA3, GATCCTCGAGCAGGAAGTTC GA; E74, GATCTCTAGCTGAATAACCGGAAGTAACT CATCCTAG; E74m, GATCTCTAGCTGAATAACCCAAG TAACTCATCCTAG; AP1, AAGCATGAGTCAGACA; and CCAAT, GATCCAAACCAGCCAATGAGAACTGCTCCA.

Generation of site-directed mutations by PCR. Introduction of specific mutations and deletions into the enhancer fragments by recombinant PCR was performed as previously described (26). Briefly, for each mutated construct a pair of overlapping primers containing the required mutations or deletion were used in two separate PCRs to produce two DNA fragments overlapping in the region to be mutated. The fragments were gel purified, and 20 ng of these primary PCR fragments was used in a further PCR together with outside primers containing *Hin*dIII restriction sites to produce a full-length DNA fragment containing the appropriate mutation or deletion. These fragments were purified, digested with *Hin*dIII, and cloned into the *Hin*dIII site of pGPY1.7. Constructs containing mutations or deletions introduced by PCR were sequenced by the dideoxy chain termination method (60). Mutations introduced into GATA-1 (GATA AGG to GATGAGG) and CACC factor (CACCC to GG GCC) binding sites were as previously described (52).

Primers used in the preparation of the constructs used in this study were as follows: pGPY1.7(650), 209H and 210H; pGPY1.7(HS II), 210H and 6; pGPY1.7(HS III), 5 and 209H; pGPY1.7(HS II)a, 33 and 34; pGPY1.7(HS II)b, 35 and 36; pGPY1.7(HS II)c, 37 and 38; pGPY1.7(HS II)d, 39 and 40; pGPY1.7(HS III)a, 17 and 18; pGPY1.7(HS III)b, 3 and 209H; pGPY1.7(HS III)c, 19 and 20; pGPY1.7(HS III)d, 10 and 13; pGPY1.7(HS III)e, 9 and 12; pGPY1.7(HS III)f, 23 and 24; pGPY1.7(HS III)g, 11 and 14; pGPY1.7(HS III)h, 11, 14, 23, and 24.

Oligonucleotide PCR primers. The following oligonucleotide PCR primers were used (all sequences are written 5′ to 3′): 209H, GCCAAGCTTCTTCATACTACCACAAGTCTC; 210H, GCCAAGCTTTCCCAGGACATGGTTAAGGTC; 3, AGCAAGCTTCGTCAAACCTTGCAAGTTTAC; 5, ACGC AAGCTTAAGGTGGAAGGCAGGTGTAG; 6, CGCAAGC TTTTCCGTCGAATCTCACAACAC; 9, GGGCCCCACCC TTTAAACATCTCAAGACAC; 10, GGGCCCGGGCCTTT AAATATCTCAAGACAC; 11, GGGCCCGGGCCTTTAAA CATCTCAAGACAC; 12, GTGTCTTGAGATGTTTAAAG GGTGGGGCCC; 13, GTGTCTTGAGATATTTAAAGGCC CGGGCCC; 14, GTGTCTTGAGATGTTTAAAGGCCCGG GCCC; 17, GTGTGGCAGGAAACACTCCAGTGTGTGTC C; 18, GGACACACACTGGAGTGTTTCCTGCCACAC; 19, AATAGGACACATCTTAAAACATACACATCTTAAA GG; 20, CCTTTAAGATGTGTATGTTTTAAGATGTGTC CTATT; 23, GGAGCTGTTTCTCATCATAACCTCCCAG AC; 24, GTCTGGGAGGTTATGATGAGAAACAGCTCC; 33, GGCCACCCTAACCCCTTTTTGGTGGATGAG; 34, CTCATCCACCAAAAAGGGGTTAGGGTGGCC; 35, TTC TGAGTTTTTGGTTGCCCCTGTTCACAC; 36, GTGTGA ACAGGGGCAACCAAAAACTCAGAA; 37, TTCACAC AAGACTCTCCCACAGAAGTCTGT; 38, ACAGACTTCT GTGGGAGAGTCTTGTGTGAA; 39, TTGCCCACAGAAG TCCACAGTGTTACCAGC; and 40, GCTGGTAACACTGT GGACTTCTGTGGGCAA.

Transfections and CAT assays. Plasmid constructs were transfected into MEL or C5 cells together with a plasmid containing the β-galactosidase gene (pHSVβ-gal) by the calcium phosphate precipitate method as previously described (15). CAT activities were measured, normalized for transfection efficiency by using the β-galactosidase activity as a control (48), and then expressed relative to the CAT activity produced in the same cell type by transfection of plasmid pLW2 (17) in which the CAT gene is under the control of the HSV IE-5 promoter.

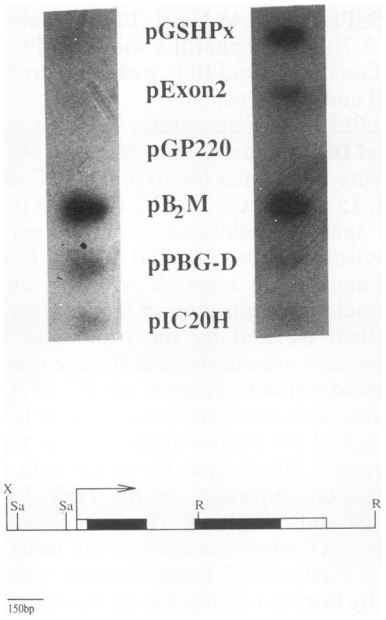
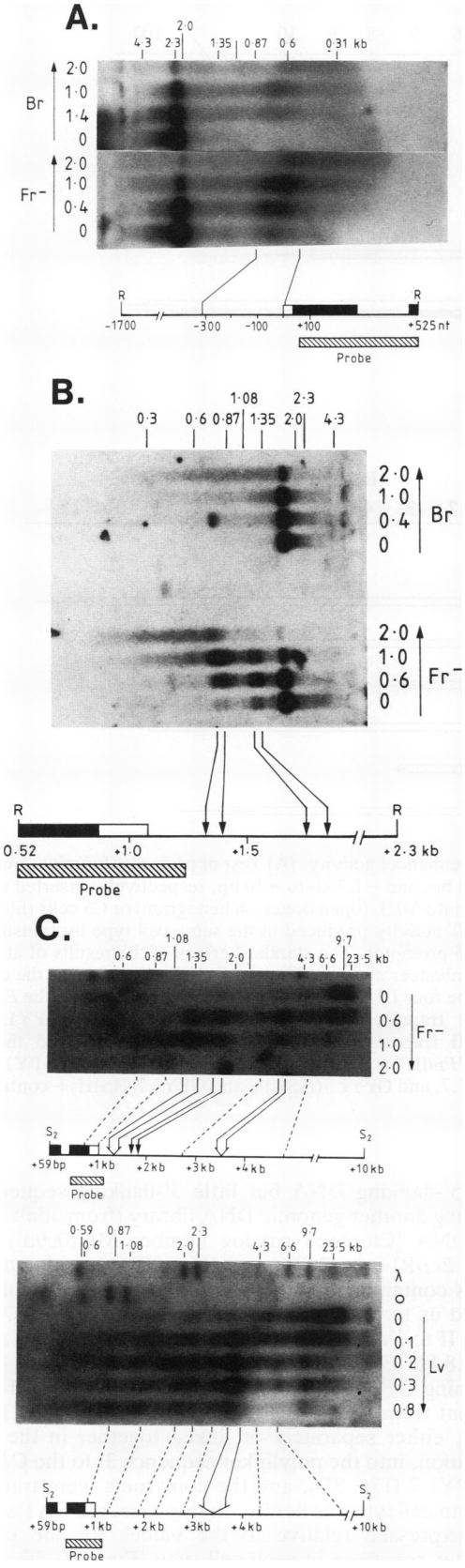


FIG. 1. Nuclear run-on experiments. The figure shows an autoradiograph of the hydridization of in vitro transcripts from MEL nuclei (right panel) or C5 epithelial cell nuclei (left panel) to the recombinant plasmids indicated. pGSHPX contains the entire 1.55-kb *Xba*I (X)-to-*Eco*RI (R) cGSHPX gene fragment. pExon2, the 706-bp *Eco*RI cGSHPX gene fragment; pGP220, the upstream 200-bp *Sau*3A (Sa) cGSHPX gene fragment (see line diagram of the cGSHPX gene, with the exons boxed and coding regions filled in and the transcription start site marked with an arrow). Plasmids containing part of the β_2M cDNA (pβ_2M) (65) and the entire PBG-D cDNA (pPBG-D) (11) and PIC20H were included as controls.

RESULTS

Nuclear run-on experiments. In order to confirm that the up-regulation of cGSHPX mRNA in erythroid cells is due to transcriptional control, in vitro transcription experiments were performed. To allow the possibility of future functional studies, MEL cells were compared with an epithelial cell line (C5) that expresses only low levels of cGSHPX mRNA, since both cell lines are capable of being transfected at a reasonable efficiency. Thus, nuclei from MEL or C5 cells were incubated with radioactive nucleoside triphosphates, and the nuclear RNA was extracted and hydridized to filters containing various fragments of the cGSHPX gene cloned in the plasmid vector pIC20H (see map of cGSHPX gene for details [Fig. 1]). The controls used included plasmid vector pIC20H and plasmids containing a fragment of the β_2-microglobulin (β_2M) cDNA or of the complete cDNA encoding the heme pathway enzyme, porphobilinogen deaminase (PBG-D). The results (Fig. 1) show a very much higher level of transcripts on the cGSHPX gene in MEL nuclei compared with those in C5 nuclei, whereas transcription on the β_2M and PBG-D genes is similar in MEL and C5 cells. The lower hydridization to the cGSHPX exon 2 probe compared with the entire gene probably simply reflects the fact that the former contains a shorter sequence to which transcripts can hybridize. As would be expected, no hydridization is found



to a probe upstream of the main transcription initiation site (pGP220).

Mapping of DHSS. To locate potential *cis* control regions that might be responsible for up-regulating the transcription of the cGSHPX gene in erythroid cells, DHSS in the vicinity of the cGSHPX gene were mapped by the technique of indirect end labelling. To map any DHSS 5′ to the gene, DNA purified from DNase I-treated nuclei was digested with *Eco*RI, and the fragments were electrophoresed in agarose before Southern blotting and hydridization with a 469-bp *Sst*II-to-*Eco*RI probe covering part of the first exon, intron, and second exon of the cGSHPX gene, as shown in Fig. 2A. This experiment showed cleavage of the 2.2-kb genomic *Eco*RI fragment in DNase-treated nuclei from MEL cells and, to a lesser extent, brain tissue (a low-expression tissue) to give a diffuse band at about 0.6 to 0.9 kb that maps to a diffuse 300-bp DHSS (site I) over the promoter.

To identify any DHSS 3′ to the cGSHPX gene, DNA from DNase-treated nuclei was digested with *Eco*RI, and the resulting Southern blot was hydridized with a 706-bp fragment containing mostly exon 2 sequences (Fig. 2B). These experiments showed DNase I cleavage of the 1.8-kb *Eco*RI genomic band in MEL nuclei in the region of two prominent DHSS resolved as doublets (sites II and III), mapping to about +1.4 and +1.8 kb from the transcription initiation site. However, DHSS II and III appear to be absent in brain tissue (Fig. 2B) (the apparent minor band in the 0.4 U of DNase track was spurious hydridization and not reproducible). Other experiments using an *Sst*II digest with the same exon 2 probe allowing detection of any DHSS within the 9-kb region 3′ to the cGSHPX gene (Fig. 2C) confirm the presence of DHSS II and III in MEL cells and reveal another 3′ DHSS (IV) at about +3.5 kb that is also present in liver, another high-expression tissue, together with other more diffuse DHSS (Fig. 2C). Other experiments have failed to detect DHSS IV in brain tissue (data not shown). Overall, these results suggest that *cis* control regions responsible for the high expression of the cGSHPX gene in erythroid cells lie in the 3′-flanking DNA from +1 to +3.5 kb.

The upstream 1.7 kb of the cGSHPX gene acts as a relatively non-tissue-specific promoter. To attempt to identify *cis* control regions regulating the tissue-specific transcription of the cGSHPX gene, we have tested the 5′-flanking DNA for

FIG. 2. Mapping DHSS around the cGSHPX gene in different tissues. Nuclei from MEL cells (Fr^-) or mouse liver (Lv) or brain tissue (Br) were treated with the indicated amounts of DNase I (units per milligram of DNA) prior to DNA isolation, as described in Materials and Methods, and then digested with the appropriate restriction enzyme before electrophoresis and Southern blotting. (A) 5′ DHSS were mapped by *Eco*RI digestion and hybridization with the *Sst*II-*Eco*RI probe shown (hatched bar). The positions of size markers are given at the top of the photograph. The line diagram shows the cGSHPX gene (with exons marked as boxes and the coding regions filled in) with the positions of the *Eco*RI sites (R) in the genomic DNA at −1.7 kb and +525 bp relative to the transcription initiation site and the region to which the DHSS maps marked. (B) Mapping of DHSS in the 3′-flanking DNA. DNA from DNase I-treated nuclei was digested with *Eco*RI, and the resulting blot was hydridized with the *Eco*RI exon 2 fragment probe shown (hatched box). The map indicates the positions of the genomic *Eco*RI sites at +525 bp and +2.3 kb and the positions to which the major DHSS map are shown. (C) DNA from DNase I-treated nuclei was digested with *Sst*II, and the blot was hydridized with the same exon 2 probe used for Fig. 2B (hatched). The positions of the *Sst*II sites at +59 bp and about +9 to 10 kb are shown.

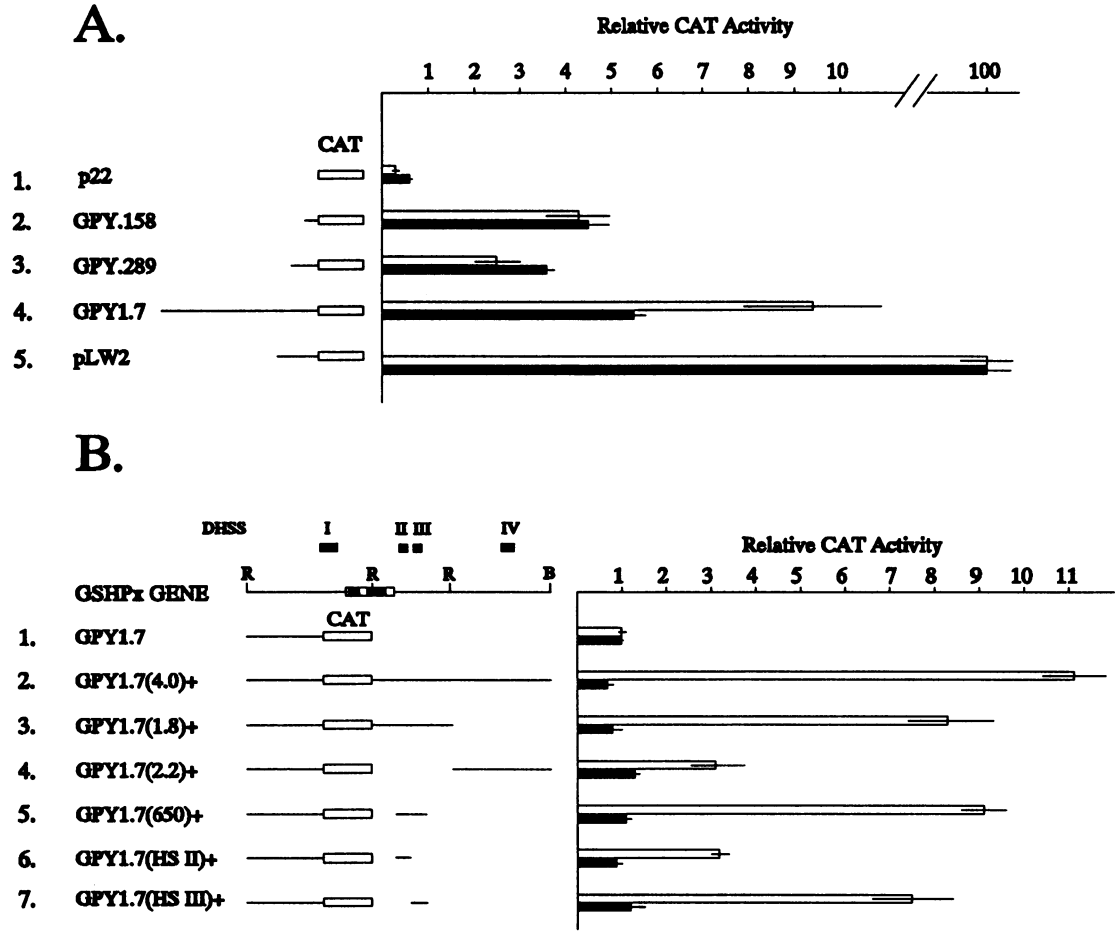


FIG. 3. Test of cGSHPX 5′ and 3′ sequences for tissue-specific promoter and enhancer activity. (A) Test of promoter function. pGPY.289, pGPY.158, and pGPY.1.7 contain cGSHPX gene sequences from −289 bp, −158 bp, and −1.7 kb to +10 bp, respectively, inserted upstream of the promoterless CAT gene in plasmid p22. These constructs were transfected into MEL (open boxes on histogram) or C5 cells (filled boxes on histogram), CAT activities were measured and expressed relative to the CAT activity produced in the same cell type by transfection of plasmid pLW2 (17) in which the CAT gene is under the control of the HSV IE-5 promoter. The standard errors of the results of at least six transfections are indicated by the error bars. (B) Test of 3′ DHSS regions for enhancer activity. The top line shows a map of the cGSHPX gene with the exons boxed and the coding regions filled in and the positions of the four DHSS marked above. The positions of the *Eco*RI (R) and *Bam*HI (B) restriction enzyme sites used to generate constructs for transfection studies are also marked. pGPY1.7(4.0)+, pGPY1.7(1.8)+, and pGPY1.7(2.2)+ contain the 4-kb partial *Eco*RI-*Bam*HI fragment, the 1.8-kb *Eco*RI fragment and the 2.2-kb *Eco*RI-*Bam*HI fragment, respectively, cloned in the natural orientation into the *Hin*dIII site 3′ to the CAT gene in pGPY1.7. pGPY1.7(650)+ contains a 650-bp subfragment of the 1.8-kb *Eco*RI fragment inserted into pGPY1.7, and GPY1.7(HSII)+ and GPY1.7(HSIII)+ contain either the DHSS II or DHSS III regions alone, respectively.

promoter activity when linked to the CAT reporter gene and transfected into MEL or C5 cells. In all our transfection experiments, variations in transfection efficiency were controlled by measurement of the level of β-galactosidase produced from a β-galactosidase expression plasmid cotransfected with the cGSHPX/CAT constructs. We have tested three promoter constructs containing sequences from −1.7 kb, −289 bp, or −158 bp to +10 bp inserted upstream of the CAT gene (the cGSHPX translation initiation codon is at +38 bp). The results (Fig. 3A) show that the −289 and −158 bp promoters function equally well in both MEL and C5 cells compared with the activity of the HSV IE-5 promoter in pLW2. However, the longer 1.7-kb cGSHPX promoter fragment functions 1.5 to 2 times as efficiently in MEL cells compared with C5 cells.

Identification of an erythroid cell-specific 3′ enhancer. Our original genomic DNA recombinant (21) contained about 14 kb of 5′-flanking DNA but little 3′-flanking sequence. By screening another genomic DNA library (from adult DBA/2J liver DNA [Clontech catalog number ML1009d]) with a 706-bp *Eco*RI exon 2 fragment (Fig. 2B), we isolated recombinants containing about 5 kb of 3′-flanking sequence. This allowed us to test fragments of the 3′-flanking DNA in the DHSS II to IV regions for tissue-specific enhancer activity. The 1.8-kb *Eco*RI fragment from +525 bp to +2.3 kb containing DHSS II and III and the 2.2-kb *Eco*RI-*Bam*HI fragment from +2.3 to +4.5 kb containing DHSS IV were cloned, either separately or linked together in the natural orientation, into the polylinker sequence 3′ to the CAT gene in pGPY1.7 (Fig. 3B), and the constructs were transfected into both cell types as before. In this case, the CAT activities were expressed relative to the values for the pGPY1.7 promoter construct in each cell type (Fig. 3B). The results show that the entire 4-kb region containing DHSS II to IV

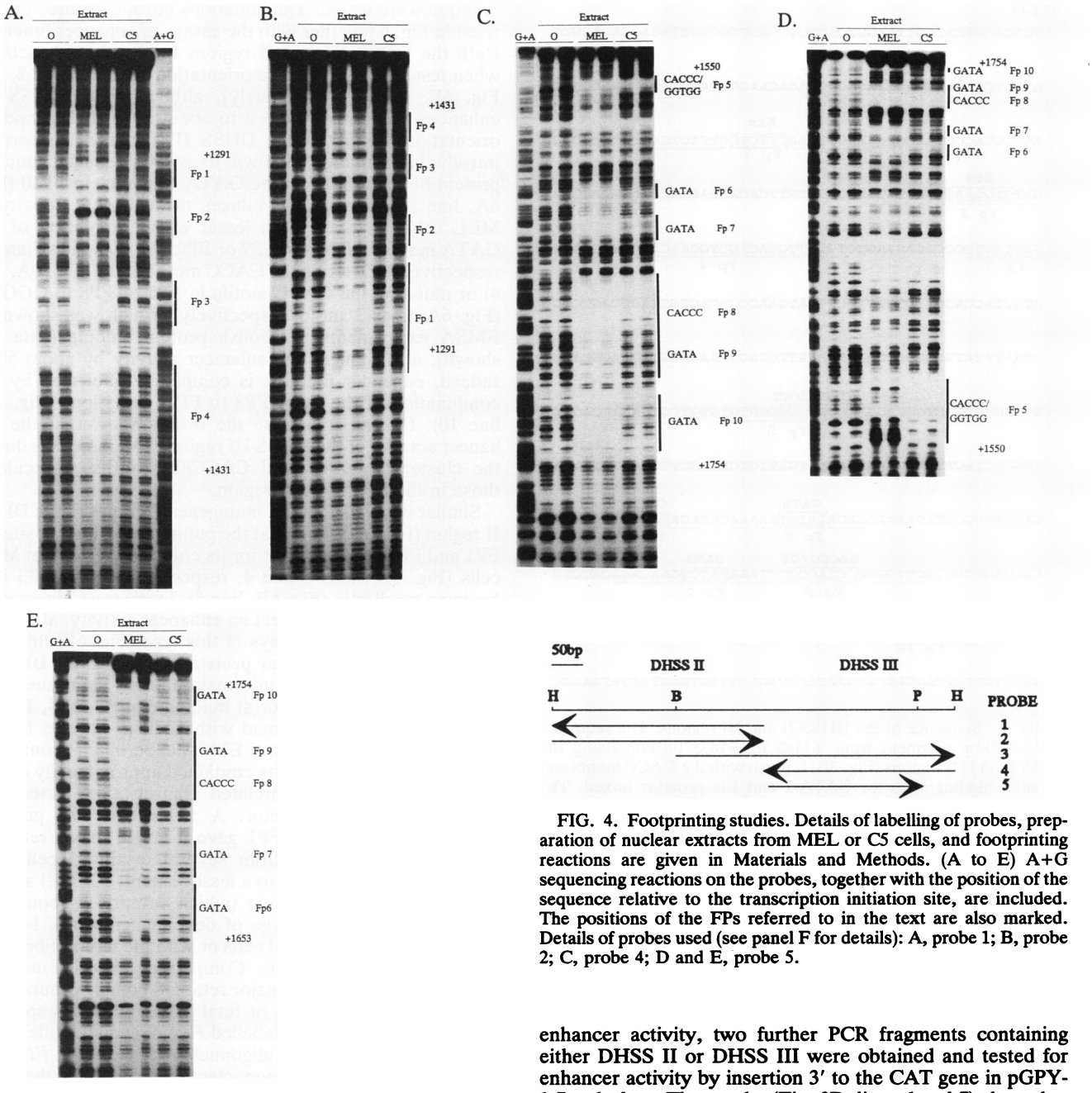


FIG. 4. Footprinting studies. Details of labelling of probes, preparation of nuclear extracts from MEL or C5 cells, and footprinting reactions are given in Materials and Methods. (A to E) A+G sequencing reactions on the probes, together with the position of the sequence relative to the transcription initiation site, are included. The positions of the FPs referred to in the text are also marked. Details of probes used (see panel F for details): A, probe 1; B, probe 2; C, probe 4; D and E, probe 5.

gave an 11-fold increase in CAT activity in MEL cells but not in C5 cells (Fig. 3B, line 2). Using the 1.8-kb *Eco*RI and 2.2-kb *Eco*RI-*Bam*HI fragments separately revealed that this effect is due mainly to sequences within the 1.8-kb *Eco*RI fragment containing DHSS II and III (which gives an eightfold increase [Fig. 3B, line 3]), rather than the 2.2-kb *Eco*RI-*Bam*HI fragment containing DHSS IV (which gives only a threefold increase [Fig. 3B, line 4]). By further experiments, the erythroid cell-specific enhancer was located more precisely to a 650-bp fragment containing DHSS II and III, which retained the full enhancer effect of the 1.8-kb fragment when tested in either orientation (only the natural orientation is shown in Fig. 3B, line 5). To define whether DHSS regions II, III, or both are required for full enhancer activity, two further PCR fragments containing either DHSS II or DHSS III were obtained and tested for enhancer activity by insertion 3′ to the CAT gene in pGPY-1.7 as before. The results (Fig. 3B, lines 6 and 7) show that the DHSS III region possesses most of the enhancer activity, although DHSS II has some activity.

Footprinting studies. To determine whether binding of erythroid cell-specific nuclear proteins to the enhancer are involved in its activity, in vitro footprinting studies were performed with various subfragments of the 650-bp fragment region suitable for identifying any protein-binding sites over the entire DHSS II and III region (Fig. 4F). By using probes 1 and 2 labelled on complementary strands of the DHSS II region (Fig. 4F), four FPs were found with MEL nuclear extracts but not with C5 extracts (Fig. 4A and B). The positions of these FPs are shown alongside the sequence of the DHSS II and III region in Fig. 5: two (FP1 and FP2) contain potential binding sites for the Ets-related family of transcription factors. There is also a further weak FP over

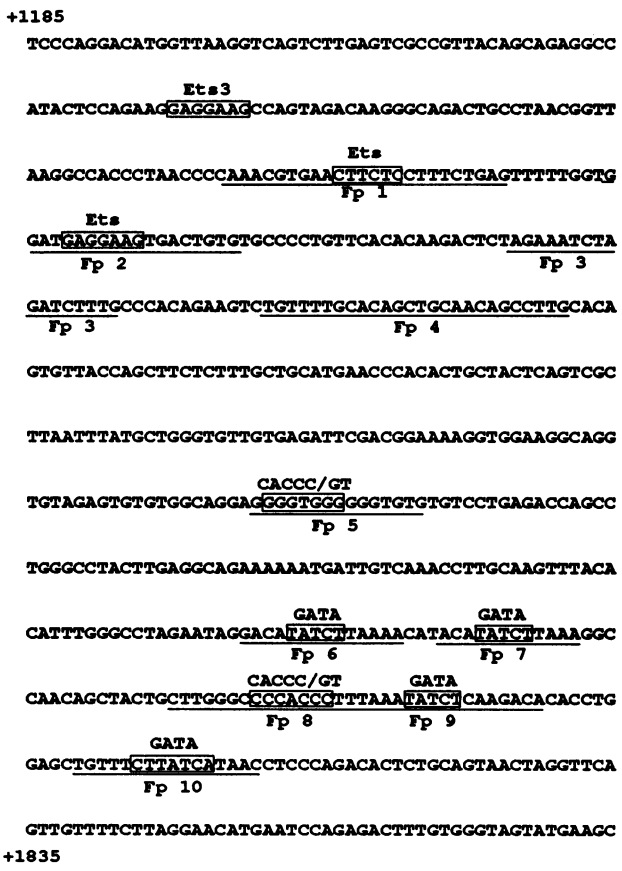


FIG. 5. Sequence of the DHSS II and III regions. The sequence of the 650-bp fragment from +1185 to +1835 bp containing the DHSS II and III regions (Fig. 3B) is given with the CACC motifs and potential binding sites for GATA-1 and Ets proteins boxed. The locations of the FPs identified in Fig. 4 are underlined.

another possible Ets binding site (marked ets3 in Fig. 5) (data not shown). In similar studies with probes from the DHSS III region labelled on each strand (probes 4 and 5, Fig. 4F), six further FPs were found with MEL extracts (FP5 to FP10, Fig. 4C to E [Fig. 4E is a longer run of the gel shown in Fig. 4D in order to resolve FP8 to FP10 better). In contrast, only FP5 and FP8 are found in C5 cells (Fig. 4C to E). The positions of FP5 to FP10 are shown with the sequence of the DHSS III region in Fig. 5. This reveals that FP5 and FP8 contain CACC/GT motifs which are known to bind various nuclear factors involved in gene regulation in erythroid and other contexts, whereas the erythroid cell-specific FP6, FP7, FP9, and FP10 contain potential binding sites for the GATA family of transcription factors. In view of the tissue specificity of the FPs in these regions and the fact that they contain consensus binding sequences for GATA-1 (WGATAR), we assume that this is the GATA family member binding to these sites. Finally, footprinting experiments with probe 3 (Fig. 4F) using both MEL and C5 extracts failed to reveal any further FPs between DHSS II and III (data not shown).

Site-directed mutagenesis. To test directly the roles of the various nuclear protein binding sequences within the erythroid cell-specific enhancer, sequences within the various FPs in the DHSS II and III regions were mutated individually and in various combinations, and their effects on enhancer function were tested. The mutations introduced are summarized in Fig. 6 together with the assays of enhancer function. Both the DHSS III and II regions have enhancer activity when tested in the opposite orientation (Fig. 6A, line 2, and Fig. 6B, line 2, respectively), although with DHSS III enhancer activity is reduced to about 50% in the opposite orientation. As far as the DHSS III region is concerned, introducing mutations known to prevent specific nuclear protein binding (52) into the GATA sequence in FP10 (Fig. 6A, line 8) substantially reduces the enhancer activity in MEL cells, as does, to a lesser extent, mutation of the GATA motifs in FP6 and FP7 or FP9 (Fig. 6A, lines 5 and 7, respectively). Deleting the CACC motif in FP5 (Fig. 6A, line 4) or mutating the CACC motifs in FP5 or FP8 to GGGCC (Fig. 6A, lines 3 and 6, respectively), which was shown by EMSA experiments to abolish protein binding (data not shown), also reduces the enhancer activity by about 50%. Indeed, enhancer function is completely abolished by the combination of the three FP8 to FP10 mutations (Fig. 6A, line 10). Overall, therefore the results show that the enhancer activity of the DHSS III region in MEL cells is due to the cluster of GATA and CACC/GT motifs, particularly those in the FP8 to FP10 region.

Similar detailed deletion-mutagenesis studies of the DHSS II region (Fig. 6B) show that the putative Ets-binding sites in FP1 and FP2 are important for its enhancer activity in MEL cells (Fig. 6B, lines 3 and 4, respectively), with FP3 also having a small role (Fig. 6B, line 5). In contrast, deletion of FP4 has no detectable effect on enhancer activity, at least in transient transfection assays of this type (Fig. 6B, line 6).

The nature of the nuclear proteins binding to the DHSS II region. To obtain further information about the tissue specificity of the nuclear factor(s) binding to FP1 to -4, EMSA experiments were performed with nuclear extracts from a variety of cell lines. Control EMSA experiments confirmed that all the nuclear extracts contained approximately equivalent amounts of an unrelated transcription factor, the CCAAT box binding factor. A probe from the putative Ets-like binding site in FP1 gave a single major retarded complex with extracts from several erythroid cell lines (MEL, K562, HEL, and, to a lesser extent, KMOE) as well as normal 14-day fetal liver (which comprises about 80% erythroid cells at this stage of development [24]), but not with C5 or HeLa epithelial cells or with the mouse fibroblast line, STO (data not shown). Competition experiments (Fig. 7A and B) show that the major retarded complex found with the *FP1* probe and MEL or fetal liver extracts is specific, since it is inhibited by unlabelled *FP1* oligonucleotide itself. It is also inhibited by an oligonucleotide from the *FP2* and *ets3* regions as well as oligonucleotides containing the binding sites recognized by some Ets family members, such as the *PU1* oligonucleotide, which binds PU1 specifically (42–44) and the *E74* oligonucleotide, recognized by Fli-1 (45) and Erg (56). However, the *PEA3* oligonucleotide, which is recognized by Ets-1, Ets-2, and Elk-1 (32, 55, 70, 73) competes less efficiently (Fig. 7A and B). The specificity of these competition experiments is also demonstrated by the fact that a mutated *E74* oligonucleotide (*E74m*), which does not bind Ets proteins (45), fails to compete with *FP1* (Fig. 7A and B). These EMSA experiments strongly suggest that the Ets protein binding to *FP1* is PU1. Although *FP1* does not contain a perfect Ets binding site consensus sequence, an identical sequence involved in the regulation of the CD11b promoter has been shown to bind PU1 (50).

In contrast, the Ets-like binding site in *FP2* gives two retarded complexes with fetal liver extracts (Fig. 7D): one

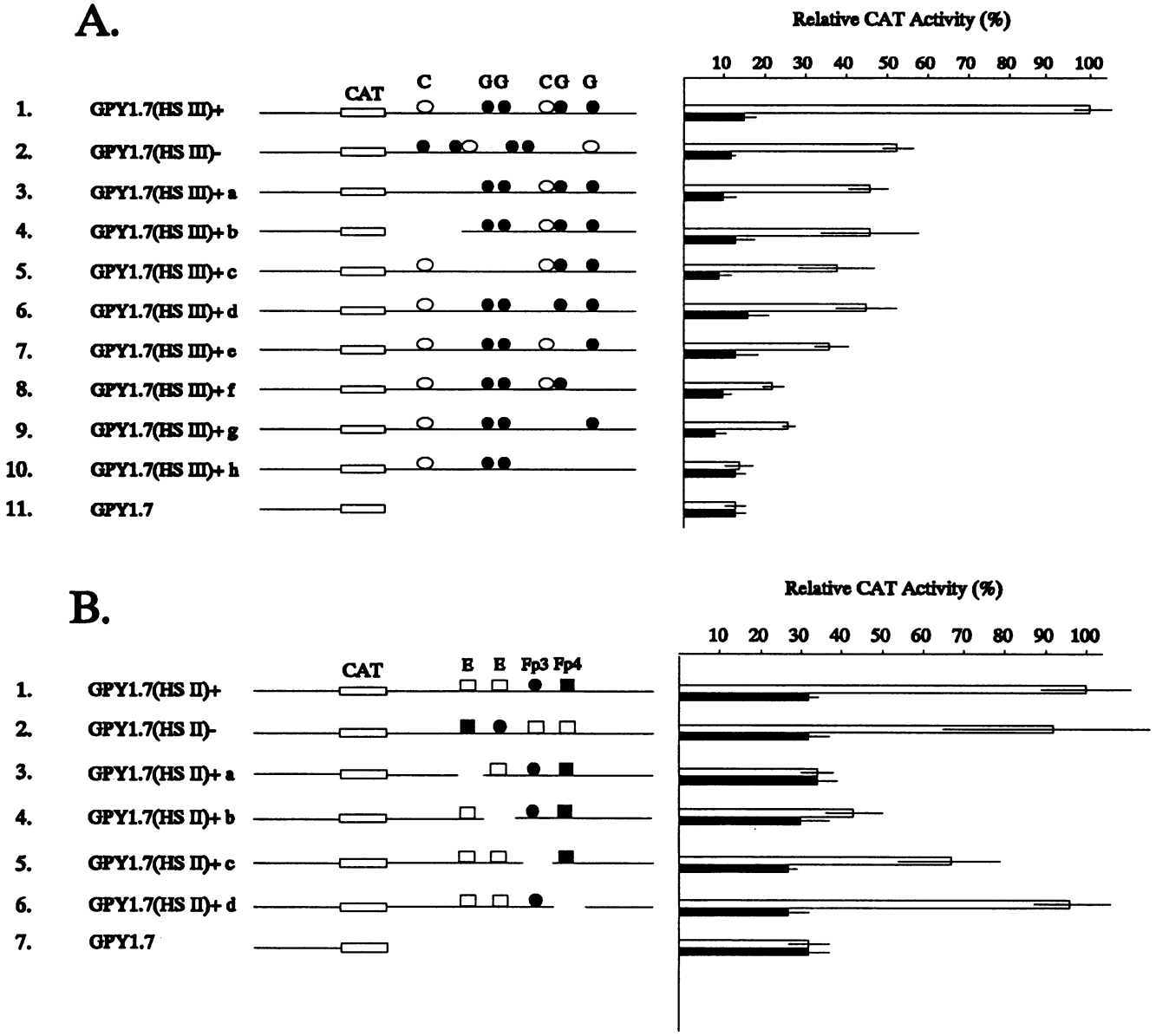


FIG. 6. Site-directed mutagenesis of the erythroid cell-specific enhancer. Mutations were introduced into the DHSS III (A) and DHSS II (B) regions and then tested for enhancer activity when placed 3′ to the CAT gene under the control of the 1.7-kb cGSHPX promoter fragment in GPY1.7. The details of the precise mutations and how they were constructed are given in Materials and Methods. Constructs were tested by transient transfection into MEL cells (open boxes on histograms) or C5 cells (solid boxes), and CAT activities are expressed relative to the value for GPY1.7. Error bars give the standard errors of the results from at least six transfections. (A) C (open ovals), CACC motif; G (filled circles), GATA binding site; (B) E (open rectangles), Ets binding site; Fp3 (filled circle) and Fp4 (filled square), FP3 and FP4 regions, respectively.

migrates at the same position as the complex obtained with the *FP1* oligonucleotide and is inhibited effectively by the *FP1* oligonucleotide; the other, more slowly migrating complex, which is less obvious with MEL extracts (Fig. 7C), is not inhibited by the *FP1* or *PU1* oligonucleotides, whereas it is inhibited by *PEA3* and *E74* oligonucleotides (Fig. 7D). Both complexes are inhibited by the *ets3* oligonucleotide (Fig. 7D). We conclude, therefore, that normal fetal liver contains Ets proteins capable of binding to the *FP1* and *FP2* regions of the cGSHPX enhancer, one of which is almost certainly PU1.

Similar EMSA experiments performed with *FP3* probes confirmed our previous footprinting data indicating binding with extracts from two erythroid cell lines, MEL and K562 (and at lower levels with HEL extracts), but not with the non-erythroid cell extracts tested (Fig. 8A). Although the sequence of *FP3* showed no obvious similarity to the consensus binding sequence for the erythroid/megakaryocyte-specific transcription factor GATA-1 (WGATAR), in fact competition experiments demonstrated that the *FP3* retarded complex was effectively inhibited by an oligonucleotide containing a GATA-1 binding site from the mouse α-globin promoter, but not by one containing a mutated GATA-1 binding site or an unrelated AP1 binding site (Fig.

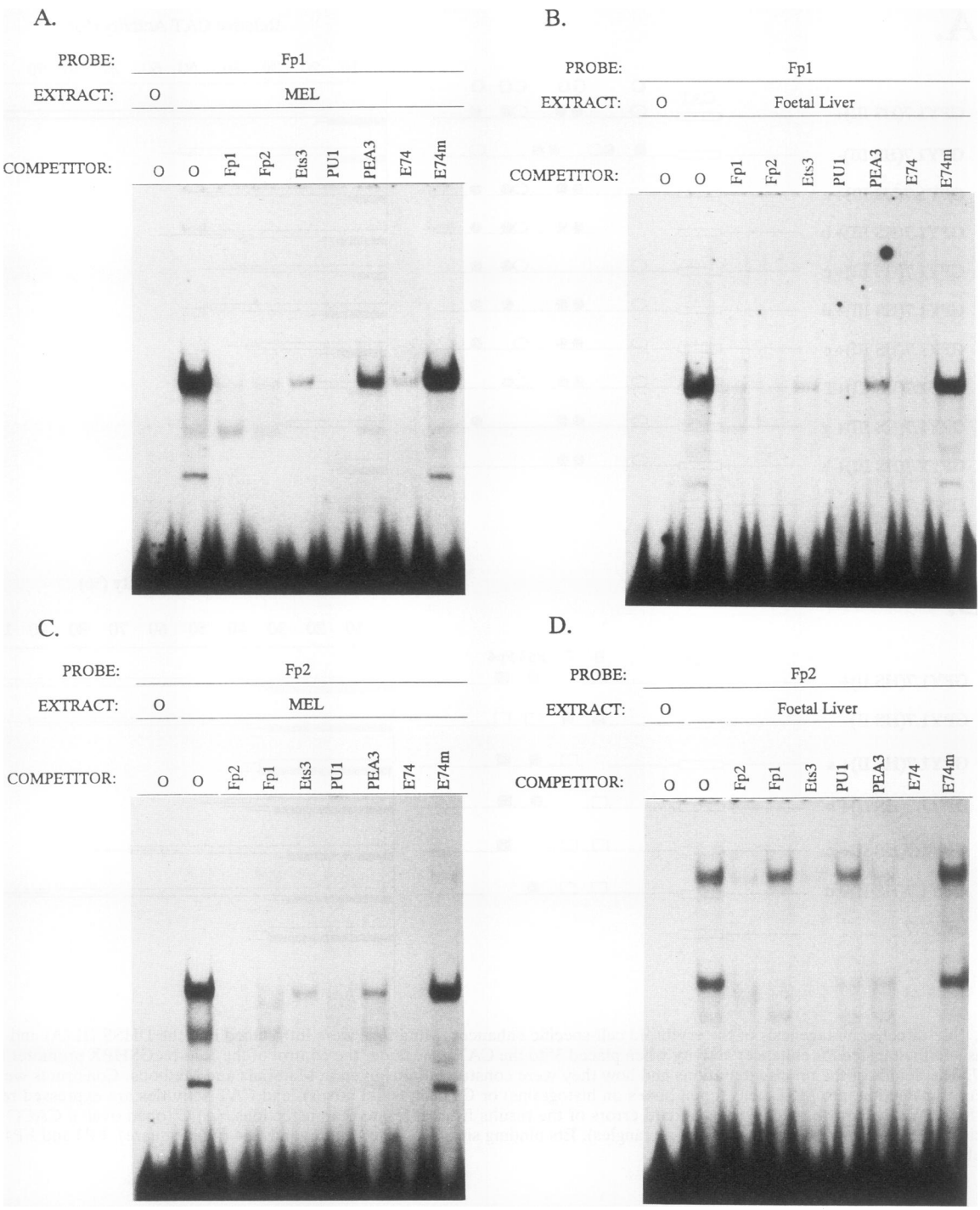


FIG. 7. Nuclear proteins binding to the FP1 and FP2 regions. Labelled FP1 (A and B) or FP2 (C and D) probes were incubated with MEL (A and C) or fetal liver (B and D) nuclear extracts either alone or in the presence of a 200-fold excess of various unlabelled competitors: FP1, FP2, and Ets3, oligonucleotides from the DHSS II region (Fig. 5); PU1, PEA3, and E74, oligonucleotides known to bind different Ets proteins (see text for details); E74m, a mutated version of E74 that has lost its ability to bind Ets proteins. The sequences of the various oligonucleotides are given in Materials and Methods together with the details of how the EMSA experiments were performed.

8B). Reciprocal experiments showed that unlabelled *FP3* oligonucleotide competed efficiently with a labelled GATA-1-binding oligonucleotide (Fig. 8B). There is little doubt, therefore, that the erythroid cell-specific *FP3* is attributable to binding of a member of the GATA family of transcription factors. However, the position of migration of the retarded complexes with the FP3 and GATA-1 probes are not exactly the same (Fig. 8B), which might suggest that the factor finding to FP3 is not be GATA-1. Information on the DNA-binding specificities of GATA family members has

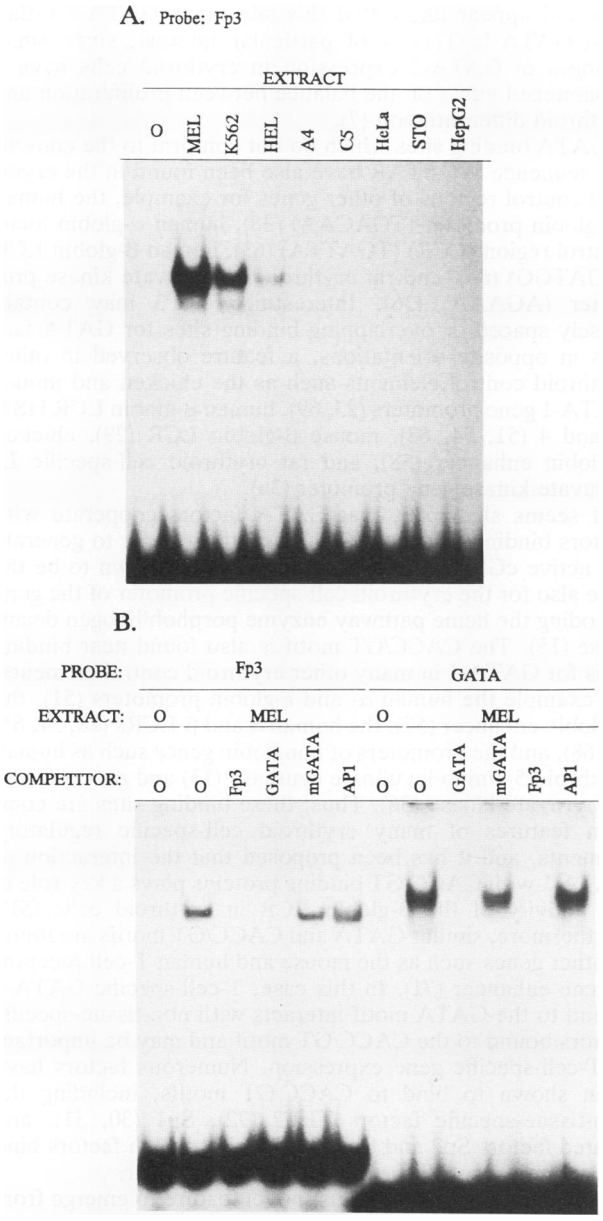


FIG. 8. Nuclear proteins binding to the FP3 region. (A) A labelled oligonucleotide from the FP3 region was analyzed by EMSA with various erythroid (MEL, K562, and HEL) and non-erythroid cell lines (see Materials and Methods for details). (B) A labelled FP3 oligonucleotide or a GATA binding probe (AG ATAAG) from the mouse α-globin gene were incubated with MEL nuclear extracts either alone or in the presence of a 200-fold excess of unlabelled competitor oligonucleotide and subjected to EMSA.

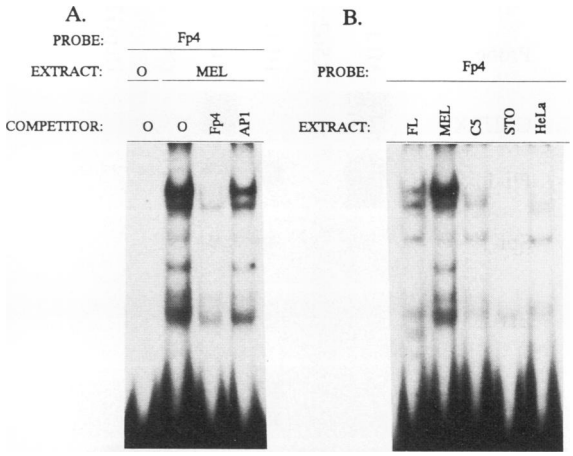


FIG. 9. Nuclear proteins binding to the FP4 region. A labelled oligonucleotide from the FP4 region was incubated with nuclear extracts from MEL cells (A) or other cell types either alone (B) or in the presence of 200-fold excess of unlabelled FP4 oligonucleotide or an AP1-binding oligonucleotide (A) and then analyzed by EMSA. The sequences of the oligonucleotides used are given in Materials and Methods.

recently been published (33, 41); intriguingly, the sequence of the *FP3* region contains a 12-nucleotide match to an oligonucleotide found to bind GATA-2 and GATA-3 rather than GATA-1 (oligonucleotide CC28 [33]).

Finally, EMSA experiments with oligonucleotides from the FP4 region (Fig. 9B) show a series of retarded complexes with MEL and fetal liver nuclear extracts, with the highest-molecular-weight complex seemingly erythroid cell specific. Competition experiments (Fig. 9A) show this complex to be specific, but the nature of the proteins binding to its remains unknown. However, this region of the GSHPX enhancer seems not to be functionally important in transient transfection experiments (Fig. 6B).

***ets* mRNAs expressed in erythroid cells.** To determine which members of the *ets* gene family are expressed in the various erythroid cell lines and might therefore be involved in the role of the FP1 and FP2 regions in erythroid cell-specific cGSHPX enhancer activity, total cellular RNAs from MEL, K562, HEL, and normal fetal liver were analyzed by Northern (RNA) transfer experiments for the presence of specific *ets* mRNAs, as well as cGSHPX mRNA and 7S RNA (as a loading control) (Fig. 10). *erg* mRNA was not detectable in any of the RNAs analyzed (data not shown). However, although none of the other four *ets* mRNAs examined is expressed in C5 or liver-derived HepG2 cells, both fetal liver and MEL cells express significant amounts of *Fli-1* mRNA and *Spi-1* mRNA (which encodes PU1), but little or no *ets-1* or *ets-2* mRNAs, whereas HEL cells express detectable amounts of all four *ets* mRNAs (Fig. 10). In contrast, K562 cells express only very low levels of the *ets-2* mRNA and possibly a trace of *Spi-1* (PU1) mRNA (Fig. 10). This result probably explains why the GPY1.7 (HSII)+ construct containing the DHSS II region (Fig. 3B)

Key to competitors: mGATA, mutated GATA sequence (AGAT GAG); AP1, AP1 binding probe. The sequences of the various oligonucleotides used are given in Materials and Methods. (C) Possible GATA binding sequences in the FP3 region.

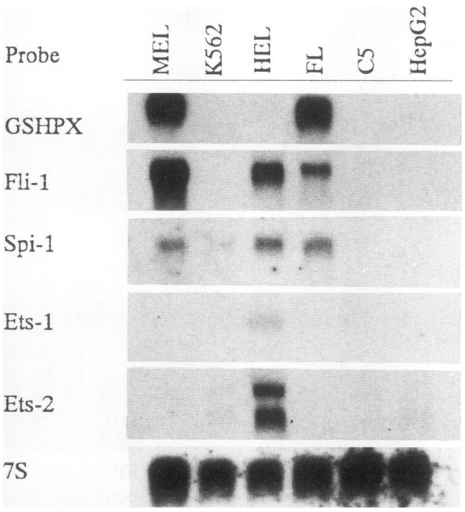


FIG. 10. *ets*-related mRNAs expressed in normal mouse fetal liver cells and various erythroid and nonerythroid cell lines. Total cellular RNAs were electrophoresed in agarose-formaldehyde gels, transferred to nylon filters, and then hydridized with a GSHPX gene probe, a variety of *ets* cDNA probes, or a 7S RNA probe as a loading control.

does not possess enhancer function after transfection into K562 cells (data not shown). Indeed, it seems significant that the only erythroid cell line that expresses high levels of cGSHPX mRNA is the one that also contains high levels of both GATA-1 (Fig. 8A) and PU1 or Fli-1 (Fig. 10).

DISCUSSION

These experiments show that the up-regulation of cGSHPX mRNA expression in erythroid (MEL) cells is controlled at the level of transcription, as evidenced by nuclear run-on experiments and the presence of tissue-specific DHSS 3′ to the gene. The cGSHPX promoter itself seems to function in a non-tissue-specific manner when tested in transient transfection experiments, although there is some suggestion that upstream sequences, between −0.289 and −1.7 kb, may confer a small erythroid cell-specific increase in promoter activity. However the most important conclusion from our studies is that the up-regulation of cGSHPX gene transcription in erythroid cells is mediated primarily by a tissue-specific enhancer 3′ to the gene located in a region that shows two DHSS (DHSS II and III) in the chromatin of erythroid cells.

Detailed analysis of the enhancer in terms of nuclear protein binding sites and site-directed mutagenesis has revealed some novel features as well as similarities with the control elements regulating the erythroid cell-specific expression of other genes. The 3′ region of the cGSHPX gene enhancer contains a cluster of four binding sites for the GATA family of transcription factors and two sites for CACC/GT factors. The function of these GATA sites is most likely attributable to GATA-1, which we and others have implicated in the functioning of erythroid cell-specific promoters and enhancers regulating a variety of globin and nonglobin genes (15, 52) (reviewed in reference 49). In addition, there is another GATA binding site of unusual sequence further upstream, and from recent data identifying the DNA-binding specificities of different family members by the method of PCR-mediated random site selection (33, 41), it would appear likely that this site binds GATA-2 rather than GATA-1. This is of particular interest, since small changes in GATA-2 expression in erythroid cells have a pronounced effect on the balance between proliferation and erythroid differentiation (7).

GATA binding sites which do not conform to the consensus sequence WGATAR have also been found in the erythroid control regions of other genes for example, the human $^{A}\gamma$-globin promoter (TGACAA) (38), human α-globin locus control region (LCR) (TGATTA) (63), human β-globin LCR (AGATGG) (64), and rat erythroid L′ pyruvate kinase promoter (AGAAA) (36). Interestingly, FP3 may contain closely spaced or overlapping binding sites for GATA factors in opposite orientations, a feature observed in other erythroid control elements such as the chicken and mouse GATA-1 gene promoters (23, 69), human β-globin LCR HS2, 3, and 4 (51, 54, 68), mouse β-globin LCR (29), chicken β-globin enhancer (58), and rat erythroid cell-specific L′ pyruvate kinase gene promoter (36).

It seems significant that GATA factors cooperate with factors binding to the CACC/GT motif in order to generate the active cGSHPX enhancer, as we have shown to be the case also for the erythroid cell-specific promoter of the gene encoding the heme pathway enzyme porphobilinogen deaminase (15). The CACC/GT motif is also found near binding sites for GATA-1 in many other erythroid control elements, for example the human α- and ε-globin promoters (51), the β-globin enhancer (57), the human α and β LCRs (28, 54, 63, 64, 68), and the promoters of nonglobin genes such as human erythroid 5-amino levulinate synthase (13) and rat erythroid L′ pyruvate kinase (36). Thus, these binding sites are common features of many erythroid cell-specific regulatory elements, and it has been proposed that the interaction of GATA-1 with CACC/GT binding proteins plays a key role in the activity of the β-globin LCR in erythroid cells (51). Furthermore, similar GATA and CACC/GT motifs are found in other genes such as the mouse and human T-cell receptor α gene enhancer (71). In this case, T-cell-specific GATA-3 bound to the GATA motif interacts with non-tissue-specific factors bound to the CACC/GT motif and may be important in T-cell-specific gene expression. Numerous factors have been shown to bind to CACC/GT motifs, including the non-tissue-specific factors TEF2 (72), Sp1 (30, 51), and related factors Sp2 and Sp3 (22, 31), but which factors bind to erythroid control regions in vivo is not clear.

However, perhaps the most novel feature to emerge from these studies is the identification of a cluster of two, or possibly three, functionally important Ets binding sites just upstream of the GATA/CACC cluster, within the DHSS II region. The role of these Ets-binding sites has, of necessity, been established only in erythroid cell lines, some of which have aberrant activation of *ets* genes as a result of retroviral transformation (5, 42, 44, 62). Nevertheless, we presume that they are functionally important in normal erythroid development in view of their location next to the GATA-CACC/GT cluster. Which members of the Ets family are involved in the regulation of cGSHPX in normal erythroid cells via binding to these sites is not certain. However, likely candidates are PU-1 and Fli-1, since they are expressed in normal fetal liver and MEL cells. The involvement of Ets family members in the regulation of erythroid genes has not previously been reported. However, there are potential Ets binding sites in the mouse and human GATA-1 gene promoters (46, 69) and in the DNase I hypersensitive region within intron 2 of the mouse β-globin gene (19). In the latter case, a factor binds to an Ets consensus sequence, and its tissue

distribution suggests that this factor is PU-1 (18). This site is closely associated with a binding site for GATA-1. It is possible, therefore, that Ets family members are involved in the regulation of other genes in erythroid cells. Recently, it has been shown that Ets proteins and GATA-1 appear to be involved in the regulation of megakaryocyte-specific gene expression (37). Interestingly, Ets-1 and Fli-1 were shown to transactivate the promoter of the megakaryocyte-specific gene GP IIb, whereas Ets-2 and PU-1 do not (37, 74). Thus, different Ets proteins seem to be involved in the regulation of tissue-specific expression in the various hematopoietic lineages. Since both erythroid cells and megakaryocytes express GATA-1 and Fli-1, other control mechanisms must be involved in distinguishing erythroid cell-versus megakaryocyte-specific patterns of gene expression, for example, other tissue-specific activators or repressors.

A question we have been unable to address is how cGSHPX mRNA is up-regulated in other highly expressing tissues, such as liver and kidney. The presence of a predominant 3′ DHSS (DHSS IV) in both liver and kidney cells might suggest that sequences in that region are responsible. However, this region has only a small (threefold) enhancer activity in erythroid cells, which also show DHSS IV. Our intention to test the DHSS IV region functionally by transfection experiments with liver or kidney cell lines has been thwarted by the fact that none of the 13 cell lines we have been able to test shows levels of cGSHPX mRNA comparable with those found in vivo (data not shown).

ACKNOWLEDGMENTS

This work has been supported by the Cancer Research Campaign and by Medical Research Council Training Awards to S.R. and I.C.

REFERENCES

1. **Affara, N., J. Fleming, P. S. Goldfarb, E. Black, B. Thiele, and P. R. Harrison.** 1985. Analysis of chromatin changes associated with the expression of globin and non-globin genes in cell hybrids between erythroid and other cells. Nucleic Acids Res. **13**:5629–5644.
2. **Affara, N., P. S. Goldfarb, P. S. Yang, and P. R. Harrison.** 1983. Patterns of expression of erythroblast non-globin mRNAs. Nucleic Acids Res. **11**:931–945.
3. **Andrews, N. C., and D. V. Faller.** 1991. A rapid micropreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. Nucleic Acids Res. **19**:2499.
4. **Becker, P., R. Renkawitz, and G. Schutz.** 1984. Tissue-specific DNase I hypersensitive sites in the 5′-flanking sequences of the tryptophan oxygenase and the tyrosine aminotransferase genes. EMBO J. **3**:2015–2020.
5. **Ben-David, Y., E. B. Giddens, K. Letwin, and A. Bernstein.** 1991. Erythroleukaemia induction by Friend murine leukaemia virus: insertional activation of a new member of the *ets* gene family, Fli-1, closely linked to c-ets-1. Genes Dev. **5**:908–918.
6. **Berry, M. J., L. Banu, and P. R. Larsen.** 1991. Type 1 iodothyronine deiodinase is a selenocysteine-containing enzyme. Nature (London) **349**:438–440.
7. **Briegel, K., K. C. Lim, C. Plank, H. Beug, J. D. Engel, and M. Zenke.** 1993. Ectopic expression of a conditional GATA-2/estrogen receptor chimera arrests erythroid differentiation in a hormone-dependent manner. Genes Dev. **7**:1097–1109.
8. **Chambers, I., J. Frampton, P. Goldfarb, N. Affara, W. McBain, and P. R. Harrison.** 1986. The structure of the mouse glutathione peroxidase gene: the selenocysteine in the active site is encoded by the 'termination' codon, TGA. EMBO J. **5**:1221–1227.
9. **Chiu, D., B. Lubin, and S. B. Shohet.** 1982. Peroxidative reactions in red cell biology, p. 115–160. *In* W. A. Pryor (ed.), Free radicals in biology, vol. V. Academic Press, Inc., New York.
10. **Chomczynski, P., and N. Sacchi.** 1987. Single step method of RNA isolation by guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162**:156–159.
11. **Chretien, S. A., D. Dubart, N. Beaupain, B. Raiche, J. Grandchamp, M. Rosa, M. Goossens, and P.-H. Romeo.** 1988. Alternative transcription and splicing of the human porphobilinogen deaminase gene result either in tissue-specific or in housekeeping expression. Proc. Natl. Acad. Sci. USA **85**:6–10.
12. **Chu, F.-F., J. H. Doroshow, and R. S. Esworthy.** 1993. Expression, characterization, and tissue distribution of a new cellular selenium-dependent glutathione peroxidase, GSHPx-GI. J. Biol. Chem. **268**:2571–2576.
13. **Cox, T., M. J. Bawden, A. Martin, and B. K. May.** 1991. Human erythroid 5-aminolevulinate synthase: promoter analysis and identification of an iron-responsive element in the mRNA. EMBO J. **10**:1891–1901.
14. **Flohe, L.** 1982. Glutathione peroxidase brought into focus, p. 223–254. *In* W. A. Pryor (ed.), Free radicals in biology, vol. V. Academic Press, Inc., New York.
15. **Frampton, J., M. Walker, M. Plumb, and P. R. Harrison.** 1990. Synergy between the NF-E1 erythroid-specific transcription factor and the CACC factor in the erythroid-specific promoter of the human porphobilinogen deaminase gene. Mol. Cell. Biol. **10**:3838–3842.
16. **Fritton, H. P., A. E. Sippel, and T. Igo-Kemenes.** 1983. Nuclease-hypersensitive sites in the chromatin domain of the chicken lysozyme gene. Nucleic Acids Res. **11**:3467–3485.
17. **Gaffney, D. F., J. McLauchlan, J. L. Whitton, and J. B. Clements.** 1985. A modular system for the assay of transcription regulatory signals: the sequence TAATGARAT is required for herpes simplex virus immediate early gene activation. Nucleic Acids Res. **13**:7847–7863.
18. **Galson, D. L., J. O. Hensold, T. R. Bishop, M. Schalling, A. D. D'Andrea, C. Jones, P. E. Auron, and D. E. Housman.** 1993. Mouse β-globin DNA-binding protein B1 is identical to a proto-oncogene, the transcription factor *Spi-1*/PU.1, and is restricted in expression to hematopoietic cells and the testis. Mol. Cell. Biol. **13**:2929–2941.
19. **Galson, D. L., and D. E. Housman.** 1988. Detection of two tissue-specific DNA-binding proteins with affinity for sites in the mouse β-globin intervening sequence 2. Mol. Cell. Biol. **8**:381–392.
20. **Garcia, G. E., and T. C. Stadtman.** 1991. Selenoprotein A component of the glycine reductase complex from *Clostridium purinolyticum*: nucleotide sequence of the gene shows that selenocysteine is encoded by UGA. J. Bacteriol. **173**:2093–2098.
21. **Goldfarb, P. S., J. O'Prey, N. Affara, Q.-S. Yang, and P. R. Harrison.** 1983. Isolation of non-globin genes expressed preferentially in mouse erythroid cells. Nucleic Acids Res. **11**:3517–3530.
22. **Hagen, G., S. Muller, M. Beato, and G. Suske.** 1992. Cloning by recognition site screening of two novel GT box binding proteins: a family of Sp1 related genes. Nucleic Acids Res. **20**:5519–5525.
23. **Hannon, R., T. Evans, G. Felsenfeld, and H. Gould.** 1991. Structure and promoter activity of the gene for the erythroid transcription factor GATA-1. Proc. Natl. Acad. Sci. USA **88**:3004–3008.
24. **Harrison, P. R., D. Conkie, and J. Paul.** 1973. Role of cell division and nucleic acid synthesis in erythropoietin-induced maturation of foetal liver cells *in vitro*, p. 341–364. *In* M. Balls and F. S. Billett (ed.), The cell cycle in development and differentiation. Cambridge University Press, Cambridge.
25. **Heyworth, C. M., T. M. Dexter, O. Kan, and A. D. Whetton.** 1990. The role of hemopoietic growth factors in self-renewal and differentiation of IL-3-dependent multipotential stem cells. Growth Factors **2**:197–211.
26. **Higuchi, R., B. Krummel, and R. K. Saiki.** 1988. A general method of *in vitro* preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. Nucleic Acids Res. **16**:7351–7367.
27. **Hill, K. E., R. S. Lloyd, J.-G. Yang, R. Read, and R. F. Burk.** 1991. The cDNA for rat selenoprotein P contains 10 TGA

codons in the open reading frame. J. Biol. Chem. **266**:10050–10053.
28. **Ikuta, T., and Y. W. Kan.** 1991. *In vivo* protein-DNA interactions at the β-globin gene locus. Proc. Natl. Acad. Sci. USA **88**:10188–10192.
29. **Jiminez, G., K. B. Gale, and T. Enver.** 1992. The mouse β-globin locus control region: hypersensitive sites 3 and 4. Nucleic Acids Res. **20**:5797–5803.
30. **Kadonaga, J. T., K. R. Carner, F. R. Masiarz, and R. Tijan.** 1987. Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. Cell **51**:1079–1090.
31. **Kingsley, C., and A. Winoto.** 1992. Cloning of GT box-binding proteins: a novel Sp1 multigene family regulating T-cell receptor gene expression. Mol. Cell. Biol. **12**:4251–4261.
32. **Klemsz, M. J., S. R. McKercher, A. Celada, C. Van Beveren, and R. A. Maki.** 1990. The macrophage and B cell-specific transcription factor PU.1 is related to the *ets* oncogene. Cell **61**:113–124.
33. **Ko, L. J., and J. D. Engel.** 1993. DNA-binding specificities of the GATA transcription factor family. Mol. Cell. Biol. **13**:4011–4022.
34. **Krummel, B.** 1990. DNAseI Footprinting, p. 184–188. *In* PCR protocols: a guide to methods and applications. Academic Press, Inc., New York.
35. **Kulesz-Martin, M., A. E. Kilkenny, K. A. Holbrook, V. Digernes, and S. H. Yuspa.** 1983. Properties of carcinogen altered mouse epidermal cells resistant to calcium induced terminal differentiation. Carcinogenesis **4**:1367–1377.
36. **Lacronique, V., D. Boquet, S. Lopez, A. Kahn, and M. Raymondjean.** 1992. *In vitro* and *in vivo* protein-DNA interactions on the rat erythroid-specific L' pyruvate kinase gene promoter. Nucleic Acids Res. **20**:5669–5676.
37. **Lemarchandel, V., J. Ghysdael, V. Mignotte, C. Rahuel, and P.-H. Romeo.** 1993. GATA and Ets *cis*-acting sequences mediate megakaryocyte-specific expression. Mol. Cell. Biol. **13**:668–676.
38. **Mantovani, R., G. S. Furga, J. Gilman, and S. Ottolenghi.** 1989. The deletion of the distal CCAAT box region of the γ^A-globin gene in black HPFH abolishes the binding of the erythroid-specific protein NFE-3 and of the CCAAT displacement protein. Nucleic Acids Res. **17**:6681–6691.
39. **Martin, P., and T. Papayannopoulou.** 1982. HEL cells: a new human erythroleukaemia cell line with spontaneous and induced globin expression. Science **216**:1233–1235.
40. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
41. **Merika, M., and S. H. Orkin.** 1993. DNA-binding specificity of GATA family transcription factors. Mol. Cell. Biol. **13**:3999–4010.
42. **Moreau-Gachelin, F., D. Ray, N. J. deBoth, M. J. M. van der Feltz, P. Tambourin, and A. Tavitan.** 1990. Spi-1 oncogene activation in Rauscher and Friend murine virus-induced acute erythroleukaemias. Leukaemia **4**:20–23.
43. **Moreau-Gachelin, F., D. Ray, P. Tambourin, A. Tavitan, M. J. Klemsz, S. R. McKercher, A. Celada, C. Van Beveren, and R. A. Maki.** 1990. The PU.1 transcription factor is the product of the putative oncogene Spi-1. Cell **61**:1166.
44. **Moreau-Gachelin, F., A. Tavitan, and P. Tambourin.** 1988. Spi-1 is a putative oncogene in virally induced murine erythroleukaemias. Nature (London) **331**:277–280.
45. **Nibbs, R. J. B., K. Itoh, W. Ostertag, and P. R. Harrison.** 1993. Differentiation arrest and stromal cell-independent growth of murine erythroleukemia cells are associated with elevated expression of *ets*-related genes but not with mutation of *p53*. Mol. Cell. Biol. **13**:5582–5592.
46. **Nicolis, S., C. Bertini, A. Ronchi, S. Crotta, L. Lanfranco, E. Moroni, B. Giglioni, and S. Ottolenghi.** 1991. An erythroid specific enhancer upstream of the gene encoding the cell-type specific transcription factor GATA-1. Nucleic Acids Res. **19**:5285–5291.
47. **Okano, H., J. Okamura, K. Yagawa, H. Tasaka, and S. Motomura.** 1981. Human erythroid cell lines derived from a patient with acute erythremia. J. Cancer Res. Clin. Oncol. **102**:49–55.
48. **O'Prey, J., J. Chester, B. J. Thiele, S. Janetzki, S. Prehn, J. Fleming, and P. R. Harrison.** 1989. The promoter structure and complete sequence of the gene encoding the rabbit erythroid cell-specific 15-lipoxygenase. Gene **84**:493–499.
49. **Orkin, S. H.** 1990. Globin gene regulation and switching: circa 1990. Cell **63**:665–672.
50. **Pahl, H. L., R. J. Scheibe, D.-E. Zhang, H.-M. Chen, D. Galson, R. A. Maki, and D. G. Tenen.** 1993. The proto-oncogene PU.1 regulates expression of the myeloid-specific CD11b promoter. J. Biol. Chem. **268**:5014–5020.
51. **Philipsen, S., D. Talbot, P. Fraser, and F. Grosveld.** 1990. The β-globin dominant control region: hypersensitive site 2. EMBO J. **9**:2159–2167.
52. **Plumb, M., J. H. Frampton, M. Wainwright, M. Walker, K. Macleod, G. Goodwin, and P. Harrison.** 1989. GATAAG: a cis-control region binding an erythroid-specific nuclear factor with a role in globin and non-globin gene expression. Nucleic Acids Res. **17**:73–92.
53. **Plumb, M. A., and G. H. Goodwin.** 1988. Detection of sequence-specific protein-DNA interactions by the DNA footprinting technique. *In* J. M. Walker (ed.), Methods in molecular biology, vol. 4. Humana Press, Clifton, N.J.
54. **Pruzina, S., O. Hanscombe, D. Whyatt, F. Grosveld, and S. Philipsen.** 1991. Hypersensitive site 4 of the human β-globin locus control region. Nucleic Acids Res. **19**:1413–1419.
55. **Rao, V. N., and E. S. P. Reddy.** 1992. A divergent Ets-related protein, Elk-1, recognizes similar c-*ets*-1 proto-oncogene target sequences and acts as a transcriptional activator. Oncogene **7**:65–70.
56. **Reddy, E. S. P., and V. N. Rao.** 1991. erg, an *ets*-related gene, codes for sequence-specific transcriptional activators. Oncogene **6**:2285–2289.
57. **Reddy, P. M. S., and C. J. Shen.** 1991. Protein-DNA interactions *in vivo* of an erythroid-specific human β-globin locus enhancer. Proc. Natl. Acad. Sci. USA **88**:8676–8680.
58. **Reitman, M., and G. Felsenfeld.** 1988. Mutational analysis of the chicken β-globin enhancer reveals two positively acting domains. Proc. Natl. Acad. Sci. USA **85**:6267–6271.
59. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
60. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
61. **Schuckelt, R., R. Brigelius-Flohe, M. Maiorino, A. Roveri, J. Reumkens, W. Strassburger, F. Ursini, B. Wolf, and L. Flohe.** 1991. Phospholipid hydroperoxide glutathione peroxidase is a seleno-enzyme distinct from the classical glutathione peroxidase as evident from cDNA and amino acid sequencing. Free Radical Res. Commun. **14**:343–361.
62. **Sels, F. T., S. Langer, A. S. Schulz, J. Silver, M. Sitbon, and R. W. Friedrich.** 1992. Friend murine leukaemia virus is integrated at a common site in most primary spleen tumors of erythroleukaemic animals. Oncogene **7**:643–652.
63. **Strauss, E. C., N. C. Andrews, D. R. Higgs, and S. H. Orkin.** 1992. In vivo footprinting of the human α-globin locus upstream regulatory element by guanine and adenine ligation-mediated polymerase chain reaction. Mol. Cell. Biol. **12**:2135–2142.
64. **Strauss, E. C., and S. H. Orkin.** 1992. *In vivo* protein-DNA interactions at hypersensitive site 3 of the human β-globin locus control region. Proc. Natl. Acad. Sci. USA **89**:5809–5813.
65. **Suggs, S. V., R. B. Wallace, T. Hirose, E. H. Kawashima, and K. Itakura.** 1981. Use of cloned synthetic oligonucleotides as hybridisation probes: isolation of cloned cDNA sequences for human β_2-microglobulin. Proc. Natl. Acad. Sci. USA **78**:6613–6617.
66. **Sunde, R. A.** 1990. Molecular biology of selenoproteins. Annu. Rev. Nutr. **10**:451–474.
67. **Takahashi, K., M. Akasaka, Y. Yamamoto, C. Kobayashi, J. Mizoguchi, and J. Koyama.** 1990. Primary structure of human

plasma glutathione peroxidase deduced from cDNA sequences. J. Biochem. **108**:145–148.

68. **Talbot, D., S. Philipsen, P. Fraser, and F. Grosveld.** 1990. Detailed analysis of the site 3 region of the human β-globin dominant control region. EMBO J. **9**:2169–2178.
69. **Tsai, S.-F., E. Strauss, and S. H. Orkin.** 1991. Functional analysis and *in vivo* footprinting implicate the erythroid transcription factor GATA-1 as a positive regulator of its own promoter. Genes Dev. **5**:919–931.
70. **Wasylyk, B., C. Wasylyk, P. Flores, A. Begue, D. Leprince, and D. Stehelin.** 1990. The c-*ets* proto-oncogenes encode transcription factors that cooperate with c-fos and c-jun for transcriptional activation. Nature (London) **346**:191–193.
71. **Winoto, A., and D. Baltimore.** 1989. A novel and inducible T cell specific enhancer located at the 3′ end of the T cell receptor α-locus. EMBO J. **8**:729–733.
72. **Xiao, J. H., I. Davidson, M. Macchi, R. Rosales, M. Vigneron, A. Staub, and P. Chambon.** 1987. *In vitro* binding of several cell-specific and ubiquitous nuclear proteins to the GT-1 motif of the SV40 enhancer. Genes Dev. **1**:794–807.
73. **Xin, J.-H., A. Cowie, P. Lachance, and J. A. Hassell.** 1992. Molecular cloning and characterisation of PEA3, a new member of the *ets* oncogene family that is differentially expressed in mouse embryonic cells. Genes Dev. **6**:481–496.
74. **Zhang, L., V. Lemarchandel, P. H. Romeo, Y. Ben-David, P. Greer, and A. Bernstein.** 1993. The *Fli-1* proto-oncogene, involved in erythroleukemia and Ewing's sarcoma, encodes a transcriptional activator with DNA-binding specificities distinct from other Ets family members. Oncogene **8**:1621–1630.
75. **Zinoni, F., A. Birkmann, T. C. Stadtman, and A. Bock.** 1986. Nucleotide sequence and expression of the selenocysteine-containing polypeptide of formate dehydrogenase (formate-hydrogen-lyase-linked) from *Escherichia coli*. Proc. Natl. Acad. Sci. USA **83**:4650–4654.
76. **Zinoni, F., J. Heider, and A. Bock.** 1990. Features of the formate dehydrogenase mRNA necessary for decoding of the UGA codon as selenocysteine. Proc. Natl. Acad. Sci. USA **87**:4660–4664.