Expression of class 3 aldehyde dehydrogenase (ALDH-3) is constitutive or inducible, depending on the tissue. ALDH-3 induction occurs both during neoplastic development and after exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD). In order to study the regulation of ALDH-3 gene expression, ALDH-3 genomic sequences have been obtained from normal rat genomic DNA. Two overlapping genomic fragments (ALDH-UTR-1 and ALDH-NL2) contain the entire ALDH-3 gene along with considerable 5'- and 3'-flanking sequences. The rat ALDH-3 gene spans approximately 9 kilobases in length and consists of eleven exons; ten coding and one 5'-noncoding. The region 5' to exon one contains several putative transcription factor binding elements which may be important in the TCDD inducibility of this gene. These include a xenobiotic response element (XRE), a drug response element (DRE), LAP and Ap1 binding sites, and one Sp1 site. There are considerable differences in organization between the rat and human class 3 ALDH genes. Primer extension and RNase protection analysis indicate that both basal and TCDD-inducible expression of the ALDH-3 gene utilize the same multiple transcription start sites.

Aldehyde dehydrogenase (ALDH) catalyzes the oxidation of a large variety of aldehydes to their corresponding acids. Among the three classes of mammalian aldehyde dehydrogenases, class 3 aldehyde dehydrogenase is of considerable interest because its expression can be either constitutive or inducible depending on the tissue. Class 3 ALDH is induced in liver and urinary bladder during carcinogenesis (Lindahl, 1992). This phenomenon occurs early in tumorigenesis when the normally repressed class 3 ALDH gene becomes activated in putative tumor cells. This results in overt tumor cells possessing high levels of class 3 ALDH activity. Independently, class 3 aldehyde dehydrogenase is also induced acutely in normal rat liver by treatment with 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) or polycyclic aromatic hydrocarbons (Dunn et al., 1989; Deitrich et al., 1977; Torronen et al., 1981; Mar- lses et al., 1987).

In contrast, class 3 ALDH is constitutively expressed at very high levels in the cornea of the eye (Evches and Lindahl, 1989). The enzyme is also expressed at moderate levels in the stomach (Yin et al., 1989) and lung (Dunn et al., 1988). For the stomach and lung, the activity is also inducible, as activity increases after exposure to inducing agents.

Presently, the exact mechanism for the modulation of class 3 ALDH gene expression is unknown. Considerable evidence exists that the TCDD and polycyclic aromatic hydrocarbon-mediated induction of class 3 ALDH activity may occur via an Ah-receptor-mediated pathway (Nebert et al., 1990; Lindahl, 1992; Takimoto et al., 1991; Vasiliou et al., 1992). This mechanism involves binding of the ligand (TCDD or polycyclic aromatic hydrocarbon) to a cytoplasmic Ah receptor, followed by translocation of the ligand-receptor complex from the cytosol to the nucleus, and ultimately activation of the class 3 ALDH gene. However, the inducibility of the ALDH-3 gene by TCDD does not correlate well with the tissue distribution of the Ah receptor (Carlstedt-Duke, 1979), and the time course of TCDD induction of the class 3 ALDH gene in liver is distinctly different from that of other TCDD-inducible genes (Dunn et al., 1988; Takimoto et al., 1992). Therefore, control of class 3 ALDH gene regulation likely includes both Ah receptor-mediated and Ah receptor-independent mechanisms.

Complementary cDNAs encoding class 3 ALDHs have been cloned from HTC rat hepatoma cells and TCDD-treated rat liver (Jones et al., 1988; Dunn et al., 1988). Nucleotide sequencing indicates the coding regions of these 2 cDNAs are identical, signifying that the two distinct conditions induce the same mRNA. One full-length class 3 ALDH cDNA (pALDH) or ALDH-3 cDNA contains 178 nucleotides of 5'-noncoding sequence, 1359 bp of coding region, and 248 bp of 3'-noncoding sequence including the polyadenylation signal. Interestingly, in the absence of any manipulation of this cDNA or its vector, Escherichia coli transformed with pALDH produce class 3 ALDH protein that is enzymatically active (Jones et al., 1988).

Rat class 3 aldehyde dehydrogenase has a cloned human homologue in the form of the human stomach ALDH-3 gene (Hsu et al., 1992). The degree of identity between the coding regions of rat ALDH-3 and human stomach ALDH-3 has
been shown to be 81% at both the nucleotide and amino acid levels. This is similar to homologies seen among class 1 or class 2 ALDH genes and proteins from various species (Hemphill and Lindahl, 1989). Some nucleotide sequence 5' to the putative human class 3 ALDH coding region has been examined for possible regulatory elements, and putative transcriptional control elements were identified (Hsu et al., 1992).

At this time, we report the cloning and characterization of the rat class 3 ALDH gene. Southern analysis and sequence data indicate that striking structural differences exist between the human and rat ALDH-3 genes. The promoter region of the ALDH-3 gene was sequenced, and multiple start sites were mapped from RNAs from untreated and TCDD-treated rat hepatoma cells.

**Materials and Methods**

**Libraries and Screening** A λ EMBL2 rat genomic library was a gift of Dr. Perry Churchill, the University of Alabama, A λ DASH rat genomic library was obtained from Stratagene. cDNA to the rat class 3 ALDH (pTALDH) was that described by Jones et al. (1989) as cloned by this laboratory. Probes were labeled with [α-32P]dATP (Du Pont-New England Nuclear) by random priming (United States Biochemical Corp.). The hybridization conditions were those described in Sambrook et al. (1989). The hybridization temperature was 65 °C. Filters were washed posthybridization in 1 x, 0.1 x SSCP, and 0.1 x SSCP + 5 x SSC. Hybridization was followed by autoradiography at -80 °C for 24 h. Random priming of double-stranded DNA with [α-32P]dCTP and end-labeling of oligonucleotides with [γ-32P]dATP was carried out by the methods of Sambrook et al. (1989).

**Southern Blot Analysis**—DNA from 3 class 3 ALDH genomic clones was isolated via CsCl gradient purification. Approximately 2 μg of cloned DNA were digested with restriction enzymes (New England Biolabs) according to the manufacturer's recommendations, then electrophoresed on autoradiography at 65 °C. Filters were washed posthybridization in 1 x, 0.1 x SSCP, and 0.1 x SSCP + 5 x SSC. Hybridization was followed by autoradiography at -80 °C for 24 h. Aspecific binding was determined by prehybridization with excess competitor DNA.

**Cloning the 5' End of Class 3 ALDH cDNAs**—The 5' end of ALDH-3 cDNA was prepared by reverse transcription of the cDNA followed by nest polymerase chain reaction (PCR) (Oohara et al., 1988; Saiki and Gelfand, 1989). The first-strand cDNA synthesis was performed using 20 μg of total RNA purified from superfused H4IE C3V cells (Takimoto et al., 1991). The primer for first-strand synthesis was a synthetic oligonucleotide (primer 3) with sequence 5'-TGCAGCGATCGAGTCTTGCCGGAG-3' which corresponds to nucleotides +221 to +244 of pTALDH (Jones et al., 1988).

The conditions were the same as those described for primer extension. After first-strand synthesis, a poly(A) tail was added to the 3' end of the cDNA by terminal deoxynucleotransferase. The first-round PCR was performed in a 100-μl reaction containing 10 mm Tris-HCl (pH 8.8), 50 mm KCl, 15 mm MgCl2, 3 mM dNTPs, 0.1 mg/ml bovine serum albumin, 0.2 mm each dATP, dGTP, dCTP, and dTTP, 100 pmol of primers (dT(24-26) and primer 3), and the polyadenylated cDNA product. The reaction conditions were as follows: 35 cycles of denaturing at 94 °C for 1 min, annealing at 42 °C for 1 min, extension at 72 °C for 1 min, and final extension at 72 °C for 10 min. The second-round PCR was performed under the same conditions as the first-round PCR except that primer 3 was replaced by primer 2, with the first-round PCR product as substrate. The second-round PCR products were blunted by DNA polymerase Klenow fragment and cloned into the phosophorylated HindII site of pGEM3 (Promega). Four clones obtained from this procedure were sequenced by the methods described below for DNA sequencing.

**RESULTS AND DISCUSSION**

**Transcription Start Sites**—The rat liver ALDH-3 gene can be induced under two distinct conditions, transformation-associated expression and TCDD-inducible expression. Moreover, different rat hepatoma cell lines express different levels of ALDH-3 in the absence of inducers (Lin et al., 1988; Takimoto et al., 1991). For example, H4IE C3V cells normally do not produce ALDH-3 mRNA. Untreated HTC cells produce considerable ALDH-3 mRNA. TCDD-treated H4IE C3V cells produce significant ALDH-3 mRNA. However, maximal levels of ALDH-3 mRNA attained in these cells is only 3-fold that of untreated HTC cells. TCDD-treated HTC cells produce about 2–3-fold the ALDH-3 mRNA of untreated HTC cells (Lin et al., 1988; Takimoto et al., 1991). In order to determine transcription start sites for the ALDH-3 gene under the uninduced and induced conditions, RNAs from untreated and TCDD-treated H4IE C3V and HTC cells, as well as RNAs from both cell lines treated with both TCDD and cycloheximide, were used. Cycloheximide, an inhibitor of protein synthesis, markedly enhances the TCDD-induced transcription of the ALDH-3 gene resulting in superinduction (Takimoto et al., 1991). Simultaneous treatment of EC3 or HTC cells with TCDD and cycloheximide for 24 h resulted in an approximately 15-fold greater accumulation of ALDH-3 mRNA than TCDD treatment alone. In all RNA preparations, the same pattern of primer extension products was obtained (Fig. 1A). This result indicates that the initiation of basal, TCDD-induced,
Class 3 ALDH Gene Organization

12532

Fig. 1. Primer extension analysis. A, for the induction, cells were treated with 10 nm TCDD (TCDD-induced) or both 10 nm TCDD and 10 μg/ml cycloheximide (Superinduced) for 24 h (Takimoto et al., 1991). The total RNAs were prepared, and primer extension analysis was performed with primer 2 (corresponding to nucleotides +174 through +197 of pTALDH). Normal liver total RNA was isolated and used as a negative control as well as yeast tRNA. The extended products were electrophoresed on a 6% sequencing gel along with the control sequence of bacteriophage M13mp18. B, the primers were designed on the basis of the previously reported ALDH-3 cDNA (pTALDH) (Jones et al., 1988). The asterisk represents the translational start site determined by Jones et al. (1988).

and superinduced transcription occurs at the same multiple sites. The main initiation site appears to be 45 bp upstream from the translation start site, whereas positions nearer to the translation start site appear to be used as minor initiation sites.

Structure of the 5' End of the ALDH-3 Transcripts—The longest class 3 ALDH cDNA, pTALDH, consists of 173 bp 5' to the translation start site (Jones et al., 1988). Primer extension analysis was performed with two different primers corresponding to nucleotides +36 to +59 (primer 1) and +174 to +197 (primer 2) of the ALDH-3 cDNA (Fig. 1). No significant extended product was detected with primer 1 from RNAs obtained from either untreated, TCDD-, or TCDD plus cycloheximide-treated cells. Strong signals were obtained with primer 2 from all RNAs (Fig. 1A). However, the lengths of the extended products obtained with primer 2 were shorter than predicted based on the sequence of pTALDH. If the majority of ALDH-3 mRNAs have the same 5' end sequence as pTALDH, then the extended products with primer 2 should be approximately 197 bases long (Fig. 1B). However, the extended products obtained with primer 2 were 69 bases in length or shorter. Since primer 2 was designed to correspond to the nucleotides inside the coding region of pTALDH, and the N-terminal amino acid sequence of the protein encoded by pTALDH was determined directly (Hempel et al., 1989), it is likely that the ALDH-3 mRNA from H4IIE C3V and HTC cells possess a 5' end structure different from the cDNA carried as pTALDH. This possibility is supported by the fact that the relative intensities of the signals were consistent with the abundances of ALDH-3 mRNA in each RNA preparation.

The 5' end structure of the ALDH-3 mRNA was further confirmed by an RNase protection assay (Fig. 2). An RNA probe was prepared from the 5' end EcoRI-Stul fragment of pTALDH (Fig. 2A). If the majority of ALDH-3 mRNA contains the region corresponding to the 5' end EcoRI-Stul fragment of pTALDH, then the protected product would be approximately 213 bases in length (RNA 1). However, no significant signal was observed which would indicate a product of this size. However, strong signals were visualized for products 84 bp in length and shorter (RNA 2 in Fig. 2A). Since the distance from the Stul site to the translation start site is 39 bp long, this result is in good agreement with both the previously described results for primer extension analysis and

Fig. 2. RNase protection assay. A, RNA probe contained the 213 bases of insert DNA and 18 bases of the vector polylinker. RNA 1 and RNA 2 indicate the proposed 5' end structure of ALDH-3 mRNA either by pTALDH (Jones et al., 1988) or by our results, respectively (see text). B, total RNAs were prepared from induced, superinduced, and uninduced H4IIE cells, as described in the legend for Fig. 1. The protected products were electrophoresed on a 6% sequencing gel along with the control sequence of bacteriophage M13mp18 DNA. The gel was dried and exposed to x-ray film with an intensifying screen at 70 °C for 15 h.
the amplification of the 5' end of ALDH-3 cDNA (84 – 39 = 45). Again, as in the previous experiments, the intensities of the signals observed were consistent with the amounts of ALDH-3 mRNA in each RNA preparation.

In order to further establish the 5' region of the major ALDH-3 mRNAs, we generated cDNAs by reverse transcription of isolated mRNA. ALDH cDNAs were then amplified by nested PCR. A primer 3' to primer 2, corresponding to pTALDH nucleotides +221 to +224 (primer 3, Fig. 1B), was obtained and utilized as an outer primer for the first round of PCR amplification. Since ALDH-3 mRNA was most abundant in the superinduced cells, and since primer extension analysis with the superinduced RNA preparation gave the same lengths of the extended products as the uninduced and TCDD-induced preparations, the superinduced ALDH-3 mRNA was used as a starting material. The amplified PCR products were cloned into pGEM3, and the four clones obtained were sequenced in both directions. Since the amplified DNA fragments each contained differing lengths of poly(A) tail at their 5' end, the total lengths of the inserts varied between 118 and 143 bp. Each clone, however, contained the same 69-bp fragment, within which a 24-bp portion was identical with that of primer 2. Additionally, the length of the PCR products was in good agreement with those obtained by primer extension analysis. The sequence of the remaining 45 bp (69 – 24 = 45 bp) appears identical with the sequence between +138 and +197 of pTALDH indicating that the 5' end of the ALDH-3 mRNA is encoded within the same portion of the gene as that of pTALDH mRNA.

From these results, we conclude that the principal ALDH-3 mRNAs in the uninduced, TCDD-induced, and TCDD/cycloheximide-superinduced cells are the shorter mRNAs. Moreover, since pTALDH was isolated from a library of HCT hepatoma cells (Jones et al., 1988), and RNA preparations from HTC cells were used in this study, it appears unlikely that different cells express different lengths of mRNAs in any detectable amount. Thus, it is likely that the class 3 ALDH cDNA carried in pTALDH is representative of a much rarer lengthened transcript that is present in very small amounts in HTC cells.

Structure of the ALDH-3 Gene—An EMBl2 rat genomic library was screened with a BglII-NheI fragment from pTALDH, and three clones were identified. pTALDH was used to screen because even though it may represent a rare transcript it is the longest ALDH-3 cDNA available. Southern analysis using oligonucleotides specific for both the 5' and 3' ends of the coding region of pTALDH established that the 5' and 3' ends of the coding sequences of the gene were present within the boundaries of one of the three isolated clones (ALDH-NL2). Subsequent sequence analysis confirmed this information, and the clone NL2 was positively identified as the class 3 ALDH gene. However, upon further characterization, it became apparent that NL2, although containing all of the protein-coding information of the ALDH-3 gene, lacked the majority of the pTALDH-UTR sequence and its associated 5' ALDH-3 gene regulatory domains. Further screening of the EMBl2 library yielded no clones. Therefore, a λ Dash rat genomic library was screened for the missing pTALDH-UTR and its associated 5'-flanking domain.

A three-tiered screening approach utilizing three different DNA probes was utilized. The probes were as follows: probe 1 was a SacI-BamHI fragment of approximately 1500 bp from the most extreme 5' region of the previously isolated partial ALDH-3 genomic clone NL2. This probe contained the ALDH-3 exon 2 that includes the translational start site and 7 bp of the 173-bp pTALDH upstream sequence. In addition, probe 1 contained approximately 800 bp of DNA subsequently identified as intron one. Probe 2 was an 800-bp EcoRI-BglII fragment from the 5' end of the ALDH-3 (pTALDH) cDNA. This probe was used to screen the clones selected in the first screening and to identify them as ALDH clones. Probe 3 was a 140-bp EcoRI-PvuII fragment from the untranslated region of the ALDH-3 (pTALDH) cDNA. This DNA was not contained within the genomic clone NL2. Probe 3 was used in all tertiary screens to select for only those clones that contained DNA 5' to the most 5' DNA within clone NL2. Of 18 clones selected by probe 2, 3 showed positive hybridization with probe 3.

The three ALDH-3 gene clones selected were initially digested with BamHI and EcoRI and analyzed by Southern blotting. One clone, ALDH-UTR-1, contained the largest insert (approximately 16 kb) and demonstrated the strongest hybridization. This clone was further characterized via restriction digestion and Southern analysis.

ALDH-3 clones ALDH-NL2 and ALDH-UTR-1 together contain the entire ALDH-3 gene. Both clones were subjected to extensive characterization (Fig. 3). The rat ALDH-3 gene coding region spans approximately 10 kb and consists of 11 exons ranging in size between 85 bp and 268 bp with an average of 167 bp. The intron/exon junctions were determined by direct DNA sequencing, and acceptor splice signals conform to accepted consensus sequences (Table I). With the exclusion of exon 1 which is noncoding, only 3 of the remaining 10 exons (3, 7, and 9) interrupt the coding sequence within a codon. The nucleotide sequences of the exons are in complete agreement with that of the cDNA pTALDH.

Structure of the ALDH-3 5'-Flanking Region—Using appropriate oligonucleotides as probes, restriction fragments containing the genomic DNA corresponding to the ALDH-3 cDNA 5' UTR and the translational start site were identified. Oligonucleotide 1 corresponded to positions –38 through –60 in the ALDH-3 cDNA UTR, and oligonucleotide 2 corresponded to the ALDH-3 translational start site position +1 through +18. Southern hybridization on HindIII-digested ALDH-UTR-1 indicated that the translational start site (oligonucleotide 2) was in a single 4- to 5-kb HindIII fragment.

Sequence analysis of these two fragments confirmed that the translational start site plus the first 5 bp of the UTR and the remaining UTR sequence are encoded by two distinct exons separated by an intron of approximately 3 kb (Fig. 4). An additional 1115 bp of class 3 ALDH genomic sequence upstream of the UTR has been sequenced. A DNA analysis program was used for open reading frame analysis, and no open reading frames of any significant length were located in the sequence 5' to exon 1, or in the intron residing between exon 1 and exon 2.

Based on the primer extension and RNase protection studies, exon 1 of the rat ALDH-3 gene is 40 bp in length and contains nucleotides corresponding to –6 through –45 in the ALDH-3 (pTALDH) cDNA UTR. Exon 2 is 167 bp long and consists of ALDH-3 (pTALDH) cDNA UTR nucleotides +1 through +5 along with coding sequence to +335 and includes the translation start site at position 174 (Fig. 4). Therefore, exon 1 is a non-peptide coding exon, and exon 2 encodes the first 54 codons of class 3 ALDH mRNA. The remaining 128 nucleotides corresponding to pTALDH UTR –7 through –128 are found immediately 5' to the transcriptional start site located in exon 1 (marked with an * in Fig. 4). Thus, the translational start site for the mRNA that produced the pTALDH cDNA is 128 bp 5' to the major transcriptional
Class 3 ALDH Gene Organization

Fig. 3. Restriction maps of rat liver ALDH-3 genomic clones ALDH-UTR-1 and ALDH-NL2 indicating the following: region of overlapping DNA, 5'-untranslated region sequenced (see Fig. 4), and the approximate location of the rat ALDH-3 gene.

Table I

<table>
<thead>
<tr>
<th>Exon</th>
<th>bp</th>
<th>Donor sites</th>
<th>Intron length</th>
<th>Acceptor sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>3'</td>
<td>(161)</td>
<td>GTAGGAG</td>
<td>(3)</td>
<td>tatatatccag</td>
</tr>
<tr>
<td>Exon 2</td>
<td>(157)</td>
<td>GGC AAG Gly Lys</td>
<td>(1.5)</td>
<td>ttctttctag</td>
</tr>
<tr>
<td>Exon 3</td>
<td>(232)</td>
<td>GCT GCA G Ala Ala</td>
<td>(0.6)</td>
<td>ttgcag</td>
</tr>
<tr>
<td>Exon 4</td>
<td>(86)</td>
<td>GAC CAG Asp Gin</td>
<td>(0.3)</td>
<td>ttgctgaag</td>
</tr>
<tr>
<td>Exon 5</td>
<td>(209)</td>
<td>TGC AG Cys Arg</td>
<td>(0.7)</td>
<td>acttcccag</td>
</tr>
<tr>
<td>Exon 6</td>
<td>(121)</td>
<td>AAA GAT Lys Asp</td>
<td>(0.8)</td>
<td>ctcttctcag</td>
</tr>
<tr>
<td>Exon 7</td>
<td>(139)</td>
<td>TAC ATA G Tyr Ile</td>
<td>(0.7)</td>
<td>cctgcctcag</td>
</tr>
<tr>
<td>Exon 8</td>
<td>(100)</td>
<td>GAG AAG Gly Lys</td>
<td>(0.8)</td>
<td>ctttatata</td>
</tr>
<tr>
<td>Exon 9</td>
<td>(167)</td>
<td>GTG GTG G Gly Val</td>
<td>(0.3)</td>
<td>tgcagttgag</td>
</tr>
<tr>
<td>Exon 10</td>
<td>(131)</td>
<td>GCC AGG Ala Lys</td>
<td>(0.2)</td>
<td>gtcttcacag</td>
</tr>
</tbody>
</table>

*Refers to intron-exon boundary location in ALDH-3 cDNA.
* Refers to amino acid boundary at exon-intron junction.

ALDH-3 start site identified previously (located at +1 in Fig. 4) The intron/exon junctions for exon 1 and 2 conform to the accepted consensus sequences specific for intronic donor and acceptor splice signals (Breathnach and Chambon, 1981).

Structurally, the region 5' to exon 1 contains several putative regulatory elements (Fig. 4). A TATA box motif (TTAAT) is found at position -24 to -29. A CCAAT motif (GGCCCATCT) is present upstream of exon 1 at position -193 to -206. It may be that the pTALDH cDNA is a minor transcript that results when the CCAAT motif is utilized independently of the TATA sequence to align the transcriptional machinery. In addition, the region 5' to exon 1 contains numerous other putative transcription factor binding elements including the following: two xenobiotic response elements (XRE), a drug response element (DRE), a LAP (NF-IL6) binding site, Ap1 binding sites, and one Sp1 site (Fig. 4). Also present in this region are several direct and inverted repeat elements.

At positions -80 to -86 and -369 through -388 reside putative xenobiotic response elements that share very close homology to those XREs described by Fujisawa-Sehara et al. (1987) and Denison et al. (1988) residing upstream of the mouse and rat CYPIA1 genes. The core sequence of the putative class 3 ALDH-XREs consists of a short alternate
to the XRE, a putative drug-responsive element (DRE) where the consensus sequence for the XRE located at -369 by Nebert and Jones. This DRE shares the mology to that of the CYPlAl gene. Interestingly, the two

The XRE at position -498 to -506. An Sp1 site (GGGGCG) is located at position -42 to -47. Finally, Ap1 sites reside at positions -228 to -234, -279 to -285, -339 to -345, -598 to -604, -779 to -785, -904 to -910, and -932 to -938.

The abundance of putative transcription factor elements suggests that the ALDH-3 5'-flanking sequence is able to interact with a variety of different proteins, each of which may modulate gene expression differently. Whether the regulatory proteins that bind to these regions Ap1, Sp1, LAP, CAAT, XRE, DRE all actually do so remains to be determined. However, DNase I footprinting of the Sp1 site indicate that this element may be functional.

The organization of the 5' region of the rat class 3 ALDH gene is substantially different from the human class 3 ALDH gene reported by Hsu et al. (1992). Although rat and human ALDH-3 share a high degree of similarity at the protein level, these two isozymes appear to be very different at the level of gene organization (Table II and Fig. 5). First, in the human class 3 ALDH gene, the short 5'-untranslated region is reported to be contiguous with the translational start codon in a single first exon. This is clearly not the case in rat, as an intron of approximately 3 kb resides between the bulk of the UTR (exon 1) and the translational start site (exon 2). Additionally, the human gene has clearly defined CAATT (-117 to -114) and TATA (-79 through -76) regulatory sequences, whereas these sequences are much less defined in the corresponding region of the rat ALDH-3 gene. The TATA and CAAT motifs for the rat ALDH-3 gene are located 5' to the noncoding exon 1. Based upon published sequence, the 5'-flanking region of the human gene lacks both the XRE and DRE regulatory sequences found in the 5'-flanking region of the rat ALDH-3 gene (Hsu et al., 1992).

In the human class 3 ALDH gene, exons 1, 2, 3, 4, 7, 8, 9, and 10 encode the same amino acids as rat exons 2, 3, 4, 5, 8, 9, and 11. Human exons 5, 6, and 9 differ from rat exons 6, 7, and 10 in one amino acid. Thus, rat ALDH-3 possesses 10 coding and 1 noncoding exons; all 10 human ALDH-3 exons are coding.

For the rat ALDH-3 gene, the presence of an XRE may explain Ah receptor-mediated inducibility, and its presence is consistent with proposed models for class 3 gene regulation (Nebert et al., 1990; Lindahl, 1992). In addition, the presence and distribution of the XREs identified in the ALDH-3 gene are compatible with the organization of two other TCDD-inducible genes Nmo-1 and Gst-1 (Nebert et al., 1990).

Table II

<table>
<thead>
<tr>
<th>Exon</th>
<th>Rat</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Noncoding</td>
<td>Not present</td>
</tr>
<tr>
<td>2</td>
<td>1-54*</td>
<td>1-54</td>
</tr>
<tr>
<td>3</td>
<td>55-131</td>
<td>55-131</td>
</tr>
<tr>
<td>4</td>
<td>132-160</td>
<td>132-160</td>
</tr>
<tr>
<td>5</td>
<td>161-230</td>
<td>161-230</td>
</tr>
<tr>
<td>6</td>
<td>231-270</td>
<td>231-269</td>
</tr>
<tr>
<td>7</td>
<td>271-317</td>
<td>270-317</td>
</tr>
<tr>
<td>8</td>
<td>317-372</td>
<td>317-372</td>
</tr>
<tr>
<td>9</td>
<td>373-406</td>
<td>373-406</td>
</tr>
<tr>
<td>10</td>
<td>406-449</td>
<td>406-450</td>
</tr>
<tr>
<td>11</td>
<td>450-453</td>
<td>450-453</td>
</tr>
</tbody>
</table>

* Based on the rat gene organization.

* Codons included in exon.
The class 3 ALDH cDNA used to isolate the ALDH-3 gene was cloned from the rat hepatoma cell line HTC (Jones et al., 1988). Therefore, of some concern is the possibility that the cDNA does not represent the organization of the ALDH-3 gene in normal cells. However, we believe that the ALDH-3 gene organization reported here is correct for the following reasons. Although the class 3 cDNA was from a tumor cell line, both genomic clones NL2 and UTR-1 were cloned from normal rat liver genomic libraries. Second, Southern analysis of HindIII- or PstI-digested genomic DNA from normal rat liver and several rat hepatoma cell lines, including HTC, show identical restriction patterns when the full-length cDNA is the probe (Lin et al., 1988). Moreover, comparing NL2, UTR-1, and genomic DNA HindIII digests identifies the same two restriction fragments containing the UTR and translation start site (Lin et al., 1988).

Based on the structural organization of the 5'-flanking region of the rat ALDH-3 gene, it is possible that multiple mechanisms of transcriptional control may regulate expression of this gene. The presence of multiple regulatory elements upstream of the transcription start site offer the opportunity for both constitutive and inducible expression, perhaps in a tissue-specific manner. Also intriguing is the presence of potential regulatory elements in the first intron of the rat ALDH-3 gene which are similar to those found upstream of the human ALDH-3 gene. Whether any of these elements, either independently or in combination with those in the 5'-flanking region, are involved in the regulation of this gene is currently under investigation.

REFERENCES


