MOLECULAR AND CELLULAR BIOLOGY, June 1996, p. 3125–3137 Vol. 16, No. 6
0270-7306/96/$04.00+0

The Activity of the Highly Inducible Mouse Phenylalanine Hydroxylase Gene Promoter Is Dependent upon a Tissue-Specific, Hormone-Inducible Enhancer

DANIELA M. FAUST, ANNE-MARIE CATHERIN, SANDRINE BARBAUX,[1]† LAURENT BELKADI, TEREZA IMAIZUMI-SCHERRER, AND MARY C. WEISS*

Unité de Génétique de la Différenciation, Unité de Recherche Associée Centre National de la Recherche Scientifique 1149, Département de Biologie Moléculaire, Institut Pasteur, Paris, France

Received 29 February 1996/Accepted 27 March 1996

Expression of the phenylalanine hydroxylase gene in livers and kidneys of rodents is activated at birth and is induced by glucocorticoids and cyclic AMP in the liver. Regulatory elements in a 10-kb fragment upstream of the mouse gene have been characterized. The promoter lacks TAATA and CCAAT consensus sequences and shows only extremely weak activity in transitory expression assays with phenylalanine hydroxylase-producing hepatoma cells. No key elements for regulation of promoter activity are localized within 2 kb of upstream sequences. However, a liver-specific DNase I-hypersensitive site at kb −3.5 comprises a tissue-specific and hormone-inducible enhancer. This enhancer contains multiple protein binding sites, including sites for ubiquitous factors (NF1 and AP1), the glucocorticoid receptor, and the hepatocyte-enriched transcription factors hepatocyte nuclear factor 1 (HNF1) and C/EBP. Mutation revealed that the last two sites are critical not only for basal activity but also for obtaining a maximal hormone response. Efficient transcription from the highly inducible promoter shows absolute dependence upon the enhancer at kb −3.5, which in turn requires HNF1 and C/EBP as well as hormones. The regulatory region of the mouse phenylalanine hydroxylase gene differs totally from that of humans, even though the genes of both species are expressed essentially in the liver. Furthermore, the phenylalanine hydroxylase gene of mice shows an expression pattern very similar to those of the rodent tyrosine aminotransferase and phosphoenolpyruvate carboxykinase genes, yet each shows a different organization of its regulatory region.

The gene for phenylalanine hydroxylase (PAH) (EC 1.14.16.1) encodes the enzyme responsible for the conversion of phenylalanine to tyrosine (29). The enzymatic activity requires tetrahydrobiopterin as a cofactor, which in turn is synthesized and regenerated by multiple enzyme-catalyzed steps (14, 35). Mutations in the human PAH gene resulting in a deficiency of enzymatic activity are responsible for the inherited metabolic disease of either phenylketonuria or nonphenylketonuria hyperphenylalaninemia, depending on the serum phenylalanine levels accumulated (40); so far, these mutations have all been found within the coding portion of the gene (20). In humans, PAH is expressed only in the liver, and it is first detected during the first trimester of fetal development (references cited in reference 49). No information concerning hormone inducibility of the human gene is available.

In rodents, PAH is present mainly in the liver but also is found in the kidneys; it is first expressed only after birth, and the gene is induced by glucocorticoids and cyclic AMP (cAMP) (26, 27, 34, 44, 45, 51). The hormone inducibility, distribution in tissues, late expression, and negative regulation by TSE1 (21) permit grouping of the rodent PAH with the neonatal hormone-inducible tyrosine aminotransferase (TAT) and phosphoenolpyruvate carboxykinase (PEPCK) genes. The regulatory elements of the last two genes from rats (28, 37) have been shown to possess different types of modular organization. Characterization of the regulatory region of a rodent PAH gene would complete the scheme of organization of this class of genes and provide a possibly informative comparison with the human PAH gene.

The human gene was cloned many years ago (18), but the regulation of its expression has been only incompletely defined. A 9-kb fragment upstream of the human gene is sufficient for controlling late onset of liver- and kidney-specific reporter gene expression in transgenic mice (49). Hormone responsiveness of the transgene has not been studied. Human PAH transcripts initiate at several sites, and the first 300 bp of the promoter is devoid of TAATA and CCAAT box consensus sequences (30). Expression assays of the human PAH promoter and upstream sequences revealed significant activity only in the incompletely characterized HepG2 and PLC (Alexander) human hepatoma cells. Full activity was retained in the proximal promoter sequences; however, further analysis revealed only binding sites for unidentified proteins that show ubiquitous expression (50). Moreover, these sequences activate a heterologous promoter in transient-expression assays, but not in a cell-type-specific manner. Hence, the basis for liver-specific expression of the human PAH gene remains unexplained.

We have analyzed the regulatory region of the PAH gene of mice, where PAH-expressing tissues are readily available. Four DNase I-hypersensitive sites specific for PAH-expressing tissues were identified, and their function was tested by transfection. As recipient cells, we have used well-differentiated rat hepatoma cells of adult phenotype which accumulate PAH mRNA at levels similar to those in mouse liver tissue and respond at the transcriptional level to hormone induction. Here we show that as for the human gene, several transcription

* Corresponding author. Mailing address: Department of Molecular Biology, Institut Pasteur, 25 rue du Dr. Roux, 75724 Paris Cedex 15, France. Phone: 33-1-45-68-8500. Fax: 33-1-40-61-3231. Electronic mail address: mweiss@pasteur.fr.

† Present address: Unité d'Immunogénétique Humaine, Département d'Immunologie, Institut Pasteur, Paris, France.

start sites can be detected, and neither TAATA nor CCAAT homologies exist. Characterization of the regulatory elements associated with the DNase I-hypersensitive sites in the 10 kb upstream of the transcription start sites has revealed that expression from the extremely weak promoter of the mouse PAH gene is strictly dependent not only upon a hormone-responsive and tissue-specific enhancer located at kb −3.5 but also upon the presence of hormones. Furthermore, activity of this enhancer requires the hepatocyte-enriched transcription factors hepatocyte nuclear factor 1 (HNF1) and C/EBP.

MATERIALS AND METHODS

Isolation of the mouse PAH regulatory region, subcloning, and sequencing. Screening of a genomic BALB/c mouse DNA library in λ-EMBL3 (Clontech) with a 0.45-kb *Nco*I-*Pvu*II restriction fragment of the 5′ part of the cDNA coding for mouse PAH (pmPAH20 [22]) resulted in isolation of clone λ-17, which included 15 kb of sequences upstream of exon 1. Fragments were subcloned into M13tg130/131 (Amersham) or pBluescript KSII (Stratagene) by standard protocols. Sequence analysis was carried out by the dideoxy method (Sequenase kit; Amersham). The following sequences were determined: nucleotides −2047 to +225 of the promoter region and the noncoding portion of exon 1, named P2, and nucleotides 1 to 1170 localized at kb −4.4 to −3.2 and containing the enhancer called HSIII. Sequence comparisons were carried out by using the Pasteur Institute's Scientific Information Service computer facilities.

Subcloning into eukaryotic expression vectors. Fragments of the PAH regulatory region or oligonucleotides were inserted 5′ of the reporter gene or of the TK (thymidine kinase) promoter driving reporter gene expression into vectors pBLCAT5 or pBLCAT6 (4) or pGL2basic and pGL2enhancer (Promega). Deletions were created by using the pBluescript ExoIII/Mung DNA sequencing system (Stratagene).

In vitro mutagenesis. Mutations in the HSIII enhancer sequence were introduced by using a Chameleon double-stranded, site-directed mutagenesis kit (Stratagene) with the HSIII (positions 760 to 1170) fragment as the template. Oligonucleotides used carried the following mutations: f2-m1, mutation in the C/EBP binding site consensus sequence of f2 according to the mutation of the albumin distal element I (DEI) (24); f3-m2, mutation in the HNF1 binding site homology sequence of f3 (described as mutation PE.34 in references 8 and 47); f5-m1, as for f2-m1 but the mutation also interfered with the putative HNF1 binding site of f5. The effect of these mutations on the abolition of the binding of the corresponding transcription factors was tested by gel mobility shift assays and competition experiments.

Primer extension analysis. Oligonucleotide PAH10 (positions +175 to +201) was 5′ end labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase and hybridized at 65°C for 15 min with 1 to 2 μg of poly(A)$^+$ RNA prepared from mouse organs or cultured cells treated or not with hormones. The primer was extended by using reverse transcriptase rTth (Cetus) for 15 min at 65°C. The products were separated on 6% polyacrylamide–7 M urea gels together with the sequencing reaction products of the corresponding region and detected by autoradiography.

S1 nuclease mapping. Single-stranded DNA of a 1-kb *Nco*I-*Pvu*II genomic fragment (positions −772 to +225) was hybridized with oligonucleotide PAH10. The primer was extended by using the Klenow fragment of DNA polymerase in the presence of [α-^{32}P]dCTP and [α-^{32}P]dGTP. The purified probe was hybridized at 40°C overnight in 80% formamide–0.4 M NaCl–1 mM EDTA–40 mM piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) (PIPES) with poly(A)$^+$ RNA (1 μg) isolated from mouse liver or BW cells treated or not with hormones. Following digestion with 100 U of S1 nuclease (Boehringer Mannheim) at 37°C for 1 h, RNA-DNA heteroduplexes were subjected to electrophoresis on DNA sequencing gels.

Analysis of DNase I-hypersensitive sites. Nuclei were prepared from freshly excised liver, kidney, and spleen tissues of adult BALB/c female mice, treated for 10 min at 4°C with DNase I (grade I; Boehringer Mannheim) by using between 0 and 32 U per μg of DNA; lysed, and digested with proteinase K. DNA was subjected to Southern blot analysis (13) by using probes labeled by random priming (Megaprime; Amersham). For treatment with hormones, mice were injected either once with dexamethasone (DEX) (10 μg/100 g of body weight) or twice with dibutyryl-cAMP (5 mg/100 g followed by a dose of 3 mg/100 g at a 1-h interval) 3.5 h before sacrifice.

Cell lines and culture. FGC4 cells (2) descend from H4IIEC3 cells of the Reuber H35 rat hepatoma and possess a well-differentiated phenotype. H5 cells are also derived from H4II; they are dedifferentiated variant cells which fail to express the entire group of liver-specific functions (16). BW mouse hepatoma cells (43) produce serum proteins but not the specialized enzymes of the liver such as PAH. Mouse fibroblast cells of the L-cell line cl1D were chosen as cells of nonhepatic origin (19). Cells were cultured in modified Ham's F12 medium (15) containing 5% fetal calf serum.

Transfections and transient-expression assays. Cells (5×10^5 to 8×10^5) were transfected by the calcium phosphate procedure by using 10 μg of test plasmid and 5 μg of internal control RSVlacZ plasmid. The precipitate was washed off after overnight incubation, and the medium was renewed daily until harvesting. Cells were harvested 24 h after the wash for luciferase assays and after 48 h for chloramphenicol acetyltransferase (CAT) assays. In some cases hormones were added to the medium [DEX, 1 μM for 48 h; 8-(4-chlorophenylthio)-cAMP, 0.1 mM for 24 h]. The luciferase reporter gene was used only for analysis of the activity of the non-tissue-specific promoter, because luciferase activity declined sharply if the cells were harvested more than 24 h after washing away of the precipitate and tissue-specific promoters (albumin) and enhancers (PAH), in contrast to viral regulatory sequences, require 48 h for optimal expression. CAT activity was determined by the method of Seed and Sheen (41), luciferase was assayed by the method of De Wet et al. (17), and β-galactosidase activity was measured by the standard colorimetric method.

Nuclear protein extracts. Nuclei of cultured cells were prepared as described elsewhere (9). Nuclear extracts used for in vitro footprinting assays were obtained as described previously (31), and extracts used in gel mobility shift experiments were prepared essentially as indicated in reference 9, except that $(NH_4)_2SO_4$ precipitation and dialysis were omitted. The protein concentrations were determined by the method of Bradford (6). Adult mouse liver extract was kindly provided by S. Power (Institut Pasteur, Paris, France). Purified recombinant HNF1 protein (6hHNt) (consisting of the rat HNF1 DNA binding domain, corresponding to amino acids 1 to 281, fused to the 6 NH_2-terminal histidine residues of the Novagen pET-14b vector and expressed in *Escherichia coli*) was kindly given by T. Chouard (Institut Pasteur), and bacterially expressed constitutively active glucocorticoid receptor protein (GR) expressed from plasmid T7EX556 (23) was provided by C. Muchart (Institut Pasteur).

DNase I footprints. Three overlapping fragments corresponding to nucleotides 760 to 950, 850 to 1108, and 950 to 1170 of the HSIII enhancer sequence served for the preparation of single-stranded probes labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP. Probes (1 ng) were incubated with 20 μg of nuclear proteins extracted from FGC4 cells or mouse liver tissue in a final volume of 20 μl as described elsewhere (9). DNase I (grade I; Boehringer Mannheim) was used at 10 to 40 U/ml for FGC4 protein-containing samples and at 2 to 16 U/ml for mouse liver protein-containing samples. Mixtures for reactions with HNF1 or GR protein contained 20 μg of bovine serum albumin. Binding of the GR was allowed to occur for 45 min at 25°C before addition of DNase I (2.5 to 10 U/ml). The reaction products were separated on 6% polyacrylamide–7 M urea gels with G+A chemical sequencing reactions run in parallel. Autoradiograms were obtained from dried gels.

Oligonucleotides. Gel-purified oligonucleotides (Genset, Paris, France) used for gel mobility shift assays were annealed in order to form double-stranded probes. The following oligonucleotides were used. f2, f3, f4, f5, and f6 correspond to the sequences of the footprint regions in the PAH enhancer HSIII (see Fig. 6C). A TCGA overhang at their 5′ ends allowed multimerization and cloning into the *Sal*I site of pBLCAT5. PE corresponds to the HNF1 binding site of the rat albumin promoter proximal element (5′-TGGTTAATGATCTACAGTTA). The C/EBP probe was the binding site for C/EBP in the rat albumin promoter DEI element (5′-GGTATGATTTTGTAATGGGGTA), the HNF3 probe was the binding site for HNF3 of the transthyretin promoter (5′-TTGACTAAGTC AATAATCAGAATCAG), HNF4 probe was the binding site for HNF4 of the apolipoprotein CIII promoter (5′-GGTCAGCAGGTGACCTTTGCCCAGCG), the CRE probe was the cAMP response element of the rat PEPCK promoter (5′-GGGCCCTTACGTCAGAGGCGAGA), the AP1 probe was the AP1 binding site of the mouse α1-antitrypsin promoter (5′-GATCAGGGGCCATGTGA CTCATTACACCAG), the NF1 probe was the binding site of NF1 of the adenovirus 2 positions +20 to +49 (5′-TATTTTGGATTGAAGCCAATATGATA ATGA), and the GRE probe was the glucocorticoid response element of papillomavirus (5′-TGGAGAAAAAATTCGTAGCTAGAACAGACTGTTCTGAG AT).

Gel mobility shift assays. Double-stranded oligonucleotides were labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase. Incubation of the labeled probes (0.4 ng) with 10 μg of nuclear protein extracted from FGC4, H5, or cl1D cells for 10 min at 4°C in a final volume of 14 μl was followed by electrophoresis in 6% polyacrylamide gels in 0.25× TBE (22 mM Tris, 22 mM boric acid, 0.5 mM EDTA) and autoradiography. The reaction conditions were as described previously (8). For competition assays, a 12.5- to 125-fold molar excess of nonlabeled oligonucleotides was added to the reaction mixtures.

Transactivation studies. H5 cells were transfected with 0 to 3 μg of expression plasmids for HNF1 or C/EBP and 10 μg of pBLCAT5; pBLCAT5-HSIII, -f2, -f3, or -f5; or a control plasmid containing a known binding site. The concentration of expression vector promoter was kept constant by addition of the appropriate amount of the corresponding vector DNA. In plasmids pBLCAT5-f2, -f3, and -f5, a trimer of the corresponding oligonucleotide was inserted 5′ of the TK promoter; pD1 contains a monomer of the C/EBP binding DE1 oligonucleotide and the TAATA box of the albumin promoter; and ΔTK-H (given by S. Power) contains a dimer of the albumin HNF1 binding site in PE 5′ of a TK minimal promoter.

Nucleotide sequence accession numbers. The EMBL nucleotide sequence database accession numbers for the promoter and HSIII enhancer regions of the mouse PAH gene are X97252 (P2) and X97253 (HSIII enhancer).

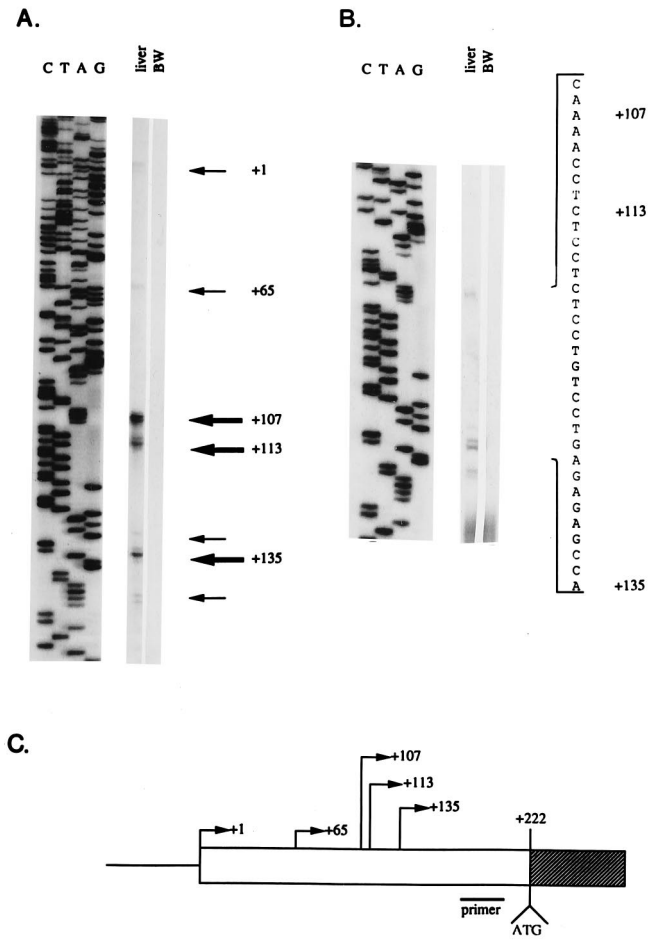


FIG. 1. Localization of the PAH transcription start sites. S1 mapping or primer extension reactions were carried out with 1 μg of poly(A)$^+$ RNA of adult mouse liver or BW (PAH$^-$) mouse hepatoma cells. The products were separated on sequencing gels and detected by autoradiography. Sequencing reactions for the corresponding region were run in parallel. (A) S1 mapping analysis. The single-stranded labeled probe (positions −772 to +201) was hybridized with poly(A)$^+$ RNA and treated with S1 nuclease. The S1-resistant RNA-DNA heteroduplexes were separated on the gel. (B) Primer extension reactions. Labeled oligonucleotide PAH10 (for position, see panel C) was used as a primer for reverse transcription, and extended products were analyzed. (C) Positions of the transcription initiation sites. The noncoding part of exon 1 was determined to span 222 bp (open box); the coding part (amino acids 1 to 20) is represented as a shaded box. The line labeled "primer" indicates the position of the oligonucleotide PAH10.

RESULTS

Multiple transcription initiation sites of the mouse PAH gene. Although the transcription initiation sites have been characterized for the human PAH gene (30, 49), the lack of sequence homology upstream of the coding region between the human and mouse genes made it necessary to reexamine this question in order to identify the promoter region for the mouse. Both nuclease S1 mapping and primer extension were carried out. As shown in Fig. 1A, the S1 nuclease-protected fragments were multiple, indicating the existence of several transcription initiation sites. Major sites were found at positions +107, +113, and +135, while minor sites were detected mainly at positions +1 and +65. The same spectrum of bands was observed in total as well as poly(A)$^+$ RNA from the liver tissue of control or DEX-treated animals. The bands at positions +1 and +65 were not enriched in nuclear compared with cytoplasmic RNA, implying that they do not represent processing intermediates (data not shown). In contrast, none of the protected fragments was produced with RNA of PAH-negative BW mouse hepatoma cells (Fig. 1A). These observations were confirmed by the primer extension assay (Fig. 1B), using as a primer an oligonucleotide located in the 5′ untranslated region (Fig. 1C). The strongest band was at position +135, with the +107 band clearly visible and the +113, +65, and +1 bands emerging upon longer exposure. No difference was observed with RNA from hormone-treated cells or organs, and no extension products were detected with primers localized in the region between positions +1 and −121 (data not shown). No fragments were found with RNA from PAH$^-$ BW hepatoma cells (Fig. 1B).

DNase I-hypersensitive sites in the region flanking the PAH gene. To identify hypersensitive sites in the chromatin surrounding the transcription start sites of the PAH gene, two probes were used. The first (probe A in Fig. 2C) covers part of the first exon and of the first intron: it hybridizes with an 8.3-kb *Sac*I fragment of genomic DNA (Fig. 2A). To search for hypersensitive sites in the first intron, DNA was digested with *Eco*RI or *Eco*RV (data not shown). The second probe is situated at kb −8.6 to −8.4 (probe B; Fig. 2C) and hybridizes with a 3.5-kb fragment of *Xba*I-digested DNA (Fig. 2B).

Nuclei prepared from liver (treated or not with DEX or cAMP), kidney, or spleen tissue were incubated with DNase I, and DNA was extracted and analyzed by Southern blots. The first probe permitted the identification of three hypersensitive sites in 7.2 kb of 5′ flanking sequences. Hypersensitive site I (HSSI) corresponds to the region of transcription initiation sites, while HSSII and -III are observed at approximately kb −1.3 and −3.5. These sites were found in liver samples, and treatment of the animal with DEX or cAMP did not influence the presence or the intensity of the bands obtained (Fig. 2A). In kidney samples, only HSSI was observed, while none of the three sites was present in spleen tissue, where the PAH gene is not expressed (data not shown). Hybridization of the same probe with *Eco*RV- or *Eco*RI-digested DNA revealed that no hypersensitive sites exist within the proximal 8 kb of the first intron. The second probe (Fig. 2B) identified an additional hypersensitive site (HSSIV) in liver and kidney samples but not in spleen samples. Figure 2C shows a scheme of the 5′-flanking sequences of the PAH gene, the positions of relevant restriction sites, the probes, and the approximate locations of the hypersensitive sites. It should be noted that a region around kb −8 to −7, defined by the *Xba*I-*Sac*I fragment, has not been screened for hypersensitive sites.

PAH expression in cells used for functional characterization of the PAH regulatory region. FGC4 cells (2) are well-differentiated rat hepatoma cells that express numerous liver-specific serum proteins as well as the specialized metabolic enzymes of adult liver, including PAH. These cells (as well as the related H4II and Fao cells) accumulate PAH mRNA at about 50% of the level in adult mouse liver tissue. Treatment with DEX causes a significant increase, treatment with cAMP leads to a slight increase, and treatment with both hormones results in a further increase in the amounts of PAH mRNA. In addition, run-on experiments demonstrated that DEX stimulation occurs at the transcriptional level (reference 22 and data not shown). Furthermore, this family of hepatoma cell lines expresses the set of transcription factors preferentially expressed in the liver (liver-enriched transcription factors) at levels sufficient to permit functional studies of liver-specific promoters (10, 37, 47).

H5 cells are dedifferentiated variants of the same hepatoma line as FGC4. While these cells are of hepatic origin, they do

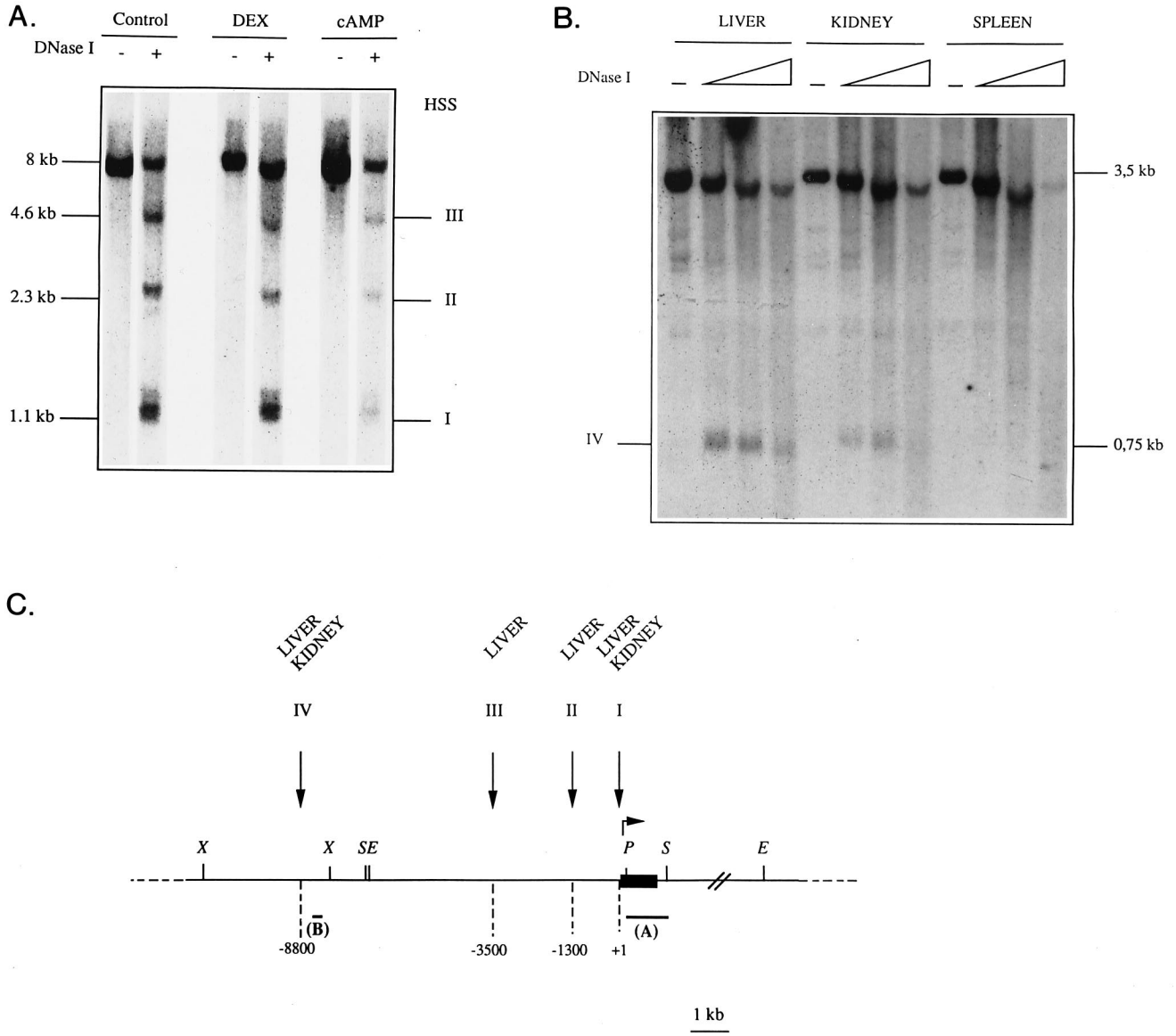
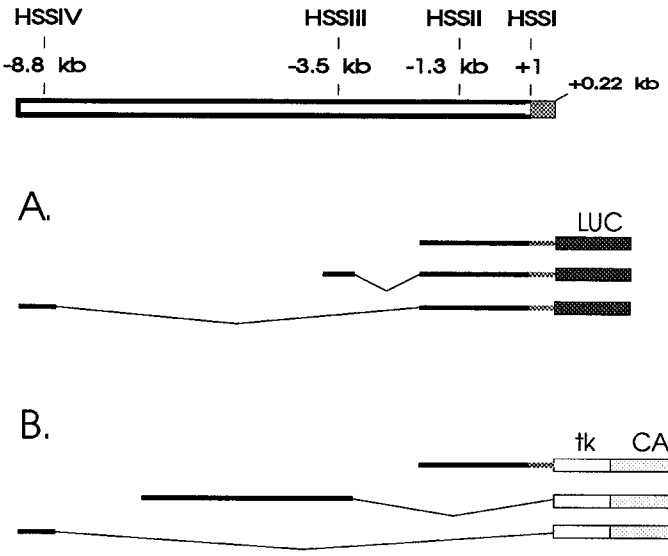


FIG. 2. DNase I-hypersensitive sites in the region surrounding the transcription start sites of the PAH gene. Nuclei from the PAH-expressing liver and kidney tissues, as well as spleen tissues (PAH^-), were treated with DNase I, and genomic DNA (15 μg per lane) was subjected to Southern blot analysis. (A) *Sac*I-digested DNA from liver chromatin treated with DNase I (+) or not treated (−) was hybridized with probe A [for position, see bar designated "(A)" in panel C]. Three bands in DNase I-treated samples (I, II, and III) indicate the presence of DNase I-hypersensitive sites within the 8.3-kb *Sac*I fragment. No differences between DEX- or cAMP-treated and untreated animals was observed. (B) *Xba*I-digested DNA from chromatin of liver, kidney, or spleen tissues treated with increasing amounts of DNase I (0 to 32 U/μg of DNA) was hybridized with probe B. A DNase I-hypersensitive site (IV) was detected in liver and kidney samples. (C) Localization of the DNase I-hypersensitive sites. The restriction sites around exon 1 (black box with transcription start site +1 indicated by the bent arrow) relevant for the analysis are indicated: *X*, *Xba*I; *S*, *Sac*I; *E*, *Eco*RI; *P*, *Pst*I. The hybridization probes correspond to the black bars designated "(A)" and "(B)". The approximate positions of hypersensitive sites I to IV are indicated in base pairs. Sites I and IV were found in both liver and kidney tissues, and sites II and III were liver specific.

not express hepatic functions (16) or the liver-enriched transcription factors (1, 10); no PAH transcripts can be detected, even after hormone treatment (data not shown). Finally, cl1D cells are mouse L-cell fibroblasts chosen as a control to distinguish responses due to ubiquitous factors from those that could be restricted to cells of hepatic origin.

Extremely weak activity of the PAH promoter. The PAH promoter has been analyzed as a 2-kb (P2) and as a 7-kb (P7) fragment of upstream sequences, both extending to position +222. In addition, a series of 5′ deletions was prepared from the P2 fragment. When these fragments were tested for promoter activity in the promoterless pBLCAT6 vector in each of the three cell lines, they failed to show reliably measurable activity (data not shown). We therefore turned to luciferase vectors, which allow reporter gene expression to be assayed with greater sensitivity. The P2 fragment (Fig. 3A), as well as its 5′ deletions (data not shown), showed similar profiles of weak activity in the three cell lines. In this test, the sequences around HSSII do not contribute to transcriptional activity. Deletions beyond bp −58 lead to a progressive loss of activity, with no further activity detected upon deletion of transcription start sites.

Our failure to detect even moderate promoter activity led us to verify that no cloning artifact had modified the structure of



	FGC4	H5	cl1D
	fold promoter-less vector		
P2	11.0 ± 2.0	10.6 ± 0.6	24.0
HSIII-P2	13.2 ± 2.2	12.0	31.2
HSIV-P2	16.5 ± 1.7	12.0	38.4
	fold tk		
P2	6.1 ± 0.4	2.3 ± 0.4	1.9 ± 0.4
HSIII (3.8 kb)	9.5 ± 1.8	5.3 ± 0.4	0.9 ± 0.2
HSIV	11.0 ± 2.0	4.3 ± 0.7	3.6 ± 0.5

FIG. 3. Activity of the PAH promoter and enhancer regions. The constructs schematically represented by the lines were tested in transient-expression assays with the different cell types. The results are represented as mean values with standard deviations for at least three to five independent experiments; when the standard deviation is not shown, only two experiments were carried out. (A) The PAH fragments were inserted into the pGL2 vector containing the luciferase gene (LUC) as a reporter gene. The PAH promoter fragment P2 corresponds to the sequence from bp −2047 to +222, fragment HSIII corresponds to the minimal HSIII enhancer sequence defined in Fig. 5, and the HSIV enhancer is contained in a 680-bp fragment located at kb −9.3 to −8.6. Enhancer fragments were inserted 5′ of P2. The values are expressed as fold activation relative to values for the promoterless vector. Activities of P2 in comparison with that of the RSV LTR were 0.09% in FGC4, 0.2% in H5, and 0.001% in cl1D cells. (B) The PAH fragments were placed 5′ of the TK promoter of pBLCAT5. Here, the HSIII enhancer is contained in a 3.8-kb fragment spanning kb −7 to kb −3.2. Results are expressed as fold activation in comparison with that for the TK promoter alone, which is set at 1. The relative CAT activities obtained with pBLCAT5 were 13 ± 3.6 for FGC4, 2.9 ± 0.8 for H5, and 42 ± 7.6 for cl1D. HSIII and HSIV were tested in both orientations, and no significant differences were observed.

our starting material, the genomic fragment of clone λ-17. Comparison of Southern blot patterns and PCR products obtained from λ-17 and from DNA of different mouse strains revealed no differences (data not shown). Moreover, the 450-bp sequence upstream of the ATG of the rat PAH gene (33) presents a higher degree of homology than does the human promoter to the corresponding mouse sequences. We are therefore confident that the promoter region of the mouse PAH gene has been correctly identified.

We conclude that the TAATA- and CCAAT-less PAH proximal promoter and up to 2 kb of flanking sequences lack important regulatory elements endowed with activity in the transient-expression assay. Furthermore, the similarities observed for the three cell lines imply that any transcription-promoting activity present is dependent on ubiquitous and not on tissue-specific transcription factors.

Enhancer activity of PAH upstream regulatory elements. The P2 fragment was inserted upstream of the TK promoter of the pBLCAT5 vector. In addition, the upstream regions comprising HSSIII or HSSIV were analyzed as a 3.8-kb fragment (HSIII) and a 680-bp fragment (HSIV) (Fig. 3B).

P2 confers a sixfold increase in TK promoter activity in FGC4 cells and a two- to threefold activation in H5 or cl1D cells. A substantial enhancement by sequences in the region of HSSIII was revealed: fragment HSIII (3.8 kb) provokes a 10- to 12-fold stimulation of the TK promoter in FGC4 cells, compared with a 5-fold stimulation in H5. In contrast, these sequences are totally inactive in cl1D cells. A 680-bp fragment containing the region (kb −9.3 to kb −8.6) surrounding HSSIV showed significant enhancer activity in all three cells lines, two- to threefold higher in FGC4 than in H5 or cl1D.

Surprisingly, none of the enhancer fragments showed the same activity when placed upstream of the PAH promoter (Fig. 3A). Neither a small distal fragment containing HSSIII (shown below to possess full HSIII enhancer activity; 760/1170 in Fig. 5) nor HSIV had a significant stimulatory effect when inserted upstream of P2 in the luciferase vector. Thus, the enhancer activity of these PAH upstream fragments is revealed only in the context of a heterologous promoter or, as shown below, upon the addition of hormones.

PAH promoter activity requires hormone induction of its tissue-specific enhancer. Preliminary experiments revealed that the P7-containing plasmids but neither the P2- nor the HSIV-containing plasmids were inducible by hormones and that this response was limited to cells of hepatic origin (data not shown). Subsequent experiments therefore concentrated on the upstream region containing HSSIII and on FGC4 and H5 cells.

Figure 4 shows the profile of induction obtained with constructs containing HSSIII in either the P7 fragment or the P7Δ1.7 fragment (an internal deletion of the region between kb −3.2 and kb −1.5 of P7) and placed in the context of either the PAH promoter (P7Δ1.7*) or the TK promoter (pBLCAT5 series). This profile was characteristic for each of the cell lines: in FGC4 cells, DEX stimulates activity 10- to 80-fold, cAMP stimulates activity 2- to 4-fold, and the addition of the two compounds results in a synergistic response of 20- to 150-fold induction. In contrast, in H5 cells the DEX response is more modest (8- to 13-fold), there is no response to cAMP, and the two hormones together behave like DEX alone. The presence of the 3.8-kb HSIII fragment upstream of the PAH promoter (P7Δ1.7*) resulted in barely measurable basal activity, accompanied by an 80- to 150-fold induction, but only in differentiated FGC4 cells. In the presence of hormones, high CAT activity levels are obtained for P7Δ1.7*, the maximum activity corresponding to around one-fifth of the values observed for the control plasmid RSV-CAT. These results demonstrate that the PAH promoter shows an absolute dependence not only

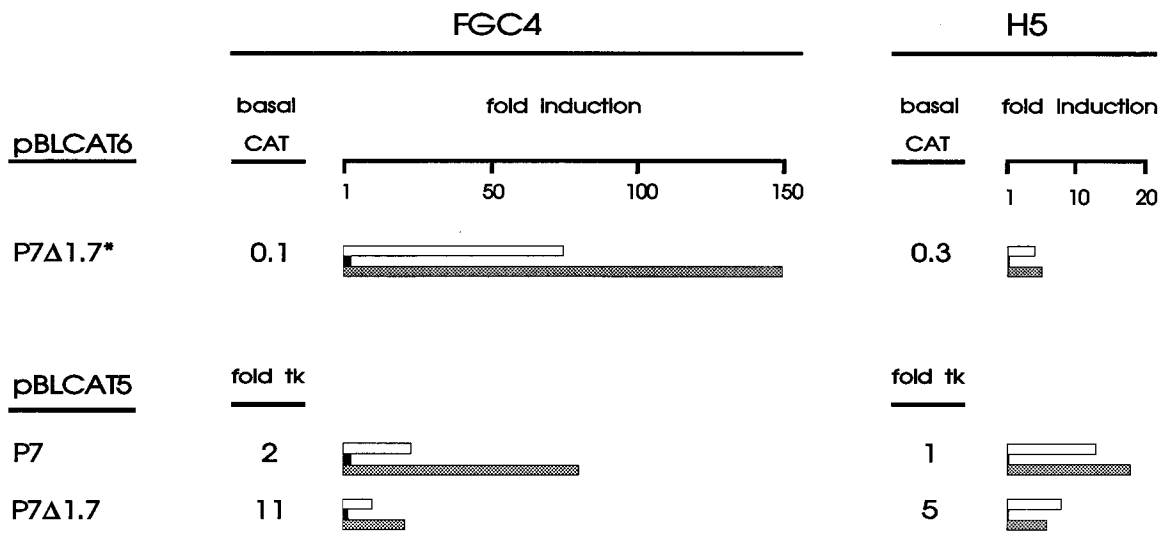
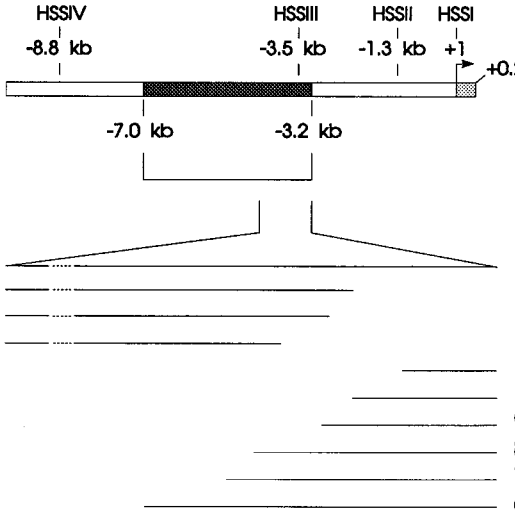


FIG. 4. Inducibility of HSIII-containing constructs in FGC4 and H5 cells. In P7Δ1.7* the 3.8-kb HSIII fragment was inserted 5′ of the PAH promoter; the basal activity shown is relative CAT activity (to be compared with the TK values given in the legend to Fig. 3). Values from assays of the hormone-treated cells are given as means of the fold increase over the basal activity. For the pBLCAT5 constructs, basal activity is presented as the fold increase over the activity of the TK promoter alone. ☐, DEX; ■, cAMP; ▨, DEX plus cAMP.

upon the tissue-specific enhancer located in the HSIII fragment and harboring the hormone response elements but also upon hormones for significant expression, explaining why only negligible activities were obtained with the PAH promoter in P2, HSIII-P2, or HSIV-P2 in the absence of hormones (Fig. 3A).

Full HSIII enhancer activity is localized within a 410-bp fragment. Starting with the plasmid pBLCAT5-HSIII (3.8 kb), restriction fragments with various 5′ borders were created and tested. The smallest which retained the same basal and induced activities as those of HSIII (3.8 kb) was chosen for sequencing and creation of further 3′ and 5′ deletions in order to define the hormone-responsive region: it corresponds to a 1,170-bp sequence between kb −4.4 and kb −3.2 (Fig. 5). Removal of 218 to 500 bp from its 3′ end resulted in a 2- to 3-fold reduction of basal activity. Basal activity throughout most of the 5′-deletion series was essentially constant, similar to that of the entire fragment, with at most a variation of ±30 to 40%. However, a significant drop was observed when only 144 bp of the fragment remained. Concerning the hormone response of these fragments, even the smallest of the 3′ deletions results in reduced inducibility and response is completely lost when more than 259 bp has been deleted (Fig. 5; compare 1/847 and 1/911). Full inducibility of the 5′ deletions requires retention of at least 360 bp in FGC4 and of 410 bp in H5 cells (corresponding to constructs 810/1170 and 760/1170, respec-



	FGC4				H5			
	fold induction				fold induction			
	basal	d	c	d+c	basal	d	c	d+c
HSIII (3.8 kb)	10	17	4	46	5	11	1	13
1/1170								
1/952	3	3	1	15	2	5	1	5
1/911	4	7	3	18	3	4	1	4
1/847	3	2	2	3	1	1	1	1
1026/1170	4	1	1	2				
950/1170	10	2	2	2				
905/1170	6	1	2	2	3	1	1	1
810/1170	10	13	5	28	2	2	1	2
760/1170 = HSIII	13	11	3	26	4	5	1	5
640/1170	12	8	3	28	5	6	1	6

FIG. 5. Characterization of the HSIII enhancer region. The HSIII region and its variants with 3′ and 5′ deletions were tested for their potential to activate the TK promoter in the presence and absence of hormones in FGC4 and H5 cells. Note that the fragment extending from kb −4.4 to kb −3.2 has been sequenced and redesignated 1/1170 for precise definition of the deletions. Fragment 760/1170, the smallest fragment retaining full enhancer activity, is designated HSIII in subsequent figures and the text. The lines indicate the regions included in each construct. Except for 640/1170 in H5 cells, constructs were tested at least twice. Values are given as multiples of TK promoter activity for basal activity and as fold stimulation of basal activity for hormone-treated samples. d, DEX; c, cAMP.

tively). Consequently, the hormone-responsive elements must lie within the region from nucleotides 810 to 911 and sequences 3′ contribute to basal enhancer activity. We conclude that the entire HSIII enhancer is localized in the 410-bp sequence from positions 760 to 1170, hereafter designated HSIII.

Binding of nuclear proteins to multiple sites in the HSIII enhancer. A DNase I footprint analysis was carried out to identify sequences in the HSIII enhancer capable of binding nuclear proteins from PAH-expressing cells. Strand-specific probes were created to cover the region of nucleotides 760 to 1170. Figure 6 gives examples of the results obtained upon incubation of two of these probes with nuclear protein extracts from FGC4 cells and adult mouse liver tissue or with the HNF1 or glucocorticoid receptor (GR) protein. With the probe spanning nucleotides 850 to 1108 (Fig. 6A), four footprints were observed with FGC4 proteins (f2, f4, f5, and f6), and these were also obtained with mouse liver proteins. An additional region (f3) weakly bound liver extract proteins.

To identify possible HNF1 binding sites in the HSIII enhancer, footprint reactions were carried out with the bacterially expressed protein; it bound to two regions which colocalize with f3 and f5. However, the binding to f5 required larger amounts of protein than did the binding to f3, suggesting a difference in the efficiencies of binding to these two sites. With FGC4 extracts, the only footprint obtained in the 5′-adjacent fragment (nucleotides 760 to 950) is f2 (Fig. 6B). Weak liver protein binding to an additional region further 5′ was observed (data not shown). Since the sequences between positions 810 and 911 harbor the hormone response elements, a bacterially expressed derivative of glucocorticoid receptor protein was incubated with the sequence of nucleotides 760 to 950 as a probe (Fig. 6B). GR protein bound to three sequences located around and partially overlapping f2. With the sequence of nucleotides 950 to 1170 as a probe, no further footprints were observed with the FGC4 extracts, but the mouse liver extract contains proteins binding to nucleotides 983 to 997 and to nucleotides 1131 to 1150 (data not shown). Essentially identical footprint positions were observed with probes corresponding to the second strand (data not shown). The localizations of the regions protected from DNase I are indicated in Fig. 6C, and the positions of the oligonucleotides (f2 to f6) chosen for further analysis are marked.

Identification of transcription factors interacting with HSIII enhancer sites. Sequences of the HSIII enhancer binding nuclear proteins were used as double-stranded oligonucleotide probes in gel mobility shift assays. The pattern of complexes obtained with cellular extracts from FGC4, H5, or cl1D cells was different for each cell line with f2, f3, or f5, whereas identically migrating complexes were formed between both f4 or f6 and all extracts. Binding to the probes was efficiently inhibited by competition with a 50-fold excess of the unlabeled oligonucleotide (data not shown).

Competitions for FGC4 protein binding with oligonucleotides corresponding to known binding sites of ubiquitous or liver-enriched transcription factors (Fig. 7) indicated that several of these factors are potentially involved in binding to the elements of the HSIII enhancer. The formation of complexes obtained with both f4 and f6 was inhibited by competition with NF1 binding sites, present also in regulatory regions of other liver-specific genes, such as the albumin gene (9); f4 contains an NF1 consensus sequence, and an NF1 half-site is found within f6. Most of the complexes formed with f2 and f5 were abolished by competition with the C/EBP binding site of the albumin promoter. The f2 sequence harbors a perfect consensus C/EBP binding site, and a region of partial C/EBP homology is localized in f5. In both cases, the formation of the same complexes was also inhibited by competition with a cAMP response element consensus sequence, although no CRE consensus sequence could be identified by sequence analysis and neither f2 nor f5 competed with the CRE for protein binding (data not shown). The competition with the CRE could thus be due to sequence similarities between the C/EBP and the CRE sites. An AP1 consensus sequence was identified in f2 by sequence comparison, but competition was not obtained. Consequently, it is likely that C/EBP or related proteins are responsible for complex formation with f2 and f5.

The formation of three slowly migrating complexes obtained with f3 was inhibited by competition with the albumin HNF1 binding site. Another, faster-migrating complex was not formed in the presence of an AP1 binding site or a CRE oligonucleotide as competitor. The f3 sequence does not correspond to a perfect HNF1 binding site, but 11 of 15 bp match the consensus sequence. By using the more comprehensive criteria established by Tronche et al. (46), it would be recognized as a binding site of moderate affinity. Homology was found with an AP1 consensus site but not with the CRE sequence. The comparison between the complexes obtained with f3 or the HNF1 binding site of the rat albumin promoter (PE) and proteins of FGC4 or H5 cells is shown in Fig. 8A. Among the slowly migrating complexes, identical bands were observed for FGC4 with both probes, corresponding to the expected HNF1/vHNF1 homo- and heterodimers. In contrast, only the vHNF1 homodimer formed complexes with f3 and PE in H5 extracts. Competition with increasing amounts of unlabeled oligonucleotide revealed that the competition by f3 is considerably weaker than is the competition by PE.

Binding of HNF1 to f5 could not be demonstrated unambiguously with cellular extracts (Fig. 7 and 8B): the HNF1/vHNF1 complexes formed with f5 are barely visible, and f5 did not compete efficiently for the formation of complexes with the PE probe. In addition, the homology between the f5 sequence and the HNF1 consensus binding site is too weak to be recognized in a homology search. However, purified HNF1 indeed bound to f5 at the highest concentration of protein tested. Comparison of the amounts of protein required for similar amounts of complexes leads to an estimation of the different binding efficiencies: f5 requires roughly 50-fold and f3 requires approximately 10-fold more HNF1 than does the strong HNF1 binding site PE. That f3 is apparently a better HNF1 site than f5 is surprising in view of the stronger footprint observed with cellular extracts at f5 (Fig. 6A). However, the footprints were obtained with extracts prepared in the presence of the detergent Nonidet P-40, which could interfere with DNA-protein binding under the conditions used in these experiments. Hence, the proteins bound to f5 in the footprint reactions could be different from HNF1. Moreover, liver extracts prepared without the detergent (kindly provided by S. Cereghini, Institut Pasteur) produced a stronger footprint in the f3 region (data not shown) and smaller amounts of HNF1 protein were required for the footprint of f3 than for the footprint of f5 (Fig. 6A). Taken together, these results indicate that both HNF1 binding sites in the HSIII enhancer of the PAH gene are weak but probably real sites.

Transactivation of HSIII by HNF1 and C/EBP. To investigate the possible role of C/EBP and HNF1 in the activity of the HSIII enhancer, H5 cells were transfected with pBLCAT5-HSIII and increasing amounts of expression vectors for HNF1 or C/EBP. H5 cells do not produce HNF1 (8, 10), contain vHNF1 as well as proteins of the C/EBP family (1, 39), and are of hepatic origin; they do not express the endogenous PAH gene. To evaluate the relative efficiencies of transactivation, two types of controls were employed: one or two copies of a

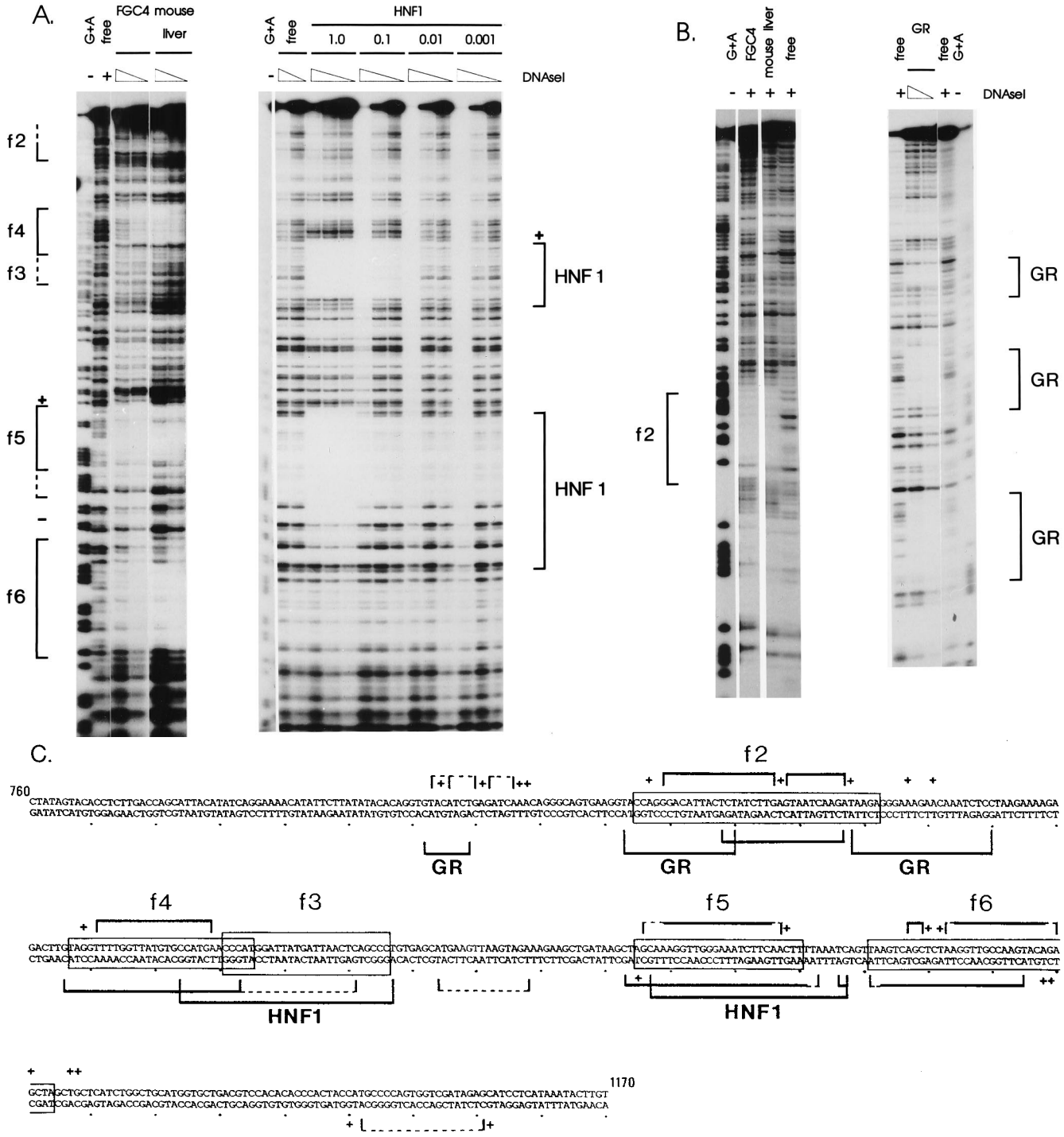


FIG. 6. DNase I footprint analysis of the HSIII enhancer. Nucleotides 760 to 1170 were previously identified by transfection data to comprise the complete enhancer activity. Results obtained with the labeled upper strands of two of the probes are shown. The positions of footprints obtained are indicated by brackets, and the precise locations of footprints within the fragment are given in panel C. (A) The probe, spanning nucleotides 850 to 1108, was incubated with 20 μg of nuclear proteins extracted from FGC4 cells or from adult mouse liver (ML) or with decreasing amounts (1 to 0.001 μg) of a bacterially expressed purified HNF1 protein. Reaction mixtures were treated with increasing amounts of DNase I (10 to 40 U/ml for FGC4, 2 to 16 U/ml for ML, and 2.5 to 10 U/ml for HNF1) before separation on sequencing gels and autoradiography. (B) DNase I footprints with a probe spanning nucleotides 760 to 950 and 20 μg of nuclear proteins of FGC4 cells or mouse liver (ML). DNase I concentrations are as described for panel A. Incubation of the probe with a bacterially expressed derivative of the glucocorticoid receptor protein (GR) revealed three regions of GR binding around or partially overlapping the f2 footprint. (C) Sequence of the HSIII enhancer (nucleotides 760 to 1170), positions of DNase I footprint regions, and oligonucleotides chosen for further analysis. The brackets indicate the DNase I footprints obtained with the upper-strand (⌜¬) and lower-strand (L_) probes. Broken lines represent protections obtained only with liver nuclear extracts, and solid lines represent those observed with both FGC4 and liver extracts. Sequences corresponding to oligonucleotides f2 (positions 855 to 891), f3 (positions 950 to 975), f4 (positions 926 to 955), f5 (positions 1014 to 1039), and f6 (positions 1050 to 1083) are boxed.

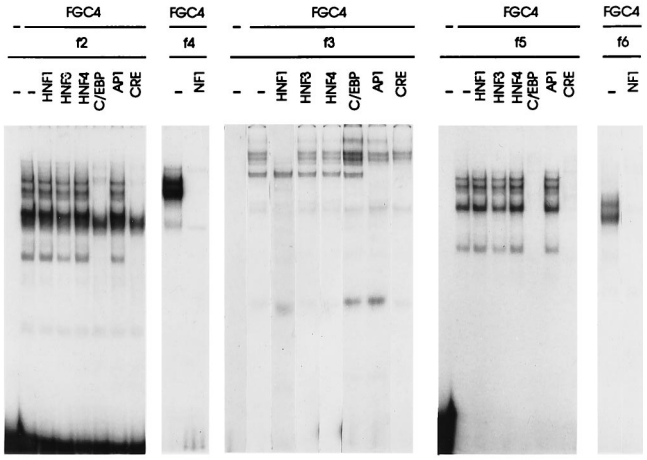


FIG. 7. Gel mobility shift analysis with oligonucleotides spanning regions of in vitro footprints of the HSIII enhancer. Labeled oligonucleotides (0.4 ng) f2, f3, f4, f5, and f6 (Fig. 6C) were incubated with nuclear protein extracts (10 μg) of FGC4 cells to allow complex formation. The reaction mixtures were separated on native polyacrylamide gels, and autoradiograms are shown. For competition experiments, known binding sites for transcription factors HNF1, HNF3, HNF4, C/EBP, AP1, and NF1 and a cAMP response element, CRE, were added in a 50-fold excess as unlabeled oligonucleotide. −, no addition.

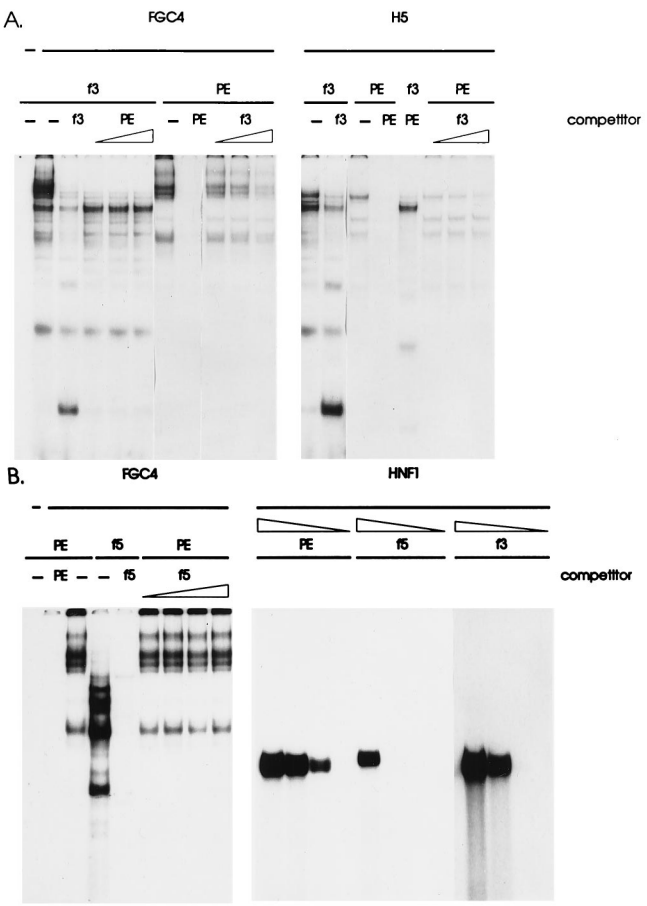


FIG. 8. HNF1 binding to the HSIII enhancer. Gel shift analysis was carried out as described in the legend for Fig. 7. Competition experiments were carried out with increasing amounts of unlabeled oligonucleotide (12.5- to 125-fold excess). (A) Labeled oligonucleotide f3 or PE (corresponding to the HNF1 binding site of the albumin promoter proximal element) was incubated with extracts of FGC4 or H5 cells. (B) Labeled oligonucleotide PE, f5, or f3 was incubated with FGC4 cell extract or with decreasing amounts (1 to 0.001 μg) of the HNF1 protein used for DNase I footprint analysis (Fig. 6A). −, no addition.

known C/EBP or HNF1 binding site upstream of a minimal promoter and PAH f2, f3, or f5 trimers inserted into pBLCAT5.

Comparably strong transactivations by C/EBP were obtained for f5 and for the albumin C/EBP site (maximum of about 200-fold). The f2-containing plasmid was stimulated at roughly 10% of this level but showed a significantly higher level of activity in the absence of C/EBP. In contrast, C/EBP induces only a threefold increase in activity of the HSIII fragment and no further stimulation was obtained upon the addition of DEX. The presence of C/EBP family members in H5 cells, as well as the existence of multiple sites in HSIII available for the fixation of other regulatory proteins, could account for the unexpectedly weak transactivation observed.

An approximately 10-fold transactivation by HNF1 was obtained for a dimer of the strong albumin HNF1 site and for the trimer of f3. In addition, f5, which contains an HNF1 binding site even weaker than the site in f3 (see Fig. 8B), is also transactivated in a similar fashion. The activity of the HSIII fragment was increased 6- to 7-fold, and the factor of activation was the same in the presence as it was in the absence of DEX. These results strongly suggest that the HNF1 binding sites of the PAH enhancer are functional and that HNF1, as well as C/EBP, exerts not a synergistic but rather an additive effect on the DEX response.

Mutations of the HNF1 and C/EBP binding sites impair HSIII enhancer function. To determine the contributions to the HSIII enhancer activity of the corresponding binding sites, mutations which abolish binding of the factors were created in the f2 (C/EBP), f3 (HNF1), and f5 (C/EBP and HNF1) sites in the context of HSIII. Each of these mutations was tested in pBLCAT5 in FGC4 and H5 cells. In addition, since expression from the PAH promoter is indissociable from inducibility, the consequences upon hormonal response were evaluated.

In FGC4 cells, mutations of f2, f3, and f5 reduce basal activity by 25 to 50% (Fig. 9). However, the repercussions upon hormonal induction are even stronger: the activities obtained upon DEX treatment with the f2 and f3 mutants are only one-quarter of that observed with the intact HSIII enhancer. The mutation in f5 has less influence upon the hormonal stimulation: DEX inducibility is reduced only by one-half, while the synergistic response with cAMP is retained. The reduction of hormonal inducibility by abolishing HNF1 or C/EBP binding demonstrates cooperation among these factors to achieve maximal activity of the enhancer.

In H5 cells, abolition of the binding of the tissue-specific factors has less influence. The three mutations reduce basal activity by approximately 30%, while the remaining DEX inducibility is still about 50% that of the unmutated HSIII enhancer (Fig. 9). These results indicate that the C/EBP and the HNF1 sites do contribute to activity of the HSIII enhancer in H5 cells via C/EBP-related proteins and vHNF1. Although in gel mobility shift assays f3 forms complexes which can be abolished by competition with an AP1 binding site (Fig. 7), interaction with HNF1 family members seems to be more relevant in the cells tested, since the HNF1-site-specific mutation has a stronger effect in FGC4 than in H5 cells.

The comparison of transfection results obtained with FGC4 and H5 cells raises the question of why the endogenous PAH gene is inactive in H5 cells. With one exception, the differences observed are quantitative rather than qualitative. The exception concerns cAMP inducibility and the synergy of the two

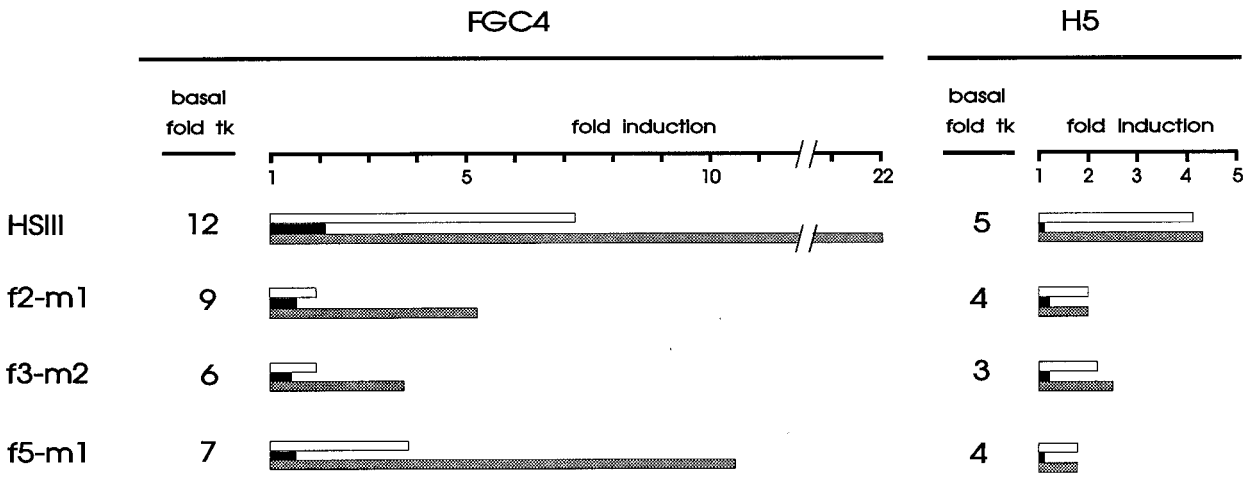


FIG. 9. Effect of mutations of the C/EBP and HNF1 binding sites in the HSIII enhancer. Mutation f2-m1 abolishes C/EBP binding, f3-m2 abolishes HNF1 binding, and f5-m1 abolishes binding of C/EBP and HNF1. These mutations were introduced into the 410-bp HSIII fragment and tested in pBLCAT5 by transfection of FGC4 and H5 cells. Basal activity corresponds to the fold activation of the TK promoter. Hormonal induction is expressed as fold activation of the basal activity. ▭, DEX; ▬, cAMP; ▨, DEX plus cAMP.

hormones. In H5 cells, cAMP responsiveness is known to be deficient (reference 23 and data not shown) and HNF1 is totally absent (8, 10). It can therefore be suggested that both HNF1 and full hormone responsiveness are critical elements for expression of the PAH gene.

Figure 10 summarizes the results obtained from analysis of the HSIII deletions, DNase I footprints, and gel mobility shift assays: they indicate that the enhancer spans 410 bp and contains at least 10 different protein binding sites, some of them probably complex and involving binding of ubiquitous and of tissue-specific transcription factors as well as binding of proteins conferring response to hormonal stimulation.

DISCUSSION

This work has revealed a surprising lack of similarity between the human and mouse PAH regulatory regions. This is an unexpected finding in view of the conservation during evolution of the liver specificity of expression of the gene. The similarities between the human (50) and mouse genes are essentially limited to the fact that both present multiple sites of initiation of transcription and absence from the proximal promoter of TAATA and CCAAT consensus sequences.

Three major differences between the human and mouse regulatory regions can be distinguished. First, the human promoter and upstream regions, up to kb −9, are active in a transitory expression assay in PLC (Alexander) human hepatoma cells. Successive 5′ deletions give rise to reductions and recoveries of activity, with 50% of the initial activity remaining when only the sequence up to bp −121 is retained. While the reporter gene activities were not compared with those of known viral promoters, the values were around 100-fold those obtained with a high-background promoterless CAT vector (25) and therefore must be considerable (50). In contrast, no construct of the mouse PAH promoter and upstream regions, up to kb −7 and including successive 5′ deletions,

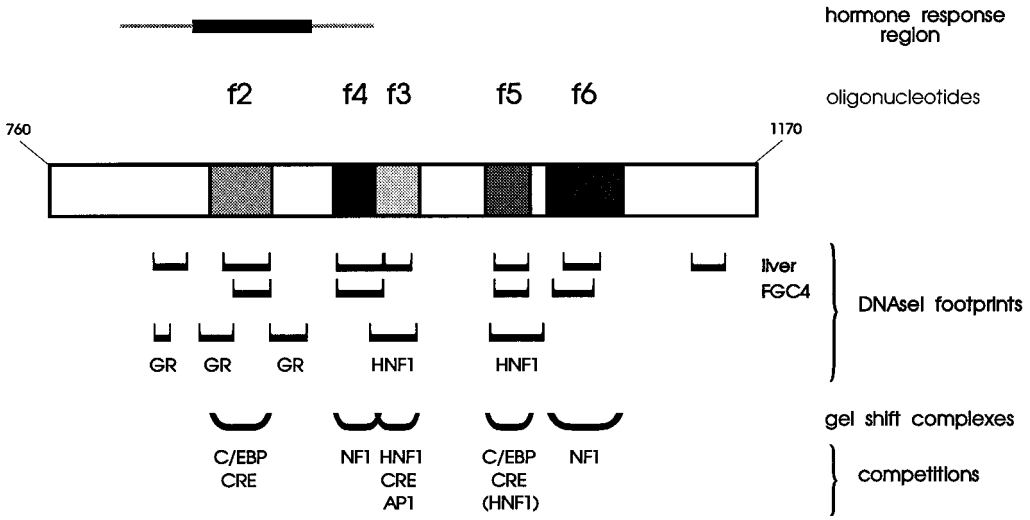


FIG. 10. Schematic representation of the HSIII enhancer based upon transient-expression assays of 5′ and 3′ deletions, DNase I footprint analysis, gel mobility shift assays, and competition experiments.

shows significant activity, irrespective of the origin of the host cell.

Second, the human PAH promoter shows cell type specificity of expression: no activity is detected in nonhepatic cells (50). The extremely weak activity of the mouse PAH promoter is the same in rat hepatoma cells that express the PAH gene as it is in cells that do not express this gene and is slightly higher in mouse L-cell fibroblasts.

Third, analysis of the bp −121 fragment of the human proximal promoter revealed the existence of binding sites for unidentified ubiquitous proteins, designated the B and A boxes. Mutation of these sequences reduces promoter activity, and a single copy of the B or A box activates a heterologous promoter, but not in a cell-type-specific fashion (50). Sequences homologous to the B and A boxes do not exist in the mouse promoter, and partial homologies found at bp −651 to −628 and bp −518 to −507, respectively, do not include the nucleotides identified as crucial for binding (50).

It could be argued that the lack of similarity in the two genes is not unexpected, since their timings of expression are different: the human gene is first expressed during the first trimester of development, and the mouse gene is expressed only just before birth. However, it is informative to recall the expression pattern of the reporter gene under the control of the human PAH regulatory region from kb −9 to bp +140 in transgenic mice (49). The transgene obeys the rules of the mouse PAH gene: it is expressed only at birth and is present in both liver and kidney tissues. It has been concluded that the upstream region of the human PAH gene contains all information required to respond to mouse regulatory factors and that differences between the two species are due to cellular environment and not to sequences governing tissue-specific expression (49). Therefore, the possibility that the human PAH regulatory region will turn out to present an overall structure similar to that of the mouse gene cannot be excluded. The possibility that the human gene could be hormone inducible is of potential interest for those PAH mutations for which residual enzymatic activity is retained in the mutant protein.

The TAATA- and CCAAT-less promoters may tend to precede housekeeping genes, but they are not unprecedented for tissue-specific genes, such as the macrophage scavenger receptor gene (36). A number of liver-specific genes of this type have been described, including those for xanthine dehydrogenase oxidase (12), prothrombin (11), serine dehydratase (42), HNF4 (52), and human PAH (50). While the information to compare the relative promoter activities of members of this group of genes is not available, it can be noted that both prothrombin and HNF4 promoters present only modest activities in transfection tests. The prothrombin gene (11), like the mouse PAH gene (as shown here), contains a functional HNF1 site within an enhancer, whereas sites for this factor are more frequently encountered in proximal promoters (46), including that in the HNF4 gene (52). The genes coding for tyrosine or tryptophan hydroxylase, whose proteins are closely related to PAH, show tissue-specific expression and are hormone inducible. However, the promoter structures of these three genes are different (5, 32).

Upstream enhancer elements involved in the regulation of mouse PAH expression were localized within the sequences surrounding DNase I-hypersensitive sites HSSIII (kb −3.5) and HSSIV (kb −8.8). However, both enhancers showed activity only when upstream of a heterologous promoter and not upstream of the PAH promoter. For the HSIII enhancer, this activity was found to be restricted to cells of hepatic origin. Furthermore, response elements for both glucocorticoids and cAMP were identified within the HSIII region. Importantly, in the presence of hormones and the HSIII enhancer it was finally possible to detect PAH promoter activity: the HSIII enhancer confers on the PAH promoter a 150-fold induction following treatment with hormones. This finding defines the HSIII enhancer as a conditional one, dependent on hormones for activity when associated with its own promoter but not when upstream of a heterologous promoter.

DNase I footprint analysis and gel mobility shift assays provided evidence for binding of the ubiquitous factors NF1, glucocorticoid receptor, HNF1, and C/EBP. With one exception, the binding sites for the liver-enriched transcription factors do not correspond to consensus binding sites and are, therefore, potentially weak sites. The significance of such weak sites could be questioned. Their importance was, however, revealed by the fact that mutation of each led to reduced basal activity and to a severe decrease in hormone inducibility. The fact that even when the PAH promoter is flanked by the HSIII enhancer, hormones are critical for PAH promoter activity reinforces this conclusion.

The structure of the PAH regulatory region is different in some aspects from those of two other hormone-responsive genes first expressed in the neonatal liver, the TAT and PEPCK genes. In the PEPCK gene, promoter and enhancer are entirely overlapping in a 600-bp segment; this fragment shows a high level of activity in the absence of hormones. Induction by cAMP is mediated by proximal elements, and glucocorticoid inducibility is mediated by more-distal sites (28). The regulatory elements of the TAT gene are spread out: one hormone response element-containing enhancer, located at kb −2.5, binds the glucocorticoid receptor and requires for maximal activity the collaboration of HNF3, and another enhancer at kb −3.6 requires CREB and HNF4. It is important to note that both enhancers confer significant activity in the absence of hormones in transfection tests with hepatoma cells (37). In the PAH gene, responsiveness to both hormones is embedded in a single upstream enhancer, but this enhancer shows no activity in the absence of hormones when flanking its own promoter. DEX and cAMP response elements must be located close to each other, since none of the HSIII enhancer deletions showed selective loss of either DEX or cAMP inducibility. In contrast to the well-conserved binding sites found in the regulatory elements of both the TAT and the PEPCK genes, the PAH HSIII enhancer is composed of sites exhibiting only weak homology to high-affinity sites. In spite of the differences in the regulatory regions of these three genes, they do share the property of cooperation between liver-enriched and hormone response element-binding proteins for activity (PAH) or for maximal hormone-induced activity (PEPCK and TAT).

A real surprise concerning the PAH gene and its regulation came from the study of mice whose HNF1 genes had been disrupted. Of all liver-specific genes examined in the deficient animals, only the PAH gene fails to be expressed (38). This result indicates an absolute requirement for HNF1 for activation of expression of the PAH gene. The HNF1 sites we have identified in the HSIII enhancer are therefore potentially important not only for the maintenance of expression of the PAH gene but also for its activation during the course of development. Although C/EBPα-deficient mice have been generated, the expression of PAH has not been analyzed (48). It is intriguing that C/EBPα expression first occurs shortly before birth and is maintained at a high level in resting hepatocytes (3) and other terminally differentiated cells, such as adipocytes (7).

Taking these data together, we propose that in the mouse PAH gene high-affinity sites are replaced with a multiplicity of

weaker sites, resulting in a highly responsive transcription unit (see also the discussion in reference 46) that requires the cooperation of multiple transcription factors for activity. The strict dependence of the strongly inducible promoter upon both the upstream HSIII enhancer and the presence of hormones is consistent with this notion. We suggest that activation of expression of the PAH gene requires the simultaneous participation of most, if not all, of the four elements implicated in tissue specificity and hormonal regulation: HNF1, C/EBP, glucocorticoids, and cAMP. Future experiments will aim to define the possible contribution of the HSIV enhancer, as well as the mechanisms of hormone inducibility, of the synergism observed with the combination of two hormones, and of the roles of each of the players in the activation as well as the maintenance of expression of the PAH gene.

ACKNOWLEDGMENTS

We are grateful to Silvia Cereghini and to Susan Power for providing the liver protein extracts and the plasmid ΔTK-H, for their interest in this work, and for many stimulating discussions. We acknowledge the generous gift of the bacterially expressed HNF1 protein from Tanguy Chouard and of glucocorticoid receptor protein from Christian Muchart. We thank Marco Pontoglio and Moshe Yaniv for communicating unpublished results and for critical reading of the manuscript. We thank Céline Mulet for technical help and Solange Papelard for her help in preparing the manuscript.

Laurent Belkadi was supported by a fellowship from La Ligue Française contre le Cancer. This work was supported by the Biotechnology program of the EEC under contract number BIOT CT 93-0103.

REFERENCES

1. **Angrand, P. O., C. Coffinier, and M. C. Weiss.** 1994. Response of the phosphoenolpyruvate carboxykinase gene to glucocorticoids depends on the integrity of the cAMP pathway. Cell Growth Differ. **5:**957–966.
2. **Angrand, P. O., S. Kallenbach, M. C. Weiss, and J. P. Rousset.** 1990. An exogenous albumin promoter can become silent in dedifferentiated hepatoma variants as well as intertypic hybrids. Cell Growth Differ. **1:**519–526.
3. **Birkenmeier, E. H., B. Gwynn, S. Howard, J. Jerry, J. I. Gordon, W. H. Landschulz, and S. L. McKnight.** 1989. Tissue-specific expression, developmental regulation, and genetic mapping of the gene encoding CCAAT/enhancer binding protein. Genes Dev. **3:**1146–1156.
4. **Boshart, M., M. Kluppel, S. Schmidt, G. Schütz, and B. Luckow.** 1992. Reporter constructs with low background activity utilizing the cat gene. Gene **110:**129–130.
5. **Boularand, S., M. C. Darmon, P. Ravassard, and J. Mallet.** 1995. Characterization of the human tryptophan hydroxylase gene promoter. J. Biol. Chem. **270:**3757–3764.
6. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248–254.
7. **Cao, Z., R. M. Umek, and S. L. McKnight.** 1991. Regulated expression of three C/EBP isoforms during adipose conversion of 3T3-L1 cells. Genes Dev. **5:**1538–1552.
8. **Cereghini, S., M. Blumenfeld, and M. Yaniv.** 1988. A liver-specific factor essential for albumin transcription differs between differentiated and dedifferentiated rat hepatoma cells. Genes Dev. **2:**957–974.
9. **Cereghini, S., M. Raymondjean, A. G. Carranca, P. Herbomel, and M. Yaniv.** 1987. Factors involved in control of tissue-specific expression of albumin gene. Cell **50:**627–638.
10. **Cereghini, S., M. Yaniv, and R. Cortese.** 1990. Hepatocyte dedifferentiation and extinction is accompanied by a block in the synthesis of mRNA coding for the transcription factor HNF1/LFB1. EMBO J. **9:**2257–2263.
11. **Chow, B. K. C., V. Ting, F. Tufaro, and R. T. A. MacGillivray.** 1991. Characterization of a novel liver-specific enhancer in the human prothrombin gene. J. Biol. Chem. **266:**18927–18933.
12. **Chow, C. W., M. P. Clark, J. E. Rinaldo, and R. Chalkley.** 1995. Multiple initiators and C/EBP binding sites are involved in transcription from the TATA-less rat XDH/XO basal promoter. Nucleic Acids Res. **23:**3132–3140.
13. **Church, G., and W. Gilbert.** 1984. Genomic sequencing. Proc. Natl. Acad. Sci. USA **81:**1991–1995.
14. **Citron, B. A., M. D. Davis, S. Milstien, J. Gutierrez, D. B. Mendel, G. R. Crabtree, and S. Kaufman.** 1992. Identity of 4α-carbinolamine dehydratase, a component of the phenylalanine hydroxylation system, and DCoH, a transregulator of homeodomain proteins. Proc. Natl. Acad. Sci. USA **89:**11891–11894.
15. **Coon, H. G., and M. C. Weiss.** 1969. A quantitative comparison of formation of spontaneous and virus produced viable hybrids. Proc. Natl. Acad. Sci. USA **62:**852–859.
16. **Deschatrette, J., and M. C. Weiss.** 1974. Characterization of differentiated and dedifferentiated clones from a rat hepatoma. Biochimie **56:**1603–1611.
17. **De Wet, J., K. V. Wood, M. DeLuca, D. R. Helinski, and S. Subramani.** 1987. Firefly luciferase gene: structure and expression in mammalian cells. Mol. Cell. Biol. **7:**725–737.
18. **DiLella, A. G., S. C. M. Kwok, F. D. Ledley, J. Marvit, and S. L. C. Woo.** 1986. Molecular structure and polymorphic map of the human phenylalanine hydroxylase gene. Biochemistry **25:**743–749.
19. **Dubbs, D. R., and S. Kit.** 1964. Effect of halogenated pyrimidines and thymidine on growth of L cells and sublines lacking thymidine kinase. Exp. Cell Res. **33:**19–28.
20. **Eisensmith, R. C., and S. L. C. Woo.** 1991. Phenylketonuria and the phenylalanine hydroxylase gene. Mol. Biol. Med. **8:**3–18.
21. **Faust, D. M., M. Boshart, T. Imaizumi-Scherrer, G. Schütz, and M. C. Weiss.** 1994. Constancy of expression of the protein kinase A regulatory subunit R1α in hepatoma cell lines of different phenotypes. Cell Growth Differ. **5:**47–53.
22. **Faust, D. M., T. Imaizumi-Scherrer, M. C. Fulchignoni-Lataud, A. M. Catherin, I. Iost, and M. C. Weiss.** 1990. Activation of phenylalanine hydroxylase expression following genomic DNA transfection of hepatoma cells. Differentiation **44:**74–79.
23. **Freedman, L. P., S. K. Yoshinaga, J. N. Vanderbilt, and K. R. Yamamoto.** 1989. In vitro transcription enhancement by purified derivatives of the glucocorticoid receptor. Science **245:**298–301.
24. **Friedman, A. D., W. H. Landschulz, and S. L. McKnight.** 1989. CCAAT/enhancer binding protein activates the promoter of the serum albumin gene in cultured hepatoma cells. Genes Dev. **3:**1314–1322.
25. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2:**1044–1051.
26. **Haggerty, D. F., F. Chiappelli, R. Kern, S. Scully, and M. Lynch.** 1983. Regulation by glucocorticoids of rat-liver phenylalanine hydroxylase in vivo. Biochem. Biophys. Res. Commun. **115:**965–970.
27. **Haggerty, D. F., P. L. Young, G. PopJak, and W. H. Carnes.** 1973. Phenylalanine hydroxylase in cultured hepatocytes. I. Hormonal control of enzyme levels. J. Biol. Chem. **248:**223–232.
28. **Imai, E., J. N. Miner, J. A. Mitchell, K. R. Yamamoto, and D. K. Granner.** 1993. Glucocorticoid receptor-cAMP response element-binding protein interaction and the response of the phosphoenolpyruvate carboxykinase gene to glucocorticoids. J. Biol. Chem. **268:**5353–5356.
29. **Kaufman, S.** 1987. Phenylalanine-4-monooxygenase from rat liver. Methods Enzymol. **142:**3–29.
30. **Konecki, D. S., Y. Wang, F. K. Trefz, U. Lichter-Konecki, and S. L. C. Woo.** 1992. Structural characterisation of the 5′ regions of the human phenylalanine hydroxylase gene. Biochemistry **31:**8363–8368.
31. **Lavery, D. J., and U. Schibler.** 1993. Circadian transcription of the cholesterol 7α hydroxylase gene may involve the liver-enriched bZIP protein DBP. Genes Dev. **7:**1871–1884.
32. **Lewis, E. J., C. A. Harrington, and D. M. Chikaraishi.** 1987. Transcriptional regulation of the tyrosine hydroxylase gene by glucocorticoid and cyclic AMP. Proc. Natl. Acad. Sci. USA **84:**3550–3554.
33. **McDowall, I. L., and M. J. Fisher.** 1995. The immediate 5′-flanking region of the rat phenylalanine hydroxylase-encoding gene. Gene **153:**289–290.
34. **McGee, M. M., O. Greengard, and W. E. Knox.** 1972. The quantitative determination of phenylalanine hydroxylase in rat tissues—its developmental formation in the liver. Biochem. J. **127:**669–674.
35. **Mitnaul, L. J., and R. Shiman.** 1995. Coordinate regulation of tetrahydrobiopterin turnover and phenylalanine hydroxylase activity in rat liver cells. Proc. Natl. Acad. Sci. USA **92:**885–889.
36. **Moulton, K. S., H. Wu, J. Barnett, S. Parthasarathy, and C. K. Glass.** 1992. Regulated expression of the human acetylated low density lipoprotein receptor gene and isolation of promoter sequences. Proc. Natl. Acad. Sci. USA **89:**8102–8106.
37. **Nitsch, D., M. Boshart, and G. Schütz.** 1993. Activation of the tyrosine aminotransferase gene is dependent on synergy between liver-specific and hormone-responsive elements. Proc. Natl. Acad. Sci. USA **90:**5479–5483.
38. **Pontoglio, M., J. Barra, M. Hadchouel, A. Doyen, C. Kress, J. Poggi Bach, C. Babinet, and M. Yaniv.** 1996. Hepatocyte nuclear factor 1 (HNF1) inactivation results in hepatic dysfunction phenylketonuria and renal Fanconi syndrome. Cell **84:**575–585.
39. **Rey-Campos, J., T. Chouard, M. Yaniv, and S. Cereghini.** 1991. vHNF1 is a homeoprotein that activates transcription and forms heterodimers with HNF1. EMBO J. **10:**1445–1457.
40. **Scriver, C. R., S. Kaufman, and S. L. C. Woo.** 1988. Mendelian hyperphenylalaninemia. Annu. Rev. Genet. **22:**301–321.
41. **Seed, B., and J. Y. Sheen.** 1988. A simple phase-extraction assay for chloramphenicol acetyltransferase activity. Gene **67:**271–277.
42. **Su, Y., R. Kanamoto, H. Ogawa, and H. C. Pitot.** 1992. Regulatory ele-

ments for the tissue-specific expression of the rat serine dehydratase-encoding gene. Gene **120:**301–306.

43. **Szpirer, C., and J. Szpirer.** 1975. A mouse hepatoma cell line which secretes several serum proteins including albumin and α-fetoprotein. Differentiation **4:**85–91.
44. **Tourian, A.** 1973. Phenylalanine hydroxylase activity in foetal hepatic organ culture. Biochim. Biophys. Acta **309:**44–49.
45. **Tourian, A., D. M. Treiman, and J. S. Carr.** 1972. Developmental biology of hepatic phenylalanine hydroxylase activity in foetal and neonatal rats synchronized as to conception. Biochim. Biophys. Acta **279:**484–490.
46. **Tronche, F., F. Ringeisen, M. Blumenfeld, M. Yaniv, and M. Pontoglio.** Analysis of the distribution of targets for a tissue-specific transcription factor in the vertebrate genome. Submitted for publication.
47. **Tronche, F., A. Rollier, I. Bach, M. C. Weiss, and M. Yaniv.** 1989. The rat albumin promoter: cooperation with upstream elements is required when binding of APF/HNF1 to the proximal element is partially impaired by mutation or bacterial methylation. Mol. Cell. Biol. **9:**4759–4766.
48. **Wang, N., M. J. Finegold, A. Bradley, C. N. Ou, S. V. Abdelsayed, M. D. Wilde, L. R. Taylor, D. R. Wilson, and G. J. Darlington.** 1995. Impaired energy homeostasis in C/EBPα knockout mice. Science **269:**1108–1112.
49. **Wang, Y., J. L. DeMayo, T. M. Hahn, M. J. Finegold, D. S. Konecki, U. Lichter-Konecki, and S. L. C. Woo.** 1992. Tissue- and development-specific expression of the human phenylalanine hydroxylase/chloramphenicol acetyltransferase fusion gene in transgenic mice. J. Biol. Chem. **267:**15105–15110.
50. **Wang, Y., T. M. Hahn, S. Y. Tsai, and S. L. C. Woo.** 1994. Functional characterization of a unique liver gene promoter. J. Biol. Chem. **269:**9137–9146.
51. **Yeoh, G. C. T., E. Edkins, K. Mackenzie, S. Fuller, J. F. B. Mercer, and H. H. M. Dahl.** 1988. The development of phenylalanine hydroxylase in rat liver: in vivo and in vitro studies utilizing fetal hepatocyte cultures. Differentiation **38:**42–48.
52. **Zhong, W., J. Mirkovitch, and J. E. Darnell, Jr.** 1994. Tissue-specific regulation of mouse hepatocyte nuclear factor 4 expression. Mol. Cell. Biol. **14:**7276–7284.