Molecular analysis of two closely related mouse aldehyde dehydrogenase genes: identification of a role for Aldh1, but not Aldh-pb, in the biosynthesis of retinoic acid

Lily C. HSU*, Wen-Chung CHANG*, Ines HOFFMANN† and Gregg DUESTER††

*Department of Biochemical Genetics, Beckman Research Institute of the City of Hope, Duarte, CA 91010, U.S.A., and †Gene Regulation Program, Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037, U.S.A.

INTRODUCTION

Retinoic acid and its derivatives have been the focus of considerable attention due to their role as potent modulators of gene expression, particularly regulation of vertebrate homeobox genes, and their role in cell differentiation and development [1,2]. Their action is mediated by nuclear receptors (e.g. retinoic acid receptors and retinoid X receptors), which are DNA-binding transcriptional regulators [3]. Retinoic acid deficiency or excess has been linked with developmental disorder symptoms [4]. Maintenance of homeostatic retinoic acid levels in tissues is thus essential to secure normal embryogenesis, organogenesis and adult life.

Retinoic acid homeostasis is achieved by multiple-step regulation of retinoid metabolism in retinoic acid-responsive cells, including the uptake of plasma retinol (the precursor of retinoic acid), and the biosynthesis and degradation of retinoic acid. The major biosynthetic pathway of retinoic acid from retinol (vitamin A) involves two sequential steps: (a) a reversible dehydrogenation into retinal catalysed by either cytosolic retinyl dehydrogenases, which are members of the alcohol dehydrogenase family, or by microsomal retinol dehydrogenases, which are members of the short-chain dehydrogenase/reductase family; and (b) an irreversible oxidation of retinal to retinoic acid catalysed by cytosolic retinal dehydrogenases, which are members of the aldehyde dehydrogenase (ALDH) family, as reviewed previously [5]. Activity assays in vitro have identified several ALDHs able to oxidize retinal, including human class I aldehyde dehydrogenase (ALDH1) [6,7], mouse ALDH1 (previously known as Ahd-2) [8,9], rat ALDH1 ([10] also called RALDH [11] or RalDH-I [12]), mouse RALDH2 (previously called V2 [13]) and its rat homologue RalDH-II [14] and perhaps also its human homologue (called ALDH11) [15], and, finally, mouse V1 enzyme, localized in the embryonic ventral retina, which is distinct from ALDH1 and RALDH2 [16]. Members of the ALDH family are involved in the oxidation of a wide variety of exogenous and endogenous aldehydes in addition to retinal [6,17,18].

Although ALDH1 was the first human ALDH to be purified and characterized [19], its physiological role has long been unclear, owing to its broad substrate specificity (e.g. acetaldehyde, retinal, aldophosphamide and 11-hydroxythromboxane B2), as well as its ability to bind various molecules (e.g. androgens and daunorubicin) [6,20–26]. It has been shown that purified human liver ALDH1 has extremely high activity for oxidation of free all-trans-retinal to all-trans-retinoic acid, with a $K_m$ value of 0.06 µM at pH 7.5 [7]. Evidence supporting retinoic acid synthesis as a major role for ALDH1 includes studies showing that the catalytic efficiency ($V_{max}/K_m$) of human ALDH1 for retinal oxidation is about 700 times higher than that for either acetaldehyde or aldophosphamide oxidation, and studies showing that ALDH1 is the sole isozyme among six purified liver ALDHs that can mediate NAD+-dependent retinal oxidation [6].

A mouse ALDH1 homologue (previously called Ahd-2) has been identified upon the basis of its substrate preferences and its coding sequence identity with human ALDH1 [27,28]. Six cytosolic ALDH isoenzymes were identified in adult mouse liver, and of these it was determined that ALDH1 mediates 95% of NAD+-dependent retinal oxidation in liver, and has a $K_m$ value of 185 µM.

Mammalian class I aldehyde dehydrogenase (ALDH1) has been implicated as a retinal dehydrogenase in the biosynthesis of retinoic acid, a modulator of gene expression and cell differentiation. As the first step towards studying the regulation of ALDH1 and its physiological role in the biosynthesis of retinoic acid, mouse ALDH1 cDNA and genomic clones have been characterized. During the cloning process, an additional closely related gene was also isolated and named Aldh-pb, owing to its high amino acid sequence identity (92%) with the rat phenoobarbitol-inducible ALDH protein (ALDH-PB). Aldh1 spans about 45 kb in length, whereas Aldh-pb spans about 35 kb. Both genes are composed of 13 exons, and the positions of all the exon/intron boundaries are conserved with those of human ALDH1. The promoter regions of Aldh1 and Aldh-pb demonstrate high sequence similarity with those of human ALDH1 and rat ALDH-PB. Expression of Aldh1 and Aldh-pb is tissue-specific, with mRNAs for both genes being found in the liver, lung and testis, but not in the heart, spleen or muscle. Expression of Aldh-pb, but not Aldh1, was also detected at high levels in the kidney. Aldh1 and Aldh-PB encode proteins of 501 amino acids with 90% positional identity. To examine the relative roles of these two enzymes in retinoic acid synthesis in vivo, Xenopus embryos were injected with mRNAs encoding these enzymes to assay the effect on conversion of endogenous retinal into retinoic acid. Injection of ALDH1, but not ALDH-PB, mRNA stimulated retinoic acid synthesis in Xenopus embryos at the blastula stage. Thus our results indicate that Aldh1 can function in retinoic acid synthesis under physiological conditions, but that the closely related Aldh-pb does not share this property.

Key words: promoter, retinal, retinoid signalling, vitamin A, Xenopus.

Abbreviations used: ALDH, aldehyde dehydrogenase; ALDH-PB, phenobarbitol-inducible ALDH; ALDH1, mammalian class I ALDH; ALDH1, Aldh1 and Aldh-pb, genes encoding human class I ALDH, mouse class I ALDH and the mouse homologue of rat ALDH-PB respectively; MMR, Marc’s modified Ringer’s solution; poly*(A)+, polyadenylated; UTR, untranslated region.

1 To whom correspondence should be addressed (e-mail duester@burnham-inst.org).
containing exon 12. All hybridizations were carried out essentially as described previously [33] in the presence of 50% (v/v) formamide at 42 °C for 18 h, followed by 2 × 15 min washes in 0.1 × SSC [15 mM NaCl/1.5 mM sodium citrate (pH 7.0)] and 0.1% (w/v) SDS at 55 °C.

PCR cloning of mouse ALDH1 and ALDH-PB cDNAs

DNA purified from a mouse testis cDNA library (Clontech Laboratories, Palo Alto, CA, U.S.A.) and RNA from mouse liver were used as templates to clone the ALDH1 and ALDH-PB cDNAs. PCR cycles were performed at 94 °C for 45 s, 56 °C for 45 s and 72 °C for 1 min. The 5'- and 3'-portions of the cDNAs were amplified with coding-region-specific and λgt11 vector-end-specific probes. The PCR products were separated on a gel, Southern-blotted and hybridized with coding-region probes. The longest positive bands were eluted from the gel, subcloned into pBluescript vector and sequenced. GenBank accession numbers are M36069 for mouse Aldh1 and U96401 for mouse Aldh-pb.

Restriction maps and exon/intron junctions of Aldh1 and Aldh-pb genes

DNA was prepared from phage clones using Qiagen Lambda DNA Prep columns (Qiagen, Valencia, CA, U.S.A.). The restriction maps of each genomic clone were determined by single- and double-restriction digestion, and by hybridization with exon-specific probes and vector-end probes (T3- and T7-map primers). The restriction endonuclease fragments containing exonic sequences were isolated by electrophoresis and subcloned into pBluescript + vector. The exonic restriction fragments of suitable sequencing size were further subcloned and/or sequenced. The exon–intron junctions were identified by comparing the genomic sequence with the cDNA sequence. PCR between exons was performed to estimate or confirm intron sizes. Introns of less than 5.5 kb were amplified in a DNA thermal cycler (PTC-100; MJ Research, Watertown, MA, U.S.A.) using the following programme: 94 °C for 45 to 60 s, 52 °C for 1 to 3 min. The 5'- and 3'-untranslated regions (UTR) sequences of either Aldh1 or Aldh-pb were used as specific hybridization probes. Probe labelling and hybridization were performed by standard methods [33].

Tissue distribution analysis

Northern blot analysis was performed with a mouse Multiple Tissue Northern blot (Clontech). The 3'-untranslated region (UTR) sequences of either Aldh1 or Aldh-pb were used as specific hybridization probes. Probe labelling and hybridization were performed by standard methods [33].

Retinoic acid bioassay in Xenopus embryos

Full-length mouse ALDH1 and ALDH-PB cDNAs were subcloned into plasmid pSP65 (Promega, Madison, WI, U.S.A.) in the sense direction for in vitro transcription from the SP6 promoter. Plasmids were linearized and subjected to in vitro transcription with SP6 RNA polymerase, 5'-capping with 7-methylguanosine, and 3'-polyadenylation with Escherichia coli poly(A) polymerase to produce full-length mRNAs [33]. mRNAs were brought to a concentration of 0.2 µg/µl in water, and 0.4 µg was used for in vitro translation in a rabbit reticulocyte lysate (Stratagene) to verify the ability to produce full-length proteins. Xenopus laevis embryos were produced by artificial fertilization and staged as described previously [34]. Embryos were placed in

Figure 1 Organization of mouse Aldh1 and Aldh-pb genes

Restriction maps of the Aldh1 (A) and Aldh-pb (B) genes are shown as rectangular bars. A scale in kb is shown between the maps for each gene. The corresponding exons of the two genes are assigned with the same numbers and represented by vertical solid bars, and introns are represented by the open areas between the solid bars. Above each gene are indicated the overlapping phage genomic clones containing the genomic regions (see the Materials and methods section for details) defined by arrows. Restriction maps for each of the indicated enzymes are shown above the open bars: S, SstI; B, BamHI.

for free retinal of 0.7 µM [8]. In addition, cells transfected with mouse ALDH1 cDNA can convert retinal into retinoic acid, supporting its role as a retinal dehydrogenase [29]. Mouse ALDH1 was also found in the embryonic and adult dorsal retina, an organ rich in retinoic acid [30]. In addition, a rat ALDH1 homologue isolated from either kidney (also called RALDH) [11,31] or liver (also called RaDH-1) [10,12,32] demonstrates high amino acid sequence identity with mouse ALDH1, and has high activity for retinal oxidation.

The kinetic studies provide in vitro support for human ALDH1, as well as its mouse and rat homologues, in the pathway of retinoic acid biosynthesis. In the present study we have begun a genetic examination of ALDH1 function and regulation by cloning and characterizing the mouse Aldh1 gene. A closely related gene, Aldh-pb [a gene encoding the mouse homologue of rat phenobarbitol-inducible ALDH], was also isolated and characterized. Both Aldh1 and Aldh-pb were examined for their ability to function in retinoic acid synthesis in vitro using a Xenopus embryo assay. Our results provide firm support for ALDH1 as a retinoic acid-synthesizing enzyme under physiological conditions, and show that the closely related ALDH-PB cannot perform this function.

MATERIALS AND METHODS

Genomic cloning

A mouse genomic library in FIXII prepared from DNA of mouse strain 129 (Stratagene Cloning Systems, La Jolla, CA, U.S.A.) was screened with a probe containing the human ALDH1 coding region in order to clone the mouse Aldh1 gene. Genomic walking with intron probes derived from an isolated Aldh1 clone (f15) was subsequently carried out to obtain overlapping Aldh1 clones (f6, f17 and f1). During the process of cloning Aldh1, we also identified two overlapping clones (f5 and f16) of a novel mouse gene, Aldh-pb. A 5'-end overlapping clone for Aldh-pb, f5-2, was isolated from the library with a human ALDH1 exon 1 probe, and a 3'-end overlapping clone, f11, was identified by genomic walking using as a probe a barH1 fragment from f16 containing exon 12. All hybridizations were carried out essentially as described previously [33] in the presence of 50% (v/v) formamide at 42 °C for 18 h, followed by 2 × 15 min washes in 0.1× SSC [15 mM NaCl/1.5 mM sodium citrate (pH 7.0)] and 0.1% (w/v) SDS at 55 °C.

PCR cloning of mouse ALDH1 and ALDH-PB cDNAs

DNA purified from a mouse testis cDNA library (Clontech Laboratories, Palo Alto, CA, U.S.A.) and RNA from mouse liver were used as templates to clone the ALDH1 and ALDH-PB cDNAs. PCR cycles were performed at 94 °C for 45 s, 56 °C for 45 s and 72 °C for 1 min. The 5'- and 3'-portions of the cDNAs were amplified with coding-region-specific and λgt11 vector-end-specific probes. The PCR products were separated on a gel, Southern-blotted and hybridized with coding-region probes. The longest positive bands were eluted from the gel, subcloned into pBluescript vector and sequenced. GenBank accession numbers are M36069 for mouse Aldh1 and U96401 for mouse Aldh-pb.

Restriction maps and exon/intron junctions of Aldh1 and Aldh-pb genes

DNA was prepared from phage clones using Qiagen Lambda DNA Prep columns (Qiagen, Valencia, CA, U.S.A.). The restriction maps of each genomic clone were determined by single- and double-restriction digestion, and by hybridization with exon-specific probes and vector-end probes (T3- and T7-map primers). The restriction endonuclease fragments containing exonic sequences were isolated by electrophoresis and subcloned into pBluescript + vector. The exonic restriction fragments of suitable sequencing size were further subcloned and/or sequenced. The exon–intron junctions were identified by comparing the genomic sequence with the cDNA sequence. PCR between exons was performed to estimate or confirm intron sizes. Introns of less than 5.5 kb were amplified in a DNA thermal cycler (PTC-100; MJ Research, Watertown, MA, U.S.A.) using the following programme: 94 °C for 45 to 60 s, 52 °C for 1 to 3 min. The 5'- and 3'-untranslated regions (UTR) sequences of either Aldh1 or Aldh-pb were used as specific hybridization probes. Probe labelling and hybridization were performed by standard methods [33].

Tissue distribution analysis

Northern blot analysis was performed with a mouse Multiple Tissue Northern blot (Clontech). The 3'-untranslated region (UTR) sequences of either Aldh1 or Aldh-pb were used as specific hybridization probes. Probe labelling and hybridization were performed by standard methods [33].

Retinoic acid bioassay in Xenopus embryos

Full-length mouse ALDH1 and ALDH-PB cDNAs were subcloned into plasmid pSP65 (Promega, Madison, WI, U.S.A.) in the sense direction for in vitro transcription from the SP6 promoter. Plasmids were linearized and subjected to in vitro transcription with SP6 RNA polymerase, 5'-capping with 7-methylguanosine, and 3'-polyadenylation with Escherichia coli poly(A) polymerase to produce full-length mRNAs [33]. mRNAs were brought to a concentration of 0.2 µg/µl in water, and 0.4 µg was used for in vitro translation in a rabbit reticulocyte lysate (Stratagene) to verify the ability to produce full-length proteins. Xenopus laevis embryos were produced by artificial fertilization and staged as described previously [34]. Embryos were placed in
Molecular analysis of two related mouse aldehyde dehydrogenase genes

### Table 1  Splice junction sites of the mouse Aldh1 gene (a) and the mouse Aldh-pb gene (b)

Exon sequences are in capital letters; intron sequences are in lower case letters. Introns are positioned by applying the gt/ag rule. Exon 1 starts from the putative transcriptional initiation site. The codon phases of the intron boundaries are indicated: O, introns inserted between codons; I, introns inserted after the first nucleotide of a codon; II, introns inserted after the first nucleotide of a codon. Amino acids encoded at the splice sites are numbered from the Ser at the 2nd position C-terminal side from the initiator Met.

(a)

<table>
<thead>
<tr>
<th>Exon</th>
<th>Exon size (bp)</th>
<th>Sequence at exon–intron junction</th>
<th>5’ Splice donor</th>
<th>Intron (kbp)</th>
<th>3’ Splice donor</th>
<th>Codon phase</th>
<th>aa interrupted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>(119)</td>
<td>−,CAT,ACC,AAG  gtaggtcact</td>
<td>1. (10.0)</td>
<td>ataatcagac</td>
<td>ATC,TTT,ATA</td>
<td>O</td>
<td>Lys21/Ile</td>
</tr>
<tr>
<td>2.</td>
<td>(105)</td>
<td>−,GGG,GAC,AAG  gtaaatatttc</td>
<td>2. (7.5)</td>
<td>tttgccccag</td>
<td>GCT,GAT,GTT</td>
<td>O</td>
<td>Lys56/Ala</td>
</tr>
<tr>
<td>3.</td>
<td>(141)</td>
<td>−,CTG,CTA,GCT  gtaggtacct</td>
<td>3. (1.9)</td>
<td>ttttttttttcg</td>
<td>ACA,ATC,GAA</td>
<td>O</td>
<td>Ala103/Thr</td>
</tr>
<tr>
<td>4.</td>
<td>(130)</td>
<td>ATA,CCA,AGT,G  gtagttatcc</td>
<td>4. (1.5)</td>
<td>tttgccccag</td>
<td>TGG,AAT,TTT</td>
<td>I</td>
<td>Asp147</td>
</tr>
<tr>
<td>5.</td>
<td>(62)</td>
<td>−,ATC,ATC,CCC  gtagttttcc</td>
<td>5. (1.4)</td>
<td>tttgccccag</td>
<td>TGG,AAT,TTT</td>
<td>O</td>
<td>Pro167/Trp</td>
</tr>
<tr>
<td>6.</td>
<td>(129)</td>
<td>−,ATA,AAA,GAG  gtagttttcc</td>
<td>6. (1.2)</td>
<td>ttttattctag</td>
<td>GCA,GGG,TTT</td>
<td>O</td>
<td>Glu210/Ala</td>
</tr>
<tr>
<td>7.</td>
<td>(114)</td>
<td>−,TCA,ACA,CAG  tgtagggcagg</td>
<td>7. (4.8)</td>
<td>tgtattttcg</td>
<td>GTT,GGC,GAA</td>
<td>O</td>
<td>Glu248/Val</td>
</tr>
<tr>
<td>8.</td>
<td>(103)</td>
<td>GAT,GCC,GAC,T  gtagttcacg</td>
<td>8. (3.3)</td>
<td>tttagttcag</td>
<td>TG,GAC,ATC,G</td>
<td>I</td>
<td>Leu283</td>
</tr>
<tr>
<td>9.</td>
<td>(185)</td>
<td>−,GCG,CCT,AGT  gtaaagttacg</td>
<td>9. (3.6)</td>
<td>tttagtttcag</td>
<td>ATG,ACC,AAC</td>
<td>O</td>
<td>Glu344/Ala</td>
</tr>
<tr>
<td>10.</td>
<td>(165)</td>
<td>−,AAA,GAG,GAG  gtagttttcc</td>
<td>10. (2.2)</td>
<td>tttagttttag</td>
<td>ATA,TTT,GGA</td>
<td>O</td>
<td>Glu399/Thr</td>
</tr>
<tr>
<td>11.</td>
<td>(158)</td>
<td>GG,GTG,GTG,TG  gtagtttctgg</td>
<td>11. (3.2)</td>
<td>tttagttttag</td>
<td>GTT,GAC,ATG</td>
<td>II</td>
<td>Trp452</td>
</tr>
<tr>
<td>12.</td>
<td>(75)</td>
<td>GA,AGA,GAA,CT  gtaagttaag</td>
<td>12. (2.7)</td>
<td>ccccttttag</td>
<td>G,GTT,GAA,CA</td>
<td>II</td>
<td>Leu477</td>
</tr>
<tr>
<td>13.</td>
<td>(&gt;570)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b)

<table>
<thead>
<tr>
<th>Exon</th>
<th>Exon size (bp)</th>
<th>Sequence at exon–intron junction</th>
<th>5’ Splice donor</th>
<th>Intron (kbp)</th>
<th>3’ Splice donor</th>
<th>Codon phase</th>
<th>aa interrupted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>(119)</td>
<td>−,CAT,ACC,AAG  gtaggtcact</td>
<td>1. (0.38)</td>
<td>atctccacag</td>
<td>ATC,TTT,ATA</td>
<td>O</td>
<td>Lys21/Ile</td>
</tr>
<tr>
<td>2.</td>
<td>(105)</td>
<td>−,GGG,GAC,AAG  gtaaatatttc</td>
<td>2. (4.8)</td>
<td>tgtttttccag</td>
<td>GCT,GAT,GTT</td>
<td>O</td>
<td>Lys56/Ala</td>
</tr>
<tr>
<td>3.</td>
<td>(141)</td>
<td>−,CTG,CTA,GCT  gtaggtacct</td>
<td>3. (3.1)</td>
<td>ttttttttttcg</td>
<td>ACA,ATC,GAA</td>
<td>O</td>
<td>Ala103/Thr</td>
</tr>
<tr>
<td>4.</td>
<td>(130)</td>
<td>ATA,CCA,AGT,G  gtagttatcc</td>
<td>4. (1.4)</td>
<td>ttttttttttcg</td>
<td>ATG,AGG,ACC</td>
<td>A</td>
<td>Asp147</td>
</tr>
<tr>
<td>5.</td>
<td>(62)</td>
<td>−,ATC,ATC,CCC  gtagttttcc</td>
<td>5. (1.7)</td>
<td>taaattttccag</td>
<td>TGG,AAT,GGT</td>
<td>O</td>
<td>Pro167/Trp</td>
</tr>
<tr>
<td>6.</td>
<td>(129)</td>
<td>−,ATA,AAA,GAG  gtagttttcc</td>
<td>6. (1.1)</td>
<td>taacttttag</td>
<td>GCA,GGG,TTT</td>
<td>O</td>
<td>Glu210/Ala</td>
</tr>
<tr>
<td>7.</td>
<td>(114)</td>
<td>−,TCA,ACA,CAG  tgtagggcagg</td>
<td>7. (5.2)</td>
<td>ttctttttccag</td>
<td>GTT,GGC,GAA</td>
<td>O</td>
<td>Glu248/Val</td>
</tr>
<tr>
<td>8.</td>
<td>(103)</td>
<td>GAT,GCC,GAC,T  gtagttcacg</td>
<td>8. (1.0)</td>
<td>tctctttccag</td>
<td>TG,GAC,ATC,G</td>
<td>I</td>
<td>Leu283</td>
</tr>
<tr>
<td>9.</td>
<td>(185)</td>
<td>−,GCG,CCT,AGT  gtaaagttacg</td>
<td>9. (6.4)</td>
<td>tctctttccag</td>
<td>ATG,ACC,AAC</td>
<td>O</td>
<td>Glu344/Ala</td>
</tr>
<tr>
<td>10.</td>
<td>(165)</td>
<td>−,AAA,GAG,GAG  gtagttttcc</td>
<td>10. (2.1)</td>
<td>tctctttccag</td>
<td>ATA,TTT,GGA</td>
<td>O</td>
<td>Glu399/Thr</td>
</tr>
<tr>
<td>11.</td>
<td>(158)</td>
<td>GG,ATG,ATG,AG  gtagtttctgg</td>
<td>11. (3.3)</td>
<td>tgttttttag</td>
<td>GTT,GAC,ATG</td>
<td>II</td>
<td>Trp452</td>
</tr>
<tr>
<td>12.</td>
<td>(75)</td>
<td>GA,AGA,GAA,CT  gtaagttaag</td>
<td>12. (2.7)</td>
<td>cctcttttag</td>
<td>G,GTT,GAA,CA</td>
<td>II</td>
<td>Leu477</td>
</tr>
</tbody>
</table>

### RESULTS

#### Organization of the mouse Aldh1 and Aldh-pb genes

Four overlapping mouse genomic clones (f1, f17, f6 and f15) that together contain the entire coding sequence of Aldh1 were cloned and characterized (Figure 1A). Aldh1 spans 45 kb and contains 13 exons. During the process of cloning Aldh1, we also identified...
two overlapping clones (f5 and f16) of a novel mouse gene, Aldh-pb, which exhibited the homologous splice positions as contained in exons 3–12 of Aldh1. Genomic walking was performed to obtain 5'– and 3'–clones (f5-2 and f11 respectively), and it was determined that Aldh-pb spans 35 kb and contains 13 exons (Figure 1B). Southern blot analysis of BamHI- or SstI-treated total genomic DNA with each respective cDNA probe revealed band patterns consistent with the restriction maps of both cloned genes.

The intron/exon boundaries of Aldh1 and Aldh-pb were mapped by sequencing the exons and portions of the adjacent introns, and comparing with the cDNA sequences (see below).

All of the intron/exon boundaries within both genes follow the GT/AG rule except for intron 9 of Aldh1, which has a GC/AG boundary sequence (Table 1a and 1b). For both genes, eight of the twelve codon interruptions of the exons are in phase O (between codons).

Cloning of full-length cDNAs encoding mouse Aldh1 and Aldh-pb

In order to confirm the predicted amino acid sequences and assignments of intron/exon boundaries, we performed PCR cloning of Aldh1 and ALDH-PB cDNAs using as the template either mouse testis cDNA library DNA or mouse liver total DNA.

Figure 2 Nucleotide and deduced amino acid sequences of mouse ALDH-PB cDNA and sequence alignment with mouse ALDH1

The ALDH-PB and ALDH1 nucleotide and deduced amino acid sequences obtained in this laboratory are shown at the top and the bottom of the alignment respectively. Nucleotide 1 is the A of the longest cloned ALDH-PB cDNA. Amino acid residue 1 was assigned to the serine at the second codon downstream of the translation initiator methionine codon. Indicated are nucleotide identities (dots) and deletions (dashes). Amino acid residues in ALDH1 that differ from those of ALDH-PB are shown in parentheses. A stop codon is indicated by an asterisk. The potential poly(A)+ signal (AATAAA) is indicated by a single underline.

Figure 3 Comparison of 5'-flanking region sequences of mouse Aldh-pb, mouse Aldh1, rat ALDH-PB and human ALDH1 genes

Nucleotide identities (dots) and deletions (dashes) are indicated. The Aldh-pb sequence is numbered as –1 from the immediate 5'-end of the initiation codon ATG. The known transcriptional initiation site for human ALDH1 is underlined [41]. Potential regulatory sequences are indicated, i.e. the CCAAT box, Oct1 and ATA (TATA box).
RNA. Figure 2 shows the full-length cDNAs for mouse ALDH-PB and ALDH1, both of which are just over 2.0 kb in length. Both cDNA sequences have consensus polyadenylation signals (AATAAA) within 20 bases of the polyadenylated [poly(A)]\(^\_\) tail. Our ALDH1 cDNA coding sequence is consistent with a previously published ALDH1 sequence [28], except at seven positions, namely nt 66, 67, 161, 182, 303, 701 and 1418, which result in amino acid discrepancies at residues 10, 86 and 457. In addition, there are three nucleotide discrepancies in the 3'-UTR of the two sequences. For both ALDH1 and ALDH-PB, the nucleotide sequences of the cDNAs isolated here were identical with their corresponding exonic sequences from genomic clones. Further analysis may reveal whether there is any significance in the nucleotide polymorphisms observed in the ALDH1 cDNAs described both previously and in the present study.

The percentages of positional identities between the amino acid sequences deduced from the mouse Aldh-pb, mouse Aldh1, rat ALDH-PB [38], rat ALDH1 [10,12,31] and human ALDH1 [39,40] genes are 83\% to 96\% (Table 2). Mouse ALDH1 shares 96\% of its amino acid sequence with rat ALDH1 and 87\% identity with human ALDH1, thus demonstrating that it is the mouse homologue for this gene. The novel ALDH-PB cDNA sequence shows a high percentage of positional identity (90–92\%) in the coding region and 66–73\% identity at the less-restricted 3'-UTR with mouse and rat ALDH1. The sizes of the 5'- and 3'-UTRs of the novel ALDH-PB transcripts are 44 and 510 nt respectively, similar to those of other class I ALDH transcripts. A very high level of overall amino acid sequence identity was noted between mouse ALDH-PB and rat ALDH-PB (92\%), as well as high sequence identity between these two proteins in small pockets (for example, residues 116–125 of ALDH-PB and rat ALDH-PB have an 8-out-of-10 match, but either of these proteins display only a 4-out-of-10 match with the homologous regions of mouse or human ALDH1, and a 5-out-of-10 match with rat ALDH1). From these findings, as well as additional similarities listed below, we conclude that the new gene identified in our cloning attempts, Aldh-pb, is the mouse homologue of rat ALDH-PB. A human homologue for ALDH-PB has not been described.

### 5'-flanking regions of the Aldh1 and Aldh-pb genes

The nucleotide sequences of the 5'-flanking regions (about 300 nt) for Aldh1 and Aldh-pb were aligned with those of rat ALDH-PB [38] and human ALDH1, for which the transcription-initiation site has been previously determined [41] (Figure 3). The 5'-flanking region of the Aldh-pb gene (between nt \(-156\) and \(-1\) relative to the initiation codon) is closely related to the other rodents genes, showing 81–90\% sequence identity, and has 66\% sequence identity with the human ALDH1 5'-flanking region. Between nt \(-157\) and \(-300\), Aldh-pb demonstrates essentially no sequence similarity with the other three genes (23–29\% identity), although sequence similarity does exist between the other three genes in this region (50–77\% identity). In each sequence a perfect match to the CCAAT box is conserved, located about 100–120 bp upstream of the initiation codon. Also, an octameric binding motif, Oct1 (ATGCAAAT), is conserved in all except rat ALDH-PB, and an ATA motif (TATA box) is well conserved among all except mouse Aldh-pb, which shows a 1-bp mismatch in a critical T residue.

**Aldh1 and Aldh-pb tissue-specific expression patterns**

Expression of Aldh1 and Aldh-pb was examined by Northern blot analysis of adult mouse tissues using gene-specific 3'-UTR fragments as probes (Figure 4). A 2.0 kb mRNA that varied in intensity among tissues was seen for each respective probe. A mouse \(\beta\)-actin cDNA probe, which identifies mRNAs of 2.0 and 1.8 kb, was used as a control to confirm the integrity of the mRNA samples (central panel).

![Figure 4](image-url)  
**Figure 4** Tissue expression patterns of mouse Aldh1 and Aldh-pb

Shown are Northern blot results from multiple adult mouse tissues with an ALDH1-specific probe (top panel) and an ALDH-PB-specific probe (bottom panel). Indicators of the molecular integrity of the mRNA samples (central panel).
Figure 5  Retinoic acid bioassay following injection of ALDH1 or ALDH-PB mRNAs into Xenopus embryos

Xenopus embryos at the 2-4 cell stage were injected with mRNA for either mouse ALDH1 (A) or mouse ALDH-PB (B) (46 nl of RNA at 0.2 ng/nl in both cases), or were uninjected (C). This was followed by incubation to blastula stage 8 and bioassay for retinoic acid by incubation as explants on top of the F9-RARE-\(\beta\)-galactosidase activity (Figures 5A, B). Embryos injected with mouse ALDH-PB displayed no detectable \(\beta\)-galactosidase activity (Figure 5B). Uninjected embryos at blastula stage 8 also displayed no detectable \(\beta\)-galactosidase activity with this bioassay, indicating that Xenopus embryos have no detectable endogenous retinoic acid at this stage (Figure 5C); endogenous retinoic acid was detected by stage 15 (results not shown). ALDH1 and ALDH-PB mRNAs were subjected to in vitro translation in a rabbit reticulocyte lysate, and both produced proteins of the correct molecular mass (approx. 55 kDa) with almost equal efficiency (Figure 5D). Overall, we found that injection of comparable amounts of ALDH1 and ALDH-PB mRNAs over a 5-fold concentration range resulted in retinoic acid detection in most ALDH1-injected embryos (35-82%), but in none of the ALDH-PB-injected embryos or uninjected embryos (Table 3).

DISCUSSION

In this study we have cloned and characterized the cDNAs and genes encoding mouse ALDH1 and ALDH-PB. Upon the basis of gene structure and sequence identity, ALDH1 and ALDH-PB are both class I ALDHs with a close evolutionary relationship. The coding regions of the two genes predict proteins sharing 90% amino acid sequence identity. The gene structures of both Aldh1 and Aldh-pb consist of 13 exons with 12 introns interrupting their coding regions at exactly the same positions as seen in the human cytosolic class I ALDH gene Aldh1 [41]. Many of the corresponding introns of the two mouse genes are very similar in size, also indicating a close relationship to each other. Although Aldh1 and Aldh-pb appear to be very closely related genes, significant differences in their expression pattern and function described in the present study provide them with unique identities.

Mouse Aldh1 and Aldh-pb were both expressed at high levels in liver, lung and testis, but Aldh-pb, unlike Aldh1, was also highly expressed in kidney and demonstrated a low level of expression in the brain. The expression pattern of mouse Aldh-pb thus matches that previously reported for rat ALDH-PB, except that the latter is expressed at a low level in liver [38]. Our observation of a high level of expression for both mouse genes in the liver, lung and testis, but no expression for either gene in the heart, spleen or muscle, is supportive of a role in epithelial function, since the former tissues, but not the latter, have large populations of epithelial cells.

A role for mouse ALDH1 in the oxidation of retinal to retinoic acid has previously been proposed by in vitro analysis of enzyme activity [8]. This is further supported by studies showing that human and rat ALDH1 homologues also catalyse retinoic acid synthesis in vitro [6,11,32]. However, in vitro enzyme assays may not reflect physiological conditions, making it unclear whether enzymes identified in vitro are playing a significant role in retinoic acid synthesis in vivo. We overexpressed mouse ALDH1 and ALDH-PB mRNAs in Xenopus embryos to provide an in vivo test of the role of these enzymes in retinoic acid synthesis. We found that endogenous retinoic acid was undetectable in Xenopus

Table 3  Retinoic acid bioassay

Following injection embryos were incubated until they reached stage 8 (blastula). At that time they were assayed for retinoic acid detection using the F9-RARE-\(\beta\)-galactosidase bioassay. The number of embryos \(n\) used for scoring is shown in parentheses. The bioassay is qualitative, and some variation in the intensity of \(\beta\)-galactosidase expression was noted in embryos injected with ALDH1 mRNA that showed a positive response. The lack of response in some ALDH1-injected embryos might be due to embryonic damage during injection. Retinoic acid could not be detected in embryos that were uninjected \((n = 45)\).

<table>
<thead>
<tr>
<th>Injection of mRNA (nl)</th>
<th>mRNA species added</th>
<th>Mouse ALDH1</th>
<th>Mouse ALDH-PB</th>
</tr>
</thead>
<tbody>
<tr>
<td>46</td>
<td></td>
<td>82 (17)</td>
<td>0 (49)</td>
</tr>
<tr>
<td>23</td>
<td></td>
<td>77 (61)</td>
<td>0 (60)</td>
</tr>
<tr>
<td>13.8</td>
<td></td>
<td>35 (23)</td>
<td>0 (25)</td>
</tr>
<tr>
<td>9.2</td>
<td></td>
<td>60 (45)</td>
<td>0 (26)</td>
</tr>
</tbody>
</table>
embryos at blastula stage 8, as also shown by others [42], providing a large time window during which embryos could be manipulated and examined for effects on retinoic acid synthesis with essentially no background detection. We found that injection of mRNA for mouse Aldh1, but not Aldh-ph, induced easily detectable retinoic acid synthesis in Xenopus embryos using a bioassay that monitors diffusion of retinoic acid from individual embryos to lacZ reporter cells. Thus we have provided evidence that ALDH1 can in fact catalyse retinoic acid synthesis in vivo using normal endogenous concentrations of substrate (retinal), coenzyme (NAD⁺) and other factors that might effect enzyme activity.

The retinoic acid bioassay described here is made possible by the fact that Xenopus eggs and embryos contain large amounts of the vitamin A substrate all-trans-retinal, the immediate precursor of the major bioactive retinoid all-trans-retinoic acid [43,44]. In contrast, all-trans-retinol, rather than all-trans-retinal, is the abundant form of vitamin A in mammalian embryos [45]. In addition, Xenopus embryos contain significant amounts of 4-oxo-retinal, the immediate precursor for the bioactive retinoid 4-oxo-retinoic acid [42]. It is not known if 4-oxo-retinal is a substrate for ALDH1, but the ability of this enzyme to utilize both all-trans-retinal and 9-cis-retinal [11,46] suggests that its active site might also accommodate 4-oxo-retinal, resulting in metabolism to 4-oxo-retinoic acid. The reporter cells detect the sum of all active carboxylated retinoids, including all-trans-retinoic acid, 4-oxo-retinoic acid, 9-cis-retinoic acid and didehydro-retinoic acid [47]. Thus it cannot be certain which bioactive retinoids are being produced by ALDH1 in this bioassay, but it is clear that conversion from an aldehyde form into a carboxylic acid form is essential. Our results also indicate that ALDH-PB is inactive in the production of all the carboxylated retinoids mentioned above, which collectively account for the vast majority of bioactive retinoids observed in vertebrate systems.

On the basis of our findings we conclude that mouse ALDH1 can participate in retinoic acid synthesis in vivo, but that the closely related ALDH-PB does not possess this property. This is likely to define a major difference in the physiological functions of these two ALDH genes. Another indication that ALDH1 functions in retinoic acid synthesis in vivo comes from our observation here of a lack of Aldh1 expression in the adult mouse kidney, combined with our previous report that adult mouse kidney has undetectable retinoic acid using the bioassay described herein [48]. This is further supported by studies in the rat, which has abundant levels of both retinoic acid and ALDH1 in the adult rat kidney [11,49]. Together, these findings suggest a direct link between expression of ALDH1 and production of retinoic acid in the adult kidney. The high level of Aldh-ph expression we observed in the adult mouse kidney, a tissue that lacks retinoic acid, also further strengthens the argument that this gene does not function in retinoic acid synthesis.

Our studies have provided additional information about the promoter regions of class I ALDH genes. A minimal promoter region for human ALDH1 (nt –133 to −1 in Figure 3) has been defined [50], and this is highly conserved with the rat ALDH-PB promoter [38], as well as with the mouse Aldh1 and Aldh-ph 5′-flanking regions described here. Therefore it is very likely that the promoters for the two mouse genes are located in the corresponding regions. Functional analysis of human ALDH1 in different cell lines [50] indicates that this proximal promoter region directs cell-type-specific expression, and contains two major cis-regulatory elements, a CCAAT box conserved in mouse Aldh1, mouse Aldh-ph and rat ALDH-PB, and an Oct1 motif conserved only in Aldh1 and Aldh-ph. It has been suggested that a 16-bp deletion in the rat ALDH-PB promoter, which destroys the Oct1 motif, might account for this gene being expressed at low levels in rat liver [51]. In support of this hypothesis, we have shown that mouse Aldh-ph has conserved the Oct1 motif, and demonstrates high expression in the liver. Since deletion of the CCAAT box in human ALDH1 results in a significant decrease in promoter activity [50], it is likely that the CCAAT box plays a significant role in the transcriptional control of all these genes. Unlike the other promoters described above, which have a putative ATA motif (TATA box) located 20–40 bp downstream from the CCAAT box, the Aldh-ph promoter has a T + G base change at this location, suggesting that the ATA motif might not be a primary regulatory element. This is consistent with studies showing that deletion of the ATA box did not significantly decrease human ALDH1 promoter activity [50].

In total, 23 conserved amino acid residues have been identified from sequence alignment of multiple ALDH isoenzymes of different classes and phylogenetic origins [52]. The three-dimensional structures of rat cytosolic ALDH3 [53] and human mitochondrial ALDH2 [54] in complex with NAD⁺ have shown that the 23 strictly conserved residues are located in the catalytic pocket (Cys²⁹⁸, Gly⁴⁸⁷ and Phe⁴⁹⁷), at turns in the structure (Gly¹⁴⁸, Gly¹⁵⁶, Gly²²₃, Gly²³₁, Gly²⁷₉, Gly²⁹⁹, Gly³⁰⁷, Gly³⁷₂, Gly⁴⁸⁵, Gly⁴⁸⁶, Gly⁴⁸⁷, Gly⁴⁹³, Pro⁵¹₅, Pro⁵¹₆ and Pro⁵⁰₆), or at locations that stabilize adjacent elements of secondary structure (Arg⁶⁸, Lys¹⁹₂, Thr¹⁶₄, Asn¹⁶₁, Asn¹⁶₄ and Ser¹⁷¹). Therefore it was proposed that other classes of ALDH might have similar structural folds, like ALDH2 and ALDH3. Our study shows that these 23 strictly conserved residues are found in both ALDH1 and ALDH-PB. It is thus very likely that these two mouse enzymes, which share 90 % amino acid sequence identity, exhibit similar folding structures and substrate-binding pockets. However, it has been suggested that different exon/intron regions of closely related ALDH genes are under different evolutionary pressures; i.e. human ALD7H and ALDH8 [55]. We now describe a similar observation between mouse Aldh1 and Aldh-ph, in which exons 4 and 12 show 74–80 % positional identity, which is lower than the overall sequence identity of 90 %. The ALDH three-dimensional structures indicate that these two exons encode sequences corresponding to an α-helix located on the surface of the molecule and a turn in the structure respectively. It is possible that these two exons encode protein domains or folds related to the specific functions of individual isoenzymes. Perhaps one or more amino acid differences at these locations can explain our observation that ALDH1, but not ALDH-PB, functions in retinoic acid synthesis.

In summary, our studies define two closely related mouse ALDH genes called Aldh1 and Aldh-ph, demonstrate that mouse ALDH-PB and rat ALDH-PB appear to form a separate branch of class I ALDH that is distinct from human and rodent ALDH1, and show that ALDH1, but not ALDH-PB, functions in retinoic acid synthesis in vivo. We are grateful to M. Wagner for the F9-RARE-lacZ reporter cell line, and to A. Yoshida for critical discussions. G. D. acknowledges support from NIH Grant AA09731.

REFERENCES

Molecular analysis of two related mouse aldehyde dehydrogenase genes


© 1999 Biochemical Society

Received 3 December 1998/15 January 1999; accepted 4 February 1999