MOLECULAR AND CELLULAR BIOLOGY, May 1990, p. 1841–1852
0270-7306/90/051841-12$02.00/0

Vol. 10, No. 5

Analysis of the Upstream Elements of the Xenobiotic Compound-Inducible and Positionally Regulated Glutathione *S*-Transferase Ya Gene

K. ERIC PAULSON,[1] JAMES E. DARNELL, JR.,[1]* TOM RUSHMORE,[2] AND CECIL B. PICKETT[2]

Laboratory of Molecular Cell Biology, The Rockefeller University, 1230 York Avenue, New York, New York 10021,[1] and Merck Frosst Centre for Therapeutic Research, Pointe-Claire-Dorval, Quebec, Canada H9R 4P8[2]

Received 7 November 1989/Accepted 10 January 1990

In situ hybridization and other data showed that all hepatocytes express glutathione-*S*-transferase (GST) Ya mRNA but that specifically pericentral cells can be induced 15- to 20-fold with 3-methylcholanthrene (3-MC). In order to identify DNA sequences involved in inducible expression (pericentral hepatocytes) and constitutive expression (all hepatocytes), the upstream regions of the GST Ya gene were further analyzed by transient transfection and DNA-binding studies to identify the nature of proteins involved in regulating this gene. The sequences from −980 to −650 were necessary and sufficient for cell-specific and inducible expression. Within this enhancer region, four nuclear protein-binding sites were identified. One site required for inducible expression was bound by a protein(s) induced by 3-MC. Two other sites were bound by proteins similar or identical to the constitutive hepatocyte nuclear factors HNF1 and HNF4. The fourth site was shown to be bound by a non-liver-specific nuclear protein that is also important in the function of the albumin gene enhancer.

Liver cells, or hepatocytes, are thought to be part of the same cell lineage and are characterized by the coordinate cell-specific transcription of genes encoding a variety of proteins and enzymes. Many of these proteins and their mRNAs are produced equally by all hepatocytes, with albumin being a typical example (3, 35). However, a number of proteins and their mRNAs show region-specific differences within the liver acini. For example, glutamine synthetase (GS) is expressed in adult rodents only in the immediate cell layer surrounding the central vein (3, 14, 24). In addition, heightened pericentral expression is characteristic of several enzymes involved in detoxification including some P450 isoenzymes, epoxide hydrolase, and some glutathione-*S*-transferase (GST) isoenzymes (30, 36, 41). Antibody staining reveals that some of these enzymes appear to be present in all hepatocytes at a low level. After animals are treated with xenobiotic metabolites, however, these proteins are found in high levels in the pericentral 50% of hepatocytes. In other cases, it appears that expression occurs only after induction and that expression is pericentral. Pericentral expression following induction in all cases appears by antibody staining to be sharply demarcated. Finally, some proteins and their mRNAs (for example, several enzymes of the urea cycle) are strictly localized to the periportal area. Such all-or-none positional expression for several different sets of proteins seems unlikely to be due to gradients of metabolites, as previously postulated (22). In fact, we observed localized differentiation of the pericentral proteins major urinary protein (MUP) and GS in cultured fetal hepatocytes where no gradients for soluble metabolites exist (3). Therefore, region-specific distribution patterns most likely represent positional regulation of gene expression within the same cell lineage due to other signals, for example, extracellular cell surface contacts.

The GSTs are a family of isoenzymes composed of homo- and heterodimers of at least seven subunits. Several of these subunits are induced by xenobiotic compounds (35). In this article, we extend the histochemical analysis of the Ya subunit of GST to the transcriptional level. Using in situ hybridization, we demonstrate that the GST Ya mRNA is present in all hepatocytes in uninduced mice and is specifically induced by 3-methylcholanthrene (3-MC) in the pericentral region of the liver. In order to determine the genetic elements important for both basal expression (all hepatocytes) and inducible expression (pericentral hepatocytes), we dissected the upstream region of the GST Ya gene by using transient transfections in hepatoma cells and identified binding sites for proteins potentially important in both basal and inducible expression.

MATERIALS AND METHODS

Cells. HepG2 cells were maintained as monolayer cultures and grown in Ham F12 medium supplemented with 5% heat-inactivated fetal calf serum (Hyclone), 0.5× minimal essential amino acids (GIBCO Laboratories), garamycin (Schering Corp.) at 25 μg/ml, and insulin (Eli Lilly & Co.) at 0.5 U/ml. HeLa, 293, WISH, and FS2 cells were maintained as monolayer cultures and grown in Dulbecco modified Eagle medium supplemented with 10% heat-inactivated fetal calf serum.

In situ hybridization. Liver tissues were perfused with 4% paraformaldehyde in 1× PBS in situ, cut into small pieces, and then fixed overnight in 4% paraformaldehyde in 1× PBS at 4°C. The tissue was then equilibrated overnight in 30% sucrose at 4°C and mounted in Tissue-Tek O.C.T. compound for frozen sectioning. Sections (5 μm) were hybridized as described by Kuo et al. (24) with an antisense ^{35}S-labeled probe prepared from a T7 RNA transcript of a rat GST Ya cDNA (34). Following hybridization and washing, the sections were exposed on Kodak NTB-2 emulsion (typically 2 to 4 days at 4°C). After the emulsion was developed, the slides were stained with standard hematoxylin-eosin stain and examined by bright- and dark-field microscopy.

GST Ya plasmid constructions and transfections. The −4.5-kilobase-pair (kbp) GST Ya construct was prepared from the GST Ya gene by the addition of a *Bam*HI linker at the *Sma*I

* Corresponding author.

site (+66 bp) in the first intron (T. H. Rushmore, R. G. King, K. E. Paulson, and C. B. Pickett, Proc. Natl. Acad. Sci. USA, in press), followed by cleavage with *Eco*RI (−4.5 kbp) and *Bam*HI. The resulting −4.6-kbp fragment containing −4.5 kbp of upstream sequence and +66 bp of downstream sequence was cloned into the adenovirus E1B expression vector (5) by standard methods (28). The −1.55-kbp GST Ya construct was prepared from the −4.5-kbp GST Ya construct by the addition of an *Eco*RI linker at the *Xba*I site (−1.55 kbp), followed by cleavage with *Eco*RI and *Bam*HI. The resulting −1.65-kbp fragment containing −1.55 kbp of upstream sequence and +66 bp of downstream sequence was cloned into the adenovirus E1B expression vector. Additional 5′ and internal deletions were constructed from the −1.55-kbp construct by using advantageous restriction enzyme cleavage sites. Bal 31 exonuclease III was used to produce progressive 5′ deletions starting with the −1.55-kbp GST Ya construct as described previously (5). Digestion products were removed every 20 s for 4 min, and after termination of the reaction, the ends were repaired and *Eco*RI linkers were ligated (28). The products were cleaved with *Eco*RI and size-separated on a 1% low-melting-point agarose gel, isolated, and cloned by standard methods (28) into the adenovirus E1B expression vector (gift of Lee E. Babiss). After transformation of *Escherichia coli* MC1061, clones were analyzed by restriction analysis and deletion end points were determined by DNA sequencing (37). The GST Ya enhancer fragment from −980 bp to −650 bp was cloned into the β-globin promoter plasmid (5) at −341 bp in either orientation by using ligated *Hin*dIII linkers. The GST Ya 3-MC-inducible-site oligonucleotide (−910 bp to −875 bp) was cloned into the −190-bp deletion at the *Eco*RI site created by the linker.

Cultured cells were transfected at 50% confluency in Dulbecco modified Eagle medium supplemented with 10% heat-inactivated fetal calf serum, 1× nonessential amino acids, and insulin at 0.5 U/ml, using the calcium phosphate coprecipitation method as described previously (5). Expression levels of the GST Ya constructs were determined by RNase T_2 digestion of hybrids between transfected RNA (20 μg) and labeled antisense RNA probes specific to the GST Ya first exon as described previously. The RNase-resistant hybrids were analyzed on 8 M urea–6% polyacrylamide gels, followed by autoradiography. An appropriate control plasmid (5) was included in the transfections to monitor transfection efficiency.

Protein-DNA-binding assays. Nuclear extracts were prepared from cultured cells as described by Dignam et al. (10) with a final NaCl concentration of 0.4 M to extract proteins from the nuclei. Nuclear extracts were prepared from fresh tissue by the method of Gorski et al. (17). Protein-DNA binding was carried out as described by Costa et al. (6) and analyzed on a 6% polyacrylamide gel, followed by autoradiography.

Methylation protection and exo III experiments. High-specific-activity 3′-end-labeled probes were prepared as described by Costa et al. (6). Methylation protection experiments with HepG2 nuclear extracts were also performed as described by Costa et al. (6). The 3-MC-inducible protein-DNA interaction was footprinted by using partially purified protein (gift of A. S. Yee). In vitro exo III reactions were performed with 1 ng of a uniquely end-labeled probe (extending from −815 bp to −725 bp of the GST Ya upstream region) by the method of Kovesdi et al. (23). For all reactions, 4 μg of nuclear extract was used, as was 4 μg of poly(dI-dC) (Pharmacia, Inc.) as a nonspecific competitor of DNA-binding proteins.

RESULTS

In situ hybridization analysis of GST mRNA distribution. Immunohistochemical staining of 3-MC-treated and untreated animals shows a large increase in total GST protein in the pericentral hepatocytes of treated animals (41). We investigated whether the changes induced by 3-MC on GST protein distribution were reflected in changes in the GST Ya mRNA distribution. A ^{35}S-labeled antisense probe to GST Ya mRNA was used for in situ hybridization of 3-MC-treated and untreated liver sections. The GST Ya mRNA was evenly distributed throughout the liver of uninduced animals (Fig. 1B and C). However, a dramatic increase in mRNA was observed in 3-MC-induced animals (Fig. 1F). Treatment with 3-MC results in a 7- to 10-fold increase in total mRNA (34) (data not shown). Because the induction is apparently restricted only to pericentral cells (approximately 50% of all hepatocytes), the true induction is 15- to 20-fold in those cells. It is noteworthy that the induction by 3-MC was not rapid. Three hours after initial treatment with 3-MC, there was a marginal change in GST mRNA distribution (Fig. 1D and E). Only after 48 h and multiple 3-MC injections was the full pattern of induction observed (Fig. 1F). Similar kinetics were observed for total GST mRNA induction (data not shown). Whether this is due to the low solubility of 3-MC and the consequent slow delivery to the liver is not clear, but as demonstrated later, cultured cells responded more quickly.

As a control for the specificity of the in situ hybridization, an antisense RNA probe for the generally distributed transthyretin (TTR) protein was used. TTR probe gave a uniform distribution in liver sections, as expected (data not shown). In addition, an RNA probe for MUP was used as a control for pericentral expression. MUP is a constitutively expressed, predominantly pericentral protein (3), and this was also reflected in the mRNA distribution pattern (Fig. 1A).

Localization of GST Ya gene regulatory elements in HepG2 (human hepatoma) cells. In a preliminary analysis of the GST Ya upstream region, deletion from −4.5 kbp to −1.6 kbp did not affect the basal or xenobiotic compound-inducible expression of RNA from plasmid constructs transfected into a variety of hepatoma cell lines, whereas a deletion of −650 bp had no basal or inducible activity (39). Further, deletion analysis (Rushmore et al., in press) located DNA segments in the −650 bp to −750 bp region that are responsible for either induced or basal expression of the Ya mRNA. To determine the possible importance of other DNA segments upstream of −750 bp and to characterize the exact binding sites and cognate proteins, another series of deletions between −1.55 kbp and −650 bp were prepared. Transfection of these deletion mutants into HepG2 cells was carried out to test for both basal and 3-MC-inducible activity. In all experiments, a control plasmid containing a simian virus 40 (SV40) enhancer/β-globin promoter minigene was cotransfected to quantitate transfection efficiency (5). A 5′ deletion to −980 bp retained full basal activity and 3-MC-inducible activity (Fig. 2). Deletion to −875 bp had little effect on basal activity but caused a loss of about half of the induced activity, indicating a possible inducible element. Further internal deletions of the plasmid with −1.2 kbp upstream located the inducible element. A deletion between −875 bp and −650 bp retained inducibility, although no basal activity was now detectable. Expanding this deletion by only 50 additional

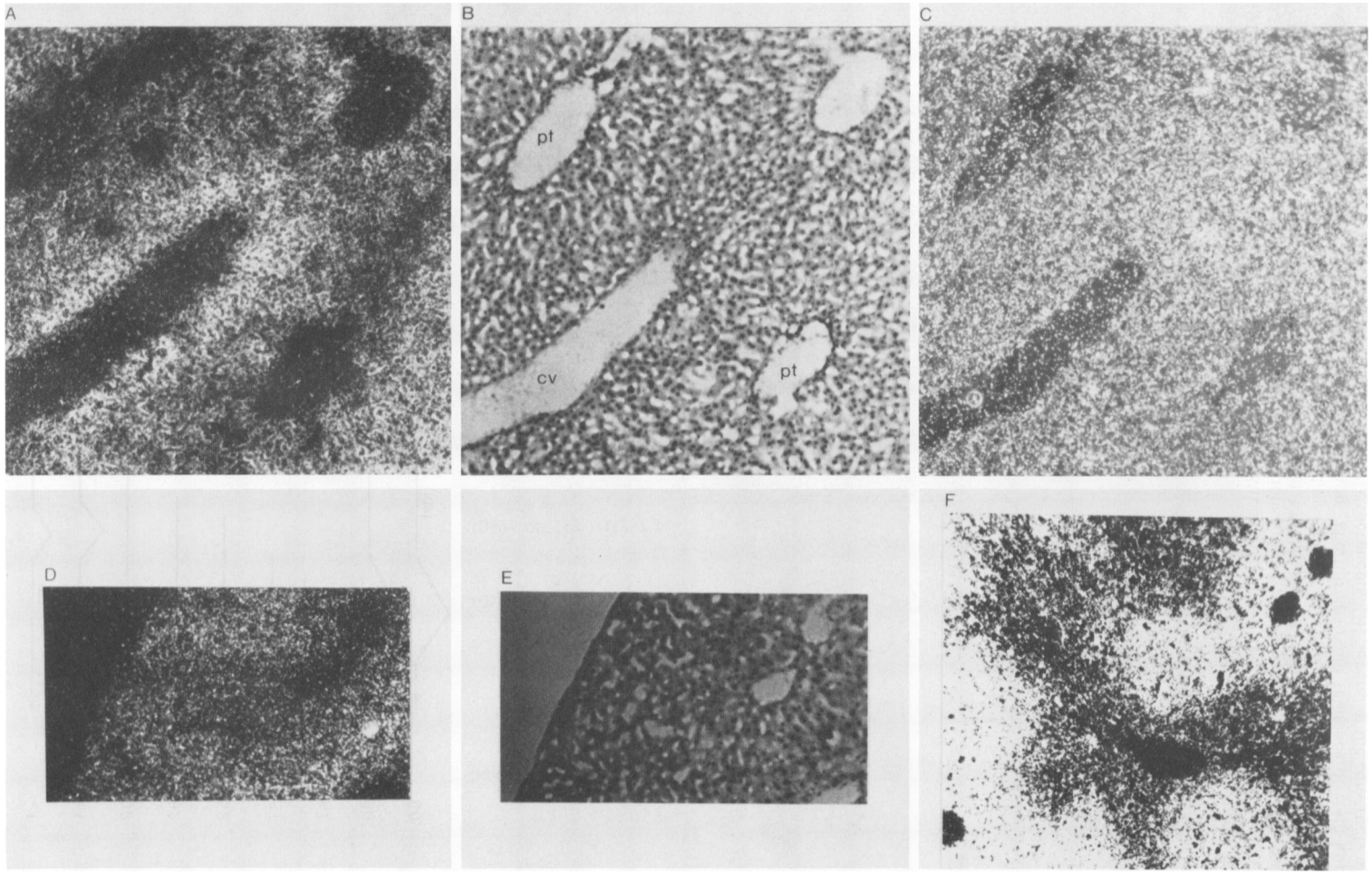


FIG. 1. In situ hybridization of 3-MC-induced and uninduced mouse liver tissue. (A) Uninduced female, dark-field exposure with the pericentral control MUP probe. (B) Bright-field of panel A. (C) Serial section following panel A; uninduced female, dark-field exposure with a GST Ya probe. (D) Female induced for 3 h with 3-MC, dark-field exposure with a GST Ya probe. (E) Bright-field of panel D. (F) Female induced for 48 h with 3-MC, dark-field exposure with a GST Ya probe. After autoradiography, the slides were routinely stained with hematoxylin-eosin stain and dehydrated. cv, Central vein; pt, portal triad. Magnification, ×100.

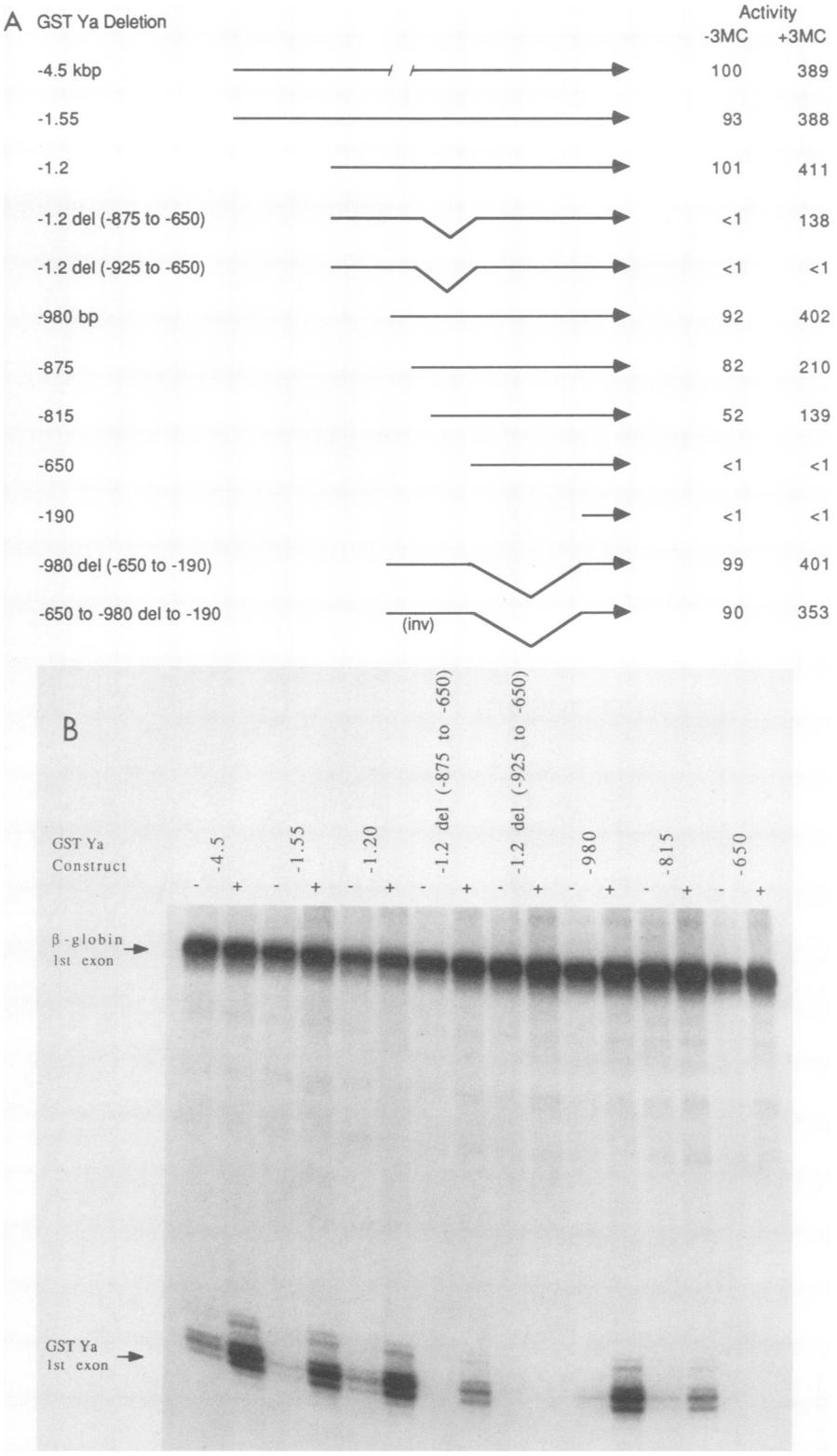
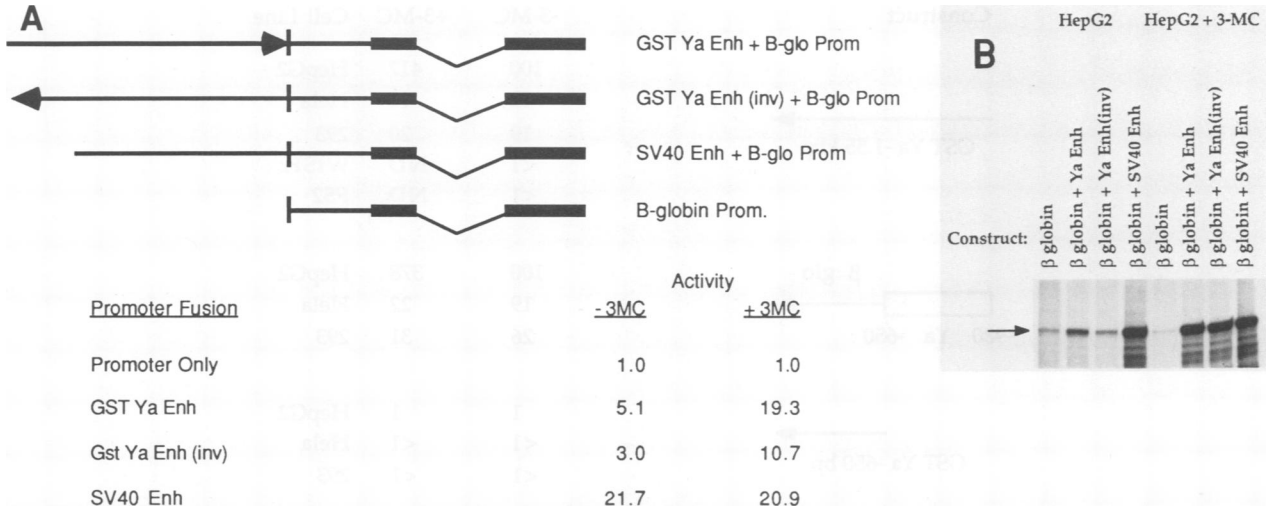


FIG. 2. GST Ya enhancer defined by deletions of the GST Ya upstream region. (A) The size of progressive upstream deletions is indicated on the left and is graphically illustrated in the center. Internal deletions are indicated by del and are also graphically illustrated by the bent bridge. The transcriptional activity of the constructs with and without 3-MC induction is represented at the right. The value of each point was determined by densitometry tracings of results similar to and including those shown in panel B. Each value was normalized to the β-globin transfection control signal before comparison for GST Ya expression. The full upstream sequence of −4.5 kbp is taken as 100% activity without induction. Each point is the average of at least two individual experiments. (B) Several of the constructs were transfected into HepG2 cells along with a control SV40 enhancer/β-globin promoter plasmid (5). Cells were transfected by the $CaPO_4$ precipitation method. Following removal of the precipitate, the cells were allowed to recover for 4 h and then induced for 12 h with 2 μM 3-MC. Following induction, cytoplasmic RNA was extracted and assayed with two separate antisense RNA probes specific for the GST Ya first exon and the mouse β-globin first exon. RNase T_2-protected products were separated on 6% polyacrylamide–urea gels, and the GST Ya and β-globin exons are indicated by arrows. Pairs of lanes are indicated with (+) or without (−) 3-MC treatment and with the GST Ya test plasmid used.



Promoter Fusion	Activity - 3MC	Activity + 3MC
Promoter Only	1.0	1.0
GST Ya Enh	5.1	19.3
Gst Ya Enh (inv)	3.0	10.7
SV40 Enh	21.7	20.9

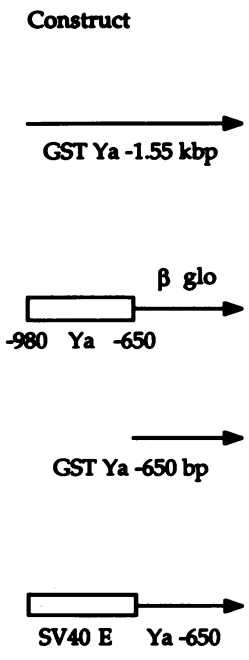
FIG. 3. Stimulation of the β-globin promoter by GST Ya enhancer sequences (−980 bp to −650 bp). (A) Recombinant plasmids of the mouse β-globin promoter (B-glo Prom) construct (5) were prepared containing the GST Ya enhancer (Enh) sequences −980 bp to −650 bp in either orientation (inv, inverted). In addition, the SV40 enhancer was cloned next to the promoter as an enhancer control. The transcriptional activity of the constructs in HepG2 cells without and with 3-MC induction is presented on the bottom. Each value was normalized to the cotransfection control signal before comparison for β-globin promoter construct expression. The basal activity of the β-globin promoter is taken arbitrarily as 1.0. (B) The various constructs were transfected into HepG2 cells along with a control plasmid; the TTR promoter-enhancer minigene (6). The cells were induced with 3-MC and assayed as described in the legend to Fig. 2. The arrow indicates the β-globin first exon.

bases (total deletion −925 bp to −650 bp) completely eliminated 3-MC inducibility. These deletions definitively demonstrated that a 3-MC-inducible element lies between −925 bp and −875 bp and also indicated that all basal elements resided between −875 bp and −650 bp. Additional deletions showed that sequences between −875 bp and −650 bp contained both basal and inducible elements. This latter result confirmed the earlier deletion analysis (Rushmore et al., in press) demonstrating both some degree of inducibility and basal activity directed by the sequences located 3′ of −875 bp.

Analysis of enhancer activity of −980 bp to −650 bp. Although the 5′ deletion data indicated that the sequences from −980 bp to −650 bp were required for full activity of the GST Ya gene, it was unclear whether additional sequences between −650 and the promoter were also required. Therefore, the −980 bp to −650 bp segment was moved next to the promoter-proximal region at −190 bp and cloned in either orientation. The −980 bp to −650 bp segment made the promoter beginning at −190 bp fully active and inducible (Fig. 2A). Therefore, it is unlikely that any sequences between −650 bp and −190 bp are required for complete enhancer activity. A further test of enhancer activity was to determine whether −980 bp to −650 bp was capable of activating a heterologous promoter. Therefore, the GST Ya enhancer was cloned in either orientation 5′ to a β-globin promoter minigene construct (5). The GST enhancer was able to activate the β-globin promoter three- to fivefold in untreated HepG2 cells and 10- to 20-fold in the presence of 3-MC (Fig. 3). This is approximately the same degree of stimulation by 3-MC that was observed in the natural gene.

Cell specificity of the GST Ya enhancer and promoter. The tissue distribution of GST Ya expression in rats is predominantly in hepatocytes and to a lesser extent in kidney cells (25) (data not shown). Therefore, we transfected a variety of cell lines with appropriate constructs to test the cell-specific function of each component of the enhancer and promoter. The fully active GST Ya gene (the −1.55 kbp construct) as well as the construct containing the GST Ya enhancer/β-globin promoter functioned maximally in HepG2 cells and exhibited a fourfold induction (Fig. 4). In HeLa and 293 cells, these constructs gave uninduced signals weaker than in HepG2 cells, and furthermore the constructs were not inducible with 3-MC. FS2 and WISH cells showed no detectable activity. The GST Ya gene deleted to −650 bp was inactive in all cell types tested. In addition, when the non-tissue-specific SV40 enhancer was used to drive the −650 bp construct, all cell types except FS2 and WISH showed activity. This result contrasts, for example, with the TTR gene, which requires liver-specific proteins in both its enhancer and promoter regions (5) and cannot be driven by the SV40 enhancer in any cell type tested. Therefore, it appears that the GST Ya promoter may use some widely available proteins but that the enhancer contains elements required for tissue specificity, 3-MC inducibility, and maximal rate of transcription.

Nuclear protein binding in the 3-MC-inducible region −925 bp to −875 bp. In order to identify the specific sites within the −925 bp to −875 bp region that were involved in 3-MC induction, we used nuclear extracts from HepG2 cells that were treated or untreated with 3-MC to examine protein-DNA contacts. Prior to preparing extracts, we determined how rapidly the transfected GST Ya gene was induced in HepG2 cells. As shown in Fig. 5A and quantitated by densitometry (data not shown), the −1.55-kbp construct was maximally induced between 4 and 8 h of 3-MC treatment. However, it is evident that even by 2 h of treatment, mRNA from the transfected DNA had increased significantly. Therefore, because 3-MC induction in HepG2 cells was rapid, we prepared nuclear extracts from both short and long 3-MC treatment times. Protein-DNA binding was tested by electrophoretic mobility shift assays, in which protein-DNA complexes are resolved from unbound labeled DNA. One or more proteins specifically bound to the DNA probe (−980 bp to −875 bp) in both 3-MC-treated and untreated extracts (Fig. 5B). The specific complexes were identified by the



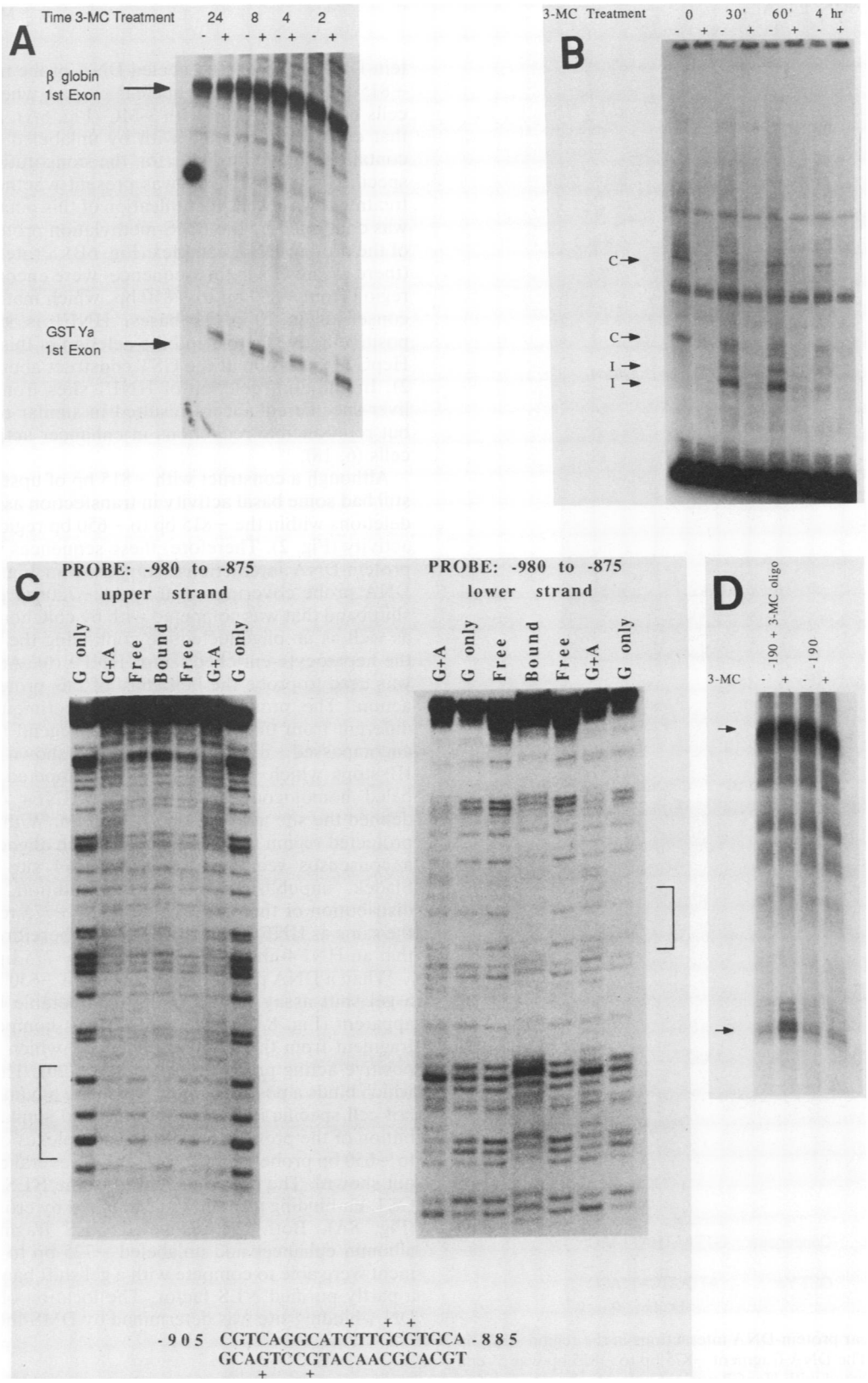
Construct	-3-MC	+3-MC	Cell Line
GST Ya -1.55 kbp	100	417	HepG2
	12	11	Hela
	19	20	293
	<1	ND	WISH
	<1	ND	FS2
β glo -980 Ya -650	100	378	HepG2
	19	22	Hela
	26	31	293
GST Ya -650 bp	1	1	HepG2
	<1	<1	Hela
	<1	<1	293
SV40 E Ya -650	122	116	HepG2
	107	115	Hela
	88	80	293
	<1	ND	WISH
	<1	ND	FS2

FIG. 4. Cell specificity of the GST Ya enhancer and promoter. Recombinant plasmids containing the GST Ya enhancer and promoter (−1.55 kbp), the GST Ya promoter (−650 bp), the GST Ya enhancer on a heterologous promoter (Ya −980 to −650 bp plus β-globin [β glo] promoter), and a heterologous enhancer (SV40E) driving the GST Ya promoter at −650 bp were transfected into several cell lines representing different cell types. Constructs were assayed with and without 3-MC induction and normalized to the appropriate cotransfection control as described in the legends to Fig. 2 and 3. Cell types are as follows: HepG2, human hepatoma; HeLa, human cervical carcinoma; 293, human embryonic kidney; FS2, human diploid fibroblast; WISH, human amnionic epithelial. ND, Not done.

ability to excess homologous cold DNA to compete with the gel shift bands. Numerous oligonucleotides encompassing other protein-binding sites were unable to compete with these bands (data not shown). Two bands were present in all extracts (identified as constitutive [C]), while two other bands were present only in 3-MC-treated extracts (identified as induced [I]). The induced bands appeared within 30 min after 3-MC treatment but were diminished after 4 h. Furthermore, extracts made after 8 or 24 h of 3-MC treatment showed only the constitutive bands. The rapid induction of the protein responsible for the gel shift bands correlated well with the rapid induction of the transfected gene, suggesting the involvement of this protein in induction of the GST Ya gene.

To locate precisely the DNA-binding site for the induced protein factor, dimethylsulfate (DMS)-methylation protection footprints were performed on the induced protein-DNA complexes. Protected G residues were found on both the upper and lower strands between −905 bp and −885 bp, that is, within the region identified as necessary in the functional analysis (−925 bp to −875 bp) (Fig. 5C). Finally, to test the functionality of this protected sequence, we prepared a double-stranded oligonucleotide which encompassed −910 bp to −875 bp and cloned it in front of the GST Ya promoter at −190 bp. As shown in Fig. 5D, the promoter plus the −910 bp to −875 bp oligonucleotide was induced by 3-MC, while the construct without the upstream element was not. Furthermore, as expected, the −910 bp to −875 bp oligonucle-

FIG. 5. Analysis of the 3-MC-inducible activation region −925 bp to −875 bp. (A) Time course of 3-MC activation of transient transfections. The fully active and inducible GST Ya construct −1.55 kbp was transfected into HepG2 cells and exposed to 2 μM 3-MC for various times (in hours). The transfections were assayed, and values were normalized to the SV40 enhancer/β-globin promoter cotransfection control as described in the legend to Fig. 2. The GST and β-globin first exons are indicated by arrows. (B) Nuclear protein-DNA interactions in the 3-MC-inducible region. The GST Ya upstream segment −980 bp to −875 bp was 5'-end labeled and incubated with untreated or 3-MC-treated HepG2 nuclear extracts in the presence (+) or absence (−) of a 40-fold molar excess of cold homologous DNA. The binding reaction was analyzed by a gel retardation assay in which protein-DNA complexes are separated from free DNA on a low-ionic-strength polyacrylamide gel. Specifically competed bands which did not vary with 3-MC treatment are labeled C for constitutive. Specifically competed bands which did vary with 3-MC treatment are labeled I for induced. (C) Analysis of the nucleotide residues involved in 3-MC-induced protein binding by DMS-methylation protection. The standard binding reaction mix was scaled up 20-fold and treated with dimethyl sulfate before being loaded on a preparative acrylamide gel. Both 5'- and 3'-end-labeled fragments (−980 bp to −875 bp) were used in binding reactions in order to identify protected residues on both strands. Bound DNA bands were isolated, cleaved with piperidine (29), and analyzed on a sequencing gel. The brackets indicate the G residues from the protein-bound bands that were protected from methylation. The bottom of the figure indicates the positions of the protected G residues (+'s). (D) Functional assay of the 3-MC-inducible protein-DNA interaction. A double-stranded oligonucleotide encompassing the methylation-protected sequence (−910 bp to −875 bp) was prepared and cloned in the front of the GST Ya promoter at −190 bp. The construct was transfected into HepG2 cells, and the transcriptional activity with (+) and without (−) 3-MC was assayed. The upper arrow indicates the β-globin first exon (transfection control) and the lower arrow indicates the GST Ya first exon.



otide was able to specifically compete with the induced gel shift bands formed with the −980 bp to −875 bp probe (data not shown). The results of these experiments identified a 3-MC-inducible positive-acting element to which a 3-MC-inducible protein factor binds.

Nuclear protein binding in the basal expression region −875 bp to −650 bp. As described earlier (Fig. 2) (Rushmore et al., in press), constructs containing the sequences between −875 bp and −650 bp are required for normal basal gene expression and in addition direct some degree of inducibility. In order to identify specific sites involved in basal or possibly 3-MC-inducible expression, we again performed gel shift experiments with DNA probes covering −875 bp to −650 bp and DMS-methylation protection footprints of specific pro-

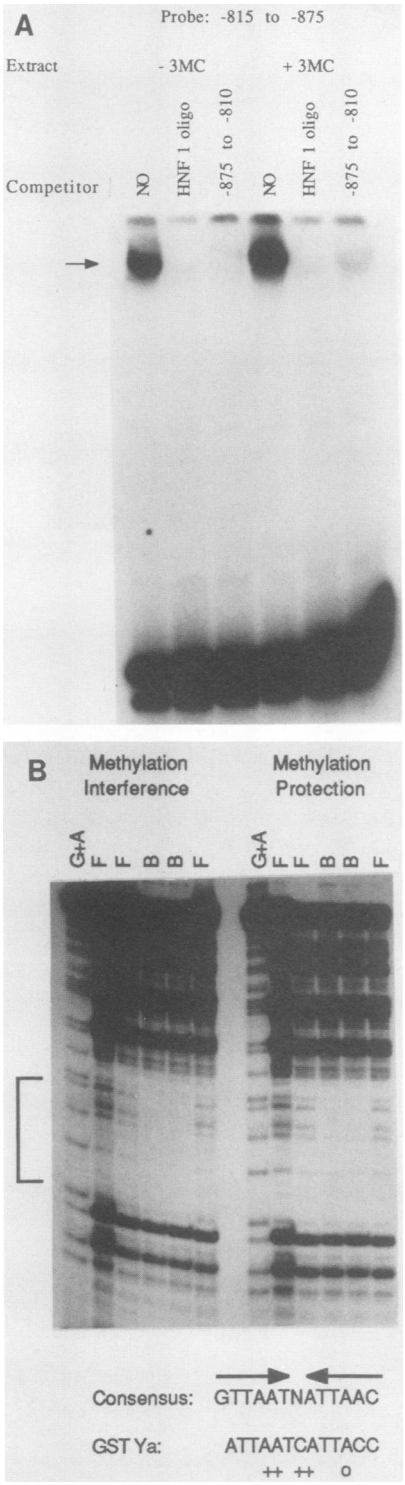


FIG. 6. Nuclear protein-DNA interactions in the region −875 bp to −815 bp. (A) The DNA fragment −875 bp to −815 bp was 5′-end labeled and incubated with HepG2 nuclear extracts. Reaction mixes also included either no cold competitor DNA (NO) or a 40-fold molar excess of cold −875 bp to −815 bp or 50-fold molar excess of cold oligonucleotide encompassing the HNF1 site from the TTR promoter. The binding reaction was assayed on a nondenaturing gel as described in the legend to Fig. 5. The specifically competed HNF1 shift is indicated with an arrow. (B) Analysis of the nucleotide residues involved in the HNF1 interaction. DMS-methylation protection was performed as described in the legend to Fig. 5. Methylation interference was performed by premethylating the −875 bp to −815 bp fragment (29), followed by a scaled-up binding reaction and analysis as described in the legend to Fig. 5. The brackets indicate the protected or interfering residues (all A's). F indicates free, unbound probe, and B indicates bound probe. Below the gels are a comparison of the consensus HNF1 site (7) and the protected sequence in the GST Ya enhancer. + indicates fully protected residues, and o indicates partially protected residues.

tein-DNA complexes. Labeled DNA probe from −875 bp to −815 bp was bound by nuclear extract whether or not the cells had been treated with 3-MC (Fig. 6A). A specific band that could be competed with by unlabeled oligonucleotide containing a binding site for the constitutive hepatocyte-specific factor HNF1 (7) was present whether or not 3-MC treatment was used. Identification of this putative HNF1 site was confirmed by the DMS-methylation protection footprint of the protein-DNA complex (Fig. 6B). Protected A residues (there are no G's in this sequence) were encompassed by the region from −860 bp to −850 bp, which matches the HNF1 consensus in 10 of 12 bases. HNF1 is known to be a positive-active factor, and the deletion of this site decreased HepG2 expression of the GST construct about twofold (Fig. 2). In addition, deletion of HNF1 sites from several other liver-specific enhancers resulted in similar or even smaller but reproducible reductions in enhancer activity in HepG2 cells (6, 18).

Although a construct with −815 bp of upstream sequence still had some basal activity in transfection assays, further 5′ deletions within the −815 bp to −650 bp region abolished all activity (Fig. 2). Therefore, these sequences were tested for protein-DNA interactions. As shown in Fig. 7A, a labeled DNA probe covering −810 bp to −720 bp produced a gel shift band that was competed with by cold homologous DNA as well as an oligonucleotide containing the binding site to the hepatocyte-enriched factor HNF4 (6). An exo III assay was used to probe the boundary of this protein-DNA interaction. The probe used in this experiment was slightly different from that described in the legend to Fig. 7A and encompassed −815 bp to −705 bp. As shown in Fig. 7B, exo III stops which were specifically competed with by unlabeled homologous DNA and the HNF4 oligonucleotide defined the site as −775 bp to −755 bp. Within this exo III-protected region were sequences which aligned closely with a consensus sequence for the HNF4 site (Fig. 7B) (F. Sladek, unpublished results). In addition, the cell type distribution of the protein binding to −775 to −755 bp was the same as HNF4 (data not shown). Therefore, we conclude that an HNF4-like factor interacts at −775 bp to −755 bp.

When a DNA probe from −725 bp to −650 bp was used in a gel shift assay, a specifically competable band was also apparent (Fig. 8A). This complex was competed with by a fragment from the albumin enhancer, which is bound by a positive-acting protein, termed NLS (20). The NLS factor, which binds a positive-acting site in the albumin enhancer, is not cell specific in its distribution, and similarly, the distribution of the protein that formed complexes on the −725 bp to −650 bp probe was also present in several cell types (data not shown). The possible identity of the NLS factor and the GST Ya-binding protein was examined by cross-competition (Fig. 8A). Both the unlabeled NLS fragment from the albumin enhancer and unlabeled −725 bp to −650 bp fragment were able to compete with a gel shift band produced by a partly purified NLS factor. The location of the NLS-like DNA-binding site was determined by DMS-methylation pro-

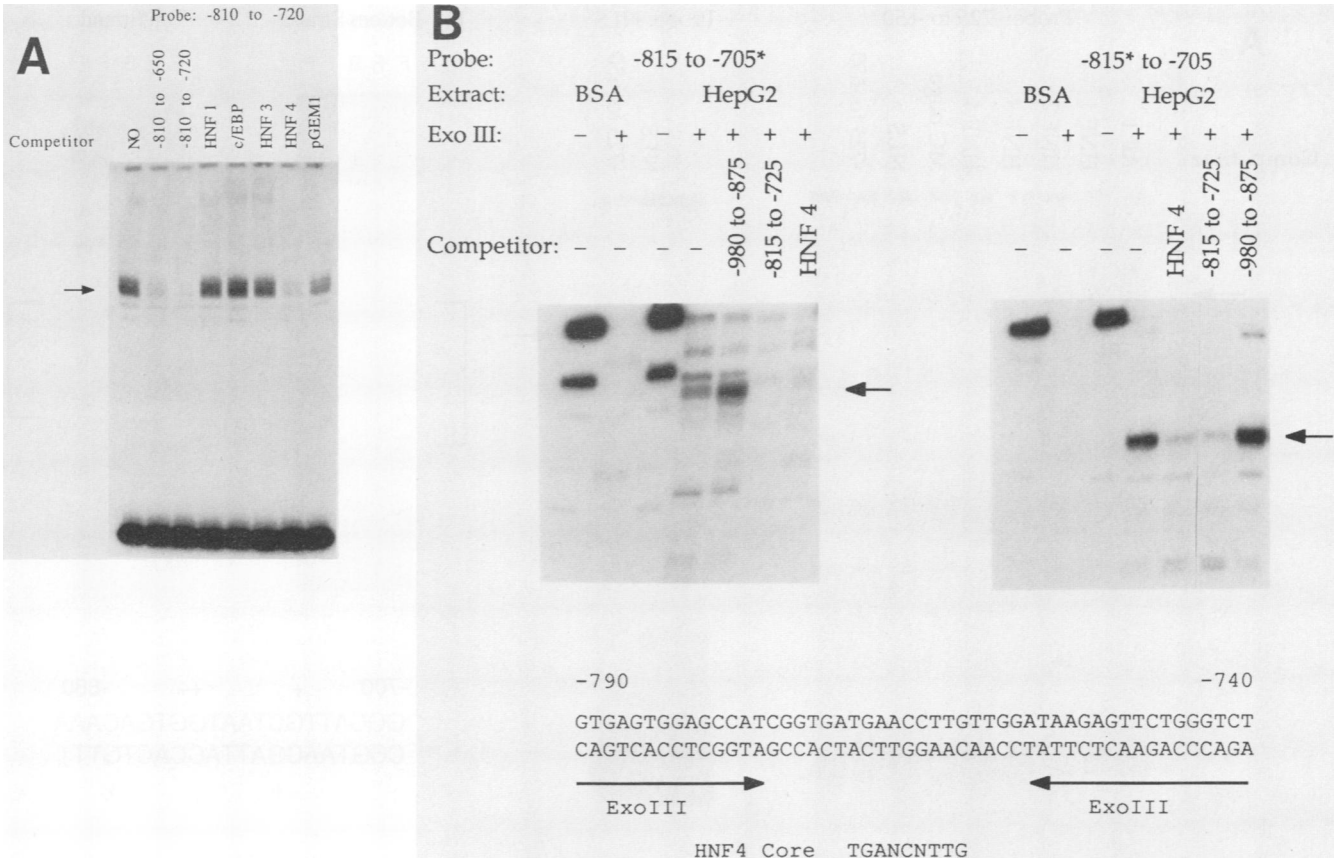


FIG. 7. Nuclear protein-DNA interactions in the region −810 bp to −720 bp. (A) The DNA fragment −810 bp to −720 bp was 5-end labeled and incubated with HepG2 nuclear extracts. Reaction mixes also included either no cold competitor DNA (NO) or 40- to 50-fold molar excess of the cold competitor DNA indicated. The binding reaction was assayed on a nondenaturing gel as described in the legend to Fig. 5. The shift band specifically competed with homologous DNA or the HNF4 oligonucleotide (−156 to −138 on the mouse TTR promoter [6]) is indicated with an arrow. (B) Exo III protection assay for the binding of HNF4-like protein to −815 bp to −705 bp. Probe fragment was 5′-end labeled on either strand as indicated (*). After incubation of 1 ng of probe with HepG2 nuclear extract in the standard binding reaction mix, the protein-DNA complexes were incubated with exo III. The digested DNA was analyzed on a 6% polyacrylamide–urea gel. Exo III stops which were specifically competed with by either excess cold homologous DNA or cold HNF4 oligonucleotide are indicated with an arrow. The bottom of the figure indicates the positions of the exo III stops and the alignment of the protected region with a consensus core sequence for HNF4 binding (F. Sladek, unpublished results). BSA, Bovine serum albumin.

tection footprinting (Fig. 8B). The footprints of the upper and lower strands showed that the site of interaction was between −700 bp and −680 bp. Comparison of this binding site and the albumin enhancer-binding site did not show any obvious homologies. However, several additional footprints of other sites may be required before a consensus sequence can be determined. It should be noted that neither the HNF4 site nor the NLS site function independently when isolated from the GST Ya enhancer; however, they can function when combined.

DISCUSSION

The experiments presented in this paper demonstrate that the induction of the GST Ya mRNA occurs specifically in pericentral hepatocytes. Furthermore, a DNA element required for 3-MC induction was detected at −925 to −875 bp deletion analysis of plasmids containing the upstream region of the GST Ya gene. Since the only change in this plasmid was upstream of the RNA start site, we conclude that this region functions at the transcriptional level in induction. In order to identify the protein(s) responsible for 3-MC induction and to locate the precise region required for induction in transfected cells, DNA-protein-binding assays were done. A 3-MC-inducible protein-DNA interaction was identified, and the protein-binding site was localized between −905 bp and −885 bp by DMS-methylation protection footprints. If this same 3-MC-inducible element is involved in inducible transcription in the animal, then we have defined a DNA region required for position-specific, inducible transcription in hepatocytes.

The inducible P450 genes have been shown to have xenobiotic-inducible elements (12, 16, 21, 32, 38). Based on DNA sequence homologies within these regions, a consensus xenobiotic-responsive element (XRE) or dioxin-responsive element (DRE) has been identified (9, 13). The sequence of the footprinted region of the GST Ya gene contains a sequence that resembles the so-called XREs or DREs.

```
GST Ya:      CAGGCATGTTGCGTGCAT
             GTCCGTACAACGCACGTA

Consensus:       GTTGCGTGA
                 C A     C
```

The binding site of the 3-MC-inducible protein in the GST Ya

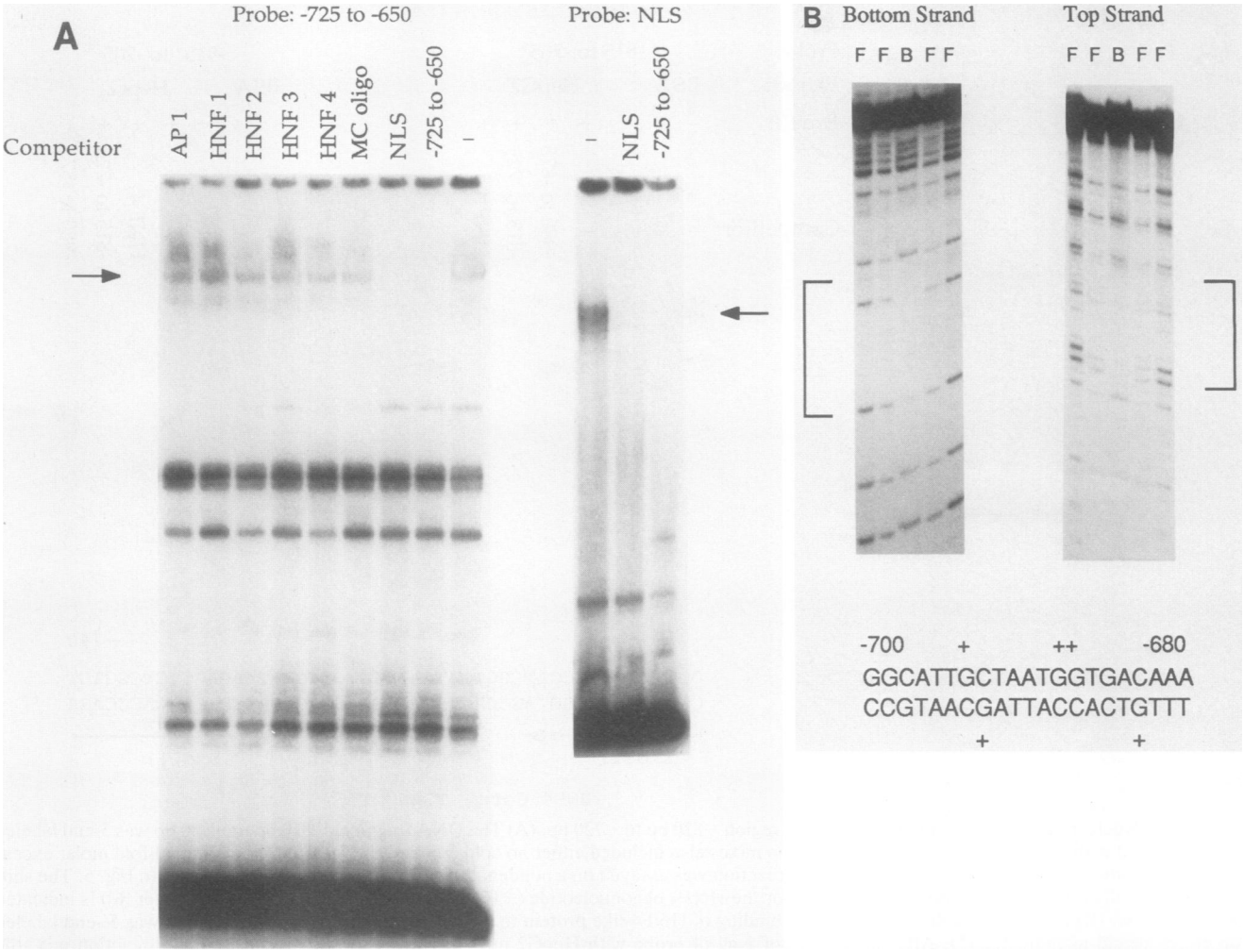


FIG. 8. Nuclear protein-DNA interactions in the region −725 bp to −650 bp. (A) The left side of the figure shows DNA fragment −725 bp to −650 bp which was 5′-end labeled and incubated with HepG2 nuclear extracts. The right side of the figure shows the NLS fragment from the albumin enhancer which was 5′-end labeled and incubated with partially purified NLS protein. Reaction mixes included either no cold competitor DNA or a 40-fold molar excess of the cold competitor DNA indicated. The binding reaction was assayed on a nondenaturing gel as described in the legend to Fig. 5. The specifically competed NLS shift is indicated with an arrow. (B) Analysis of the nucleotide residues involved in the NLS interaction. DMS-methylation protection was performed as described in the legend to Fig. 5. The brackets indicate the G residues from the protein-bound bands that were protected from methylation. F indicates free probe, and B indicates bound probe. The bottom of the figure indicates the locations of the protected G residues (+'s).

gene has several similarities to the various steroid receptor-binding sites, including an inverted repeat in the contact region (with one-base-pair mismatch, as indicated in the diagram). This characteristic was not described for XREs but is expected of a binding protein that belongs to the steroid receptor superfamily (11). Further mutagenesis studies are required to determine which bases are most important for induced protein binding and function.

The identity of the observed 3-MC-inducible protein is unknown at present. However, the kinetics of 3-MC induction of the GST Ya protein-DNA complex are very similar to those of the xenobiotic receptor or dioxin receptor, which has been postulated to be a member of the steroid hormone receptor superfamily (9, 19). Only 1 h of exposure to dioxin is required for maximal induction of a protein that binds a DRE. In addition, this binding activity is greatly diminished at 4.5 h of exposure and is completely gone by 16 h (19). It should also be noted that the study of Telakowski-Hopkins et al. (39) showed that the GST Ya gene is not responsive when transfected into cells lacking this dioxin receptor. Finally, we found that cycloheximide had no effect on 3-MC induction of the GST Ya gene (data not shown), indicating that the 3-MC-induced protein-DNA interaction may be due to a preexisting xenobiotic receptor. From the in situ hybridization results and the identification of the 3-MC-inducible site as an XRE, it might be possible that the xenobiotic compound receptor is represented only in pericentral hepatocytes.

As shown in Fig. 5, there were also constitutive proteins in HepG2 extracts which interacted with the 3-MC-inducible region. These constitutive proteins bind to the same site as the 3-MC-inducible protein, since an oligonucleotide encompassing the sequence competed with the constitutive gel shift bands. We do not know the identity of the constitutive binding proteins. However, the possibility of multiple proteins interacting at one binding site is not unprecedented, especially within the steroid receptor superfamily. The retinoic acid and thyroid hormone receptors appear to activate

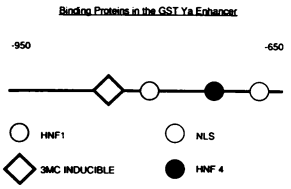


FIG. 9. Summary of the nuclear protein-binding sites regulating the function of the GST Ya enhancer. Shown schematically is the GST Ya enhancer from −950 bp to −650 bp. The different symbols represent the various protein-DNA interactions indicated.

through a common sequence element (40). Also, the thyroid hormone receptor appears to act as an antagonist to estrogen receptor binding at the same site and consequently blocks transcriptional induction by estrogen (15). We have no evidence that the constitutive binding proteins function mechanistically as in these examples. However, if the xenobiotic receptor is a member of the steroid receptor family, then there are possible mechanisms to test for the function of the constitutive proteins.

Although we have focused on the −905 bp to −885 bp region as being important in 3-MC inducibility, the region between −723 bp and −684 bp contributes to maximal induction (Rushmore et al., in press). We have been unable to identify an inducible protein-DNA interaction within this region to correlate with its inducible activity in transient transfections. Possibly an inducible gel shift could be masked by the constitutive factors which interact with the −815 bp to −650 bp region used in gel shift experiments. In addition, we did not find any sequence homology within the −723 bp to −684 bp region to the upstream 3-MC-inducible binding site or with the consensus XRE. Therefore, xenobiotic-induced protein-DNA interactions at this site may be weak or may represent either a novel class of XRE or a different type of regulation by xenobiotics (Rushmore et al., in press).

In addition to the putatively pericentral location of the induced protein factor, we also identified several nuclear factors from hepatocytes that interact with the basal or constitutive part of the GST Ya enhancer. A summary of all protein-DNA interactions in the GST Ya enhancer is shown in Fig. 9. The basal enhancer consists of at least three protein-DNA interactions, including hepatocyte-enriched factors similar or identical to HNF1 and HNF4 and a non-tissue-specific NLS-like factor. Because none of the purified factors have been shown to interact with the identified sites, we cannot be conclusive about the identity of the factors. All of the sites identified are important for full enhancer function in hepatoma cells. We have not assessed the role of these sites in other cell types, although the enhancer can function at a low level in some nonhepatoma lines. Perhaps this reflects the role of the NLS-like site, or perhaps there are additional as yet unidentified protein-DNA interactions which mediate enhancer activity. Interestingly, the arrangement of liver-specific and constitutive factor binding sites seems to be a general feature of several liver-specific enhancers and promoters studied so far (2, 4, 6, 18, 26, 27). The GST Ya enhancer also fits this pattern in its arrangement of binding proteins.

Having identified the critical sequences and protein-binding sites for both basal and 3-MC-inducible expression of the GST Ya gene, we are assessing the role of these sequences in vivo. We are producing transgenic mice to examine the role of upstream GST Ya sequences on inducible pericentral expression in vivo and the role of the constitutive elements in general expression. Our detailed understanding of the molecular interactions in the GST Ya enhancer and the use of transgenic mice should allow us to follow the developmental decisions which occur during hepatocyte positional differentiation.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant CA-16006 from the National Institutes of Health and grant MV 424 from the American Cancer Society. K.E.P. was supported by a postdoctoral fellowship from the American Cancer Society.

We thank A. S. Yee, E. Lai, and F. Kuo for helpful discussions and critical review of this manuscript. We thank O. Litvin for excellent technical assistance. We are grateful to A. S. Yee for providing partially purified 3-MC-inducible protein. We also thank A. English and L. Cousseau for the preparation of this manuscript.

LITERATURE CITED

1. **Babiss, L. E.** 1989. The cellular transcription factor E2F requires viral E1A and E4 gene products for increased DNA-binding activity and functions to stimulate adenovirus E2A gene expression. J. Virol. **63**:2709–2717.
2. **Babiss, L. E., R. S. Herbst, A. L. Bennett, and J. E. Darnell, Jr.** 1987. Factors that interact with the rat albumin promoter are present in hepatocytes and other cell types. Genes Dev. **1**: 256–267.
3. **Bennett, A. L., K. E. Paulson, R. E. Miller, and J. E. Darnell, Jr.** 1987. Acquisition of antigenes characteristic of adult pericentral hepatocytes by differentiating fetal hepatoblasts *in vitro*. J. Cell Biol. **105**:1073–1085.
4. **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.
5. **Costa, R. H., E. Lai, and J. E. Darnell, Jr.** 1986. Transcriptional control of the mouse prealbumin (transthyretin) gene: both promoter sequences and a distinct enhancer are cell specific. Mol. Cell. Biol. **6**:4697–4708.
6. **Costa, R. H., D. R. Grayson, and J. E. Darnell, Jr.** 1989. Multiple hepatocyte-enriched nuclear factors function in the regulation of transthyretin and α_1-antitrypsin genes. Mol. Cell. Biol. **9**:1415–1425.
7. **Courtois, G., S. Baumhueter, and G. R. Crabtree.** 1988. Purified hepatocyte nuclear factor 1 interacts with a family of hepatocyte-specific promoters. Proc. Natl. Acad. Sci. USA **85**:7937–7941.
8. **Damm, K., C. S. Thompson, and R. M. Evans.** 1989. Protein encoded by v-*erbA* functions as a thyroid hormone antagonist. Nature (London) **339**:593–597.
9. **Denison, M. S., J. M. Fisher, and J. P. Whitlock, Jr.** 1988. Inducible, receptor-dependent protein-DNA interactions at a dioxin-responsive transcriptional enhancer. Proc. Natl. Acad. Sci. USA **85**:2528–2532.
10. **Dignam, J. D., R. M. Lebowitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11**:1475–1489.
11. **Evans, R. M.** 1988. The steroid and thyroid hormone receptor superfamily. Science **240**:889–895.
12. **Fujisawa-Sehara, A., K. Sogawa, C. Nishi, and Y. Fujii-Kuriyama.** 1986. Regulatory DNA elements localized remotely upstream from the drug-metabolizing cytochrome P450c gene. Nucleic Acids Res. **14**:1465–1477.
13. **Fujisawa-Sehara, A., M. Yamane, and Y. Fujii-Kuriyama.** 1988. A DNA-binding factor specific for xenobiotic-responsive elements of P-450C gene exists as a cryptic form in cytoplasm: its possible translocation to nucleus. Proc. Natl. Acad. Sci. USA **85**:5859–5863.
14. **Gaasbeek, J. W., R. Gebhardt, G. H. J. Ten Voorde, W. H. Lamers, R. Charles, and A. F. M. Moorman.** 1987. Heterogeneous distribution of glutamine synthetase during rat liver

development. J. Histochem. Cytochem. **35**:49–54.

15. **Glass, C. K., J. M. Holloway, O. V. Devary, and M. G. Rosenfeld.** 1988. The thyroid hormone receptor binds with opposite transcriptional effects to a common sequence motif in thyroid hormone and estrogen response elements. Cell **54**: 313–323.
16. **Gonzalez, F. J., and D. W. Nebert.** 1985. Autoregulation plus upstream positive and negative control regions associated with transcriptional activation of the mouse P1450 gene. Nucleic Acids Res. **13**:7269–7288.
17. **Gorski, K., M. Carneiro, and U. Schibler.** 1986. Tissue-specific in vitro transcription from the mouse albumin promoter. Cell **47**:767–776.
18. **Grayson, D. R., R. H. Costa, K. G. Xanthopoulos, and J. E. Darnell, Jr.** 1988. A cell-specific enhancer of the mouse α_1-antitrypsin gene has multiple functional regions and corresponding protein-binding sites. Mol. Cell. Biol. **8**:1055–1066.
19. **Hapgood, J., S. Cuthill, M. Denis, L. Poellinger, and J.-A. Gustafsson.** 1989. Specific protein-DNA interactions at a xenobiotic-responsive element: copurification of dioxin receptor and DNA. Proc. Natl. Acad. Sci. USA **86**:60–64.
20. **Herbst, R. S., N. Friedman, J. E. Darnell, and L. E. Babiss.** 1989. Positive and negative regulatory elements in the mouse albumin enhancer. Proc. Natl. Acad. Sci. USA **86**:1533–1557.
21. **Jones, P. B. C., D. R. Galeazzi, J. M. Fisher, and J. P. Whitlock, Jr.** 1985. Control of cytochrome P1-450 gene expression by dioxin. Science **227**:1499–1502.
22. **Jungermann, K.** 1986. Zonal signal heterogeneity, p. 445–472. *In* R. G. Thuman, F. C. Kauffman, and K. Jungermann (ed.), Regulation of hepatic metabolism. Plenum Publishing Corp., New York.
23. **Kovesdi, I., R. Reichel, and J. R. Nevins.** 1986. Identification of a cellular transcription factor involved in E1A *trans*-activation. Cell **45**:219–228.
24. **Kuo, C. F., K. E. Paulson, and J. E. Darnell, Jr.** 1988. Positional and developmental regulation of glutamine synthetase expression in mouse liver. Mol. Cell. Biol. **8**:4966–4971.
25. **Lai, H.-C. J., N. L. Li, M. J. Weiss, C. C. Reddy, and C.-P. D. Tu.** 1984. The nucleotide sequence of a rat liver glutathione *S*-transferase subunit cDNA clone. J. Biol. Chem. **259**:5536–5542.
26. **Lichtsteiner, S., J. Wuarin, and U. Schibler.** 1987. The interplay of DNA-binding proteins on the promoter of the mouse albumin gene. Cell **51**:963–973.
27. **Maire, P., J. Wuarin, and U. Schibler.** 1989. The role of *cis*-acting promoter elements in tissue-specific albumin gene expression. Science **244**:343–346.
28. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
29. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
30. **Moody, D. E., L. A. Taylor, E. A. Smuckler, W. Levin, and P. E. Thomas.** 1983. Immunohistochemical localization of cytochrome P450a in liver sections from untreated rats and rats treated with phenobarbitol or 3-methylcholanthrene. Drug. Metab. Dispos. **4**:339–343.
31. **Mordacq, J. C., and D. I. H. Linzer.** 1989. Co-localization of elements required for phorbol ester stimulation and glucocorticoid repression of proliferin gene expression. Genes Dev. **3**:760–769.
32. **Neuhold, L. A., Y. Shirayoshi, K. Ozato, J. E. Jones, and D. W. Nebert.** 1989. Regulation of mouse CYP2A2 gene expression by dioxin: requirement of two *cis*-acting elements during induction. Mol. Cell. Biol. **9**:2378–2386.
33. **Pickett, C. B., and A. Y. H. Lu.** 1989. Glutathione *S*-transferases: gene structure, regulation and biological function. Annu. Rev. Biochem. **58**:734–764.
34. **Pickett, C. B., C. A. Telakowski-Hopkins, G. J.-F. Ding, L. Argenbright, and A. Y. H. Lu.** 1984. Rat liver glutathione *S*-transferases: complete nucleotide sequence of a glutathione *S*-transferase mRNA and the regulation of the Ya, Yb and Yc mRNAs by 3-methylcholanthrene and phenobarbitol. J. Biol. Chem. **259**:5182–5188.
35. **Poliard, A. M., D. Bernuau, I. Tournier, L. G. Legres, D. Schoevaert, G. Feldmann, and J. M. Sala-Trepat.** 1986. Cellular analysis by *in situ* hybridization and immunoperoxidase of alpha-fetoprotein and albumin gene expression in rat liver during the perinatal period. J. Cell Biol **103**:777–786.
36. **Ratanasavanh, D., P. Beune, G. Baffet, M. Rissel, P. Kremers, F. Guengerich, and A. Guillouzo.** 1986. Immunohistochemical evidence for the maintenance of cytochrome P450 isoenzymes, NADPH cytochrome C reductase and epoxide hydrolase in pure and mixed cultures of adult human hepatocytes. J. Histochem. Cytochem. **34**:527–533.
37. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
38. **Sogawa, K., A. Fujisawa-Sehara, M. Yamane, and Y. Fujii-Kuriyama.** 1986. Location of regulatory elements responsible for drug induction in rat cytochrome P450c gene. Proc. Natl. Acad. Sci. USA **83**:8044–8048.
39. **Telakowski-Hopkins, C. A., R. G. King, and C. B. Pickett.** 1988. Glutathione *S*-transferase Ya subunit gene: identification of regulatory elements required for basal level and inducible expression. Proc. Natl. Acad. Sci. USA **85**:1000–1004.
40. **Umesono, K., V. Giguere, C. K. Glass, M. G. Rosenfeld, and R. M. Evans.** 1988. Retinoic acid and thyroid hormone induce gene expression through a common responsive element. (London) **336**:262–265.
41. **Wolf, C. R., E. Moll, F. Oesch, A. Buchmann, W. D. Kuhlmann, and H. W. Kunz.** 1984. Characterization, localization and regulation of a novel phenobarbitol-inducible form of cytochrome P450, compared with three further P450-isoenzymes, NADPH P450-reductase, glutathione transferase and microsomal epoxide hydrolase. Carcinogenesis **5**:993–1001.