Expression of Constitutive Androstane Receptor Splice Variants in Human Tissues and Their Functional Consequences

Jatinder K. Lamba, Vishal Lamba, Kazuto Yasuda, Yvonne S. Lin, Mahfoud Assem, Emma Thompson, Stephen Strom, and Erin Schuetz

Department of Pharmaceutical Sciences, St. Jude Children's Research Hospital, Memphis, Tennessee (J.K.L., V.L., K.Y., Y.S.L., M.A., E.S.); Committee on Genetics, University of Chicago, Chicago, Illinois (E.T.); and Department of Pathology, University of Pittsburgh, Pittsburgh, Pennsylvania (S.S.)

Received March 31, 2004; accepted June 9, 2004

ABSTRACT

The constitutive androstane receptor (CAR) NR1I3 is a transcription factor that upon activation by xenobiotics induces transcription of drug-metabolizing and drug transporter genes. Our goal was to identify whether alternative splicing of CAR makes an important contribution to the generation of novel CAR proteins. The wild-type CAR mRNA (CAR.1) and splice variants (SVs) were amplified from human liver cDNAs and in a panel of cDNAs from 36 human tissues, using exon 1- and 3′-untranslated region primers, cloned and sequenced. Twenty-two unique hCAR splice variants (CAR-SVs) containing different combinations of splicing (deletion of exons 2, 4, 5, 7, partial deletion of exon 9, or insertion of 12 or 15 base pairs from introns 6 or 7) were identified. CAR mRNAs were expressed in small intestine, kidney, testis, adrenal, and brain caudate nucleus. Intestine expressed only CAR.1 mRNA, whereas spleen, heart, and prostate expressed only CAR-SVs. In vitro transcription and translation of CAR-SVs lacking exon 2 (missing ATG start site) generated CAR proteins that differed in M, from CAR.1. These CAR-SVs used a translation initiation site in exon 1, generating CAR with a unique amino-terminal sequence. Transcripts lacking part of exon 9 altered the CAR reading frame generating CAR proteins with a unique carboxy-terminal end. CAR-SVs demonstrated compromised binding to CYP2B6 NR elements and transcriptional activation of a CYP2B6 luciferase reporter. The expression of CAR in additional human cell types increases the range of tissue specific transcriptional responses regulated by this receptor, suggesting additional biological roles for CAR and CAR-SV proteins in these tissues.

The constitutive androstane receptor (CAR) NR1I3 is essential for regulating both constitutive expression and induction of some drug detoxification genes such as CYP2B in liver (Honkakoski et al., 1998; Wei et al., 2000). CAR is also highly conserved throughout evolution. CAR is activated either directly by ligands (TCPOBOP) or indirectly by some chemicals (phenobarbital) to a form that transcriptionally regulates
gene expression (Yamamoto et al., 2003). Ultimately, CAR activators facilitate translocation of CAR from the cytosol to the nucleus and specific binding to CAR response elements in the promoters of a growing number of drug detoxification genes including CYP2B, CYP3A4, and UGT1A1 (Yamamoto et al., 2003). CAR is also required for liver hypertrophy caused by treatment with drugs such as phenobarbital (Wei et al., 2000). It is possible that CAR is also responsible for phenobarbital’s ability to act as a liver tumor promoter (for review, see Schuetz, 2001). Activation of CAR by drugs like phenobarbital may be a double-edged sword, enhancing the metabolism and removal of xenobiotics and controlling liver tumor promotion by phenobarbital. Thus, it is important to understand factors regulating CAR expression.

A common paradigm for many genes, including nuclear hormone receptors, is the generation of additional spliced mRNA transcripts. It has been estimated that at least 59% of all human genes utilize alternative RNA processing to gen-

ABBREVIATIONS: CAR, constitutive androstane receptor; CITCO, 6-(4-chlorophenyl)imidazo[2,1-b][1,3]thiazole-5-carbaldehyde O-(3,4-dichlorobenzyloxime); bp, base pair(s); SV, splice variant; PCR, polymerase chain reaction; TNT, transcription and translation; EMSA, electrophoretic mobility shift assay; DBD, DNA binding domain.
erate multiple mRNA products that can have different exon compositions (Venter et al., 2001). The inclusion or exclusion of exon sequences enhances generation of protein isoforms that can vary in structure and functional properties. Nuclear hormone receptor alternative mRNAs arise by a combination of alternative promoter usage and exon splicing (Keightley, 1998; Ding and Miller, 2002). These nuclear hormone receptor splice variants can differ in their patterns of expression, gene targets, biological functions (gain- and loss-of-function), DNA binding, intracellular binding interactions with other proteins, and subcellular localization. Many receptors, both steroid and nuclear hormone, have at least one major splice variant form. This molecular diversity may allow for additional layers of signaling complexity from the nuclear hormone receptors.

Four alternatively spliced CAR splice variants have been reported in human liver (Auerbach et al., 2003; Savkur et al., 2003): an isoform with a 12-bp (4 amino acids) insertion from intron 6, an isoform with a 15-bp (5 amino acids) insertion from intron 7, an isoform with both the insertions, and an isoform with exon 7 (39 amino acids) deleted. All of these SVs alter the ligand binding domain of CAR. Compared with CAR.1, these CAR-SVs had lower affinity for CAR binding elements, decreased transactivation potential, and decreased ability to recruit coactivators (Auerbach et al., 2003; Savkur et al., 2003).

Northern blot analysis has shown that CAR is highly expressed in human and rodent liver and intestine with lower expression in heart, muscle, kidneys, and lung. The major CAR-SV identified by both groups contributed 6 to 8% of total expression in heart, muscle, kidneys, and lung. The major transcripts of CAR and would not distinguish whether additional CAR-SVs existed in these tissues. Thus, the objectives of this study were to identify all of the CAR mRNA transcripts in human tissues and to determine their functional consequences. The current report identifies a large number of CAR alternatively spliced mRNAs that are widely distributed in many human organs. The additional CAR isoforms and wider tissue distribution of CAR may increase the spectrum of gene targets regulated by CAR.

### Materials and Methods

**RNA from Liver Samples and Tissue Samples.** Total RNA was prepared from liver samples (University of Pittsburgh/St. Jude Liver Resource, Pittsburgh, PA) from three different ethnic groups (African American, n = 30; Caucasians, n = 30; and Hispanics, n = 25). The majority of the livers used in this study were from organ donors, and their histories have been previously published (Lamba et al., 2003). Rapid-Scan Human Multi-tissue and Human Brain Gene Expression Panels were purchased from OriGene Technologies, Inc. (Rockville, MD). OriGene isolated polyA+ RNA, synthesized first strand cDNA, and the cDNAs were normalized such that they all contained equivalent concentrations of β-actin cDNA.

**RT-PCR and Amplification of CAR cDNAs.** Primers for amplification of full-length CAR (Table 1) were designed with PRIMER3 (http://www.genome.wi.mit.edu/cgi-bin/primer/primer3.cgi), and their sequence homology and specificity were checked using BLASTn (http://

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence (5′-3′)</th>
<th>Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAR-F1</td>
<td>GGAGAGCGATTCGATACCAG</td>
<td>94°C × 30 s, 58°C × 30 s, 72°C × 90 s, 32 cycles</td>
</tr>
<tr>
<td>CAR-R1</td>
<td>TTCCACCACCAGTGTATCCAG</td>
<td>94°C × 30 s, 57.3°C × 30 s, 72°C × 90 s, 15 cycles</td>
</tr>
<tr>
<td>CAR-F2</td>
<td>CAGGTTGAACTGCTGCCTAAG</td>
<td>94°C × 30 s, 57°C × 30 s, 72°C × 30 s, 30 cycles</td>
</tr>
<tr>
<td>CAR-R2</td>
<td>TATCCAGGGGTTCTCAGGTG</td>
<td>94°C × 30 s, 57°C × 30 s, 72°C × 30 s, 30 cycles</td>
</tr>
<tr>
<td>CAR-Taq-F</td>
<td>CCAGCTCATCAGTCATACCA</td>
<td>94°C × 30 s, 57°C × 30 s, 72°C × 30 s, 30 cycles</td>
</tr>
<tr>
<td>CAR-Taq-R</td>
<td>GGTGACTCTCTCGAACATGAG</td>
<td>94°C × 30 s, 57°C × 30 s, 72°C × 30 s, 30 cycles</td>
</tr>
<tr>
<td>CYP2B6-F</td>
<td>CACTCATCAGCCTTGATTCG</td>
<td>94°C × 30 s, 57°C × 30 s, 72°C × 30 s, 30 cycles</td>
</tr>
<tr>
<td>CYP2B6-R</td>
<td>GTGACTCTCTCGAACATGAG</td>
<td>94°C × 30 s, 57°C × 30 s, 72°C × 30 s, 30 cycles</td>
</tr>
<tr>
<td>MRP4-F</td>
<td>GCAAGATGCTGCCGTGATAC</td>
<td>94°C × 30 s, 57°C × 30 s, 72°C × 30 s, 30 cycles</td>
</tr>
<tr>
<td>MRP4-R</td>
<td>ACCCTGTAACACTCTCTCCAG</td>
<td>94°C × 30 s, 57°C × 30 s, 72°C × 30 s, 30 cycles</td>
</tr>
<tr>
<td>MRP4-NR1</td>
<td>GATCCCTGACTTTCGTCGATCC</td>
<td>94°C × 30 s, 57°C × 30 s, 72°C × 30 s, 30 cycles</td>
</tr>
<tr>
<td>CYP2B6-NR2</td>
<td>GATCCATGGACTTTTCTCGAACCGAC</td>
<td>94°C × 30 s, 57°C × 30 s, 72°C × 30 s, 30 cycles</td>
</tr>
</tbody>
</table>

### Table 1

Sequence and conditions of oligos used for PCR amplification and EMSA

- **Oligos for amplifying the cDNA**
  - CAR-F1: GGAGAGCGATTCGATACCAG
  - CAR-R1: TTCCACCACCAGTGTATCCAG
  - CAR-F2: CAGGTTGAACTGCTGCCTAAG
  - CAR-R2: TATCCAGGGGTTCTCAGGTG
  - CAR-Taq-F: CCAGCTCATCAGTCATACCA
  - CAR-Taq-R: GGTGACTCTCTCGAACATGAG
  - CYP2B6-F: CACTCATCAGCCTTGATTCG
  - CYP2B6-R: GTGACTCTCTCGAACATGAG
  - MRP4-F: GCAAGATGCTGCCGTGATAC
  - MRP4-R: ACCCTGTAACACTCTCTCCAG

- **Oligos for EMSA**
  - CYP2B6-NR1: GATCCCTGACTTTCGTCGATCC
  - CYP2B6-NR2: GATCCATGGACTTTTCTCGAACCGAC

- **Oligos used for creating ATG mutants**
  - CAR-ATG-exon1–220-F: GGGAGGCGATTCGATTTGCTCTGGAAGAACG
  - CAR-ATG-exon220-R: GCTTCTTATGGACAGAAAGTCACCCTGGACGCCAC
  - CAR-ATG-exon4–645-F: CCCCACGGGGCGCCATGGGACACATGTTT
  - CAR-ATG-exon4–645-R: AAAATGGTGGCAAAATTGGGCTTGGTGCC
  - CAR-ATG-exon-654-F: GCCCAATGCCAACCCTTGGTGAAGCTG
  - CAR-ATG-exon-654-R: CACAACTTGTCTAACCAGGCTGACCC

- **Primers used to amplify the cDNA from hepatocytes**
  - CYP3A4-F: CCAAGGCTATCCGTCCTACGCCG
  - CYP3A4-R: TCGGGCTTATCCGTCCTACGCCG
  - CYP2B6-F: TCCGTGTAGTGGCCTGTCCAGA
  - CYP2B6-R: TCCGGATCTTACCTGAGTAATG
  - PXR-F: CGAACGAAAGAAATGGCAAGC
  - PXR-R: CCGGTGTAGTGGCCTGTCCAGA
  - GAPDH-F: ACCAGAGTTCGTCGATCC
  - GAPDH-R: TCCACACCCGTCGTGCTGTA

* Bolded nucleotide indicates the base mutated to change ATG to TTG.
www.ncbi.nlm.nih.gov). To amplify CAR from human livers, first strand cDNA was synthesized from 3 μg of total RNA from human liver samples according to the manufacturer’s instructions (Superscript Pre-amplification System for First-Strand cDNA Synthesis, Invitrogen, Carlsbad, CA). Using CAR-F1 and CAR-R1 primers (Table 1), the CAR cDNA was amplified in a total volume of 50 μl consisting of 1 μl of human liver cDNA, 10× PCR buffer, 1 mM MgCl₂, 10 μl of each primer, 0.2 mM dNTP (Invitrogen), and 2.5 U of Taq DNA polymerase (Expand High-Fidelity PCR system, Roche Diagnostics, Indianapolis, IN) using conditions in Table 1. Using the same primers, CAR was also amplified in the Rapid-Scan Human Multi-tissue and Human Brain Gene Expression Panels as per the manufacturer’s instructions. The CAR cDNAs were analyzed by agarose gel electrophoresis. Because the amount of amplified cDNA in some samples was low, a second round of nested amplification was performed using the CAR-F2/R2 primers and appropriate dilutions of the first round PCR products. The conditions used were the same as the first round amplification with minor modifications (Table 1) and inclusion of 1.5 mM MgCl₂. CYP2B6 and MRP4 were amplified in the Rapid-Scan Human Multi-tissue and Human Brain Gene Expression Panels as per the manufacturer’s instructions. The amplified CAR cDNAs were analyzed by agarose gel electrophoresis (Table 1) and sequencing. Because the amount of amplified cDNA in some samples was low, a second round of nested amplification was performed using the CAR-F2/R2 primers and appropriate dilutions of the first round PCR products. The conditions used were the same as the first round amplification with minor modifications (Table 1) and inclusion of 1.5 mM MgCl₂. CYP2B6 and MRP4 were amplified in the Rapid-Scan Human Multi-tissue and Human Brain Gene Expression Panels as per the manufacturer’s instructions. The amplified CAR cDNAs were analyzed by agarose gel electrophoresis. Because the amount of amplified cDNA in some samples was low, a second round of nested amplification was performed using the CAR-F2/R2 primers and appropriate dilutions of the first round PCR products. The conditions used were the same as the first round amplification with minor modifications (Table 1) and inclusion of 1.5 mM MgCl₂. CYP2B6 and MRP4 were amplified in the Rapid-Scan Human Multi-tissue and Human Brain Gene Expression Panels as per the manufacturer’s instructions. The amplified CAR cDNAs were analyzed by agarose gel electrophoresis.

Characterization and Cloning of hCAR Variant Isoforms. The amplified CAR cDNAs were analyzed by agarose (1%) gel electrophoresis. If a single band was observed, the amplification product was incubated with shrimp alkaline phosphatase and exonuclease I (U.S. Biochemical Corp., Cleveland, OH) for 30 min at 37°C to remove unincorporated nucleotides and primers; the enzymes were inactivated by incubation at 80°C for 15 min before sequencing. If multiple bands were observed in the agarose gel, each band was cut from the gel using the Zymoclean Gel DNA recovery kit (Zymo Research, Orange, CA), eluted in water, and sequenced by using the PCR primers. Alternately, the extracted cDNA was subcloned into pCR2.1 using the TOPO TA cloning system (Invitrogen) according to manufacturer’s instructions. Ten clones were selected at random for each band that was cloned, and after lysing the cells by repeated freezing and thawing, PCR amplification was performed using M13 reverse and M13 (–21) forward primers and the product sequenced using M13 primers. Sequencing was carried out on an ABI Prism 3700 Automated Sequencer (Applied Biosystems, Foster City, CA). GenBank accession numbers for the 22 CAR-SV nucleotide sequences are sequentially AF572806 through AF572827.

Human CAR Expression Plasmids. Once the variants were characterized, the hCAR isoforms were digested from pCR2.1 by EcoRI and cloned into pCDNA3.1 (Invitrogen) and pMSCV-IRESGFP. Mutagenesis of the [A/T]TGs in either exon 1 or 4 in the hCAR expression plasmid were achieved by site-directed mutagenesis using a mutated primer (Table 1) and the hCAR expression plasmid to assess putative ATG start sites.

In Vivo Transcription and Translation. To check if the variant isoforms of CAR encode a protein, in vitro-transcribed and translation (TNT) was carried out using the rabbit reticulocyte lysate system (Promega, Madison, WI). One microgram of each pcDNA3.1-hCAR isoform was used for the TNT. Protein expression in rabbit reticulocyte lysate system was carried out in the presence of [35S]methionine as described by the manufacturer. The protein products were resolved within 24 h of cross-clamp. Reasons for not using tissues for transplantation included traumatic damage, errors in organ harvest, brief anoxic periods, or macro- or microsteatosis. Human hepatocytes were isolated essentially as described (Strom et al., 1996, 1998). Cells were plated on collagen-coated six-well plates or 60-mm culture dishes, maintained in Modified Williams E for 48 h, and then treated with drugs for 48 h. Media was then aspirated, Trizol solution (Invitrogen) was added to lyse the cells, scraped from the plates, RNA isolated, first strand cDNAs synthesized as described above, and targets amplified using primers in Table 1.

Results

Identification of Novel hCAR Splice Variants. A recent report identified a number of CAR splice variants in the livers of Caucasian persons (Auerbach et al., 2003) using primers immediately adjacent to the translation start and stop codons in CAR. To determine whether the same pattern of CAR splice variants was observed in persons from other ethnic populations and to identify any longer alternative CAR transcripts, we designed primers residing in exon 1 and in the 3'-untranslated region of CAR. CAR was amplified from the livers of Caucasians, African Americans, and Hispanics (Fig. 1A), and the PCR products were analyzed by agarose gel electrophoresis (Fig. 1B). A CAR cDNA of the expected size was amplified in the livers of all ethnic groups along with several additional CAR transcripts. Multiple CAR splice variants were present in livers representing males and females of each ethnic group (Fig. 1B).

Using subcloning and sequencing, we characterized the additional CAR mRNA transcripts. Sixteen unique variants of hCAR were identified in the liver samples (Fig. 2). The common splicing events were complete deletion of exon 2 (140 bp), exon 4 (170 bp), exon 7 (117 bp), and partial deletion of exon 9 (154 bp). Two previously identified insertions were also observed (insertion of 12 bp from intron 6 and insertion of 15 bp from intron 7) (Auerbach et al., 2003; Savkur et al., 2003). Each of the splicing events (four deletions and two insertions) were observed (insertion of 12 bp from intron 6 and insertion of 15 bp from intron 7) (Auerbach et al., 2003; Savkur et al., 2003). Each of the splicing events (four deletions and two insertions) were observed.
insertions) occurred in multiple combinations generating 16 isoforms of hCAR in human livers.

Several CAR splice variants were unlikely to encode functional proteins because they had premature termination codons and would be more likely to be rapidly degraded by nonsense-mediated decay: SVs with an exon 4 deletion (SV4, SV9, and SV10) have a disrupted reading frame and introduce a premature termination codon. CAR-SV2 (derived from the use of a cryptic splice site in exon 2) had a deletion of 67 bp from exon 2 that introduced a premature termination codon in exon 3. SV15 had part of exon 4 deleted and also introduced a premature termination codon. In addition, SV16 had deletion of exons 3 through 9.

The other alternative mRNAs could encode unique CAR proteins. Deletion of exon 7 (Auerbach et al., 2003; Savkur et al., 2003) results in an in-frame loss of the 39 amino acids in the CAR ligand binding domain. Exon 9 is 258 bp in length and encodes 42 amino acids (the TGA stop codon is at nucleotide 128 in exon 9). Some SVs skipped the first 154 nucleotides of exon 9, resulting in the loss of the carboxy-terminal 42 amino acids. However, these same SVs utilize a cryptic AG splice site in exon 9, causing a frame shift and encoding 29 unique carboxy-terminal amino acids. Although a splice variant of mouse CAR that skips exon 8 has been previously reported (Choi et al., 1997) (Fig. 2), an orthologous hCAR-SV was not identified.

Expression Profiling of hCAR Isoforms in Different Tissues. Because tissue-specific splicing can lead to generation of unique alternatively spliced mRNAs, we screened multiple tissues for the presence of CAR and alternative CAR mRNAs. Following the first round of amplification, CAR.1 and alternative CAR mRNA transcripts were readily detected in adult and fetal liver, kidney, adrenal, small intestine, and testis (Fig. 3A). Lower CAR expression was also seen in brain. In contrast to liver where multiple mRNA transcripts were always amplified, heart, spleen, small intestine, and prostate expressed a single CAR mRNA even after a second round of amplification using nested primers (Fig. 3A). Sequencing of the CAR mRNAs revealed a novel CAR-SV21 in heart with deletion of both exons 4 and 7; however, these deletions introduce a premature termination codon. CAR-SV3 present in spleen skips exons 2 and 7, generating a unique CAR mRNA. The single unique CAR transcript present in prostate, CAR-SV22, was characterized by deletion of exon 2. Only CAR.1 was observed in small intestine.

We next screened for CAR mRNAs in regions of the human brain (Fig. 3B). CAR transcripts were only amplified in caudate nucleus and were distinct from the transcripts identified in other tissues. The caudate nucleus-specific isoforms, CAR-SV17 to SV20 (Fig. 2), all share deletion of exons 4, 5, and 7 with or without exon 2/exon 9 deletions. None of the caudate nucleus mRNAs would be expected to encode a functional protein as they have a premature termination codon. Although we detected CAR.1 in the total brain cDNAs (Fig. 3A), CAR.1 was not amplified from the brain regions present in our analysis, including caudate nucleus, and must be present in other regions of the brain not represented in this panel of cDNAs.
performed in different regions of the brain. The top panel shows first round amplicons, and the bottom panel shows second round nested amplicons. Amplification of two target genes of CAR: CYP2B6 (third panel) and MRP4 (fourth panel) was also identified as in A. The top panel shows first round amplicons, and the bottom panel shows second round nested amplicons. Amplification of two target genes of CAR: CYP2B6 (third panel) and MRP4 (fourth panel) was also performed in different regions of the brain.

Because some hCAR targets, CYP2B6 and MRP4, are present in human brain (Tyndale and Sellers, 2001; Assem et al., 2004), we screened these same samples for coexpression of these mRNAs. Although CYP2B6 was expressed in substantial amounts in caudate nucleus, it was also present in temporal lobe, cerebellum, hippocampus, and spinal cord (Fig. 3B). MRP4 was not present in caudate nucleus but was present in temporal lobe, cerebellum, hippocampus, medulla, and spinal cord, with low expression in hypothalamus. Thus, in most regions of the brain, CYP2B6 was coexpressed with MRP4 but not with CAR.

Bioinformatic Features of the Alternative Human CAR Isoforms. We identified a novel CAR SV with only deletion of exon 2 (SV22) or with exon 2 deleted in combination with other splicing changes (SV3, 4, 6, 7, 9, 10, 11, 12, 13, 15, 19, 20, and 22) (Fig. 2). Since the translation start site of hCAR is present in exon 2, it was not clear whether any alternative start site existed that could result in translation of a unique CAR protein in the absence of exon 2. We analyzed the sequence of hCAR cDNAs lacking exon 2 with an AUG evaluator program (http://l25.itba.mi.cnr.it/~webgene/wwwaug.html) that predicted translation initiation sites. The ATG at position 220 (Genbank accession no. Z30425) in exon 1 had a predicted initiation score of 0.83 compared with a score of 0.85 for the ATG in exon 2 (Fig. 4A). This score is strongly in favor of translation initiating at the ATG in exon 1. However, if the ATG in exon 1 is used as an alternative translation initiation site, it would result in full-length CAR protein only when there is deletion of exon 2. In the presence of exon 2, it encounters a premature termination codon after the first 9 amino acids. Translation initiation from the exon 1 ATG encodes six unique amino acids (Fig. 4B). In the CAR-SVs that skip exon 2, the final two nucleotides in exon 1 and the first nucleotide in exon 3 form a codon encoding an arginine (R) exactly in frame with the typical CAR protein (Fig. 4A). Thus, compared with CAR.1 that encodes a 348-amino acid protein, all CAR-SV skipping exon 2 lack the usual amino-terminal 35 amino acids but have instead six unique amino acids at the amino terminus.

CAR cDNAs have now been cloned from a variety of model organisms. The availability of these genomic sequences from diverse species allowed us to align the nucleotides in exon 1 containing the alternative ATG initiator codon. There was striking conservation between species in the exon 1 sequence, including the ATG that was conserved in baboon, chimp, and cow (Fig. 4C). However, mouse and rat do not contain this ATG. Nevertheless, high evolutionary conservation among many species of the upstream ATG suggests they could also use this alternative translation start site.

Nuclear hormone receptors have a highly conserved DNA binding domain (DBD) that interacts in a very specific manner with cognate binding elements (Rastinejad et al., 1995). Splice variants with deletion of exon 2 would have a truncated DBD missing the first zinc finger (Fig. 4D).

We next examined the features of the CAR-SVs that skip the first portion of exon 9. Although exon 9 typically encodes the 42 carboxy-terminal amino acids of CAR, many of the CAR isoforms failed to use the canonical donor/acceptor splice site for exon 9 (Fig. 5). These SVs instead used a cryptic AG splice site in exon 9 and encoded 29 unique carboxy-terminal amino acids. Comparison of the CAR sequences available in GenBank revealed that this cryptic AG splice site is evolutionarily conserved in rat and mouse CAR (data not shown). The alternative CAR transcripts lacking exon 9 are not a PCR artifact because there are two submissions in the AceView database (supporting clones AA887435 and AA918394) representing partial CAR cDNAs cloned from human kidney with the partial deletion of exon 9.

Function of CAR Splice Variants. To determine whether initiation occurred at the exon 1 ATG, we constructed expression plasmids for the CAR-SVs characterized by deletion of exon 2 alone (SV22), exon 2 deleted in combination with other splicing events (SV3, SV6, SV7, SV11, and SV12), and CAR SVs with deletion of part of exon 9 in combination with other splicing events (e.g., SV1 and SV14). Coupled in vitro transcription and translation experiments...
revealed that these CAR-SVs made a CAR protein (Fig. 6). These CAR-SV isoforms also readily expressed the SV protein in both a bacterial (BL21 strain of *Escherichia coli*) and mammalian (Cos 7 cells) expression system (data not shown). A protein of the expected molecular weight consistent with translation initiation at the exon 1 ATG was observed for each CAR-SV skipping exon 2. An additional smaller protein was also observed for most of the CAR-SVs. Based on incorporation of radiolabeled methionine into the proteins, the higher molecular weight protein was synthesized to a greater extent as compared with the smaller molecular weight protein from each cDNA. To determine whether initiation did occur at the exon 1 ATG we mutated and eliminated the putative ATG start in SV6 and SV22. The mutated CAR-SV6 and SV22 failed to express the CAR protein with the higher molecular weight, consistent with translation from the exon 1 ATG. Thus, in CAR SV mRNAs skipping exon 2, translation can start at the exon 1 ATG, and a full-length CAR protein is made. However, the proteins still expressed a smaller molecular weight protein. Based on analysis of the CAR cDNA using the AUG predictor program, it is most likely that the downstream ATGs, encoding Met 125 or Met 128 in exon 4 (Fig. 6B), are the likely sites for alternative translation initiation.

### hCAR Isoforms Demonstrate Compromised Binding to DNA and Transactivation of the CYP2B6-LUC Reporter

We used EMSAs to determine whether any of the SVs could bind to the CAR binding elements (NR1 or NR2) of the phenobarbital response element module in the CYP2B6 promoter (Sueyoshi et al., 1999). None of the SVs lacking exons 2 or 9 singly or in combination with other splicing events formed a complex with either the NR1 or NR2 in the presence or absence of RXRα (Fig. 7A). We next performed transfection experiments in HepG2 cells using a CYP2B6-LUC reporter and cotransfecting either CAR.1 or CAR-SV expression plasmids. Wild-type CAR dose-dependently activated CYP2B6-LUC reporter activity, whereas selected CAR-SVs failed to activate CYP2B6-LUC transcription either alone or in cells treated with the CAR ligand activator CITCO (Fig. 7B) (Maglich et al., 2003).

### Regulation of CAR Expression

There is growing evidence that transcription and splicing are coincident events and that coactivator recruitment to nuclear hormone receptors can alter splice site selection (Auboeuf et al., 2002). To determine whether drug treatment could alter the expression of CAR and CAR-SV mRNAs, we treated primary human hepatocytes with either PXR (rifampin) or CAR (phenobarbital and CITCO) activators and examined expression of CAR mRNAs. Treatment with phenobarbital consistently induced the amount of CAR transcripts (Fig. 8). However, CITCO, a ligand specific for human CAR, did not induce expression of either CAR or PXR. Thus, CAR appears to be differentially regulated by
its activators. The extensive number of alternatively spliced CAR transcripts that share splicing changes makes it impossible to generate primers that are unique to each CAR-SV and prevented us from determining whether there was any preferential induction of individual CAR isoforms by phenobarbital.

**Discussion**

Human genomic diversity for any gene encompasses sequence variations as well as alternatively spliced mRNAs. Indeed, alternate splicing of mRNAs is thought to be one of the mechanisms generating protein diversity. According to
recent estimates (Venter et al., 2001), between 22 and 59% of human genes are alternatively spliced. Alternative splicing can regulate the expression of the wild-type mRNA and show gain-of-function, loss-of-function, or dominant negative activity (Christmas et al., 2001; Chandrasekharan et al., 2002; Rose et al., 2003). Polymorphic alternative splicing can cause human diseases and is a frequent mechanism of human variation in cytochrome P450 expression (http://www.imm.ki.se/CYPalleles). Bioinformatics approaches to identify splice variants exist that use a variety of statistical approaches to predict the probability that these are bona fide splice sites. The results are currently inadequate because the rules governing all factors that can influence splicing are still being identified (e.g., exon splice enhancers, exon splice silencers, levels of expression of SR proteins). The only definitive way to identify all alternatively spliced mRNAs is to PCR the transcripts and directly sequence them. Moreover, because tissue-specific splicing is a common event that can further be influenced by race, development, and gene transcription, it is necessary to identify the transcripts in multiple tissues and from persons representing different ethnic populations to capture the full range of alternative transcripts for any gene.

In this report, we identified 22 novel variant isoforms of hCAR, 5 of them uniquely expressed in extrahepatic tissues and 16 transcripts amplified from human liver. Of the 22 variants, 13 had deletion of exon 2, 11 had deletion of exon 9, and 7 had both exons 2 and 9 deleted in combination with other splicing events (in frame insertion of 12 bp from intron 6 and/or insertion of 15 bp from intron 7 or deletion of exon 7). Our results are in contrast to recent reports that each identified the same three CAR-SVs with single alterations in the ligand binding domain (Auerbach et al., 2003; Savkur et al., 2003; Jinno et al., 2004) that have insertions in the hinge region (Zhang et al., 2003). In contrast to the four previously identified CAR-SVs that each had alterations in the ligand binding domain, some splice variants we identified skipped exon 2, resulting in usage of an alternative start site and yielding a unique amino-terminal sequence (AF1 domain) but lacking a portion of the DNA binding domain. CAR-SVs that had partial deletions of exon 9 and instead utilized a cryptic AG splice site in exon 9 generate CAR proteins with a unique carboxy-terminal region (AF2 domain). Although some of the CAR-SVs had premature termination codons (25% of SV in liver and 40% in all tissues) and would not encode functional proteins, many of the CAR-SV would yield novel proteins (75% in liver and
60% in all tissues). This result is consistent with a recent survey of over 1000 alternatively spliced exons that reported that the majority (~78%) of alternative spliced variants were functional. Although some CAR mRNAs identified had a single splicing event, the majority of the CAR alternative mRNAs had multiple alternative splice events in a variety of combinations.

Our results also extend the list of tissues in which CAR is present. CAR has previously been reported in liver, intestine, and adrenal gland. Using PCR, we readily detected CAR in human kidney, liver, adrenal, testis, and fetal liver and at low levels in brain, heart, spleen, small intestine, placenta, prostate, and skin. However, some of the CAR-SVs were expressed in a tissue-specific fashion, and the relative expression of the various CAR isoforms was significantly different between the tissues. For example, prostate does not express CAR.1, but exclusively expresses a CAR-SV that skips exon 2. The complexity of the CAR isoforms identified will make it very difficult, if not impossible, to quantitate individual transcripts. Although a recent study (Savkur et al., 2003) quantified by real-time PCR the amounts of CAR-SV with either the 12- and 15-bp insertions or with deletion of exon 7, our results now show that the primers utilized in that study would have quantitated multiple mRNA transcripts. Additionally, the primers used would not identify the unique CAR-SV isoforms present in some tissues.

A majority of the CAR-SVs skipped exon 2 and, thus, lost...
the native ATG codon. These CAR-SVs encode novel isoforms of CAR that utilize an alternate start site in exon 1. In these variants, 35 amino acids coded by exon 2 are replaced by 6 amino acids from exon 1. Interestingly, if the exon 1 ATG is used in the presence of exon 2, a premature termination codon is encountered after coding for nine amino acids. Thus, translation initiating from the exon 1 ATG would never produce a full-length CAR protein unless exon 2 is deleted. Upon deletion of exon 2, an in-frame full-length variant isoform is encoded with an altered NH₂ terminal end. The loss of exon 2 also deletes a portion of one of the zinc fingers in the DNA binding domain. Accordingly, transcripts skipping exon 2 had compromised ability to bind DNA and transactivate CYP2B6 transcription. Nevertheless, the facts that the alternative ATG in exon 1 is evolutionarily conserved across species and that some nuclear receptors that lack the DNA binding domain, such as small heterodimer partner, can still heterodimerize with other receptors and negatively regulate their function (Seol et al., 1996) suggest that there may be an unidentified function for these CAR-SVs.

Eleven of the CAR SVs had partial deletion of exon 9. These CAR-SVs lose the normal 42 carboxy-terminal amino acids. However, these CAR-SVs utilize a cryptic splice site in exon 9 that is 23 bp downstream of the TGA stop codon and thus encode 29 unique carboxy-terminal amino acids. This deletion also results in the loss of helices 10 to 12 in the ligand binding domain. Helix 12 contains a short transactivation function (AF-2) domain (Moras and Gronemeyer, 1998). The AF-2 domain is critical for the constitutive activity of CAR, mediating the ligand-independent interaction of CAR with coactivators. It remains to be determined how the altered carboxy terminus of these CAR-SVs affects interaction with coactivators or corepressors.

It is also possible that CAR-SVs with partial deletion of exon 9 might have an altered cellular localization compared with CAR. CAR normally resides in the cytosol but upon treatment with phenobarbital translocates to the nucleus. Carboxy-terminal amino acids in CAR are required for this translocation because deletion of the 30 C-terminal amino acids abolishes the translocation (Zelko et al., 2001). In addition, carboxy-terminal amino acids in CAR are critical to ligand-dependent and -independent transactivation of CAR (Andersen et al., 2003). It remains to be determined whether the CAR-SVs with an altered C-terminus can translocate to the nucleus.

In conclusion, the present investigation identified 22 novel splice variants of hCAR that are expressed in a tissue-specific fashion. The identification of these cDNAs is the first step toward understanding their function and whether the transcripts are polymorphically expressed. hCAR isoforms with altered N- and C-terminal ends have been identified, thus providing increased complexity and diversity of CAR proteins. Some of the CAR-SVs had alterations in functional domains, such as loss of one of the zinc fingers necessary to dock at the CAR binding element in DNA and resulting in their compromised ability to bind to or to transactivate the CYP2B6 promoter. Nevertheless, some of these CAR-SVs may have unique functions as has been shown for SV isoforms of other genes. Indeed, despite loss by some SV proteins of what are considered required functional domains, these SVs can have unique functions (Rose et al., 2003). Nuclear hormone receptors lacking DNA binding domains, such as small heterodimer partner, can still be functional (Seol et al., 1996), and nuclear hormone receptors can also have distinctive functions outside the nucleus (Li et al., 2000). Thus, it is possible and likely that some of the CAR-SVs have unique functions that remain to be identified.

Acknowledgments

We thank the Hartwell Center for DNA sequencing.

References


Sueyoshi T, Kawamoto T, Zelko I, Honkakoski P, and Negishi M (1999) The re-


Address correspondence to: Dr. Erin G. Schuetz, Department of Pharmaceutical Sciences, St. Jude Children’s Research Hospital, 332 North Lauderdale Street, Memphis, TN 38105. E-mail: erin.schuetz@stjude.org