Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical Pharmacology 65 (2003) 1087–1096
Biochemical Pharmacology
www.elsevier.com/locate/biochempharm

Interaction of upstream stimulatory factor proteins with an E-box located within the human *CYP1A2* 5′-flanking gene contributes to basal transcriptional gene activation

George V. Pickwell, Hsueh Shih, Linda C. Quattrochi*

Department of Medicine, Section of Medical Toxicology, B146, University of Colorado Health Sciences Center, 4200 East 9th Avenue, Denver, CO 80262, USA

Received 28 June 2002; accepted 10 September 2002

Abstract

Cytochrome P450 (CYP)1A2 is abundantly expressed in the liver of all vertebrate species. In most, its expression is restricted to the liver. Sequence analysis of the human *CYP1A2* 5′-flanking region from +3 to −3201 identified six E-box motifs within the 3-methylcholanthrene (MC) enhancer element (−1987 to −3201). The E-box motif is recognized by members of the basic helix-loop-helix (bHLH) family of transcription factors. Gel mobility shift and antibody supershift assays were used to examine each of the six upstream E-box motifs for their ability to bind nuclear proteins and to compete with the ubiquitously expressed bHLH protein, upstream stimulatory factor (USF), for binding. We found that USF-1 and USF-2 proteins bind to the upstream E-box motifs EB2, EB3, and EB4. Transient transfection assays in HepG2 cells were performed with different segments of the human *CYP1A2* 5′-flanking region linked to a luciferase reporter gene. Site-directed mutagenesis of one of the E-box motifs, EB2, resulted in a 60% reduction in basal reporter gene activity. Mutations in EB3 and EB4 had no effect. We found that transfection of expression vectors containing USF-1 or USF-2 cDNAs activated *CYP1A2* reporter gene activity, while a dominant-negative USF-2 expression vector blocked such activity. Chromatin immunoprecipitation assays confirmed that the interaction of USF proteins with the *CYP1A2* EB2 site occurs *in vivo*. These data support the role of USF as a constitutive transcriptional activator of human *CYP1A2*.
© 2003 Elsevier Science Inc. All rights reserved.

Keywords: CYP1A2; USF; E-box; Cytochrome P450; bHLH; HepG2

1. Introduction

Polycyclic aromatic hydrocarbons (PAHs), found in cigarette smoke and other products of combustion, as well as arylamines, such as heterocyclic amines found in char-broiled meat, undergo metabolic activation by CYP1A enzymes (CYP1A1 and CYP1A2) to produce highly reactive intermediates that possess mutagenic and carcinogenic activities (reviewed in Ref. [1]). The toxicity of environmental chemicals at a specific organ is believed to be dependent upon the relative levels of CYP enzymes present in any given tissue at any given time. For example, exposure to PAHs can result in the induction of CYP1A enzyme expression in liver and in extrahepatic tissues, increasing the potential of these enzymes to activate or detoxify chemicals, or metabolize endogenous steroids. In addition to the metabolism of these procarcinogens, hepatic CYP1A2 also plays a major role in the disposition of numerous drugs including caffeine, acetaminophen, tacrine, theophylline, imipramine, clozapine, and olanzapine [2–4].

CYP1A2 is expressed in a cell-specific manner, but is both highly and variably expressed in human liver [2,5]. Yet, there is a paucity of data concerning the molecular mechanisms by which the *CYP1A2* gene is transcriptionally regulated. By necessity, the basal expression of hepatic

* Corresponding author. Tel.: +1-303-315-3522; fax: +1-303-315-7180.
E-mail address: linda.quattrochi@uchsc.edu (L.C. Quattrochi).

Abbreviations: AhR, aryl hydrocarbon receptor; AP-1, activator protein-1; ARNT, aryl hydrocarbon receptor nuclear translocator; bHLH, basic helix-loop-helix; ChIP, chromatin immunoprecipitation assay; CYP, cytochrome P450; DMEM, Dulbecco's modified essential medium; EB, E-box; EMSA, electrophoretic mobility shift assay; FBS, fetal bovine serum; HNF, hepatic nuclear factor; MC, 3-methylcholanthrene; NF-1, nuclear factor-1; PCR, polymerase chain reaction; USF, upstream stimulatory factor; XRE, xenobiotic response element.

0006-2952/03/$ – see front matter © 2003 Elsevier Science Inc. All rights reserved.
doi:10.1016/S0006-2952(03)00037-6

CYP1A2 must be under strict control by cell-specific factors. For example, HNF-1, a liver-enriched transcription factor, has been shown to play a role in the regulation of the expression of *CYP1A2* [6]. Nonetheless, CYP1A2 has been detected recently in non-hepatic tissues, such as the human lung [7] and the olfactory mucosa of rodents [8]. The nuclear factor NF-1 has been identified as a factor contributing to the expression of *CYP1A2* in rat olfactory tissue [8]. The importance of identifying factors that play a primary role in the constitutive and/or cell-type specific expression of *CYP1A2* is underscored by the essential role this CYP enzyme plays in disposing of numerous environmental, pharmacologic, and dietary entities.

We previously characterized several transcription factor binding sites involved in the inducible expression of the human *CYP1A2* gene [9,10]. We have now extended our studies to focus on additional factors responsible for the constitutive expression of *CYP1A2*. Recently, we have identified several putative binding sites for the bHLH proteins within the MC enhancer element of the human *CYP1A2* gene. The E-box motif (consensus, CANNTG, where central dinucleotides are usually GC or CG) is recognized by members of the bHLH family of transcription factors (bHLH, bHLHZIP, bHLH-PAS). Some bHLH and bHLHZIP proteins can form both homo- and heterodimers, while others bind exclusively as either homodimers or heterodimers (reviewed in Ref. [11]). Members of this large family of transcription factors have been classified into several catagories depending upon the DNA sequence of the central two nucleotides. E-box motifs have been found in the promoters and enhancers of numerous genes encoding for proteins that participate in the control of cellular proliferation or differentiation (reviewed in Ref. [12]). Examples of E-box binding proteins are EsA, E2-2, HEB, MyoD, Myc, USF, and ARNT. AhR and ARNT are bHLH-PAS transcription factors that mediate the inducible expression of *CYP1A* genes. Although the AhR/ARNT heterodimer does not interact with the consensus E-box motif, the ARNT homodimer does interact with certain E-box sequences [13]. Indeed, Takahashi *et al.* [14] found that the sequence of the rabbit XRE overlaps with that of a USF binding site and that USF competes with the AhR/ARNT complex for binding to this element. Some bHLH proteins are expressed in a cell-specific fashion, e.g. MyoD, a regulator of muscle-specific gene expression [15], while others, for example the bHLHZIP protein, USF, are ubiquitously expressed and probably play a role in the transcription of numerous genes [12].

In this report, we examined the binding of bHLH proteins to the E-box motifs found in the MC enhancer region of the *CYP1A2* gene. We also tested these sites for functional importance in the context of full promoter sequences containing other regulatory elements, including HNF-1, AP-1, and NF-1 sites, and by *in vivo* ChIP assays.

2. Materials and methods

2.1. Reagents

FBS was from HyClone. DMEM, penicillin/streptomycin, and trypsin/EDTA were from Life Technologies, GIBCO-BRL. All other tissue culture reagents were of the highest grade commercially available from the Sigma Chemical Co. Poly[d(I-C)] was from Pharmacia. [γ-^{32}P]ATP was from Pharmacia LKB Biotechnology. For ChIP assays we employed protein G-coupled agarose beads (Pierce). Antibodies to USF-1 (sc-229X) and USF-2 (sc-862X) were from Santa Cruz Biotechnology. A goat anti-mouse IgG (H&L) antibody (Pierce) functioned as a negative control. Agarose for size fractionation of the chromatin input and immunoprecipitated DNA was Sea Plaque GTG (FMC Corp., equivalent to the present Sigma Cat. No. #A4018). The protease inhibitors used were aprotinin (CalBiochem), and pepstatin, leupeptin, and phenylmethylsulfonyl fluoride (Sigma). Proteinase K and glycine were also obtained from Sigma. We used the Qiagen QIAEX II kit for DNA desalting and concentration and for gel solubilization and DNA concentration. All other reagents and buffer components were of analytical reagent grade. For PCR, we employed Life Technologies (Invitrogen) *Taq* DNA polymerase, either native or recombinant.

2.2. Cell culture, isolation of nuclear proteins, and EMSA

HepG2 cells were grown as monolayers in DMEM supplemented with 10% FBS, 1 mM sodium pyruvate, 100 IU/mL of penicillin and 0.1 mg/mL of streptomycin (growth medium), and were maintained at 37° under a humidified atmosphere containing 5% CO_2. For the isolation of nuclear extracts, cells were seeded onto 150 mm plastic culture dishes at a density of 2.5×10^7. Isolation of nuclear extracts and EMSA were performed as described previously [9]. Cells were cultured to 75–90% confluence prior to isolation of nuclear proteins. Complementary pairs of synthetic oligonucleotides (Qiagen Operon) containing the sequences of each CYP1A2 E-box motif were synthesized, purified, annealed, and radiolabeled with [γ-^{32}P]ATP. The double-strand (ds)-oligomers containing the consensus USF binding site or the mutated binding site were purchased from Santa Cruz Biotechnology. The sequences of the oligomers (upper strand) used for EMSA studies are shown in Table 1.

EMSA was performed by incubating radiolabeled ds-oligomers at room temperature for 15 min with 5 μg of crude nuclear extracts (obtained from untreated HepG2 cells) in a binding reaction containing 2 μg poly[d(I-C)], 10 mM Tris–HCl, pH 7.5, 75 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol, and 7% glycerol. Supershift assays were performed similarly, except that 1 μg of each antibody (anti-USF-1 or anti-USF-2; Santa Cruz) was added

Table 1
Sequences of oligonucleotides used for EMSA (upper strand) or site-directed mutagenesis (forward primers)

EMSA probes	
USF-1*	5′-CACCCGGT **CACGTG** GCCTACACC-3′
mUSF-1*	5′-CACCCGGT **CAATTG** GCCTACACC-3′
EB1	5′-CTTTTTCATGAGT **CATCTG** AACCCAAT-3′
EB2	5′-CTAAATGT **CATCTG** ATATCCAAT-3′
EB3	5′-ACCTTAGC **CAGGTG** ATCAAGGTC-3′
EB4	5′-CCGGTTGC **CAGCTG** ACATATGCA-3′
EB5	5′-GTGCTAGAATTA **CAGGTG** TGAGT-3′
EB6	5′-TGGAATTA **CAGGTG** TGCACCATC-3′
Site-directed mutagenesis primers	
EB2 mutant	5′-GAGACGTGATGACTAAATGT **CAATTG** ATATCCAATGGGTCTGG-3′
EB3 mutant	5′-CGCACTACCTTAGC **CAATTG** ATCAAGGTCAACATCC-3′
EB4 mutant	5′-GTGAGTCACCGGTCC **CAATTG** ACATATGCACTTTTC-3′
HNF-1 mutant	5′-GTAAACCATGAACTTTA **TTCGGCT** GGAGAGACAGAC-3′

Sequences of the 6-bp E-box motif and HNF-1 half site are shown in bold.
* Consensus and mutant USF oligonucleotides were purchased from Santa Cruz Biotechnlogy.

after the initial 15-min incubation. The antibodies were allowed to react for an additional hour at 4°.

2.3. Plasmids and site-directed mutagenesis

The human *CYP1A2* 5′-flanking sequence from +3 to −3201 was amplified from a previous plasmid, p1A2N [16], by PCR and cloned into the pGL3 basic plasmid (Promega). This plasmid, p3.2CYP1A2, served as the template for site-directed mutagenesis, which was employed to change the minimal consensus sequences for the putative E-box sites and HNF-1 half-site (QuikChange site-directed mutagenesisis kit; Stratagene). The E-box conversions consisted of an AT substitution in the central nucleotides of the 6 bp E-box motif. Several nucleotide substitutions were generated by site-directed mutagenesis in the HNF-1 half-site, as follows: wt HNF-1: GTTAAT; mutated HNF-1: TTCGGC. Results were confirmed in transfection experiments with two to three independent plasmids, each containing the sequence-confirmed mutation. Approximately 600 bases encompassing each site were sequenced to confirm the substitutions and to determine that no other mutations occurred. Sequences of oligonucleotides used for site-directed mutagenesis are given in Table 1.

The minimal *CYP1A2* promoter construct, produced by PCR amplification of the *CYP1A2* promoter sequences from +3 to −176, was subsequently cloned into the pGL3-basic plasmid (Promega). The *CYP1A2* region from −2259 to −1970, containing either the intact E-box or mutated E-box, was then inserted upstream of this minimal promoter, and the resulting plasmid referred to as p0.3mpCYP1A2. The plasmid p0.3tkCYP1A2 contains the same sequences cloned into the luciferase reporter plasmid, pT81, as previously described [10].

USF expression vectors (psvUSF-1 and psvUSF-2) and the USF-2 dominant negative plasmid (psvUSF-2ΔB) were obtained from Dr. Michele Sawadogo (M.D. Anderson Cancer Center). The psvUSF-2ΔB plasmid functions as a dominant negative because it lacks the basic region required for DNA binding [17].

2.4. Transient transfection and luciferase assays

HepG2 cells were seeded at 2.75×10^5/mL per well in 6-well plates in growth medium with 10% serum. Cells were grown for 24 hr at 37°, and transfected the next day with test plasmids using Lipofectamine (GIBCO-BRL) according to the instructions of the manufacturer. Luciferase activities were determined ∼65 hr after transfection. Cells were washed twice with PBS before lysis, and then the lysates were harvested and assayed for luciferase activity using the Promega Luciferase Assay System as specified by the manufacturer. Light output was measured for 10 sec at 25° using a Lumat LB9501 luminometer (Berthold Systems Inc.). Test plasmids were cotransfected with the CMVβ vector (CLONTECH Laboratories, Inc.), which expresses β-galactosidase from the human cytomegalovirus immediate early gene promoter and was used for normalization of transfection efficiency. Lysates were assayed for β-galactosidase using the Promega β-Galactosidase Assay Kit according to the instructions of the manufacturer.

2.5. ChIP assays

ChIP assays were performed following the protocols of Nissen and Yamamoto [18] and, more extensively, Lambert and Nordeen [19] with modifications. HepG2 cells, grown to ∼95% confluency, were cross-linked by the addition of formaldehyde (final concentration 1%) to the medium for 10 min. The cross-linking reaction was stopped by the addition of glycine to a final concentration of 0.12 M. Crude nuclei prepared by hypotonic lysis were resuspended in SDS lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris–HCl, pH 8.0), and sonicated under conditions that reduced the DNA length to less than 2000 bp; cellular debris was removed by centrifugation. The resulting chromatin solution was then diluted with an immunoprecipitation

buffer appropriate for the protein G to be used [50 mM sodium acetate, pH 5.2, 500 mM NaCl, 0.1% SDS, 1% Nonidet P-40, 0.02% sodium azide (Pierce)], incubated with anti-human USF-1 or USF-2 antibody or non-specific anti-mouse IgG overnight at 4°, and immune complexes were collected with preadsorbed (pre-blocked) protein G agarose beads. Following a series of washes and elution of DNA–protein complexes from the beads, cross-links were reversed by heating the samples at 65° overnight followed by proteinase K digestion for 1 hr at 45°. DNA was desalted and concentrated using the Qiagen QIAEX II silica gel kit. Due to the proximity of upstream E-box sites to the E-box 2 site, the focus of interest for this study, immunoprecipitated and uncross-linked DNA was size fractionated on 1.5% Sea Plaque GTG agarose gels [20], and gel cubes corresponding to the range 200 to 500 bp, delineated by a 1 kb DNA ladder, were excised from the appropriate lanes using razor blades. DNA was eluted from the gel using the QIAEX II gel solubilization and extraction kit according to the protocol of the manufacturer. The concentrated eluate was then used in the PCR. PCRs contained specific primers for the *CYP1A2* enhancer region containing the E-box 2 site. Primer locations encompassed −2292 to −2013 relative to the transcription start site and were designated 1A2-7, 5′-CTCCTGAAAATTGTCAAGGTC-3′ (−2292 to −2272), and 1A2-4, 5′-TGAGTCATCTGAACCCAAT-3′ (−2031 to −2013). PCR conditions were preliminary denaturation, 3 min, 95° followed by 36 cycles of 95°, 30 sec, annealing 55°, 1 min, extension 72°, 2 min. As a positive control, 2–3 μg of input DNA was run on the GTG agarose and excised in a manner similar to the immuno samples. Gel cubes excised from blank lanes constituted negative controls and were processed alongside the DNA agarose cubes. A further negative control was provided by the non-specific antibody to mouse IgG. For PCR reactions, 2, 5, and 10 μL of the eluates (in 10 mM Tris–HCl, pH 8.5) from the solubilized gel concentrates were run as described. An additional positive control in the PCR runs was 5 ng of the input DNA that had not been run on a gel.

2.6. Statistical analyses

Statistical differences between values were determined by a one-way ANOVA, followed by either the Dunnett or Tukey–Kramer multiple comparisons post test. Statistics were performed using InStat Instant Statistics (GraphPad Software).

3. Results

3.1. Identification of putative binding sites for E-box binding proteins within the CYP1A2 enhancer

We previously identified a region of the *CYP1A2* MC enhancer that when deleted resulted in an ∼6-fold decrease in basal reporter gene activity [10]. Sequence analysis identified an AP-1 site and a putative E-box site within the deleted region. Since nuclear proteins from untreated cells did not interact with the AP-1 site, we believe that the observed loss of basal reporter activity was the result of deletion of the putative E-box site. Therefore, we scanned the *CYP1A2* 5′-flanking sequence from +3 to −3201 for additional E-box motifs and found six within the previously described upstream enhancer region (−1987 to −3201) [9]. This region contains a weak binding site for the AhR [9], a single binding site for HNF-1 [6], and two AP-1 sites [10], all of which have been shown to confer transcriptional activation on *CYP1A2* reporter gene plasmids. The location of the E-box motifs in the *CYP1A2* upstream enhancer in relationship to binding sites for the AhR (XRE), HNF-1, and AP-1 proteins is shown in Fig. 1. Also shown is a comparison of the *CYP1A2* E-box motifs to the binding site for the ARNT homodimer, USF homo- or heterodimers, and the Myc–Max heterodimer. An additional five putative binding sites were also identified in the promoter region (+3 to ∼−900), but have yet to be analyzed.

3.2. Binding of USF proteins to specific CYP1A2 E-box motifs

We next used EMSA to examine each of the six upstream E-box motifs (EB1-6) for their ability to bind nuclear proteins using extracts from untreated HepG2 cells. We first tested whether oligomers containing each site could compete with USF for binding to a consensus USF binding site (Fig. 2A). When nuclear proteins were incubated with a ^{32}P-labeled ds-oligomer that contained a consensus USF sequence, a retarded complex was observed after electrophoresis on a nondenaturing polyacrylamide gel. We found that EB2 and EB3 oligomers strongly competed for binding as indicated by a reduction of nuclear protein bound to the radiolabeled USF oligomer. The EB5, EB6, and EB1 (not shown) oligomers were unable to compete, while EB4 competed weakly for binding. Specificity of the DNA–protein complexes was demonstrated by preincubating nuclear extracts with 100-fold molar excess of either unlabeled USF or mutated USF oligomers. These data suggest that the EB2, EB3, and possibly EB4 motifs bind proteins related to USF. To determine if these E-box motifs can bind nuclear protein, we labeled each of the oligomers and examined binding patterns by EMSA (Fig. 2B). We observed the binding of nuclear proteins to all three oligomers. However, when the EB2 oligomer was used as a probe, we observed a more prominent DNA–protein complex relative to complexes formed at EB3 and 4 (also, Fig. 3). The DNA–protein complexes formed at these sites were abolished completely by inclusion of as little as 25-fold molar excess unlabeled USF oligomer in the reaction. Additional evidence that proteins bound to *CYP1A2* E-box motifs are related to each

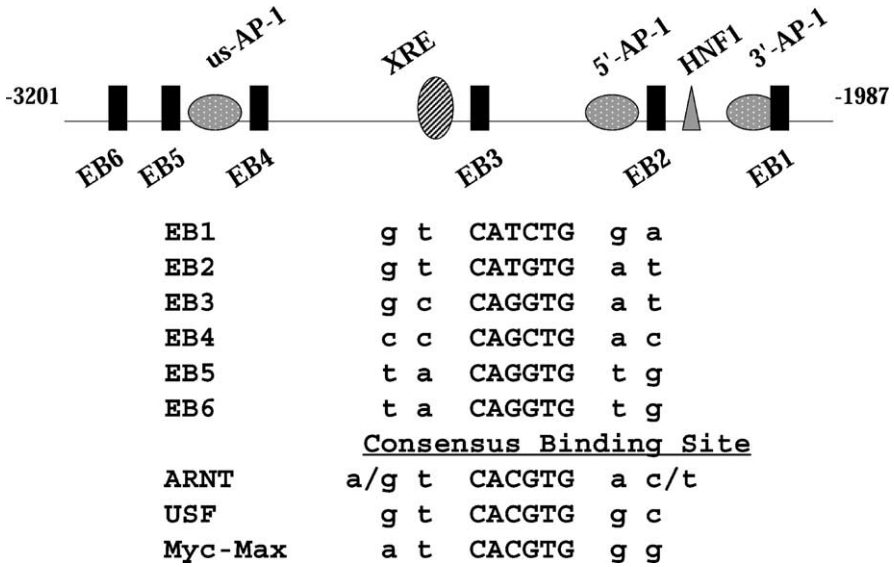

Fig. 1. Location of E-box motifs in the human *CYP1A2* 5′-flanking gene. Six putative binding sites for E-box binding proteins are shown in relationship to binding sites for the AhR (XRE) and AP-1 proteins (us-AP-1, 5′AP-1, 3′AP-1). Also shown is a comparison of the *CYP1A2* E-box motifs to the binding sites for the bHLH proteins, ARNT homodimer, USF-1 and USF-2 heterodimers, and Myc–Max heterodimers. Upper case letters denoting nucleotides represent the consensus binding sites, while lower case letters represent the flanking sequences shown to be important for specific, high-affinity binding [13].

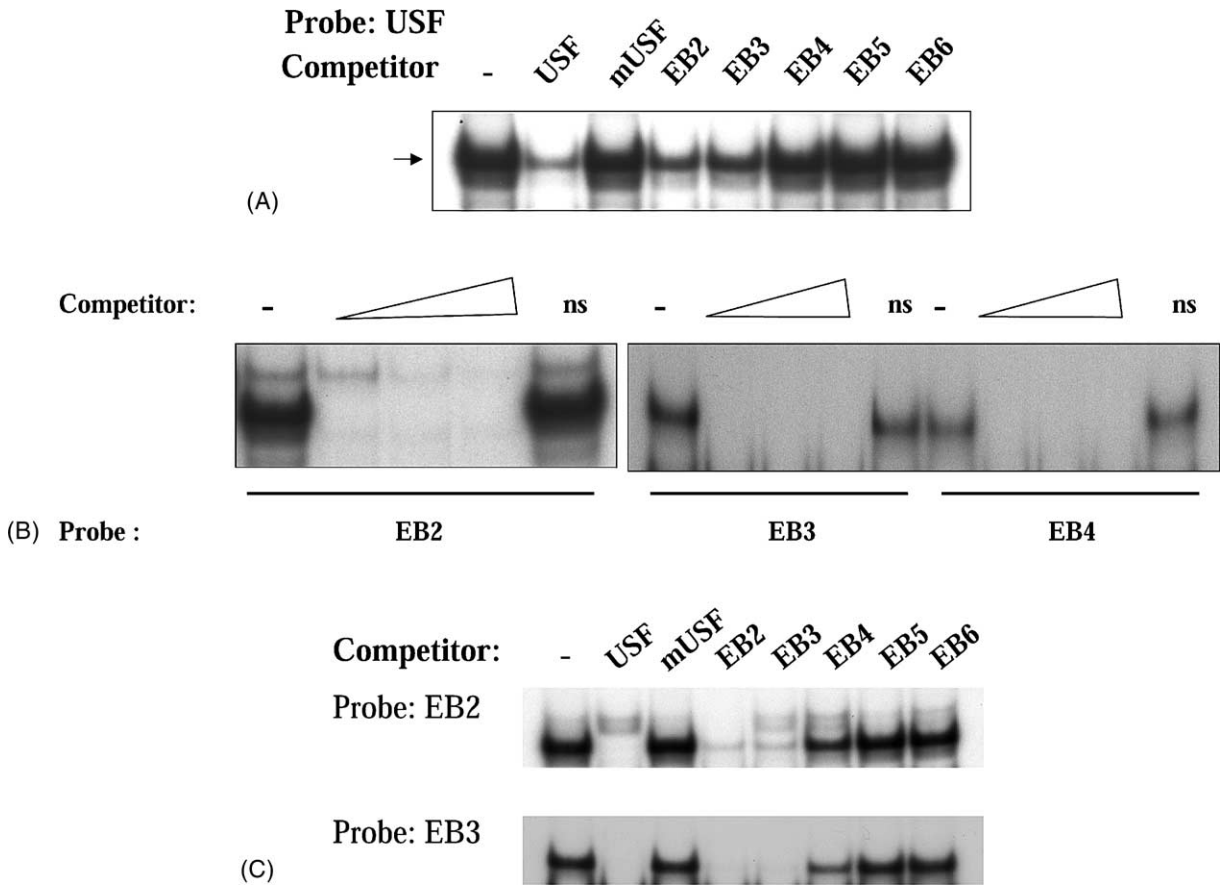

Fig. 2. Nuclear protein binding to the putative *CYP1A2* E-box motifs. The 5′-end-labeled ds-oligomers corresponding to the USF response element consensus sequence or E-box motifs were incubated with nuclear extracts of untreated HepG2 cells and DNA–protein complexes analyzed by EMSA as described under Section 2. Free probe is not shown. (A) Competition by *CYP1A2* E-box motifs for USF binding. The *arrow* indicates the position of the specific, major DNA–protein complex. (B) Binding of nuclear proteins to EB2, EB3, and EB4. A 100-fold molar excess of a non-specific oligomer containing the NFκB binding site (ns) was included to demonstrate specificity. (C) Competition among *CYP1A2* E-boxes. Competitors are at a 100-fold molar excess (A, C) or 25-, 50-, and 100-fold (B) and were incubated with nuclear extract prior to the addition of radiolabeled probe.

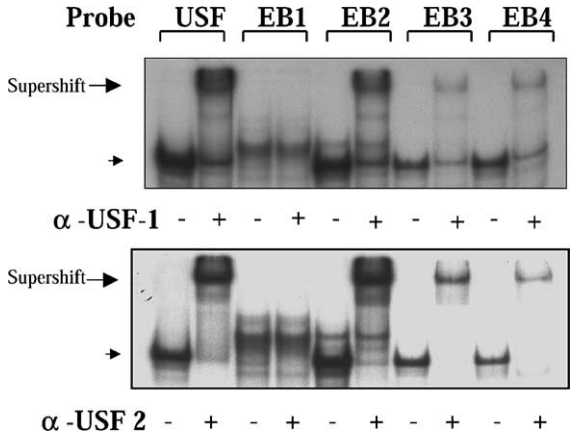

Fig. 3. Supershift assays using epitope specific antibodies to either USF-1 or USF-2. Nuclear extracts of untreated HepG2 cells were incubated with antibodies against USF-1 or USF-2. Antibodies were added after the incubation of nuclear extracts with radiolabeled probe as described under Section 2. The *upper arrow* and the *lower arrow* indicate the positions of the supershifted band and the specific DNA–protein complex, respectively.

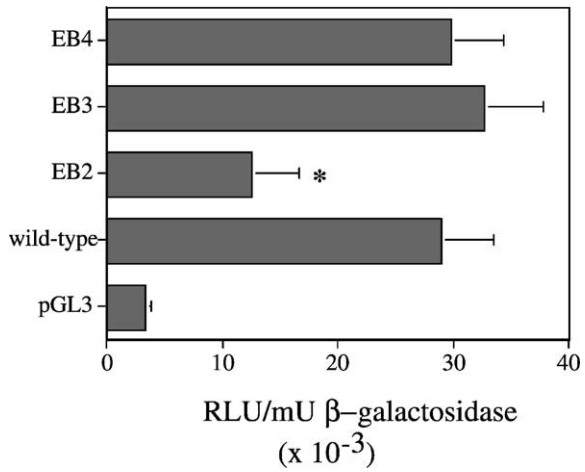

Fig. 4. Analysis of site-directed mutants of E-box motifs. The mutations were generated by site-directed mutagenesis of the luciferase reporter plasmid p3.2CYP1A2, which contains the *CYP1A2* promoter and 5′-flanking sequence from +3 to −3201. HepG2 cells were transiently transfected with p3.2CYP1A2 containing either the intact E-box or the mutated sites. All cultures were cotransfected with the CMVβ vector to normalize transfection efficiency. Values correspond to relative light units (RLU) per milliunit of β-galactosidase activity. Data points represent the mean ± SD of three independent experiments each with 6 replicate wells. The asterisk indicates statistical significance from wild-type control values ($P < 0.01$).

other was obtained by competing each oligomer against each other (Fig. 2C). Results of these experiments indicated that proteins bound to EB2 and EB3 are highly similar to each other and related to USF. Although the USF oligomer could efficiently compete with labeled EB4 oligomer (Fig. 2B), the EB4 oligomer was less effective in competing with EB2 and EB3 for binding proteins (Fig. 2C), suggesting that the proteins bound to EB4 might be different from those bound to EB2 and EB3.

We next employed supershift analysis to obtain additional evidence that USF proteins interact at the *CYP1A2* E-box motifs. We tested for the presence of both USF-1 and USF-2 in complexes formed at EB1, EB2, EB3, and EB4 by using antibodies specific for epitopes on USF-1 and USF-2 in EMSA (Fig. 3). We found that the addition of USF-1 or USF-2 antibodies to the binding reactions retarded the mobility of the major DNA–protein complexes formed at EB2, EB3, and EB4, but not EB1. These results demonstrate that both USF-1 and USF-2 interact most likely as a heterodimer with *CYP1A2* E-box motifs.

3.3. Functional analysis of CYP1A2 E-box motifs

To examine the potential role of EB2, EB3, and EB4 in the basal expression of *CYP1A2*, we employed transient transfection assays of reporter gene constructs containing native E-box (EB) or mutated (mEB) sites. These mutations were generated by site-directed mutagenesis of the luciferase reporter plasmid, p3.2CYP1A2, which contains the *CYP1A2* promoter and 5′-flanking sequence from +3 to −3201. We found that mutations in EB2 had a profound effect on basal reporter gene activity, resulting in a 60% lower level of activity than found in wild-type (Fig. 4). Mutations disrupting EB3 and EB4 had no significant effect on reporter activity when compared with wild-type. p3.2CYP1A2 also contains an HNF-1 half-site (−2119), which has been shown to mediate HNF-1-dependent transcription of a reporter plasmid in the human breast carcinoma cell line MCF-7 [6]. To test whether HNF-1 was important for basal promoter activity in the context of the full flank (+3 to −3201), we used site-directed mutagenesis to destroy this site. Comparison of reporter gene activities between mutated HNF-1 and the wild-type plasmid indicated that HNF-1 plays a minor role in transactivation of *CYP1A2*, since mutating this site diminished basal reporter gene activity by ∼10% when compared with wild-type control plasmid (data not shown).

Given the impact that EB2 had on basal reporter gene activity, we examined this site in greater detail. To test the mutated EB2 motif in the context of different promoters, a segment of the 5′-flanking region containing EB2 or mutated EB2 (mEB2) (−2259 to −1970) was isolated from p3.2CYP1A2 and subcloned into two different reporter plasmids (Fig. 5). These are as follows: a minimal *CYP1A2* promoter (p0.3mpCYP1A2) (+3 to −175) and a thymidine kinase promoter (p0.3tkCYP1A2). The results of these experiments indicated that disruption of EB2 in the context of different promoters produced the same results, i.e. an approximately 60% inhibition of basal gene activity (Fig. 5). These data are in agreement with the EB2 mutation generated within the full-flanking region from +3 to −3201 (Figs. 4 and 5).

3.4. Transactivation of the CYP1A2 promoter by USF-1 and USF-2

To further provide evidence for the role of USF in *CYP1A2* expression, we cotransfected HepG2 cells with

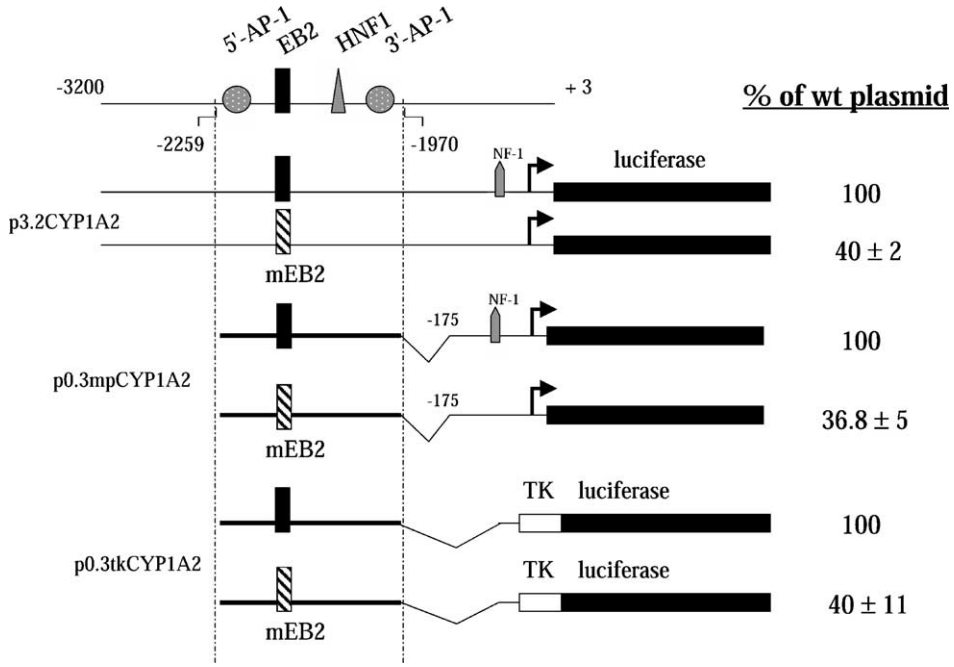

Fig. 5. Functional analysis of EB2 in the context of different promoters. HepG2 cells were transiently transfected with various *CYP1A2* reporter gene constructs containing the intact E-box site (EB2, solid bar) or mutated EB2 (mEB2, striped bar). Mutations were generated using site-directed mutagenesis of wild-type plasmids. The CMVβ vector was cotransfected as an internal control. Values represent relative light units (RLU) per unit of β-galactosidase activity. Luciferase activity is expressed as a percentage of each wild-type plasmid (% of control). Data points represent the mean ± SD of three independent experiments each with 6 replicate wells.

the reporter plasmid, p0.3mpCYP1A2, and increasing amounts of USF-1 or USF-2 expression vectors or control empty vector. We found that overexpression of USF-1 resulted in a concentration-dependent increase in reporter gene activity with an approximately 3-fold induction at the highest DNA concentration used (500 ng) (Fig. 6A). Overexpression of USF-2 at the highest concentration used (500 ng) resulted in only 1.5- to 2-fold induction of reporter activity. The small increases observed most likely are due to the relatively high level of endogenous USF binding activity in these cells (Fig. 3). Increasing amounts of expression vector (>500 ng) resulted in decreases in reporter gene activity (data not shown) possibly through squelching effects on transcription, which can occur through sequestration of general transcription factors at promoters [21]. However, transfection of 1 μg

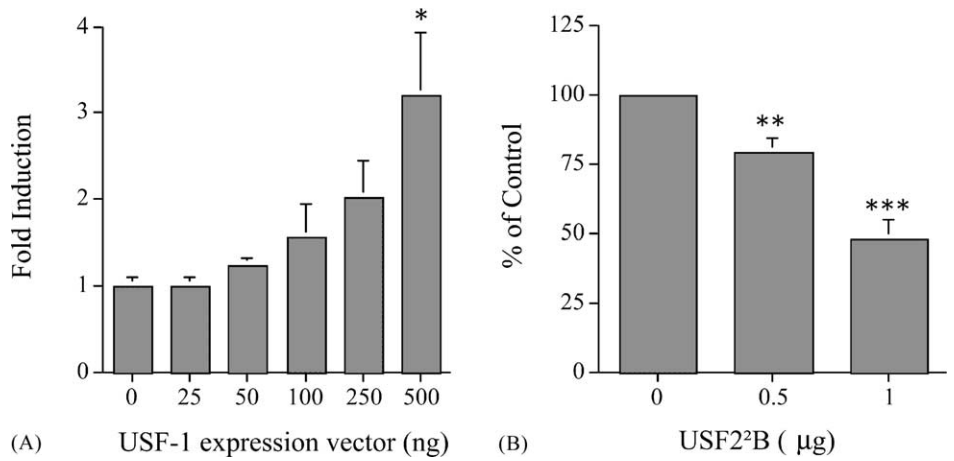

Fig. 6. Expression of exogenous USF-1 or a dominant negative USF-2 mutant in HepG2 cells. (A) Overexpression of USF-1. HepG2 cells were cotransfected with p0.3mpCYP1A2 (1 μg), CMVβgal (0.5 μg), and various concentrations of the expression plasmid, psvUSF-1, or the empty expression vector, pCMV5 (25–500 ng each). The total amount of DNA was kept constant using the empty vector. Results are expressed as the mean of the fold increase above control (i.e. cells receiving no psvUSF-1 expression vector). (B) Effect of a dominant negative USF-2. HepG2 cells were transfected with p0.3mpCYP1A2 in the absence or presence of psvUSF-2ΔB or empty vector. Luciferase activity is expressed as a percentage of p0.3mpCYP1A2 in the absence of psvUSF-2ΔB (% of control). Values in both panels represent the mean ± SD of three independent experiments each with 6 replicate wells. An asterisk(s) indicates statistical significance from wild-type control values ($^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$).

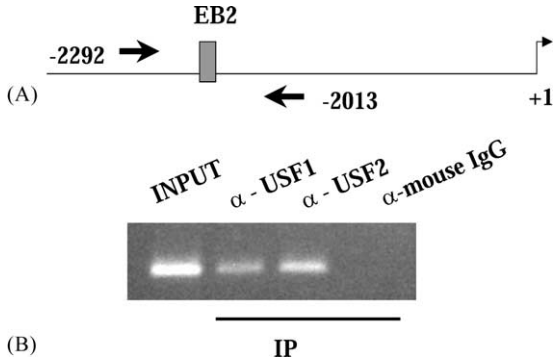

Fig. 7. *In vivo* occupancy of *CYP1A2* EB2 by USF-1 and USF-2. (A) Schematic of the annealing position of PCR primers used to specifically amplify the endogenous *CYP1A2* enhancer containing EB2 (product size 280 bp). (B) Ethidium bromide-stained agarose gel. ChIP assays were performed as described under Section 2. HepG2 cells were exposed to 1% formaldehyde for 10 min. DNA–protein complexes were sonicated, and soluble chromatin was immunoprecipitated by the addition of anti-USF-1 or -USF-2 antibodies or anti-mouse IgG (nonspecific). The purified DNA in the immunoprecipitates was analyzed by PCR using primers encompassing EB2. Aliquots of the chromatin were also analyzed before immunoprecipitation (INPUT). The ChIP assay was performed from two independent cross-linking experiments. Bands from a representative gel are shown.

psvUSF-2ΔB, a dominant negative mutant USF-2 expression vector that lacks the basic region required for DNA binding, resulted in an approximately 50% inhibition of reporter gene activity (Fig. 6B). These data provide evidence for the role of USF in the transactivation of *CYP1A2* since psvUSF-2ΔB is able to form dimers with either USF-1 or USF-2 and sequester endogenous USF.

3.5. ChIP

The *in vitro* binding data and site-directed mutagenesis studies strongly suggest that EB2 may control, in large part, the basal expression of human *CYP1A2*. Accordingly, we performed ChIP assays in HepG2 cells for direct confirmation that binding to this site occurs *in vivo*. We found that addition of anti-USF-1 or -USF-2 to cross-linked protein–DNA complexes resulted in the immunoprecipitation of the region of the *CYP1A2* promoter encompassing EB2 (Fig. 7). The CYP1A2 PCR product was not immunoprecipitated when a non-specific antibody (anti-mouse IgG) was used in identical experiments. The gel excision procedure cannot exclude other small DNA fragments of approximately 300 bp that might contain E-box sites in the immunoprecipitates, for example EB3. However, our use of PCR primers specific for the *CYP1A2* EB2 location indicates that *in vivo* USF proteins are bound to the EB2 motif.

4. Discussion

In our attempt to understand the regulation of CYP1A2 expression in the human liver, we have begun to analyze transcription factor binding sites within the 5′-flanking sequence for factors regulating basal expression. In general, the detailed molecular mechanisms underlying the basal and induced expression of CYP1A2 are poorly understood. CYP1A2 is expressed at high levels in the liver. However, humans show greater than 60-fold inter-individual differences in basal levels of hepatic CYP1A2 [22]. Studies of *CYP1A2* transcriptional regulation should contribute to an understanding of this variability. In the present study, we found that the *CYP1A2* enhancer contains several E-box motifs that are bound by the widely expressed USF. USF proteins, consisting of USF-1 and USF-2, are bHLHZIP proteins that function to transactivate numerous genes (reviewed in Ref. [11]). We demonstrated that the *CYP1A2* enhancer E-box motifs bind complexes that are recognized by epitope specific antibodies to either USF-1 or USF-2, suggesting that the complex consists of a heterodimer of these two proteins. This is consistent with the known properties of USF in which binding activity in most tissues and cell lines is due primarily to the heterodimer [23,24].

Sequence analysis of the 3201 bp of the 5′-flanking human *CYP1A2* gene identified six E-box motifs within the enhancer, three of which were shown to bind USF-1/2. One of these sites, EB2, was shown to be critical for basal expression because we demonstrated significant reduction in activity upon mutating this site even in the context of the cellular promoter and 5′-flanking sequence containing other potentially important transcription factors (e.g. HNF-1, AP-1, and NF-1) (Fig. 4). Mutations of a functional HNF-1 binding half-site within the promoter containing an intact EB2 resulted in only a minor reduction in reporter gene activity. These results are surprising because HNF-1 is enriched in hepatocytes and hepatoma cells and contributes to the transcriptional activation of numerous liver-specific genes [25]. Moreover, overexpression of HNF-1 has been shown to transactivate a CYP1A2 reporter plasmid through this site in the breast carcinoma cell line MCF-7 [6]. This discrepancy in findings could result from the fact that overexpression in MCF-7 can drive reporter gene activity since HNF-1 is in limited amounts, but deletion of the site in the context of an intact EB2 would have a minor effect in HepG2 cells where HNF-1 is relatively abundant [26]. Thus, it is feasible that both USF and HNF-1 contribute to the variable hepatic expression of *CYP1A2*, depending upon their relative concentrations within hepatocytes. Alternatively, hepatoma cell lines such as HepG2 could have altered expression of USF leading to its primary role over HNF-1 observed in our transfection experiments. This seems unlikely since Takahashi *et al.* [27] demonstrated that USF binding activity is significantly greater in rat and rabbit liver nuclear extracts than in extracts from HepG2 cells. Thus, our finding that mutating the EB2 site had a more pronounced effect on promoter activity than mutating the HNF-1 binding site cannot be explained by increased levels of USF in HepG2

cells. Nonetheless, additional studies *in vivo* or in cultured hepatocytes are required to determine if our findings are restricted to hepatoma cells. Additional putative binding sites for HNF-1 and HNF-3 have been identified recently within 4 kb of the *CYP1A2* transcriptional start site [28]. Although results from our studies do not exclude the possibility that these additional distal regulatory elements may be required for high level, tissue-specific expression of *CYP1A2 in vivo*, they do suggest that regulatory elements required for constitutive expression are contained within the first 3 kb.

Additional evidence for the role of USF in the expression of *CYP1A2* was obtained by transactivation experiments of USF expression vectors and the use of a dominant negative mutant of USF-2. Although overexpression of the dominant negative plasmid psvUSF-2ΔB in our transfection experiments did not completely repress transcription, the partial inhibition observed, i.e. approximately 50%, is consistent with reports on the effect of this plasmid on other E-box dependent promoters, and appears to be due to psvUSF-2ΔB being a relatively weak dominant negative plasmid [23]. These transfection studies were complemented by our ChIP assays demonstrating the *in vivo* occupancy of the *CYP1A2* EB2 by USF-1 and USF-2. Taken together, these studies demonstrate that EB2 bound USF functions as a potent activator of *CYP1A2* expression. While our data clearly indicate that USF proteins interact with *CYP1A2* E-box sequences, it is certainly possible that other E-box binding proteins can bind these sequences *in vivo*. Indeed, we found in EMSA experiments using the EB2 oligomer as a probe that inclusion of the USF oligomer that competes for the major DNA–protein complex resulted in the appearance of two minor complexes (see Fig. 2C, lane 2), suggesting competition for this site by different E-box binding proteins. Nucleotides flanking the consensus E-box motif are important for specific, high-affinity binding. Furthermore, the sequences flanking EB2 are identical to the E-box site recognized by the ARNT homodimer (Fig. 1), yet ARNT does not interact with this E-box motif (data not shown and [13]). Yeast one-hybrid screening should help identify USF and other bHLH proteins that interact with the *CYP1A2* E-box sites.

Our findings that the ubiquitously expressed USF proteins appear to play a critical role in the constitutive expression of *CYP1A2* in hepatoma cells are supported by the recent studies of Bengtsson *et al.* [29]. These investigators reported that a single E-box located in the promoter region of the human carboxyl ester lipase gene is important for basal transcription, and point mutations in the E-box almost completely abolish that transcriptional activity. Our findings may be further explained by studies demonstrating the interactions of USF proteins with cell-specific activators resulting in cell-specific gene expression [30]. USF proteins, although widely expressed, are able to confer cell-type specific expression on numerous genes through such interactions with coactivators (e.g. [30–32]). One could also envision a mechanism by which USF interacts with cell-specific repressors in non-hepatic cells, conferring another level of regulation on the expression of hepatocyte specific genes. The restricted pattern of CYP1A2 expression becomes more important when one considers that the route of carcinogen exposure determines how these agents are detoxified or activated. Hepatic CYP1A2 catalyzes the metabolism of numerous carcinogens. CYP1A2 is preferentially expressed in the liver and in the olfactory mucosa of rodents [8], and in humans has been considered to be exclusively expressed in the liver. However, a recent report indicated that CYP1A2 may be expressed in human lung [7]. Thus, the question of how *CYP1A2* is regulated in a cell-type specific manner may well be more complicated than previously thought.

Our studies suggest that USF proteins function as potent activators of human *CYP1A2* gene expression in the hepatoma cell line HepG2. However, their role in hepatocyte-specific control of *CYP1A2* expression has yet to be established. If the findings are also true *in vivo*, then they would suggest the possibility that the restricted expression of *CYP1A2* in hepatic and possibly other tissues (e.g. olfactory mucosa and lung) may involve cell-type specific coactivators and/or corepressors of USF. Studies to identify and characterize these factors should help establish the involvement of USF in the cell-type specific expression of *CYP1A2*. Our findings that USF proteins play a significant role in basal *CYP1A2* expression may lead to a better understanding of these processes.

Acknowledgments

This work was supported by the National Institutes of Health Grant GM54477 (L.C.Q.). DNA sequencing was performed by the University of Colorado Cancer Center DNA Sequencing and Analysis Core Facility, which is supported by the NIH/NCI Cancer Core Support Grant, CA46934.

References

[1] Gonzalez FJ, Gelboin HV. Role of human cytochromes P450 in the metabolic activation of chemical carcinogens and toxins. Drug Metab Rev 1994;26:165–83.

[2] Butler MA, Iwasaki M, Guengerich FP, Kadlubar FF. Human cytochrome $P\text{-}450_{PA}$ (P-450IA2), the phenacetin *O*-deethylase, is primarily responsible for the hepatic 3-demethylation of caffeine and N-oxidation of carcinogenic arylamines. Proc Natl Acad Sci USA 1989;86:7696–700.

[3] Fuhr U, Doehmer J, Battula N, Wolfel C, Kudla C, Keita Y, Staib AH. Biotransformation of caffeine and theophylline in mammalian cell lines genetically engineered for expression of single cytochrome P450 isoforms. Biochem Pharmacol 1992;43:225–35.

[4] Ereshefsky L. Pharmacokinetics and drug interactions: update for new antipsychotics. J Clin Psychiatry 1996;57(Suppl 11):12–25.

[5] Kadlubar FF. Biochemical individuality and its implications for drug and carcinogen metabolism: recent insights from acetyltransferase and

cytochrome P4501A2 phenotyping and genotyping in humans. Drug Metab Rev 1994;26:37–46.

[6] Chung I, Bresnick E. Identification of positive and negative regulatory elements of the human cytochrome P4501A2 (*CYP1A2*) gene. Arch Biochem Biophys 1997;338:220–6.

[7] Wei C, Caccavale RJ, Kehoe JJ, Thomas PE, Iba MM. CYP1A2 is expressed along with CYP1A1 in the human lung. Cancer Lett 2001;171:113–20.

[8] Zhang J, Zhang Q-Y, Guo J, Zhou Y, Ding X. Identification and functional characterization of a conserved, nuclear factor 1-like element in the proximal promoter region of *CYP1A2* gene specifically expressed in the liver and olfactory mucosa. J Biol Chem 2000; 275:8895–902.

[9] Quattrochi LC, Vu T, Tukey RH. The human *CYP1A2* gene and induction by 3-methylcholanthrene. A region of DNA that supports Ah-receptor binding and promoter-specific induction. J Biol Chem 1994;269:6949–54.

[10] Quattrochi LC, Shih H, Pickwell GV. Induction of the human CYP1A2 enhancer by phorbol ester. Arch Biochem Biophys 1998;350:41–8.

[11] Littlewood TD, Evan GI. Transcription factors 2: helix-loop-helix. Protein Profile 1995;2:621–702.

[12] Massari ME, Murre C. Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. Mol Cell Biol 2000;20:429–40.

[13] Swanson HI, Yang J-H. Specificity of DNA binding of the c-Myc/Max and ARNT/ARNT dimers at the CACGTG recognition site. Nucleic Acids Res 1999;27:3205–12.

[14] Takahashi Y, Nakayama K, Itoh S, Fujii-Kuriyama Y, Kamataki T. Inhibition of the transcription of *CYP1A1* gene by the upstream stimulatory factor 1 in rabbits. Competitive binding of USF1 with AhR·Arnt complex. J Biol Chem 1997;272:30025–31.

[15] Weintraub H, Dwarki VJ, Verma I, Davis R, Hollenberg S, Snider L, Lassar A, Tapscott SJ. Muscle-specific transcriptional activation by MyoD. Genes Dev 1991;5:1377–86.

[16] Postlind H, Vu TP, Tukey RH, Quattrochi LC. Response of human CYP1-luciferase plasmids to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin and polycyclic aromatic hydrocarbons. Toxicol Appl Pharmacol 1993; 118:255–62.

[17] Meier JL, Luo X, Sawadogo M, Straus SE. The cellular transcription factor USF cooperates with Varicella-Zoster virus immediate-early protein 62 to symmetrically activate a bidirectional viral promoter. Mol Cell Biol 1994;14:6896–906.

[18] Nissen RM, Yamamoto KR. The glucocorticoid receptor inhibits NFκB by interfering with serine-2 phosphorylation of the RNA polymerase II carboxy-terminal domain. Genes Dev 2000;14: 2314–29.

[19] Lambert JR, Nordeen SK. Analysis of steroid hormone-induced histone acetylation by chromatin immunoprecipitation assay. Methods Mol Biol 2001;176:273–81.

[20] Johnson TA, Wilson HL, Roesler WJ. Improvement of the chromatin immunoprecipitation (ChIP) assay by DNA fragment size. Biotechniques 2001;31:740–2.

[21] Sadowski I, Ma J, Triezenberg S, Ptashne M. GAL4-VP16 is an unusually potent transcriptional activator. Nature 1988;335:563–4.

[22] Nebert DW, McKinnon RA, Puga A. Human drug-metabolizing enzyme polymorphisms: effects on risk of toxicity and cancer. DNA Cell Biol 1996;15:273–80.

[23] Qyang Y, Luo X, Lu T, Ismail PM, Krylov D, Vinson C, Sawadogo M. Cell-type-dependent activity of the ubiquitous transcription factor USF in cellular proliferation and transcriptional activation. Mol Cell Biol 1999;19:1508–17.

[24] Viollet B, Lefrancois-Martinez A-M, Henrion A, Kahn A, Raymond-jean M, Martinez A. Immunochemical characterization and transacting properties of upstream stimulatory factor isoforms. J Biol Chem 1996;271:1405–15.

[25] Xanthopoulos KG, Mirkovitch J. Gene regulation in rodent hepatocytes during development, differentiation and disease. Eur J Biochem 1993;216:353–60.

[26] Mendel DB, Crabtree GR. HNF-1, a member of a novel class of dimerizing homeodomain proteins. J Biol Chem 1991;266:677–80.

[27] Takahashi Y, Nakayama K, Itoh S, Kamataki T. Upstream stimulatory factor 1 (USF1) suppresses induction of CYP1A1 mRNA by 3-methylcholanthrene (MC) in HepG2 cells. Biochem Biophys Res Commun 1997;240:293–7.

[28] Corchero J, Pimprale S, Kimura S, Gonzalez FJ. Organization of the CYP1A cluster on human chromosome 15: implications for gene regulation. Pharmacogenetics 2000;11:1–6.

[29] Bengtsson SH, Madeyski-Bengtson K, Nilsson J, Bjursell G. Transcriptional regulation of the human carboxyl ester lipase gene in THP-1 monocytes: an E-box required for activation binds upstream stimulatory factors 1 and 2. Biochem J 2002;365:481–8.

[30] Lanigan TM, Russo AF. Binding of upstream stimulatory factor and a cell-specific activator to the calcitonin/calcitonin gene-related peptide enhancer. J Biol Chem 1997;272:18316–24.

[31] Garcia C, Nieto A. Two progesterone-dependent endometrial nuclear factors bind to an E-box in the rabbit uteroglobin gene promoter: involvement in tissue-specific transcription. Arch Biochem Biophys 1999;362:301–8.

[32] Han SY, Kim JC, Suh JM, Chung IK. Cell type-dependent regulation of human DNA topoisomerase IIIα gene expression by upstream stimulator factor 2. FEBS Lett 2001;505:57–62.