

**MOLECULAR MECHANISMS REGULATING HUMAN *CYP4B1*
LUNG-SELECTIVE EXPRESSION**

Mark T. Poch, N. Shane Cutler, Garold S. Yost, and Ronald N. Hines

*Departments of Pediatrics and Pharmacology and Toxicology, Medical College of Wisconsin,
Milwaukee, WI 53226 (M.T.P. and R.N.H.) and Department of Pharmacology and Toxicology,
University of Utah, Salt Lake City, UT 84112 (N.S.C. and G.S.Y.)*

Running Title: Lung-Selective *CYP4B1* Regulation

Corresponding Author: Ronald N. Hines, Ph.D.
Department of Pediatrics
Medical College of Wisconsin
8701 Watertown Plank Rd.
Milwaukee, WI 53226
Tel: 414 456-4322
Fax: 414 456-6651
E-mail: rhines@mail.mcw.edu

Number of Text Pages: 34

Number of Tables: 1

Number of Figures: 10

Number of References: 39

Abstract Word Length: 249

Introduction Word Length: 401

Discussion Word Length: 1157

ABBREVIATIONS: C/EBP, CCAAT enhancer binding protein; ChIP, chromatin immunoprecipitation; Elk1, member of ETS family of transcription factors; EMSA, electrophoretic mobility shift assay; HNF3, hepatocyte nuclear factor 3; PCR, polymerase chain reaction; Sp1, sephadex protein 1; Sp3, sephadex protein 3; Sp/XKLF, sephadex protein/Krüppel-like factor family of transcription factors

Abstract

Lung-selective cytochrome P450 expression is well recognized, however, little is known regarding regulatory mechanisms. To address this knowledge gap, transient expression of *CYP4B1*/luciferase constructs was used to identify: a proximal, positively acting regulatory element, position -139 to -45 that functioned in all cells examined, a negatively acting element, position -457 to -216 that only functioned in HepG2 hepatoblastoma cells, and a distal, positively acting element, position -1087 to -1008 that functioned in A549 or BEAS-2B lung-derived cells, but not HepG2 cells or 293 kidney-derived cells. Competitive electrophoretic mobility shift assays further localized specific A549, but not HepG2, nuclear protein binding to two sites within the distal element, *CYP4B1* position -1052 to -1042 and -1026 to -1008. Several potential lung-selective transcription factor recognition sequences were identified within these elements. However, attempts to identify specific factor(s) were unsuccessful. In contrast, *in vitro* DNA/protein binding assays combined with transient expression and mutagenesis studies identified two functional Sp/XKLF sites within the proximal element (position -118 to -114 and position -77 to -73) that bound both Sp1 and Sp3 *in vitro*. Further, Sp1-dependent synergistic regulation was observed in A549 cells involving the proximal and distal regulatory elements. Chromatin immunoprecipitation assays demonstrated binding of neither Sp1 nor Sp3 to the *CYP4B1* proximal element in human liver tissue, while selective Sp1 binding was observed in human lung tissue. Thus, the composite findings are consistent with both the proximal Sp1 elements and the distal regulatory element acting to synergistically control *CYP4B1* lung-selective expression.

The cytochrome P450-dependent monooxygenases consist of a large family of heme proteins that catalyze the oxidative metabolism of both endogenous and exogenous compounds, often facilitating their elimination from the organism. Depending on the substrate, however, these same reactions can lead to reactive intermediates and contribute to a pathological response (Guengerich 1993; Nebert et al. 2002). Although the liver possesses the highest amount of cytochrome P450 metabolic activity, the lung represents a primary portal to the body and as such, is an important exposure site to many inhaled environmental protoxicants and procarcinogens (Dahl et al. 1993). As one might expect, lung tissue also can activate several procarcinogens and protoxicants, contributing to the etiology of cancer and other toxic outcomes (Ding et al. 2003). Gaining a better understanding of mechanisms controlling the selective expression of pulmonary cytochrome P450 enzymes will be critical to fully appreciating their role in pulmonary disease.

Several cytochrome P450 enzymes are selectively expressed in the lung, including CYP2A13, 2F1, 2S1, 3A5 and 4B1. Recent studies have begun to reveal mechanisms controlling the induction of pulmonary *CYP3A5* by glucocorticoids (Hukkanen et al. 2003) and *CYP2S1* by 2,3,7,8-tetrachlorodibenzo-p-dioxin (Rivera et al. 2002). However, there is a paucity of knowledge regarding constitutive lung-selective regulatory mechanisms.

In animal models, CYP4B1 bioactivates several protoxicants and procarcinogens, including 2-aminoanthracene (Smith et al. 1995), 2-aminofluorene (Vanderslice et al. 1985), and valproic acid (Rettie et al. 1995). Combined with its lung-selective expression pattern, such observations resulted in considerable interest in CYP4B1. Yet, the activity and function of human CYP4B1 is controversial. Inter-individual differences in *CYP4B1* expression have been associated with bladder cancer susceptibility (Imaoka et al. 2000) and elevated CYP4B1 mRNA was demonstrated in lung carcinoma versus normal lung tissue (Czerwinski et al. 1994). Imaoka *et al.* (2001) also demonstrated elevated lauric acid ω -hydroxylase and 2-aminofluorene-dependent *umu* gene

DMD #4523

expression using microsomes from human CYP4B1 transgenic mice that were inhibited by a CYP4B1 antibody. Similar activities also were observed using a CYP4B1-NADPH cytochrome P450 oxidoreductase fusion protein expressed in yeast. In contrast, Zheng *et al.* (2003) reported that human CYP4B1 does not appear to be an active enzyme due to a unique amino acid substitution in the meander region. Irrespective of this controversy, CYP4B1 mRNA is detectable in lung. Thus, this system remains a viable model to better understand molecular mechanisms controlling cytochrome P450 lung-selective expression and was the focus of this study.

Materials and Methods

Materials. A549 lung epithelial carcinoma cells, BEAS-2B immortalized human bronchial epithelial cells, and 293 adenovirus 5 transformed human kidney epithelial cells were obtained from the American Type Culture Collection (Manassas, VA). HepG2 hepatoblastoma cells were a gift from Dr. Barbara Knowles (Jackson Laboratories, Bar Harbor, ME). The plasmid, p14-2 (Yokotani et al. 1990), containing the human *CYP4B1* gene, was provided by Dr. Y. Fujii-Kuriyama (Osaka Prefectural Institute of Public Health, Osaka, Japan). Sp1 and Sp3 expression vectors were generously provided by Dr. Guntram Suske (Institut für Molekularbiologie und Tumorforschung, Philipps-Universität, Marburg, Germany), an AP-4 expression vector from Dr. Michael Lehman (University of Arkansas, Fayetteville, AR) and C/EBP α and β expression vectors from Dr. Peter Johnson (National Cancer Institute, Bethesda, MD). The luciferase reporter plasmid (pGL3Basic) and luciferase reporter assay kit was purchased through Promega (Madison, WI) and the luminescent β -galactosidase assay kit from Clontech (Palo Alto, CA). Oligonucleotides were synthesized by MWG Biotech (High Point, NC). The [α - 32 P]-dCTP radioisotope (3000 Ci/mmole) was acquired from Perkin Elmer Life Sciences (Boston, MA). The Micro BCA kit for protein determination using the method of Smith *et. al.* (1985) was purchased from Pierce (Rockford, IL). Sp1, Sp3, Elk-1, and C/EBP (α , β , γ , and δ forms) antibodies were purchased from Santa Cruz Biotechnology, Inc., while the AP-4 antibody was provided by Michael Lehmann (Institut für Genetik der Freien Universität Berlin, Berlin, Germany). Restriction endonucleases and DNA modifying enzymes were purchased from New England Biolabs (Beverly, MA). Cell culture reagents and routine chemicals were supplied by Sigma Aldrich (St. Louis, MO). High purity plasmid purification kits were procured from Marligen, (Ijamsville, MD). Human lung tissue was obtained from Rocky Mountain Donor Services while human liver tissue was the generous gift of Dr. Michael Franklin (University of Utah,

Salt Lake City, UT). Lipofectamine 2000 and Platinum Taq DNA polymerase were purchased from Invitrogen (Carlsbad, CA). All other reagents were obtained from commercial sources at the purest grade available.

Luciferase reporter plasmid constructs. *CYP4B1* gene coordinates are based on human chromosome 1 contig NT_032977.7 (build 35.1) (<http://www.ncbi.nlm.nih.gov>). The transcription start site, 36 bp upstream from the ATG start codon in exon 1, was determined by 5'-RACE (data not shown) and is in agreement with the major transcription start site reported in the DBTSS database (<http://dbtss.hgc.jp>). A *BgIII/NcoI* fragment representing human *CYP4B1* position -2183 to +35 was isolated from p14-2 and cloned into the unique *BgIII/NcoI* sites in pGL3basic. The *CYP4B1* sequences in this plasmid, designated pRNH684, were verified by DNA sequence analysis (Sanger et al. 1977). Two approaches were used to prepare nested deletions of the *CYP4B1* upstream region. First, pRNH684 was digested with *XhoI*, which cleaved at a unique sequence within the multiple cloning site immediately upstream of the *CYP4B1* insert, along with a second enzyme that cleaved at a unique site within the insert. The released *CYP4B1* fragment was discarded, the remaining DNA treated with T4 DNA polymerase to create flush-ends, and then self-ligated using T4 DNA ligase. Thus, digestion of pRNH684 with *XhoI/SpeI* yielded pRNH686 (*CYP4B1* position -1087 to +35 directing luciferase expression) while digestion with *XhoI/HinDIII* yielded pRNH683 (*CYP4B1* position -457 to +35 directing luciferase expression). In a second approach, high fidelity polymerase chain reaction (PCR) DNA amplification was performed using *BamHI* linearized pRNH684 or pRNH683 as a template. One of multiple unique sense primers that included a non-homologous 5' *KpnI* site was paired with a 3' antisense primer which annealed immediately downstream of the unique *HinDIII* site at *CYP4B1* position -457 (pRNH684 as template) or the unique *NcoI* site at *CYP4B1* position +35 (pRNH683 as template). PCR products were digested with *KpnI/HinDIII* or *KpnI/NcoI*, respectively, and used to replace the *KpnI/HinDIII*

or *KpnI/NcoI* fragments in pRNH684, yielding pRNH833 (*CYP4B1* position -1006 to +35 directing luciferase expression), pRNH834 (*CYP4B1* position -890 to +35 directing luciferase expression), pRNH835 (*CYP4B1* position -749 to +35 directing luciferase expression), pRNH836 (*CYP4B1* position -660 to +35 directing luciferase expression), pRNH699 (*CYP4B1* position -216 to +35 directing luciferase expression), pRNH698 (*CYP4B1* position -139 to +35 directing luciferase expression), and pRNH697 (*CYP4B1* position -45 to +35 directing luciferase expression). Similarly, an internal deletion was prepared using a sense primer that annealed at *CYP4B1* position -217 and contained a non-homologous 5' *HinDIII* site, and the antisense primer described above that annealed downstream of the *NcoI* site at +35. The resulting amplicon was digested with *HinDIII/NcoI* and cloned into pRNH686 that had been digested with *HinDIII/NcoI*, deleting a 492 bp fragment representing *CYP4B1* position -457 to +35. The resulting plasmid, pRNH922, contains *CYP4B1* position -1087 to -457 spliced to position -217 to +35 directing luciferase expression.

Cell culture and transient expression. A549 cell lung carcinoma cells were maintained in Dulbecco's Modified Eagle's medium/nutrient mixture F12 containing 10% fetal bovine serum and the antibiotics, penicillin (5 U/ml) and streptomycin (50 µg/ml). HepG2 hepatoblastoma and 293 cells were cultured in Dulbecco's Modified Eagle's Medium supplemented with 10% fetal bovine serum and antibiotics as described by Boucher *et. al.* (1993). BEAS-2B cells were cultured in LHC-9 medium (Biofluids, Rockville, MD) using fibronectin/collagen precoated flasks per the recommendations of the American Type Culture Collection. For transfection studies, 1×10^5 cells were sub-cultured in each well of a 24-well plate. Twenty four hours later the cells were at approximately 80% confluence and were transfected with reporter constructs and if appropriate, expression vectors, using a mixture of 3 µg Lipofectamine 2000, 0.8 µg test plasmid(s) and 0.2 µg pCMVβgal. After incubation for 16 h at 37°C, transfection medium (Optimem with 5% fetal bovine serum) was replaced with growth medium and the cells were incubated an additional 48 h. The cells

were processed and luciferase assays performed following the manufacturer's recommended instructions. Luciferase activity was normalized to β -galactosidase activity to correct for transfection efficiency. At least two separate reporter plasmid preparations were used in two independent transfection assays with each experiment performed in triplicate.

Electrophoretic Mobility Shift Assay (EMSA). HepG2 and A549 nuclear protein extracts were prepared as described by Chodosh (1988) using optimized final KCl extraction concentrations of 400 mM and 600 mM, respectively. Nuclear protein extracts were assayed for protein content, divided into aliquots, and stored at -80°C until used. Based on transient expression assays with nested deletion mutants, EMSA probes were prepared using specific PCR primer pairs containing non-homologous *Bam*H1 recognition sequences on their 5'-end. Amplified *CYP4B1* sequences included positions -1087 to -1008, -1052 to -1008, -216 to -138, and -138 to -46. After digestion of the PCR products with *Bam*H1, the probes were radiolabeled using Klenow DNA polymerase (exonuclease-), unlabeled deoxynucleotides and [α - ^{32}P]-dCTP. EMSA was performed as described by Boucher *et al.* (1993) with the exception that DNA/protein binding was performed in 10 mM Hepes-HCl pH 7.9, 75 mM KCl, 0.2 mM EDTA, 0.1 mM dithiothreitol, 5% glycerol and 0.24 $\mu\text{g}/\text{ml}$ bovine serum albumin. Each 30 μl binding reaction contained 10 μg nuclear protein, 1.5 μg poly[d(I-C)] and 0.2 ng (approximately 30,000 dpm) ^{32}P -labeled probe. To more finely resolve binding elements within the proximal *CYP4B1* promoter, double-stranded oligonucleotides were prepared representing *CYP4B1* positions -142 to -119, -128 to -105, -112 to -71, and -88 to -61 and used in competitive EMSA at a 25- to 300-fold molar excess over the *CYP4B1* position -138 to -46 probe. Somewhat similarly, double-stranded oligonucleotides representing *CYP4B1* sequences from position -1068 to -1008, -1052 to -1008, -1052 to -1026, and -1042 to -1008 were prepared and used in competitive EMSA at a 10- to 100-fold molar excess over the *CYP4B1* position -1052 to -1008 probe. Finally, double-stranded oligonucleotides representing known or consensus binding sites for

the Sp/XKLF family, AP-4, Elk-1, HNF3, and C/EBP (Table 1) were prepared and used in competitive EMSA at a 150-fold molar excess over the *CYP4B1* position -1052 to -1008 probe. EMSA supershift experiments were performed using Transcruz antibodies (Santa Cruz, CA) or antibody preparations generously donated by other investigators (see above). A549 nuclear extract was incubated for 15 min at 4°C with 0.5-1.0 µg antibody either prior to or after the addition of the labeled probe.

Site-directed mutagenesis. Mutagenesis was achieved using the QuickChange™ Site-Directed Mutagenesis kit (Stratagene, La Jolla, USA) as directed by the manufacturer. The mutagenic primers (targeted nucleotides are lower case), 5'-GGA GGT TGC TGT CTtt GCC TTA TGG CAC TCA GC-3' and 5'- TGC TGC TGG CTG CAtt GAG TGG CTA GGG-3' were used with pRNH686 as a template, mutagenizing the Sp/XKLF sites at *CYP4B1* position -118 to -114 (site 2) from CCGCC to ttGCC and -77 to -73 (site 5) from GCAGG to GCAtt, respectively, and generating the following modified reporter constructs: pRNH838 (*CYP4B1* -1087 to +35 directing luciferase expression with site 2 modified), pRNH842 (*CYP4B1* -1087 to +35 directing luciferase expression with site 5 modified), pRNH843 (*CYP4B1* -1087 to +35 directing luciferase expression with site 2 and site 5 modified). Using the same approach as described above to generate the original deletion construct, the following reporter constructs were prepared: pRNH844 (*CYP4B1* -139 to +35 directing luciferase expression with site 2 modified), pRNH845 (*CYP4B1* -139 to +35 directing luciferase expression with site 5 modified), and pRNH846 (*CYP4B1* -139 to +35 directing luciferase expression with site 2 and site 5 modified). In all instances, the nucleotide changes were verified by DNA sequence analysis.

Chromatin Immunoprecipitation. Chromatin Immunoprecipitation (ChIP) was performed following the procedure of Shang *et al.* (2000) with minor modifications. Approximately 200 mg of human lung or liver tissue was pulverized in liquid nitrogen, transferred immediately to a 1%

formaldehyde solution at room temperature and fixed for 15 min. After washing the fixed tissue and resuspension in lysis buffer as described (Shang et al. 2000), the tissue was sonicated 10 times for 10 sec each at a power setting of 6 (maximum setting of 10) using a W-220 ultrasonic processor (Heat-System Ultrasonics, Inc., Plainview, NY). Sonicated samples were centrifuged for 10 min at 16,060 x g to remove particulate matter. An aliquot was taken for analysis of DNA fragmentation and reserved for use as a positive control during PCR analysis. The remaining sample volumes were divided into three equal parts and adjusted to 1 mL with cold dilution buffer (1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris-HCl, pH 8.1). To each sample, 50 μ L of salmon sperm DNA (20 μ g) and protein A agarose (0.15 mg) (UpState Biotechnology, Lake Placid, NY) was added and the diluted samples were incubated at 4°C for 2 h with rocking to minimize non-specific binding. Samples were centrifuged as before and the supernatant fractions transferred to new tubes. The samples were incubated overnight with rocking at 4°C along with either Sp1, Sp3 or no antibody (control). After adding 50 μ L of salmon sperm DNA/protein A agarose (UpState Biotechnology), the samples were incubated for an additional 1 h at 4°C before collecting the protein-A-bound antibody-complexed chromatin fragments by centrifugation as before. Supernatant fractions were discarded. The protein A agarose-bound antibody-complexed chromatin fragments were sequentially washed as described. The chromatin was eluted from the complexes by incubating with three sequential 50 μ L aliquots of elution buffer (50 mM Tris, 1% SDS, 10 mM EDTA, pH 8.0) which were then pooled. After incubating overnight at 65°C to reverse cross-linking, DNA was purified using a QIAquick PCR purification kit (Qiagen, Valencia, CA). PCR analysis was carried out using 1 μ L of the purified DNA, oligonucleotide primers, 5'-TAC CCC AAA ACC AGG CCC CAG GC-3' (*CYP4B1* position -241 to -219) and 5'-GCT GGG ATG CGC TAG AGC CTG C-3' (*CYP4B1* position -1 to +20), and Platinum Taq DNA polymerase (Invitrogen).

Data Analysis. Regulatory domains identified through functional assays were examined for putative transcription factor binding sites using the Match program in conjunction with the TRANSFAC Professional V 8.3 database (<http://www.biobase.de>) (Matys et al. 2003). Luciferase reporter assay results were assessed using a one-way ANOVA with a Dunnett's T3 post-hoc multiple comparison test (SPSS, Chicago, IL). A value of $p < 0.05$ was accepted as a significant difference.

Results

Localization of *CYP4B1* tissue-selective regulatory domains. To identify functional *CYP4B1* regulatory domains, a series of reporter constructs were prepared in which *CYP4B1* sequences from position -2183 to +35, -1087 to +35, -457 to +35, -216 to +35, -139 to +35, and -45 to +35 directed luciferase expression. Transient expression of these constructs in A549 lung carcinoma cells (Fig. 1) failed to reveal an increase in promoter activity over background (pGL3Basic) with the -45 to +35 construct. However, inclusion of *CYP4B1* sequences from position -139 to +35 resulted in a 4-fold increase in reporter activity, suggesting the presence of a proximal positive regulatory domain. An additional 2-fold increase in reporter activity was observed when a more distal domain, *CYP4B1* positions -1087 to -457, was included in the reporter construct while a significant decrease in activity was seen by extending the included *CYP4B1* sequences to position -2183.

To identify *CYP4B1* domains selectively active in lung-derived cells, the same reporter constructs were used to transiently transfect both A549 lung carcinoma and HepG2 hepatocellular carcinoma cells (Fig. 2), as well as BEAS-2B SV40 large T antigen immortalized bronchial epithelial and 293 adenovirus 5 transformed kidney cells (data not shown). Consistent with what was observed in Fig. 1, transient expression of the construct containing *CYP4B1* sequences from position -139 to +35 resulted in a 4-fold increase in A549 cell luciferase activity (Fig. 2B) while a 3-fold increase was observed in HepG2 cells (Fig. 2A). Qualitatively similar results were observed in both BEAS-2B and 293 cells (data not shown). Although no further change in luciferase activity was observed in any of the cell lines by including sequences up to position -216, extending the *CYP4B1* sequences to position -457 resulted in a significant decrease in HepG2 luciferase activity (Fig. 2A), but no change in A549 (Fig. 2B), BEAS-2B, or 293 cells (data not shown). Finally, an

additional 2-fold increase in reporter activity was observed when a more distal domain, *CYP4B1* position -1087 to -457, was included in the reporter construct and used to transfect A549 cells (Fig. 2B), but no change was observed in HepG2 cells (Fig. 2A). Results qualitatively similar to those seen with A549 cells were observed when these same transient expression assays were performed using the human bronchial epithelial cells, BEAS-2B (data not shown). In contrast, when the experiments were repeated in the kidney-derived 293 cells, up-regulation was not observed with the *CYP4B1* position -1087 to -457 domain (data not shown). These results are consistent with a ubiquitous, proximal positively acting domain between *CYP4B1* position -139 and -45, a liver-selective negatively acting domain between *CYP4B1* position -457 and -216, and a lung-selective positively acting domain between *CYP4B1* position -1087 and -457.

A series of reporter constructs in which *CYP4B1* sequences from position -1087 to +35, -1006 to +35, -890 to +35, -749 to +35, -660 to +35, and -457 to +35 directed luciferase expression were utilized for transient expression assays in A549 cells to further define the location of the distal, lung-selective *CYP4B1* regulatory element. No increase in reporter gene activity over that seen with the proximal regulatory element was observed with constructs that included *CYP4B1* sequences up to position -1008 (pRNH833) (Fig. 3). However, a significant increase was observed when sequences up to position -1087 were included. These data indicated the distal regulatory element resides between *CYP4B1* position -1087 to -1008.

Regulation by the proximal *CYP4B1* positive regulatory domain involves Sp/XKLF factors. EMSA was used to assess whether or not specific DNA/protein interactions were possible within the proximal *CYP4B1* positive regulatory domain (position -138 to -46). Using A549 cell nuclear protein extract, five DNA/protein complexes were observed, three of which represented specific DNA/protein interactions based on competition with unlabeled probe DNA (Fig. 4, complexes A, B and C, compare lanes 2 and 3). Analysis of the *CYP4B1* sequences from position

-138 to -46 using the Match program in conjunction with the TRANSFAC V 8.3 Professional database revealed six strong matches to the consensus binding site for the Sp/XKLF family of transcription factors (Fig. 5). Consistent with this observation, competitive EMSA using a 100-fold molar excess of a consensus Sp/XKLF sequence (Table 1) resulted in the elimination of all three specific DNA/protein complexes (Fig. 4, complexes A, B and C, compare lanes 2 and 4). Supershift EMSA was used to determine which of the Sp/XKLF factors might be involved in specifically binding the *CYP4B1* proximal regulatory element. Inclusion of an Sp1 monospecific antibody in the DNA/protein binding reaction resulted in a supershift of complex B (Fig. 4, lane 5) while addition of an Sp3 monospecific antibody resulted in a supershift of both complexes A and C (Fig. 4, lane 6). Addition of both antibodies resulted in a supershift of all three specific DNA/protein complexes (Fig. 4, lane 7). Thus, these data strongly implicate a role for Sp1 and/or Sp3 in regulating *CYP4B1* expression through the proximal regulatory element.

Co-transfection of A549 cells with both the *CYP4B1*-directed reporter constructs and Sp1 or Sp3 expression vectors was used to assess the functional role of the identified Sp/XKLF sites (Fig. 6). Neither Sp1 or Sp3 had any impact on the background reporter activity observed with the promoterless pGL3basic reporter vector. Co-transfection with the pRNH698 reporter construct (*CYP4B1* position -138 to +30 directing luciferase expression) and either an Sp1 or Sp3 expression vector resulted in a 3-fold increase in promoter activity relative to that observed with a blank expression vector (Fig. 6A). Co-transfection with the pRNH686 reporter construct (*CYP4B1* position -1087 to +30 directing luciferase expression) and the Sp1 expression vector resulted in a 5-fold increase in reporter activity. In contrast, co-transfection with pRNH686 and the Sp3 expression vector had no effect on luciferase activity (Fig. 6B). Similar co-transfection experiments in HepG2 cells also resulted in a 5-fold increase in activity with the pRNH686 reporter construct (data not shown).

Sp1/Sp3 binding to the proximal regulatory domain *in vivo* was determined using ChIP analysis along with the same cells lines in which transient expression assays were performed (A549, HepG2, and BEAS-2B cells), and autopsy samples of liver and lung tissue. The *CYP4B1* proximal regulatory domain was bound by Sp1, but not Sp3 in HepG2 cells (Fig. 7A). In contrast, both Sp1 and Sp3 were observed to bind the proximal regulatory domain in A549 cells with Sp1 in excess of Sp3 while in BEAS-2B cells, both Sp1 and Sp3 bound, but Sp3 was in excess of Sp1. Binding appeared much less robust in the latter cell line. A different situation was observed in the lung and liver tissue samples. The *CYP4B1* proximal regulatory domain was bound by Sp1 in lung tissue, but not liver (Fig. 7B) while Sp3 binding was not observed in either tissue. These latter results contrast to what was observed with the *CYP2F1* basal promoter using these same antibodies and tissue samples in which Sp1 binding was observed in lung, but not liver while Sp3 binding was observed in liver, but not lung (Wan et al. 2005).

Given that six putative Sp/XKLF sites were identified in the proximal element (Fig. 5), competitive EMSA was used to determine which of these sites might be responsible for the observed specific DNA/protein interactions and functional impact on the *CYP4B1* promoter. Double-stranded oligonucleotides incorporating site 1, sites 3/4 and site 6 failed to compete at even a 300-fold molar excess (Fig. 8, lanes 5-7). In contrast, concentration-dependent competition was observed with the double-stranded oligonucleotide incorporating the site 2 putative Sp/XKLF element (*CYP4B1* position -128 to -105, Fig. 8, lanes 8-10). Although not as strong, concentration-dependent competition also was observed with the double-stranded oligonucleotide incorporating the putative site 5 element (*CYP4B1* position -87 to -59, Fig. 8, lanes 11-13). Consistent with these results, mutagenesis of sites 2 and 5 from 5'-CCGCC-3' to 5'-TTGCC-3' within the *CYP4B1* -138 to -46 fragment, while leaving sites 1, 3-4 and 6 intact, eliminated the ability of this fragment to compete for specific DNA/protein binding (Fig. 8, lanes 14 and 15).

To validate the activity of Sp/XKLF sites 2 and 5 identified by competitive EMSA, site directed mutagenesis was employed to eliminate Sp1/Sp3 binding at these elements within the *CYP4B1* reporter constructs. The same base changes made within Sp/XKLF site 2 (*CYP4B1* position -118 to -114) and/or Sp/XKLF site 5 (*CYP4B1* position -77 to -73) that eliminated specific DNA/protein binding (Fig. 8) were introduced into pRNH698 (*CYP4B1* position -139 to +35 directing luciferase expression) and pRNH686 (*CYP4B1* position -1087 to +35 directing luciferase expression) resulting in pRNH844 and 838 (site 2 mutagenized), pRNH845 and 842 (site 5 mutagenized) and pRNH846 and 843 (sites 2 and 5 mutagenized), respectively. The impact of these changes on *CYP4B1* directed luciferase expression was examined after transient transfection into A549 cells (Fig. 9). When examined within the context of the proximal regulatory element alone (*CYP4B1* position -139 to -45), an approximate 2-fold loss in promoter activity was observed with the elimination of site 2 (pRNH844 vs 698, $p < 0.05$) or site 5 (pRNH845 vs. 698, $p < 0.001$) alone. However, mutagenesis of both sites 2 and 5 resulted in a complete loss of promoter activity (pRNH846 vs. 698, $p < 0.001$; pRNH846 vs. pGL3basic, $p > 0.05$).

Loss of Sp1/Sp3 binding at the proximal regulatory element also had a substantial impact on the function of the lung-selective distal regulatory element (*CYP4B1* position -1087 to -1008) (Fig. 8). Elimination of either Sp/XKLF site 2 or site 5 by itself within the context of both the proximal and distal elements decreased *CYP4B1* promoter activity by 1.5- (pRNH838 vs. 686, $p < 0.01$) or 3-fold (pRNH842 vs. 686, $p < 0.001$), respectively. The reporter vector pRNH846 with both Sp/XKLF sites 2 and 5 mutagenized decreased the reporter activity 5-fold (pRNH843 vs. 686, $p < 0.001$), but to a level that remained 2-fold above that observed with the promoterless vector (pRNH843 vs. pGL3Basic, $p < 0.001$).

Further characterization of the distal *CYP4B1* regulatory element. To evaluate specific protein binding within the *CYP4B1* distal, lung-selective regulatory element, EMSAs were

performed using *CYP4B1* sequences from position -1087 to -1008 as a probe along with nuclear protein extract prepared from A549 cells. A single specific DNA/protein complex was observed, as demonstrated by the ability of a 50-fold molar excess of unlabeled DNA probe to eliminate binding to the radiolabeled DNA probe. Competition with two overlapping double-stranded oligonucleotides further localized the domain responsible for specific binding to *CYP4B1* sequences position -1052 to -1008 (data not shown). Using *CYP4B1* sequences from -1052 to -1008 as an EMSA probe, two major and one minor specific DNA/protein complexes were observed (Fig. 10, complexes A, B and C, compare lane 2 with lanes 3 and 4). Competition with a 100-fold molar excess of an oligonucleotide representing *CYP4B1* position -1052 to -1026 eliminated specific DNA/protein complex C, but not A or B (Fig. 10, lane 5). In contrast, an oligonucleotide representing *CYP4B1* position -1042 to -1008 failed to compete for specific binding (Fig. 10, lane 6). When repeated using HepG2 nuclear extract, specific DNA protein interactions were not observed with the *CYP4B1* position -1052 to -1008 probe. These observations are consistent with two binding domains between position *CYP4B1* position -1052 and -1042 and position -1026 and -1008 in which binding to the upstream site is obligatory for subsequent binding to the downstream site. Further, these DNA/protein interactions appear lung-selective.

Analysis of the *CYP4B1* distal regulatory element, position -1052 to -1008, using the Match program in conjunction with the TRANSFAC V 8.3 Professional database revealed strong matches for putative AP-4, Elk-1, C/EBP and HNF3 sites. Competitive EMSA was employed to identify which, if any of the identified factors were capable of specifically binding the *CYP4B1* lung-selective distal element. However, competition was not observed using a 150-fold molar excess of known consensus binding sequences for AP-4 (Fig. 10, lane 7), C/EBP (Fig. 10, lane 8), Elk-1 (Fig. 10, lane 9) or HNF-3 (Fig. 10, lane 10). In addition, antibodies for AP-4, Elk-1, C/EBP α , C/EBP β , C/EBP δ and C/EBP γ failed to shift the mobility of any of the specific DNA/protein complexes

DMD #4523

observed with the *CYP4B1* position -1052 to -1008 probe (data not shown). Finally, co-transfection with pRNH686 and AP-4, C/EBP α or C/EBP β expression vectors failed to alter reporter activity (data not shown).

Discussion

This investigation used *CYP4B1* as a model system to examine lung-selective mechanisms controlling cytochrome P450 expression. Three major regulatory elements were identified, a proximal positive regulatory element located between *CYP4B1* position -118 and -73 that involves Sp1 and/or Sp3 binding at two sites, a liver-selective negative regulatory element located between *CYP4B1* position -457 and -216, and a distal, lung-selective positive element located between *CYP4B1* position -1052 and -1008 that involves as yet unidentified transcription factor(s). Although the *in vitro* DNA/protein binding assays and the transient expression data are consistent with both Sp1 and Sp3 regulating *CYP4B1* in a tissue-non-selective fashion, the *in vivo* ChIP data implicates the selective binding of Sp1 to the *CYP4B1* proximal promoter element in lung tissue and as such, is consistent with the conclusion that the proximal regulatory element also contributes to *CYP4B1* lung-selective expression.

A Match analysis of the proximal *CYP4B1* element using the TRANSFAC Professional v 8.3 database identified six potential Sp/XKLF binding sites. However, only two of the sites appear to be functional (*CYP4B1* position -118 to -114 and -77 to -73); both capable of binding the Sp/XKLF family members Sp1 and Sp3 and responding to both factors in transient expression assays. Sp1 and Sp3 are ubiquitously expressed, consistent with the observation that the *CYP4B1* proximal element functioned equally well in both lung- and liver-derived cells. These latter data also are consistent with the inability of the lung-selective member of the Sp/XKLF family, LKLF (Anderson et al. 1995), to modulate the activity of the proximal element. Major differences between Sp1 and Sp3 lie with the ability of Sp1, but not Sp3, to form multimeric complexes and in doing so, synergistically up-regulate promoter activity (Su et al. 1991; Yu et al. 2003). In addition, the Sp1 inhibitory domain is located at the N-terminus of the protein while in Sp3, it is located next to the

three conserved Cys₂His₂ zinc fingers involved in DNA binding (Suske1999). These differences in functional motifs permits Sp3 to commonly function as a repressor, although instances also exist where it serves as a strong, positively acting factor (Rao et al. 2002; Sowa et al. 1999). In the case of the *CYP4B1* proximal element, Sp1 and Sp3 can compete for binding *in vitro* and both work to positively regulate promoter activity in the transient expression assays. However, Sp1 results in a more robust response.

Within the context of the proximal regulatory element alone, the mutagenesis studies suggest the two *CYP4B1* proximal Sp1/Sp3 sites act in a largely additive and independent fashion. Thus, Sp/XKLF site 5 enhanced promoter activity 1.4-fold (Fig. 8, pRNH844 versus 698) while Sp/XKLF site 2 enhanced activity 2.1-fold (Fig. 8, pRNH845 versus 698). Together, a 3.7-fold enhancement was observed (Fig. 8, pRNH846 versus 698), essentially equal to the a predicted additive response of 3.5-fold. In contrast, when examined in the context of the full-length promoter, an enhancement somewhat greater than the predicted additive response was observed, i.e. a 5.2-fold enhancement (Fig. 8, pRNH843 versus 686) versus a predicted 4.3-fold additive response (Fig. 8, pRNH842 versus 686 and pRNH838 versus 686). Although these data are not convincing alone, synergism also is apparent in experiments where Sp1 or Sp3 expression vectors were co-transfected with *CYP4B1*/luciferase reporter constructs. These data showed that both Sp1 and Sp3 were able to further enhance promoter activity in the context of the proximal promoter alone. In contrast, Sp3 had no impact when over-expressed in the presence of the full-length construct while co-transfection with the Sp1 expression vector resulted in a further 5-fold increase in promoter activity. Thus, Sp1's ability to form multimeric complexes involving multiple sites and synergistically activate transcription through protein-protein interactions has a minimal impact on the proximal promoter element alone, but does function in the context of both the proximal and lung-selective distal element.

Cooperativity between Sp1 and lung-selective, as well as other tissue-selective transcription factors has been documented (reviewed in Li et al. 2004). In this context, tissue-specific regulation involves the specific cellular environment, relative Sp1 and Sp3 expression levels, and protein modification. Thus, a significantly higher level of Sp1 was found in thymus, lung and spleen tissue than was found in other tissues, including the liver (Saffer et al. 1991). Variation in Sp1/Sp3 ratios impacted cell specific regulation of the kinase domain receptor in endothelial cells (Hata et al. 1998), the HPV-16 promoter in epithelial cells (Apt et al. 1996) and the COL2A1 promoter in chondrocytes (Chadjichristos et al. 2002). Protein modification of Sp1 and Sp3, including phosphorylation, glycosylation, sumoylation and acetylation may also contribute to the cell specific expression of a variety of proteins. Finally, a role for Sp1 and Sp3 in regulating the lung-selective expression of another cytochrome P450, *i.e.*, *CYP2F1*, has recently been documented (Wan et al. 2005). The importance of Sp1 in regulating *CYP4B1* lung-selective expression is apparent from the combined *in vitro* and *in vivo* studies presented in the current study and likely involves one or more of the above control mechanisms.

Tissue-selective *CYP4B1* expression also is controlled by the combined effect of the lung-selective distal regulatory element localized between position -1052 to -1008 and the absence of regulation through the liver-selective negative regulatory element localized between position -457 and -216. Although the importance of the negative regulatory element is apparent by the marked decrease in reporter activity observed in transient expression assays with HepG2 cells, results with the 293 kidney-derived cell line would suggest a greater importance of the combinatorial regulation observed between the proximal and distal positively-acting elements. A more complete assessment of the negative element's impact on *CYP4B1* lung specific expression will require a more exhaustive analysis of cell lines in which *CYP4B1* is not expressed. Regulation through the *CYP4B1* distal domain appears to involve transcription factor(s) binding to both *CYP4B1* position -1052 to -1042

and -1026 to -1008, but appears to require binding at the upstream site for subsequent binding to the downstream site. Although a search of the *CYP4B1* distal regulatory element with Match using the TRANSFAC database suggested potential roles for several transcription factors known to be involved in lung-selective regulation, competitive DNA binding, antibody supershift experiments and transient reporter assays with transcription factor expression vectors eliminated all of these potential candidates. Identification of this unknown lung-specific regulatory factor(s) will require further study.

In summary, Sp1 acting through a proximal enhancer element and in combination with a distal, lung-selective enhancer, is critical for *CYP4B1* lung-selective regulation. Also contributing to *CYP4B1*'s tissue-selective expression pattern is a liver-selective repressor motif. Although the activity of the human CYP4B1 protein is controversial (Zheng et al. 2003)(Imaoka et al. 2001), studies have demonstrated *CYP4B1* expression at the level of mRNA in lung tissue. Such lung-selective expression is distinctive and worthy of the mechanistic investigation reported herein. Thus, the current data contribute substantially to our understanding of the molecular mechanisms controlling cytochrome P450 lung-selective expression.

References

- Anderson KP, Kern CB, Crable SC, and Lingrel JB (1995) Isolation of a gene encoding a functional zinc finger protein homologous to erythroid Kruppel-like factor: Identification of a new multigene family. *Mol Cell Biol* **15**:5957-5965.
- Apt D, Watts RM, Suske G, and Bernard HU (1996) High Sp1/Sp3 ratios in epithelial cells during epithelial differentiation and cellular transformation correlate with the activation of the HPV-16 promoter. *Virology* **224**:281-291.
- Boucher PD, Ruch RJ, and Hines RN (1993) Specific nuclear protein binding to a negative regulatory element on the human *CYP1A1* gene. *J Biol Chem* **268**:17384-17391.
- Boulanger A, Liu S, Henningsgaard AA, Yu S, and Redmond TM (2000) The upstream region of the Rpe65 gene confers retinal pigment epithelium-specific expression *in vivo* and *in vitro* and contains critical octamer and E-box binding sites. *J Biol Chem* **275**:31274-31282.
- Chadjichristos C, Ghayor C, Herrouin JF, la-Kokko L, Suske G, Pujol JP, and Galera P (2002) Down-regulation of human type II collagen gene expression by transforming growth factor-beta 1 (TGF-beta 1) in articular chondrocytes involves SP3/SP1 ratio. *J Biol Chem* **277**:43903-43917.
- Chodosh LA (1988) Mobility shift DNA-binding assay using gel electrophoresis, in *Current Protocols in Molecular Biology* (Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, and Struhl K eds) pp 12.2.1-12.2.10, Greene Publishing and Wiley-Interscience, New York.

- Czerwinski M, McLemore TL, Gelboin HV, and Gonzalez FJ (1994) Quantification of CYP2B7, CYP4B1, and CYPOR messenger RNAs in normal human lung and lung tumors. *Cancer Res* **54**:1085-1091.
- Dahl AR and Lewis JL (1993) Respiratory tract uptake of inhalants and metabolism of xenobiotics. *Annu Rev Pharmacol Toxicol* **33**:383-407.
- Ding X and Kaminsky LS (2003) Human extrahepatic cytochromes P450: Function in xenobiotic metabolism and tissue-selective chemical toxicity in the respiratory and gastrointestinal tracts. *Annu Rev Pharmacol Toxicol* **43**:149-173.
- Guengerich FP (1993) The 1992 Bernard B. Brodie Award Lecture: Bioactivation and detoxication of toxic and carcinogenic chemicals. *Drug Metab Dispos* **21**:1-6.
- Hata Y, Duh E, Zhang K, Robinson GS, and Aiello LP (1998) Transcription factors Sp1 and Sp3 alter vascular endothelial growth factor receptor expression through a novel recognition sequence. *J Biol Chem* **273**:19294-19303.
- Hukkanen J, Vaisanen T, Lassila A, Piipari R, Anttila S, Pelkonen O, Raunio H, and Hakkola J (2003) Regulation of CYP3A5 by glucocorticoids and cigarette smoke in human lung-derived cells. *J Pharmacol Exp Ther* **304**:745-752.
- Imaoka S, Hayashi K, Hiroi T, Yabusaki Y, Kamataki T, and Funae Y (2001) A transgenic mouse expressing human CYP4B1 in the liver. *Biochem Biophys Res Commun* **284**:757-762.
- Imaoka S, Yoneda Y, Sugimoto T, Hiroi T, Yamamoto K, Nakatani T, and Funae Y (2000) CYP4B1 is a possible factor for bladder cancer in humans. *Biochem Biophys Res Commun* **277**:776-780.

- Kriwacki RW, Schultz SC, Steitz TA, and Caradonna JP (1992) Sequence-specific recognition of DNA by zinc-finger peptides derived from the transcription factor Sp1. *Proc Natl Acad Sci USA* **89**:9759-9763.
- Li L, He S, Sun JM, and Davie JR (2004) Gene regulation by Sp1 and Sp3. *Biochem Cell Biol* **82**:460-471.
- Li RC, Ping P, Zhang J, Wead WB, Cao X, Gao J, Zheng Y, Huang S, Han J, and Bolli R (2000) PKC ϵ modulates NF- κ B and AP-1 via mitogen-activated protein kinases in adult rabbit cardiomyocytes. *Am J Physiol Heart Circ Physiol* **279**:H1679-H1689.
- Mahoney CW, Shuman J, McKnight SL, Chen H-C, and Huang K-P (1992) Phosphorylation of CCAAT-enhancer binding protein by protein kinase C attenuates site-selective DNA binding. *J Biol Chem* **267**:19396-19403.
- Matys V, Fricke E, Geffers R, Gossling E, Haubrock M, Hehl R, Hornischer K, Karas D, Kel AE, Kel-Margoulis OV, Kloos DU, Land S, Lewicki-Potapov B, Michael H, Munch R, Reuter I, Rotert S, Saxel H, Scheer M, Thiele S, and Wingender E (2003) TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res* **31**:374-378.
- Nebert DW and Russell DW (2002) Clinical importance of the cytochromes P450. *Lancet* **360**:1155-1162.
- Overdier DG, Porcella A, and Costa RH (1994) The DNA-binding specificity of the hepatocyte nuclear factor 3/forkhead domain is influenced by amino-acid residues adjacent to the recognition helix. *Mol Cell Biol* **14**:2755-2766.
- Roesler WJ, Vandenbark GR, and Hanson RW (1988) Cyclic AMP and the induction of eukaryotic gene transcription. *J Biol Chem* **263**:9063-9066.

- Rao MK, Maiti S, Ananthaswamy HN, and Wilkinson MF (2002) A highly active homeobox gene promoter regulated by Ets and Sp1 family members in normal granulosa cells and diverse tumor cell types. *J Biol Chem* **277**:26036-26045.
- Rettie AE, Sheffels PR, Korzekwa KR, Gonzalez FJ, Philpot RM, and Baillie TA (1995) CYP4 isozyme specificity and the relationship between omega-hydroxylation and terminal desaturation of valproic acid. *Biochemistry* **34**:7889-7895.
- Rivera SP, Saarikoski ST, and Hankinson O (2002) Identification of a novel dioxin-inducible cytochrome P450. *Mol Pharmacol* **61**:255-259.
- Saffer JD, Jackson SP, and Annarella MB (1991) Developmental expression of Sp1 in the mouse. *Mol Cell Biol* **11**:2189-2199.
- Sanger F, Nicklen S, and Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* **74**:5463-5467.
- Shang Y, Hu X, DiRenzo J, Lazar MA, and Brown M (2000) Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription. *Cell* **103**:843-852.
- Smith PB, Tiano HF, Nesnow S, Boyd MR, Philpot RM, and Langenbach R (1995) 4-Ipomeanol and 2-aminoanthracene cytotoxicity in C3H/10T1/2 cells expressing rabbit cytochrome P450 4B1. *Biochem Pharmacol* **50**:1567-1575.
- Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, Fujimoto EK, Goeke NM, Olson BJ, and Klenk DC (1985) Measurement of protein using bicinchoninic acid. *Anal Biochem* **150**:76-85.
- Sowa Y, Orita T, Minamikawa-Hiranabe S, Mizuno T, Nomura H, and Sakai T (1999) Sp3, but not Sp1, mediates the transcriptional activation of the p21/WAF1/Cip1 gene promoter by histone deacetylase inhibitor. *Cancer Res* **59**:4266-4270.

- Su W, Jackson S, Tjian R, and Echols H (1991) DNA looping between sites for transcriptional activation: Self-association of DNA-bound Sp1. *Genes Dev* **5**:820-826.
- Suske G (1999) The Sp-family of transcription factors. *Gene* **238**:291-300.
- Vanderslice RR, Boyd JA, Eling TE, and Philpot RM (1985) The cytochrome P-450 monooxygenase system of rabbit bladder mucosa: enzyme components and isozyme 5-dependent metabolism of 2-aminofluorene. *Cancer Res* **45**:5851-5858.
- Wan J, Carr BA, Cutler NS, Lanza DL, Hines RN, and Yost GS (2005) Sp1 and Sp3 regulate basal transcription of the human *CYP2F1* gene. *Drug Metab Dispos*
doi:10.1124/dmd.105.004069.
- Xin J-H, Cowie A, Lachance P, and Hassell JA (1992) Molecular cloning and characterization of PEA3, a new member of the *Ets* oncogene family that is differentially expressed in mouse embryonic cells. *Genes Dev* **6**:481-496.
- Yokotani N, Sogawa K, Matsubara S, Gotoh O, Kusunose E, Kusunose M, and Fujii-Kuriyama Y (1990) cDNA cloning of cytochrome P-450 related to P-450_{p-2} from the cDNA library of human placenta--Gene structure and expression. *Eur J Biochem* **187**:23-29.
- Yu B, Datta PK, and Bagchi S (2003) Stability of the Sp3-DNA complex is promoter-specific: Sp3 efficiently competes with Sp1 for binding to promoters containing multiple Sp-sites. *Nucleic Acids Res* **31**:5368-5376.
- Zheng Y-M, Henne KR, Charmley P, Kim RB, McCarver DG, Cabacungan ET, Hines RN, and Rettie AE (2003) Genotyping and site-directed mutagenesis of a cytochrome P450 meander Pro-X-Arg motif critical to CYP4B1 catalysis. *Toxicol Appl Pharmacol* **186**:119-126.

Footnotes

Supported in part by U.S. Public Health Service Grant HL60143.

Figure Legends

Fig. 1. *Localization of CYP4B1 regulatory domains in A549 lung cells.* Transient expression in A549 lung carcinoma cells was performed using luciferase reporter constructs containing nested deletions of the *CYP4B1* 5'-upstream region. Salient features of each *CYP4B1*/luciferase construct are shown to the left of the bar graph with individual numbers referring to the specific plasmid designations (see Materials and Methods). The reported activities have been normalized for transfection efficiency and protein content and represent means \pm S.D. from triplicate experiments. (***, $p < 0.001$) indicates a significant difference in activity relative to the neighboring construct of shorter length.

Fig. 2. *Comparison of CYP4B1 regulatory domains in HepG2 liver and A549 lung cells.* Transient expression in HepG2 liver carcinoma (A) and A549 lung carcinoma (B) cells was performed using luciferase reporter constructs containing nested deletions of the *CYP4B1* 5'-upstream region. Salient features of each *CYP4B1*/luciferase construct are shown to the left of the bar graph with individual numbers referring to the specific plasmid designations (see Materials and Methods). The reported activities have been normalized for transfection efficiency and protein content and represent means \pm S.D. from triplicate experiments. (**, $p < 0.01$) and (***, $p < 0.001$) indicate a significant difference in activity relative to the neighboring construct of shorter length.

Fig. 3. *Finer localization of the CYP4B1 distal regulatory domain in A549 cells.* Transient expression in A549 lung carcinoma cells was performed using luciferase reporter constructs containing nested deletions of the *CYP4B1* 5'-upstream region from position -1087 to -457 to better localize the distal regulatory domain. Salient features of each *CYP4B1*/luciferase construct are

shown to the left of the bar graph with individual numbers referring to the specific plasmid designations (see Materials and Methods). The reported activities have been normalized for transfection efficiency and protein content and represent means \pm S.D. from triplicate experiments. (***, $p < 0.001$) indicates a significant difference in activity relative to the neighboring construct of shorter length.

Fig. 4. *Identification of specific DNA/protein interactions with the CYP4B1 position -139 to -45 proximal regulatory domain.* Using A549 nuclear extract, EMSA was performed as described in Materials and Methods. Specific DNA/protein complexes (A, B and C in order of increasing electrophoretic mobility) are indicated by the solid arrows to the left of the autoradiogram. Specific DNA/protein complexes supershifted by the inclusion of either Sp1 or Sp3 antibody are depicted by the open arrows to the right of the autoradiogram.

Fig. 5. *Identification of putative Sp/XKLF binding sites within the CYP4B1 position -139 to -45 proximal regulatory domain.* A computer-assisted search of *CYP4B1* sequences from position -139 to -45 using the Match program and the TRANSFAC V8.3 database revealed six putative binding sites for members of the Sp/XKLF transcription factor family. The matches to the core Sp/XKLF binding sites are shown by the shaded boxes. Bases matching the consensus matrix are in bold typeface. Arrows indicate the orientation of the putative Sp/XKLF sites.

Fig. 6. *Sp1 and Sp3 positively regulate the CYP4B1 promoter.* Co-transfection, transient expression studies were performed in A549 cells with *CYP4B1*/luciferase constructs containing: (A) only the proximal regulatory domain (pRNH698), or (B) both the proximal and distal regulatory domains (pRNH686), along with Sp1 or Sp3 expression vectors. Salient features of each *CYP4B1*/luciferase

construct are shown to the left of the bar graph. Individual numbers refer to the specific reporter construct used for the transfection (see Materials and Methods). The reported activities have been normalized for transfection efficiency and protein content and represent means \pm S.D. from triplicate experiments. (***, $p < 0.001$) indicate a significant difference in activity relative to the *CYP4B1*/luciferase construct in the absence of the Sp1 or Sp3 expression vector.

Fig. 7. *ChIP assay of Sp1 and Sp3 binding to the CYP4B1 proximal regulatory region in lung and liver tissues and specific cell lines.* ChIP assays was performed as described in Materials and Methods using: (A) tissue from lung or liver autopsy samples, and (B) A549, HepG2, or BEAS-2B cell lines. Immunoprecipitation was done with either Sp1 or Sp3 antibodies and the isolated chromatin examined for enrichment of *CYP4B1* sequences from position -241 to +20 by PCR. Amplified DNA was fractionated by agarose gel electrophoresis and visualized by ethidium bromide staining. The mobility of included molecular size standards is indicated to the left of each photo.

Fig. 8. *Putative Sp/XKLF sites 2 and 5 within the CYP4B1 position -139 to -45 proximal regulatory domain are capable of specific protein binding.* The ability of each of the putative Sp/XKLF sites within the *CYP4B1* proximal positive regulatory domain to specifically bind A549 nuclear protein was examined by competitive EMSA with double-stranded oligonucleotides representing each site. EMSA was performed as described in Materials and Methods. Specific DNA/protein complexes (A, B and C in order of increasing electrophoretic mobility) are indicated by the solid arrows to the left of the autoradiogram. Competition was performed with a 50- to 300-fold molar excess of overlapping, double-stranded oligonucleotides representing *CYP4B1* positions -138 to -46 (lanes 3-4), -142 to -119 (lane 5), -112 to -76 (lane 6), -72 to -48 (lanes 7), -128 to -105 (lanes 8-10), and

-87 to -59 (lanes 11-13). In addition, competition with an oligonucleotide having the Sp/XKLF sites mutagenized at *CYP4B1* positions -118 to -114 (site 2) and -77 to -73 (site 5) is shown (lanes 14-15).

Fig. 9. *Functional analysis of Sp/XKLF sites 2 and 5 within the CYP4B1 position -139 to -45 proximal positive regulatory domain.* The functional consequences of inactivating Sp/XKLF site 2, site 5, or both sites 2 and 5 within the *CYP4B1* proximal positive regulatory domain was examined by transient expression analysis in A549 cells. Salient features of each *CYP4B1*/luciferase construct are shown to the left of the bar graph with individual numbers referring to the specific plasmid designations (see Materials and Methods). The putative Sp/XKLF sites are depicted by boxes. Mutagenesis of either site 2 and/or site 5 is depicted by open boxes whereas closed boxes indicate intact sites. The reported activities have been normalized for transfection efficiency and protein content and represent means \pm S.D. from triplicate experiments. (*, $p < 0.05$; **, $p < 0.01$; and ***, $p < 0.001$) indicate a significant difference in activity compared to the non-mutagenized *CYP4B1*/luciferase construct, (†††, $p < 0.001$) indicates a significant difference in activity compared to the non-mutagenized construct, but no significant difference relative to pGL3Basic, and (‡‡‡, $p < 0.001$) indicates a significant difference in activity compared to the non-mutagenized construct and relative to pGL3Basic.

Fig. 10. *Evaluation of transcription factors implicated in binding the CYP4B1, lung-selective distal regulatory element.* Using A549 nuclear extract along with *CYP4B1* sequences from position -1052 to -1008 as a probe, EMSA was performed as described in Materials and Methods. Specific DNA/protein complexes (A, B and C in order of increasing electrophoretic mobility) are indicated by solid arrows to the left of the autoradiogram. Competition was performed with either a 10- or 100-fold molar excess of unlabeled probe (lanes 3 and 4, respectively), with a 100-fold molar excess

DMD #4523

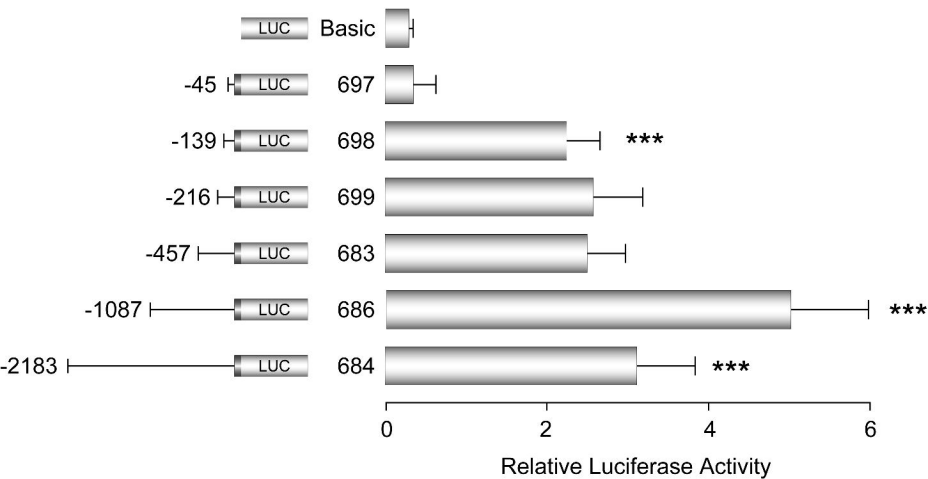
of double-stranded oligonucleotides representing *CYP4B1* position -1052 to -1026 (lane 5) and *CYP4B1* position -1042 to -1008 (lane 6), and a 150-fold molar excess of consensus sequences for AP4 (lane 7), C/EBP (lane 8), Elk1 (lane 9), or HNF3 (lane 10).

TABLE 1

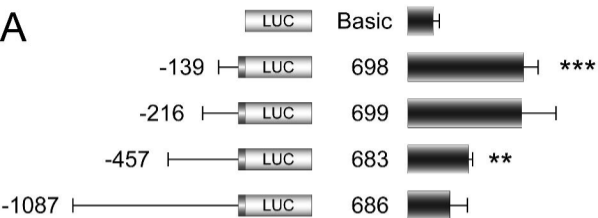
Oligonucleotides Representing Consensus and Mutagenized Consensus Transcription Factor Elements Used for Competitive EMSA

| Sequence (sense strand) ^a | Transcription Factor | Reference |
|--|----------------------|---|
| 5' GATGACTGAGGTCAG CTCAGG ACTGCATGGC 3' | AP4 | (Boulanger et al. 2000) |
| 5' CGATC GGGGCGGGG CGAGC 3' | Sp1 (Sp/XKLF) | (Kriwacki et al. 1992) |
| 5' CGATCCTT GAGGAAGT AATAAG 3' | Elk1 | (Li et al. 2000; Xin et al. 1992) |
| 5' TGCAGATT GCGCAAT CTGCA 3' | C/EBP | (Mahoney et al. 1992) |
| 5'GTTGACTA AGTCAATAAT CAGAAT 3' | HNF3 | (Overdier et al. 1994; Roesler et al. 1988) |

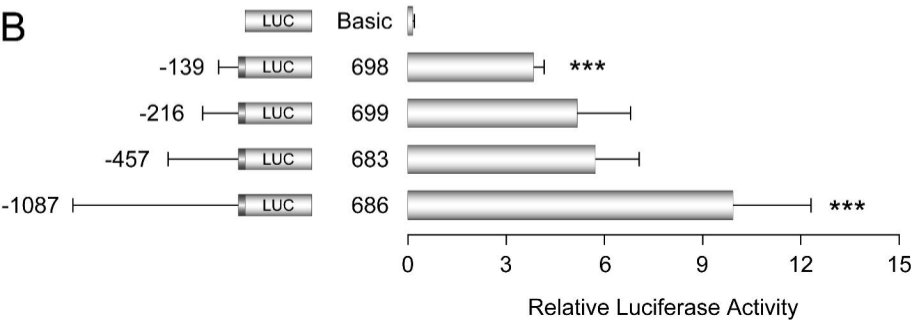
^a Core binding elements are shown in bold type.

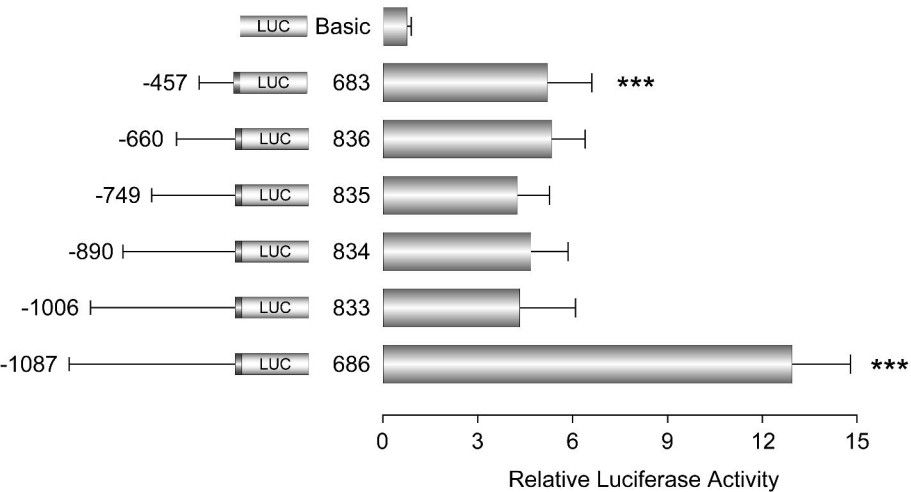


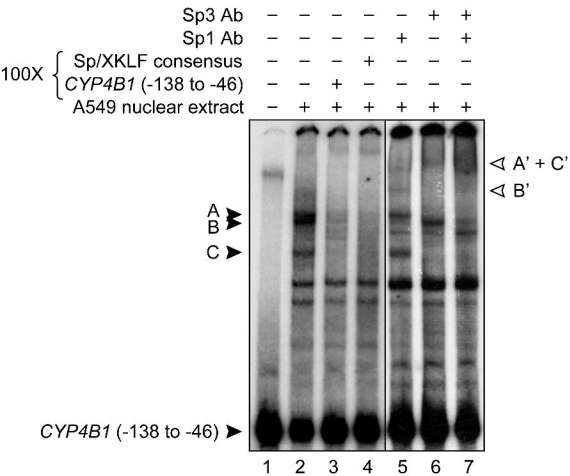
A

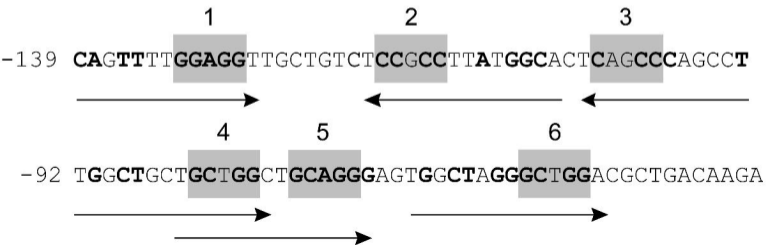


B

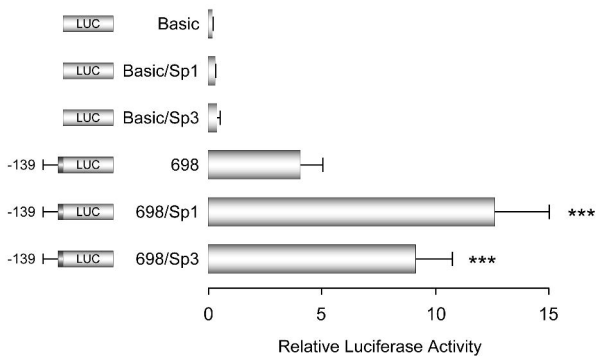




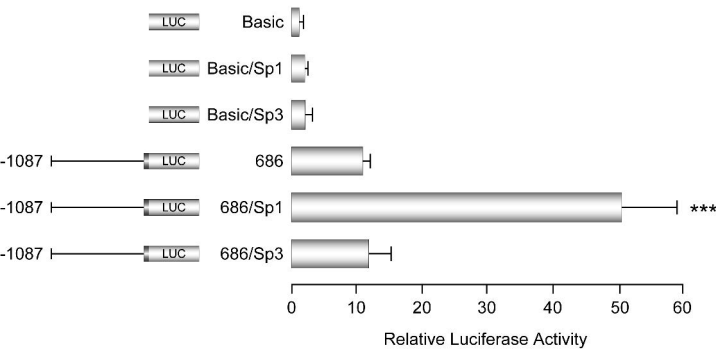




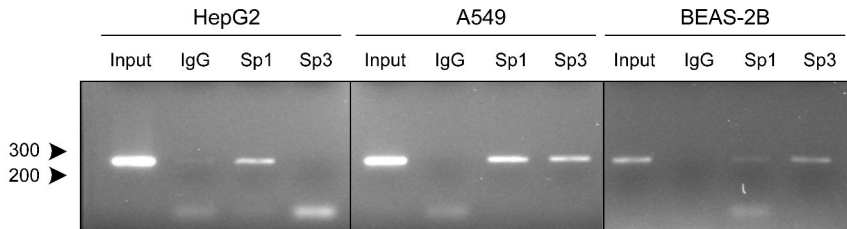
A



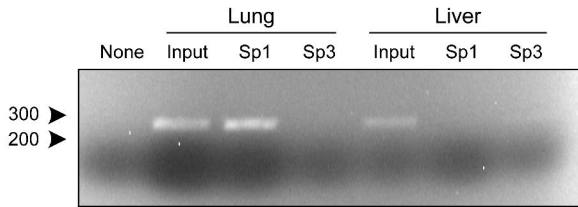
B



A



B



| | | | | | | | | | | | | | | | |
|------------|--------------------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 50 to 150X | CYP4B1 -138 to -46 S2/5m | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 50 to 300X | CYP4B1 -87 to -59 S5 | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| | CYP4B1 -128 to -105 S2 | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 300X | CYP4B1 -72 to -48 S6 | - | - | - | - | - | - | + | - | - | - | - | - | - | - |
| | CYP4B1 -112 to -76 S3/4 | - | - | - | - | - | - | + | - | - | - | - | - | - | - |
| | CYP4B1 -142 to -119 S1 | - | - | - | - | - | + | - | - | - | - | - | - | - | - |
| 50 to 100X | CYP4B1 -138 to -46 | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| | A549 nuclear extract | - | + | + | + | + | + | + | + | + | + | + | + | + | + |

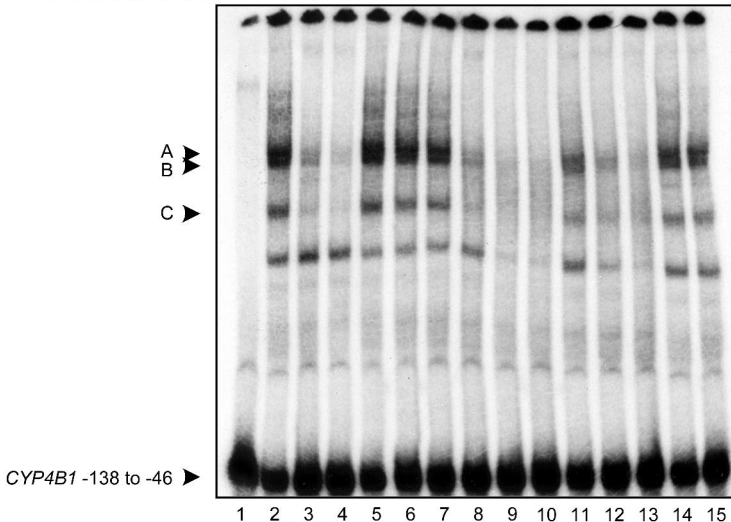
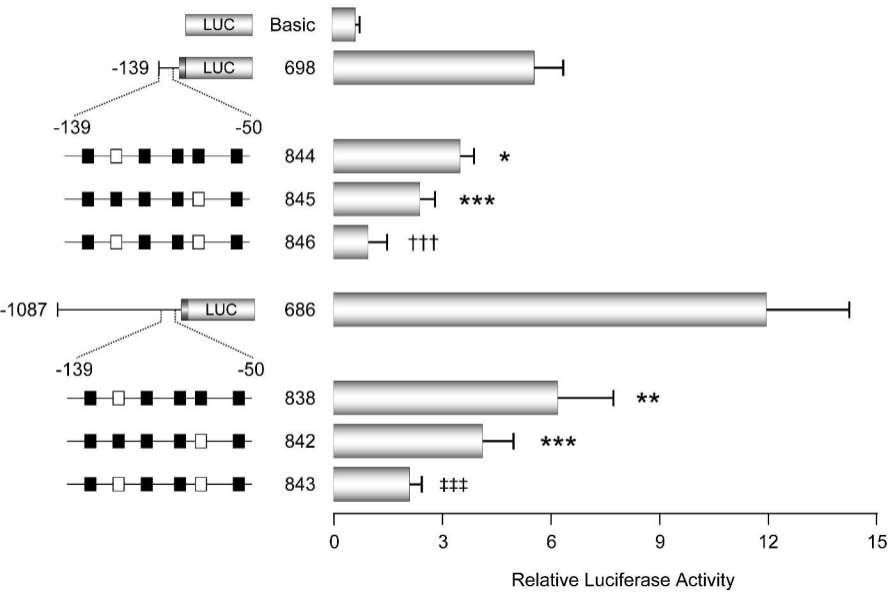



Fig. 8



| | | | | | | | | | | | | |
|------|--------------------------------|---|---|---|---|---|---|---|---|---|---|---|
| | HNF3 consensus | - | - | - | - | - | - | - | - | - | + | |
| 150X | Elk1 consensus | - | - | - | - | - | - | - | - | - | + | - |
| | C/EBP consensus | - | - | - | - | - | - | - | - | + | - | - |
| | AP4 consensus | - | - | - | - | - | - | - | + | - | - | - |
| 100X | <i>CYP4B1</i> (-1042 to -1008) | - | - | - | - | - | + | - | - | - | - | - |
| | <i>CYP4B1</i> (-1052 to -1026) | - | - | - | - | + | - | - | - | - | - | - |
| | <i>CYP4B1</i> (-1052 to -1008) | - | - |  | | | - | - | - | - | - | - |
| | A549 nuclear extract | - | + | + | + | + | + | + | + | + | + | + |

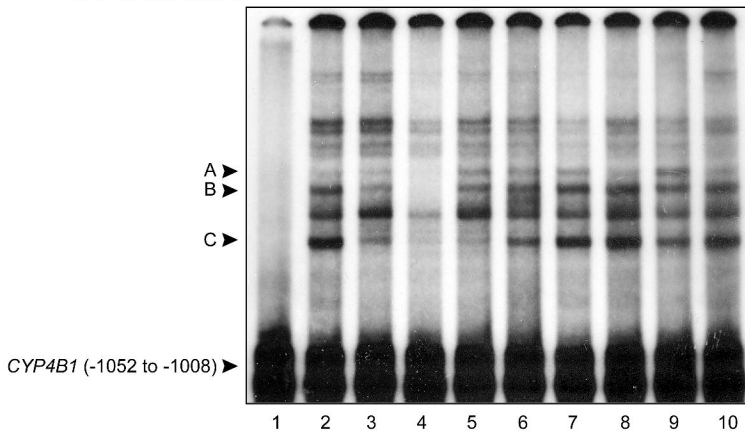


Fig. 10