THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 49, Issue of December 7, pp. 46347–46363, 2001
© 2001 by The American Society for Biochemistry and Molecular Biology, Inc.
Printed in U.S.A.

Purification of the Aldehyde Oxidase Homolog 1 (AOH1) Protein and Cloning of the *AOH1* and Aldehyde Oxidase Homolog 2 (*AOH2*) Genes

IDENTIFICATION OF A NOVEL MOLYBDO-FLAVOPROTEIN GENE CLUSTER ON MOUSE CHROMOSOME 1*

Received for publication, June 21, 2001, and in revised form, September 11, 2001
Published, JBC Papers in Press, September 18, 2001, DOI 10.1074/jbc.M105744200

Mineko Terao‡, Mami Kurosaki‡, Massimiliano Marini‡, Maria Antonietta Vanoni§, Giuliana Saltini‡, Valentina Bonetto¶, Antonio Bastone¶, Concetta Federico‖, Salvatore Saccone, Roberto Fanelli‡‡, Mario Salmona¶, and Enrico Garattini‡ §§**

*From the ‡Laboratory of Molecular Biology, Centro Catullo e Daniela Borgomainerio, the ¶Department of Molecular Biochemistry and Pharmacology, and the ‡‡Department of Environmental Toxicology, Istituto di Ricerche Farmacologiche "Mario Negri," via Eritrea, 62, Milano 20157, the §Dipartimento di Scienze Chimiche, Fisiche e Matematiche, Università dell'Insubria, via Valleggio 11, Como 22100, the ‖Dipartimento di Biologia Animale, University of Catania, via Androne, 81, Catania 95124, and the **Dipartimento di Protezione Valorizzazione Agroalimentare, University of Bologna, via F lli. Rosselli 107, Reggio Emilia 42100, Italy*

We report the cloning of the *AOH1* and *AOH2* genes, which encode two novel mammalian molybdo-flavoproteins. We have purified the AOH1 protein to homogeneity in its catalytically active form from mouse liver. Twenty tryptic peptides, identified or directly sequenced by mass spectrometry, confirm the primary structure of the polypeptide deduced from the *AOH1* gene. The enzyme contains one molecule of FAD, one atom of molybdenum, and four atoms of iron per subunit and shows spectroscopic features similar to those of the prototypic molybdo-flavoprotein xanthine oxidoreductase. The *AOH1* and *AOH2* genes are 98 and 60 kilobases long, respectively, and consist of 35 coding exons. The *AOH1* gene has the potential to transcribe an extra leader non-coding exon, which is located downstream of exon 26, and is transcribed in the opposite orientation relative to all the other exons. *AOH1* and *AOH2* map to chromosome 1 in close proximity to each other and to the aldehyde oxidase gene, forming a molybdo-flavoenzyme gene cluster. Conservation in the position of exon/intron junctions among the mouse *AOH1*, *AOH2*, aldehyde oxidase, and xanthine oxidoreductase loci indicates that these genes are derived from the duplication of an ancestral precursor.

Mammalian molybdo-iron–sulfur-flavoproteins (henceforth referred as molybdo-flavoproteins for simplicity) are widely distributed enzymes requiring a molybdo-pterin and a flavin cofactor for their catalytic activity (1–6). In mammals, this family was originally thought to consist of two members, xanthine oxidoreductase (XOR)[1] and aldehyde oxidase (AO) (1). Although XOR is the key enzyme in the catabolism of purines, oxidizing hypoxanthine to xanthine and xanthine to uric acid (7, 8), the physiological function of AO is still unknown, although the enzyme is involved in the metabolism of drugs and xenobiotics of toxicological importance (9, 10). AO and XOR have similar primary and secondary structure (1) and utilize an overlapping set of substrates (11). Recently, we suggested that the family of mammalian molybdo-flavoproteins extends beyond XOR and AO and includes at least two other members, which we provisionally named AOH1 and AOH2 (Aldehyde Oxidase Homologs 1 and 2) (12). The cDNAs coding for AOH1 and AOH2 were identified and cloned because of their remarkable nucleotide sequence similarity to both AO and XOR (12). Like AO and XOR, AOH1 and AOH2 are characterized by the presence of two highly conserved domains encoding non-identical 2Fe-2S redox clusters and the fingerprint sequence found in all molybdo-proteins (1). In addition, the length of the predicted translation products of the AOH1 and AOH2 cDNAs is similar to that of AO and XOR. The two cDNAs code for predicted polypeptides that are structurally more related to AO than to XOR. This led us to propose that the products of the AOH1 and AOH2 cDNAs represent isoenzymatic variants of aldehyde oxidase acting on a non-identical but overlapping set of substrates. AOH1 is expressed in the same tissues and cell types as AO and is synthesized predominantly in the liver. Although AO is expressed only in adult liver, AOH1 is also present in embryonic and neonatal liver. The distribution of AOH2 is strictly limited to keratinized epithelia in the oral cavity, esophagus, stomach, and skin (12).

In this article, we provide a first characterization of the native AOH1 protein, which was purified to homogeneity from mouse liver. Mass spectrometric analysis of purified AOH1 confirms the amino acid sequence deduced from the corresponding gene and cDNA. The enzyme has spectroscopic char-

* This work was supported by grants from Telethon (to M. T.), the Consiglio Nazionale delle Ricerche (Progetto Finalizzato Biotecnologie), the Associazione Italiana per la Ricerca contro il Cancro, and the EEC Training Network Program. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF322144–AF322178 (AOH1 gene); AF321780–AF321814 (*AOH2* gene).

§§ To whom correspondence should be addressed: Tel.: 39-02-3901-4533; Fax: 39-02-354-6277; E-mail: egarattini@marionegri.it.

[1] The abbreviations used are: XOR, xanthine oxidoreductase; AO, aldehyde oxidase; MALDI-MS, matrix-assisted laser desorption-time of flight mass spectrometry; ESI-MS/MS, electrospray ionization tandem mass spectrometry; 5FR1/1, 5FR1/1bis and 5FR2, 5′-flanking regions of *AOH1* exon 1 and exon 1bis, and *AOH2* exon 1, respectively; DAPI, 4′-6-diamidine-2-phenylindole; PAGE, polyacrylamide gel electrophoresis; CID, collision-induced dissociation; kb, kilobase(s); bp, base pair(s); PCR, polymerase chain reaction; RACE, rapid amplification of cDNA ends; AP1, -2, anchor primers 1 and 2; SP1, -2, specific primers 1 and 2; NP1, -2, nested primers 1 and 2.

acteristics similar to those of other molybdo-flavoproteins, contains FAD as the cofactor, and oxidizes benzaldehyde and phthalazine, two substrates of aldehyde oxidase (13, 14). In addition, we report the molecular cloning, structural characterization, and chromosomal mapping of the mouse AOH1 and AOH2 genetic loci.

EXPERIMENTAL PROCEDURES

Purification of Mouse Liver AOH1 Protein, Electrophoresis, and Western Blot Analysis—Unless otherwise stated, all the purification steps were carried out at 4 °C. Male mouse livers were homogenized in 3 volumes of 100 mM sodium phosphate buffer, pH 7.5, with an Ultraturrax homogenizer (Omni 2000, Omni International, Waterbury, CT). Homogenates were centrifuged at 100,000 × *g* for 45 min to obtain cytosolic extracts. Extracts were heated at 55 °C for 10 min and centrifuged at 15,000 × *g* to remove precipitated proteins. An equal volume of saturated ammonium sulfate was added to the supernatant, and the precipitate was collected by centrifugation at 15,000 × *g* and resuspended in 100 mM Tris-glycine buffer, pH 9.0. Solubilized proteins (equivalent to 10 g of fresh liver) were mixed with 5 ml of benzamidine-Sepharose (Amersham Pharmacia Biotech, Uppsala, Sweden) pre-equilibrated in 100 mM Tris-glycine buffer, pH 9.0. Following 2 h of incubation, the resin was washed four times with 10 ml each of the equilibration buffer to remove unbound proteins. Adsorbed proteins were eluted twice with 5-ml aliquots of equilibration buffer containing 5 mM benzamidine (Sigma Chemical Co., St Louis, MO). The eluate was concentrated to ~1 ml with Centriplus YM-100 (Millipore Corp., Bedford, MA) and diluted to 10 ml with 50 mM Tris-HCl, pH 7.4. The solution was applied to a 5/5 FPLC Mono Q column (Amersham Pharmacia Biotech) equilibrated in 50 mM Tris-HCl, pH 7.4. AOH1 protein was eluted at 0.5 ml/min with a linear gradient (30 ml) from 0 to 1 M NaCl in 50 mM Tris-HCl, pH 7.4. The purification of AOH1 was monitored by determination of phthalazine oxidizing activity, as described below, or by quantitative Western blot analysis (12). In the case of phthalazine oxidation, one unit of enzymatic activity corresponds to 1 nmol of phthalazine oxidized/min.

A specific anti-AOH1 rabbit polyclonal antibody raised against a synthetic peptide of the protein (12) was used for Western blot analysis, which was carried out with a chemiluminescence-based protocol, as already described (12). For quantitative Western blot analysis, an equivalent volume (10 μl) of protein solution, at each purification step, was loaded onto the same gel and processed for analysis. Chemiluminescence signals corresponding to AOH1 bands were quantitated with a scanning densitometer (Hoefer Scientific Instruments, San Francisco, CA). The total amount of AOH1-immunoreactive protein in the various experimental samples is expressed in arbitrary units and is calculated on the basis of the intensity of the Western blot signal in optical density (OD) multiplied by the total volume of each purification step. One arbitrary unit of immunoreactive protein corresponds to 1.0 OD (optical density) of the specific AOH1 band in each experimental sample.

Zymographic analysis of benzaldehyde oxidizing activity was performed following electrophoresis on cellulose acetate plates, as already described (12). SDS-PAGE was performed according to standard techniques (15). Proteins were measured according to the Bradford method with a commercially available kit (Bio-Rad, Richmond, VA).

Mass Spectrometry of the Purified AOH1 Protein—The mass of purified AOH1 was determined by MALDI-TOF mass spectrometry (matrix-assisted laser desorption ionization-time of flight; MALDI-MS). An aliquot of pure AOH1 in 5 mM Tris-HCl, pH 7.0, was mixed with the matrix (sinapinic acid in 50% acetonitrile/0.1% trifluoroacetic acid) and analyzed by MALDI-MS. MALDI-MS and Electrospray Ionization (ESI-MS) tandem mass spectrometric analyses of AOH1 tryptic peptides were performed according to standard protocols following *in situ* or *in gel* tryptic digestion (16, 17). Briefly, proteins or Coomassie Blue-stained gel slices were incubated with 10 mM dithiothreitol in 100 mM ammonium bicarbonate at 56 °C for 30 min to reduce disulfide bridges. Thiol groups were alkylated upon reaction with 55 mM iodoacetamide in 100 mM ammonium bicarbonate at room temperature in the dark for 20 min. Tryptic digestion was carried out overnight at 37 °C in 50 mM ammonium bicarbonate and 12.5 ng/μl trypsin (Promega, Madison, WI). Peptides were extracted twice in 50% acetonitrile/5% formic acid. The combined extracts were lyophilized and re-dissolved in 0.5% formic acid and desalted using ZipTip (Millipore). Peptides were eluted in 50% methanol/0.5% formic acid. The eluate was mixed 1:1 (v/v) with a saturated matrix solution of α-cyano-4-hydroxycinnamic acid in acetonitrile/0.1% trifluoroacetic acid 1:3 (v/v).

Mass mapping of tryptic peptides was performed with a Bruker Biflex MALDI-TOF mass spectrometer (Bruker, Bremen, Germany). Data generated were processed with the Mascot program (www.matrixscience.com) (18) allowing a mass tolerance of ≤0.4 Da. Direct sequence analysis was carried out via collision-induced dissociation (CID) on an electrospray mass spectrometer API 3000 (Applied Biosystems, San Diego, CA). The program MS-Tag (prospector.ucsf.edu) (19) was used to correlate the experimental CID spectra to the theoretical CID spectra of tryptic peptides derived from proteins present in data bases.

Metal Content of Purified AOH1 Protein—Following purification of AOH1 to homogeneity, the molybdenum content was determined on an inductive-coupled plasma mass spectrometer (model ELAN 5000, PerkinElmer Sciex, Foster City, CA) equipped with a Coolflow CFT-75 NESLAB device and a Cross-Flow nebulizer. The iron content was determined in standard conditions (20) on a PerkinElmer Zeeman 30/30 atomic absorption spectrophotometer equipped with a graphite furnace.

Absorbance Spectra, Identification of AOH1 Flavin Cofactor, and Steady-state Kinetic Measurements of Phthalazine-oxidizing Activity—The absorbance spectra of AOH1 in its native state and of the cofactors released following heat denaturation of the protein (100 °C for 10 min) were recorded at 20 °C with a Hewlett-Packard HP8453 diode array spectrophotometer interfaced to a Vectra XA personal computer (Hewlett-Packard, Palo Alto, CA).

Identification of the flavin cofactor bound to AOH1 was carried out fluorometrically as described previously (21). Briefly, aliquots of AOH1 preparations (~2–2.5 μM) were incubated at 100 °C for 10 min. After removal of the denatured protein by centrifugation, the intensity of light emitted at 524 nm on excitation with light at 450 nm was recorded before and after addition of snake venom phosphodiesterase (2 μl, 6 milliunits, Roche Molecular Biochemicals, Mannheim, Germany). An increase in fluorescence emission of ~10 is expected for a homogeneous solution of FAD being converted into FMN + AMP by phosphodiesterase. Fluorescence measurements were performed in a Jasco FP777 (OmniLab Ltd. & OmniLab Biosystems Ltd., Mettmenstetten, Switzerland) thermostatted at 20 °C.

The initial reaction velocity of AOH1-dependent phthalazine oxidizing activity was measured at 25 °C in 50 mM Hepes/KOH buffer, pH 7.5, containing 1 mM ferricyanide, 0.04 μM AOH1, and variable concentration of phthalazine (1 μM to 1 mM). Ferricyanide reduction was monitored at 420 nm using a Cary219 spectrophotometer (Varian, Palo Alto, CA). Activities are expressed as apparent turnover numbers (*i.e.* micromoles of phthalazine oxidized per second per micromole of AOH1) taking into account that 2 mol of ferricyanide (extinction coefficient of 1.04 mM^{-1} cm^{-1}) are reduced per mole of oxidized phthalazine. The AOH1 concentration is calculated considering a subunit mass of 150,000, following determination of the protein concentration with the Bradford method. Estimates of the apparent maximum velocity and K_m values were obtained by fitting the initial velocity to the Michaelis-Menten equation (Equation 1) as a function of phthalazine concentration, after visual inspection of double-reciprocal plots (Equation 2) allowed us to detect a phthalazine concentration range that did not give substrate inhibition,

$$v = (V \times S)/K + S) \quad \text{(Eq. 1)}$$

$$1/v = 1/V + (K/V) \times 1/S \quad \text{(Eq. 2)}$$

where v is the initial velocity measured at a given substrate concentration (S); V is the apparent maximum velocity extrapolated at infinite substrate concentration, and K is the apparent Michaelis constant (22). The Grafit version 4.0 (Erythacus Software Ltd., Staines, UK) was used for data analysis.

Chromosome Preparation and in Situ Hybridization—Mouse metaphase chromosomes were prepared as previously described (23). *In situ* hybridizations were performed using the following biotin- or digoxigenin-labeled probes: the ~3-kb AOH1 cDNA fragment contained in AOH1–24/1 (12); the ~2-kb AOH2 cDNA fragment contained in AA79991 (12); the *AOH1*-containing PAC 2 DNA; the *AOH2* containing PAC3 DNA; and the AO containing lambda phage DNA 9m (12).

Following labeling by nick translation, the probes were ethanol-precipitated with a 50-fold excess of yeast tRNA and salmon sperm DNA (carrier DNAs), and a 50-fold excess of mouse Cot1 (competitor DNA). The nucleic acid mixtures were resuspended in 50% formamide, 2× SSC, 10% dextran sulfate, 50 mM sodium phosphate, pH 7.0, to have a probe concentration of 5–20 ng/μl. A pre-annealing step was performed by incubating the hybridization mixtures at 37 °C for 30–60 min. Hybridization, detection with fluoresceinated avidin (Vector Laboratories Inc., Burlingame, CA), and chromosome G-banding were performed as described previously (23). In the case of co-hybridization of

TABLE I
Purification scheme of mouse liver AOH1

Ten grame of mouse liver was isolated, homogenized in 30 ml of buffer, and ultracentrifuged to obtain a cytosolic extract (*Extract*), which was processed as indicated. Enzymatic activity was measured as the ability of the various purification fractions to oxidize phthalazine and contemporaneously reduce potassium ferricyanide. The reduction of potassium ferricyanide was monitored at 420 nm. One unit of enzymatic activity is defined as 1 nmol of phthalazine oxidized/min. The amount of AOH1 immunoreactive protein was determined by quantitative Western Blot and is defined in relative units (see "Experimental Procedures"). The results are representative of six separate AOH1 preparations.

Step	Volume	Protein	Phthalazine-oxidizing activity				AOH1 immunoreactive protein			
			Total	Specific activity	Purification factor	Yield	Total	Specific activity	Purification factor	Yield
	ml	*mg*	*units*	*units/mg*	*-fold*	*%*	*units*	*units/mg*	*-fold*	*%*
Extract	25	1041	6250	6.3	1.0	100	1090	1.0	1.0	100
55 °C	20	701	6360	9.1	1.4	102	941	1.3	1.3	86
$(NH_4)_2SO_4$	20	420	4090	9.7	1.5	65	812	1.9	1.9	74
Benzamidine-Sepharose	1	0.54	520	963.0	152.8	8.3	39	72.2	72	3.6
Mono Q	0.5	0.11	110	1000	158.7	1.8	19	172.7	173	1.7

two different probes labeled with biotin (Lambda phage DNA 9m) and digoxigenin (PAC2 or PAC3 DNA), detection was performed using rhodamine-conjugated avidin (Vector Laboratories Inc.) and fluorescein conjugated anti-digoxigenin antibodies (Intergen Company, Purchase, NY).

Isolation and Characterization of Mouse AOH1 and AOH2 Genomic Clones—A genomic library from Sv/129 mice arrayed on nylon filters at high density was obtained from the Human Genome Mapping Program (HGMP, Oxford, UK). The library was sequentially screened with ^{32}P-labeled full-length mouse AOH1 and AOH2 cDNAs (12). Hybridization conditions were as indicated by the filter manufacturer. This resulted in the isolation of 14 AOH1 and 5 AOH2 hybridizing clones. Three overlapping clones, PAC1 to PAC3, were further characterized according to their hybridization profile with AOH1, AOH2, and AO cDNA probes. The distance between adjacent exons was established by long-range PCR analysis on either PAC DNA preparations or genomic DNA fragments, using specific couples of amplimers.

DNA Sequencing and Southern Blot Analysis on Genomic DNA—Appropriate DNA fragments were subcloned into the pBluescript plasmid vector (Stratagene, La Jolla, CA) and sequenced. Alternatively, DNA fragments containing one or more exons and the corresponding introns were PCR-amplified and subcloned as above or directly sequenced, according to the Sanger dideoxy chain termination method using double-stranded DNA as template and T7 DNA polymerase (Amersham Pharmacia Biotech, Uppsala, Sweden) or Sequenase (Upstate Biotechnology, Cleveland, OH). Exons and exon/intron junctions were sequenced in both directions. Oligodeoxynucleotide primers were custom-synthesized by M-Medical srl (Florence, Italy). Computer analysis of the DNA sequences was performed using the GeneWorks sequence analysis system (IntelliGenetics, San Diego, CA). A search of potential binding sites for transcription factors in the 5′-flanking region of the *AOH1* and *AOH2* genes was performed using the MatInspector algorithm and the TRANSFAC data base (24).

Southern blot analysis was performed according to standard procedures (15) on DNA extracted from PAC clones or high molecular weight genomic DNA derived from various animal species. The probes used were fragments of the mouse AO, AOH1, and AOH2 cDNAs.

Determination of the 5′-End of the Mouse AOH1 and AOH2 Transcripts and Functional Characterization of the Promoters—Total RNA was extracted from mouse liver and skin, and the poly(A^+) fraction of the RNA was selected according to standard protocols (15). 5′-RACE (rapid amplification of cDNA ends) was performed with the commercially available Marathon cDNA amplification kit (CLONTECH, Palo Alto, CA), according to the nested PCR protocol included and using the following amplimers: SP1, 5′-ACTGATCCTTTTGGAGATGGGG-3′ (complementary to nucleotides 388–409 of the AOH1 cDNA); NP1, 5′-CCTAGCAAAGGGCGTATATGATGATA-3′ (complementary to nucleotides 51–76 of the AOH1 cDNA); NP2, 5′-TAAGCTCATCTGACTCCTTAGAAGGA-3′ (complementary to nucleotides 205–230 of the AOH1 cDNA); SP2, 5′-TTGATGCTCCCCACACCTTCCA-3′ (complementary to nucleotides 380–401 of the AOH2 cDNA); NP3, 5′-GGTGTCAATGGGGAGCCAGAAGCACAAATT-3′ (complementary to nucleotides 76–105 of the AOH2 cDNA); 5′-CCTGGTGTAGAACAGTAGATTC-3′ (complementary to nucleotides 189–210 of the AOH2 cDNA). PCR products were subcloned in pBluescript, and multiple clones were sequenced to determine the 5′-end of the AOH1 and AOH2 transcripts.

HEK-293 cells obtained from the American Type Culture Collection (ATCC, Rockville, MD) were grown in Dulbecco's modified Eagle's medium containing 10% fetal calf serum (Life Technologies, Inc., Grand Island, NY). To determine the promoter activity of the 5′-flanking regions of exons 1bis and 1 of the *AOH1* and exon 1 of the *AOH2* genes, we constructed the plasmids: *p5FR1/1bis-Luc*, *p5FR1/1-Luc*, and *p5FR2-Luc*. To obtain *p5FR1/1bis-Luc*, a 0.9-kb *Xba*I-*Pst*I fragment, containing 583 bp of the 5′-flanking region *of AOH1* exon 1bis, exon 1bis and 221 bp of the corresponding intron, was subcloned in pBluescript (*p5FR1/1bis-blue*). The insert was released from p5FR1bis-blue by *Kpn*I-*Xho*I cleavage and inserted into the pGL-Basic Vector (CLONTECH) pre-digested with the same endonucleases. To construct p5FR1-Luc, a 1.7-kb *Eco*RI-*Hin*cII fragment, containing 1.2 kb of the 5′-flanking region *of AOH1* exon 1, exon 1 and 483 bp of the corresponding intron, was subcloned in pBluescript (*p5FR1/1-blue*). The plasmid was digested with *Eco*RI, blunted, and digested with *Xho*I. The resulting fragment was inserted into the pGL-Basic Vector pre-digested with *Nhe*I and *Xho*I. To prepare *p5FR2-Luc*, a 1.4-kb *Xba*I-*Apa*I fragment, containing 695 bp of the 5′-flanking region of *AOH2* exon 1, exon 1 and 588 bp of the corresponding intron, was subcloned in pBluescript (*p5FR2-blue*). The resulting plasmid was digested with *Not*I-*Apa*I, blunted, and inserted into the pGL-Basic Vector, which was pre-digested with *Xho*I and blunted. The constructs were transfected in HEK-293 cells using cationic liposomes as already described (25). Forty-eight hours following transfection, cells were harvested and lysed, and cell extracts were used for the determination of luciferase activity using an already described protocol (25).

RESULTS

Purification and Structural Characterization of the AOH1 Protein from Mouse Liver—To determine the structure and the biochemical characteristics of the novel putative molybdo-flavoenzyme, AOH1 (12), the protein was purified to homogeneity from mouse liver. Purification of the protein was followed either by determining the level of phthalazine-oxidizing activity or by quantitative Western blot analysis with mono-specific anti-AOH1 polyclonal antibodies (12). Typical results obtained from one of six AOH1 purification experiments are shown in Table I. The procedure yields ~2% of AOH1 and results in more than 150-fold enrichment of the protein relative to liver cytosolic extracts. Interestingly, similar final yields and purification factors are calculated from the immunological data described under "Experimental Procedures." This indicates that almost all mouse liver phthalazine-oxidizing activity can be accounted for by AOH1 and suggests that the majority of the enzyme is in its catalytically active form.

Several criteria indicate that our purification procedure results in the isolation of AOH1 in a pure form. The Mono Q chromatogram shown in Fig. 1*A* demonstrates that all the phthalazine-oxidizing activity elutes in correspondence with anti-AOH1 immunoreactivity (*WB inset*) and a single peak of proteins (see *Coomassie inset*). As documented in Fig. 1*B* (*left panel*), PAGE analysis of the various purification steps demonstrates progressive enrichment of a band of ~150 kDa. Purified AOH1 protein is devoid of AO-contaminating activity (*right panel* of Fig. 1*B*). In fact, upon cellulose acetate electrophoresis of the Mono Q-active fractions, only the benzaldehyde-oxidizing band α (corresponding to AOH1) (12) is visible, whereas the presence of band β (corresponding to AO) (12) is evident

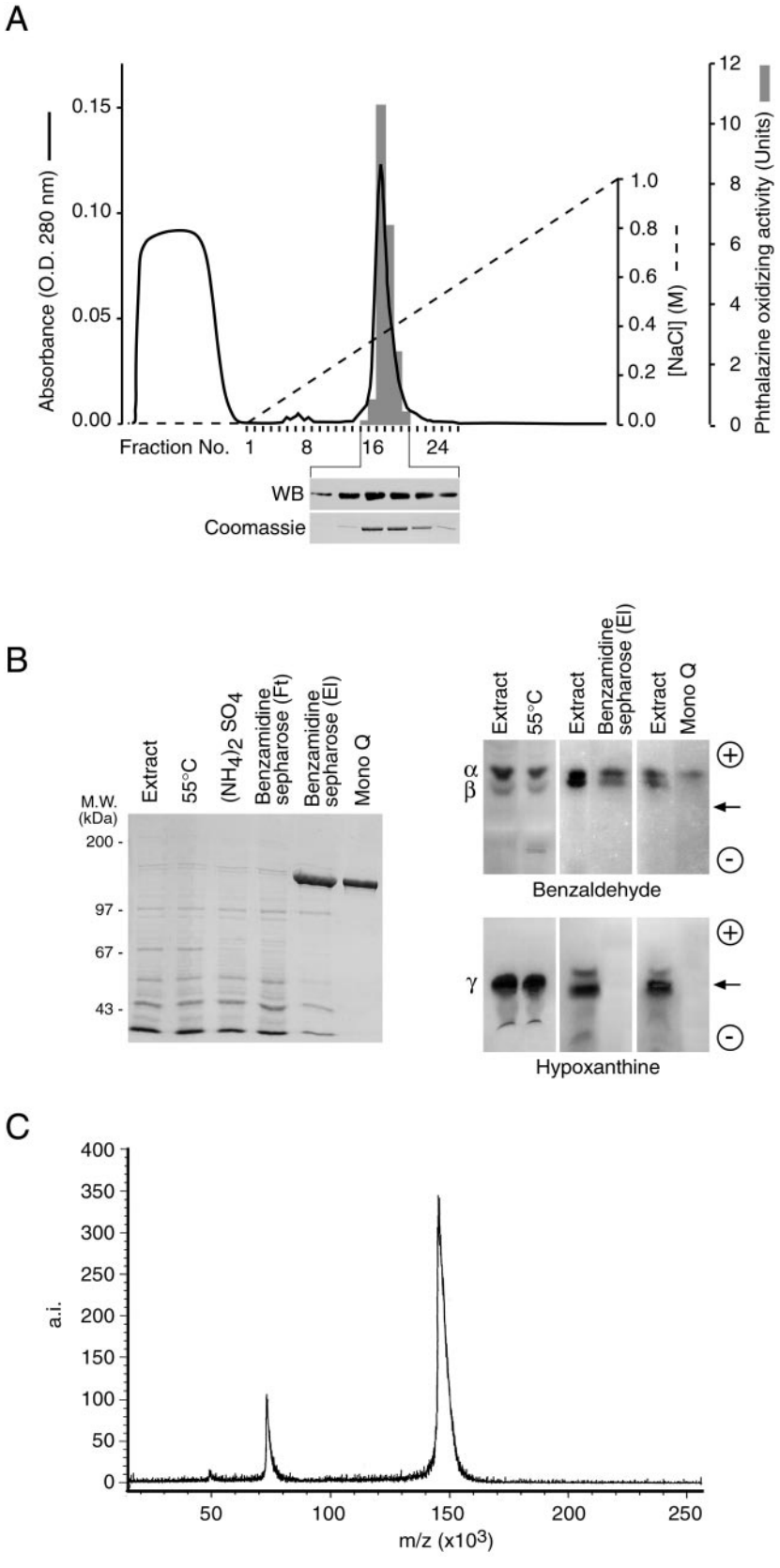

FIG. 1. **Purification and characterization of mouse liver AOH1.** *A*, Mono Q ion exchange chromatography of mouse liver AOH1. The protein fraction (500 μg) purified on benzamidine-Sepharose was applied on a Mono Q fast-protein liquid chromatography column and eluted with a linear gradient of NaCl as indicated (*dashed line*). The column was run at a flow rate of 0.5 ml/min, and 2-min fractions were collected. An aliquot of each chromatographic fraction was assayed for the presence of phthalazine oxidizing activity, and the corresponding elution profile is indicated by the *shaded columns* superimposed over the protein elution profile monitored at 280 nm. The *inset* shows the results of a Western blot performed on aliquots of the indicated chromatographic fractions using a specific anti-AOH1 polyclonal antiserum along with the corresponding Coomassie Blue staining. *B*, Coomassie Blue protein staining, following PAGE analysis (*left panel*) or zymograms of cellulose acetate electrophoresis (*right panel*), of the indicated protein preparations obtained at various steps in the purification procedure of mouse liver AOH1 are shown. Zymograms were developed in the presence of phenazine and nitro blue tetrazolium using benzaldehyde or hypoxanthine as substrates for the determination of benzaldehyde oxidase and xanthine oxidoreductase activity, respectively. *Extract* = cytosolic extract of mouse liver; *55 °C* = cytosolic extract following treatment for 10 min at 55 °C; *$(NH_4)_2SO_4$* = ammonium sulfate cut of heated cytosolic extracts; *Ft* = flow-through of benzamidine-Sepharose chromatography; *El* = pooled eluate of benzamidine-Sepharose chromatography; *Mono Q* = pooled fractions containing phthalazine-oxidizing activity from Mono Q chromatography. *C*, MALDI-MS analysis of purified AOH1 protein eluting from the Mono Q anion exchange chromatography. *m/z* = mass/charge; *a.i.* = absolute intensity.

throughout all the other purification steps. Similarly, the AOH1-containing benzamidine Sepharose and Mono Q fractions are free of XOR contamination, as demonstrated by the lack of hypoxanthine-oxidizing activity. As documented in Fig. 1*C*, MALDI-MS analysis of pure AOH1 demonstrates the presence of a singly charged molecular ion peak of 147,249 ± 129 Da (mean ± S.D.; $n = 3$) and the corresponding doubly charged counterpart. The experimentally determined mass is in excellent agreement with that calculated from the deduced amino acid sequence of the deflavo- and demolybdo-monomeric subunit encoded by the *AOH1* gene (147,379 Da).

Trypsinization of AOH1 results in the generation of numer-

TABLE II

Identification of AOH1 tryptic peptides by MALDI-MS and sequencing of selected peptides by ESI-MS/MS

AOH1 protein was trypsinized and subjected to MALDI-MS analysis. *MALDI-MS*: The analysis was limited to tryptic peptides with molecular mass equal or superior to 800 Da. The data in the *first column* are the mass values obtained experimentally ($M + H^+$ *obs*), whereas the results in the *second column* are those calculated from the tryptic fragmentation of the *AOH1* gene product ($M + H^+$ *cal*). The *third column* indicates the number of the first and last amino acid of the identified AOH1 peptides (numbering as in Fig. 8), whereas the *fourth column* shows the corresponding amino acid sequences. *ESI-MS/MS*: The *first column* indicates the mass/charge (m/z) values of the peptides whose sequences were determined. The *second column* shows the charge state of each peptide, and the *third column* indicates the determined sequence along with the number of the first and last residue in *parentheses*.

MALDI-MS $M + H^+$ obs	$M + H^+$ cal	Identification	Sequence
3376.1	3376.6	958–985	TIHNQEFDPTNLLQCWEACVENSSYYNR
2430.4	2429.8	205–224	EFQPLDPTQELIFPPELMR
2301.7	2301.7	206–224	KEFQPLDPTQELIFPPELMR
1949.0	1949.3	259–276	HPSAPLVIGNTYLGLHMK
1896.3	1896.2	1291–1308	GLSPIWAINSPATAEVIR
1869.5	1869.0	225–240	MAEESQNTVLTFRGER
1579.4	1579.8	277–290	FTDVSYPIIISPAR
1568.2	1568.7	988–999	AVDEFNQQRFWK
1564.4	1564.8	23–35	ESDELIFFVNGK
1526.3	1526.7	225–237	MAEESQNTVLTFR
1461.2	1461.6	1112–1123	QNPSGTWEEWVK
1358.3	1358.6	824–834	TGRPIRFILER
1344.0	1344.5	419–429	SSKWEFVSAFR
1223.8	1224.4	351–363	NVASLGGHIISR
1184.7	1185.3	1218–1227	YSPEGVLYTR
1120.9	1121.2	908–917	TNLPSNTAFR
1106.8	1107.2	988–996	AVDEFNQQR
1041.7	1042.2	422–429	WEFVSAFR
887.0	887.0	343–350	TLAGQQIR
825.66	825.9	948–953	ELNMYR
805.0	805.0	1051–1057	MIQVAASR

ESI-MS/MS m/z	Charge state	Sequence
1151.4	2+	EFQPLDPTQELIFPPELMR (206–224)
948.0	2+	GLSPIWAINSPATAEVIR (1291–1308)
893.2	2+	TTIAGTLNDLLELK (241–256)
612.6	2+	NVASLGGHIISR (351–362)
528.4	3+	ASKPGLLASVAAVAAQK (807–823)

ous peptides whose molecular masses were determined by MALDI-MS analysis. Five peptides were sequenced by ESI-MS/MS. A summary of these results is presented in Table II. Our data confirm the presence of 20 distinct peptides and establish the sequence of 283 amino acid residues, representing ~21% of the entire primary structure of the AOH1 protein. Interestingly, three independent AOH1 preparations ruled out the presence of molecular masses corresponding to peptides expected from the tryptic fragmentation of AO and XOR, unequivocally demonstrating the purity of our AOH1 preparations (data not shown).

Spectroscopy, Metal Analysis, and Catalytic Properties of the Mouse Liver AOH1 Protein—AOH1 exhibits an absorbance spectrum similar to that of the well-characterized iron–sulfur-containing XOR and AO molybdo-flavoproteins with maxima at 330 and 446 nm, and a pronounced shoulder at 471 nm (Fig. 2*A*, *spectrum a*). The A_{278}/A_{446} ratio is ~5.3, and the calculated extinction coefficient is 37.6 mM^{-1} cm^{-1}. These values are comparable to those reported for XORs isolated from various sources, suggesting the presence of a similar cofactor complement (26). The absorbance spectrum of the solution obtained following heat denaturation and removal of AOH1 indicates the presence of a flavin cofactor (Fig. 2*A*, *spectrum b*). The difference between the spectrum of native AOH1 and that of the released cofactors (Fig. 2*A*, *spectrum c*) is very similar to the absorbance spectrum of the deflavo-form of XOR (27–29). Phosphodiesterase treatment of the material released from AOH1 denaturation results in an increase of ~7-fold in flavin fluorescence, which is consistent with the fact that the majority of the cofactor present in solution is FAD (21). Thus, using the extinction coefficient of 11.3 mM^{-1} cm^{-1} (450 nm), the calculated values of FAD/AOH1 subunit were 1.1 and 1.4 mol in two independent experiments. The values are compatible with one molecule of FAD bound per AOH1 subunit.

Metal analysis of purified AOH1 resulted in the determination of 0.92 ± 0.05 nmol of molybdenum (mean ± S.D. of three determinations) and 4.4 ± 0.3 nmol of iron (mean ± S.D. of three determinations) per 1 nmol of enzyme monomeric subunit.

As shown in Fig. 2*B*, AOH1 catalyzes the oxidation of phthalazine with concomitant reduction of ferricyanide. Thus the reaction can be studied spectrophotometrically by monitoring the velocity of ferricyanide reduction. At concentrations of phthalazine below 50 μM, which are not inhibitory, the AOH1-catalyzed oxidation reaction exhibits hyperbolic dependence on substrate concentration. The apparent K_m value for phthalazine is 2.3 ± 0.4 μM, and the apparent enzyme turnover number is 4 ± 0.15 s^{-1}. Preincubation of purified AOH1 with 5 mM KCN results in ~86% inhibition of phthalazine-oxidizing activity (143 ± 13 *versus* 1039 ± 13 units/mg of protein; mean ± S.D., $n = 3$). A similar level of inhibition is observed in the case of the phenanthridine- and benzaldehyde-oxidizing activity associated with AOH1 (data not shown).

Chromosomal Mapping of the AOH1 and AOH2 Loci—The full-length cDNAs coding for AOH1 and AOH2 were used to define the chromosomal location of the two corresponding loci by fluorescence *in situ* hybridization (FISH). All the observed metaphases show specific AOH1 (Fig. 3, *A* and *B*) and AOH2 (Fig. 3, *E* and *F*) hybridization signals on two large chromosomes, subsequently identified, by G-banding, as the two chromosome 1. The signals map to the same region, where the *AO* gene was previously localized (25). The use of genomic probes containing the *AOH1* (Fig. 3*C*) and *AOH2* (Fig. 3*G*) loci confirm these results. To better define the position of *AOH1* and *AOH2* relative to that of the *AO* genetic locus, co-hybridization experiments using genomic probes labeled with two different fluo-

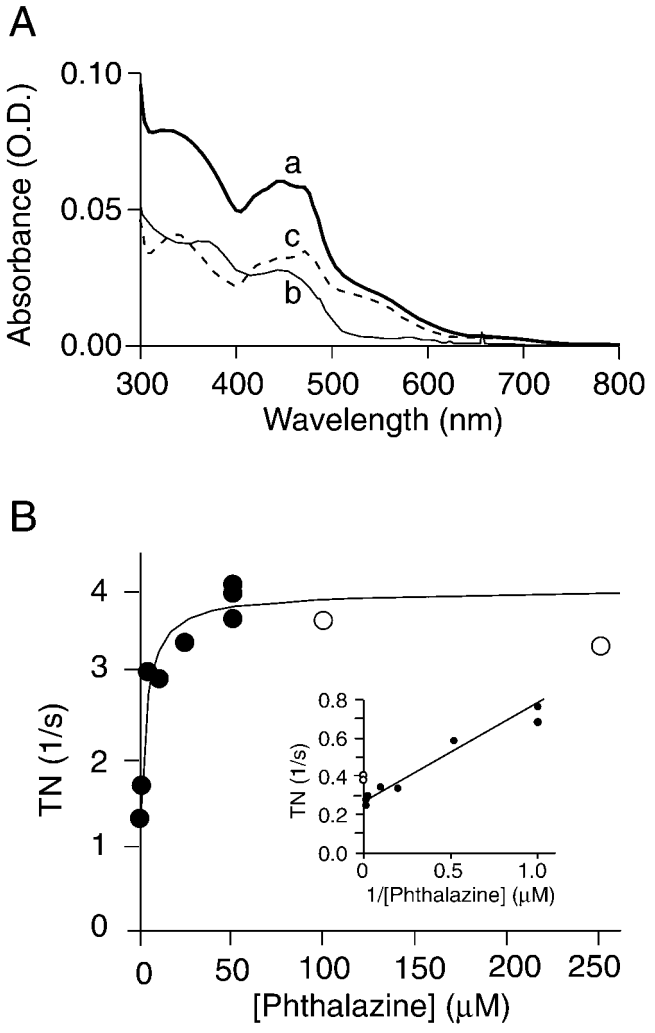

FIG. 2. **Spectroscopic and enzymatic properties of mouse liver AOH1.** *A*, absorbance spectrum of mouse liver AOH1 in native conditions (*trace a*) and of the released cofactors following treatment at 100 °C of the protein (*trace b*). *Trace c* is the subtraction spectrum of *traces a* and *b*. *B*, Michaelis-Menten plot of AOH1 phthalazine-oxidizing activity. The double-reciprocal plot used to calculate the enzyme K_m is shown in the *inset*. *Open circles* indicate the experimental values showing substrate inhibition.

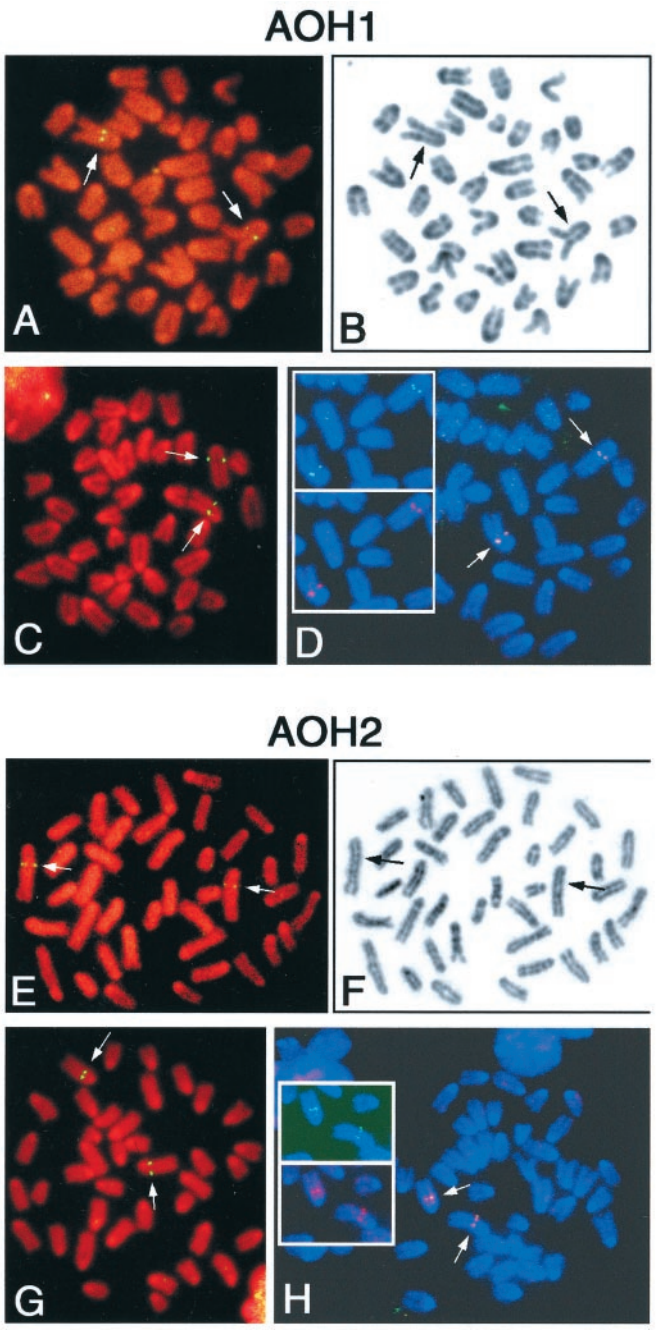

FIG. 3. **Chromosomal mapping of the mouse *AOH1* and *AOH2* genes.** Metaphase chromosomes were hybridized with biotin-labeled AOH1 (*A*) and AOH2 (*E*) cDNA probes. G-banding of the metaphase shown in *A* and *E* are represented in *panels B* and *F*. Metaphase hybridized with a biotin-labeled PAC containing the entire *AOH1* and *AOH2* genes are shown in *C* and *G*, respectively. Metaphase showing the co-hybridization of biotin-labeled *AO* gene probe and digoxigenin-labeled *AOH1* (*D*) or *AOH2* (*H*) probes were obtained by triple exposure of the metaphase with single band-pass filters specific for fluorescein, rhodamine, and 4′-6-diamidine-2-phenylindole (DAPI). The *insets* show the metaphase obtained by double exposure with fluorescein and DAPI (*green signals*), or rhodamine and DAPI (*red signals*), corresponding to the *AOH1* (*D*) or *AOH2* (*H*) genes and to the *AO* gene, respectively. In *panels A*, *C*, *E*, and *G* chromosomes were stained with propidium iodide, whereas in *D* and *H*, chromosomes were stained with DAPI. The *arrows* indicate the hybridization sites.

rescent tags were performed. Fig. 3*D* demonstrates that the *AOH1* and *AO* hybridization signals are coincident. As shown in Fig. 3*H*, a similar situation is evident also in the case of *AOH2* and *AO*. These results indicate that *AOH1*, *AOH2*, and *AO* are strictly associated on the C1–C2 bands of mouse chromosome 1. This location is different from that of the mouse *XOR* gene, which maps to chromosome 17 (30).

Cloning and Characterization of the AOH1 and AOH2 Genes—To determine the structure of the *AOH1* and *AOH2* genes, we screened a mouse PAC library with the AOH1 and AOH2 cDNAs as probes and selected three clones for further analysis. As documented by Fig. 4*A*, *PAC 1* and *PAC 2* hybridize with the full-length AOH1 cDNA probe. The two clones overlap, because they show a number of common hybridization bands upon cleavage with three distinct restriction enzymes. *PAC 2* and *PAC 3* also contain overlapping genomic fragments, which hybridize with the AOH2 full-length cDNA. Given the proximity of the *AOH1* and *AOH2* genes to the *AO* locus on chromosome 1, we performed Southern blot analysis of the selected PAC clones with probes corresponding to the 5′ and 3′ ends of the AO cDNA. These experiments demonstrate that *PAC 1* also contains the entire *AO* locus. As schematized in Fig. 4*B*, long range PCR experiments with appropriate amplimers and nucleotide sequencing demonstrate that the three genes are located on the same DNA strand in the order *AO*, *AOH1*, and *AOH2* from the 5′- to the 3′-end. *AO* is separated by ~4.5

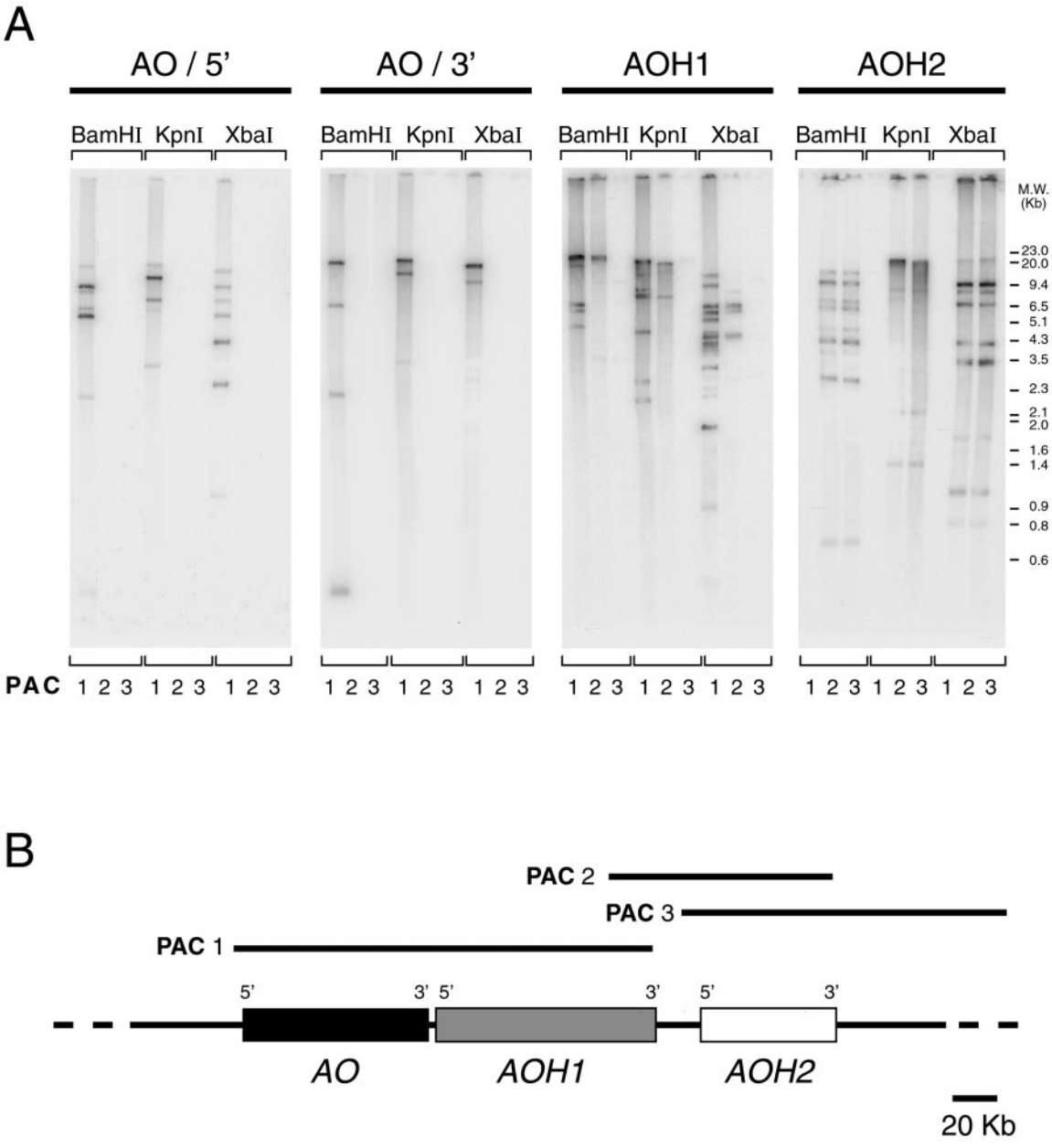

FIG. 4. **Physical map of the AOH1 and AOH2 loci.** *A*, Southern blot analysis of three selected PAC clones isolated following screening with the AOH1 or the AOH2 cDNA probes. One microgram of PAC 1, PAC 2, or PAC 3 DNA was cleaved with the indicated restriction enzymes. Equivalent aliquots of the restricted DNA were subjected to Southern blot analysis with the following radiolabeled probes: AO cDNA (GenBank™ accession number AF076216), nucleotides 1–1701 (AO/5′); AO cDNA nucleotides 3368–4347 (AO/3′); full-length AOH1 cDNA (AOH1; GenBank™ accession number AF172276); and full-length AOH2 cDNA (AOH2; GenBank™ accession number AF233581). Molecular weight DNA standards are indicated on the *right*. *B*, the relative position of the *AO*, *AOH1*, and *AOH2* genetic loci on chromosome 1 is schematized in the *lower portion of the panel*. The *thick lines* represent the PAC clones whose physical map is shown in *panel A*.

kb from *AOH1*, whereas *AOH1* and *AOH2* are separated by 20 kb.

The exon/intron structure of the *AOH1* and *AOH2* genes as well as the nucleotide sequence of the relative junctions are shown in Tables III and Table IV, respectively. The *AOH1* gene is 98 kb, consisting of 35 coding exons and has the potential to transcribe an extra non-coding leader exon (exon 1bis), as discussed below. The *AOH2* gene is 60 kb and is composed of 35 exons; exon 1 codes for the entire 5′-untranslated region of the corresponding mRNA. Except for the boundary between *AOH1* exons 1bis and 1, which is unusual (31), all of the exon/intron junctions conform to the GT/AG consensus sequence found in other eukaryotic genes.

The sequences of the *AOH1* and *AOH2* genes are different from those of the corresponding cDNAs (GenBank™ accession numbers AF076216 and AF233581) at some positions. In *AOH1*, three of the fourteen different nucleotides give rise to the following predicted amino acid changes in the corresponding protein product: $Ser^{261} \rightarrow Asn$, $Leu^{530} \rightarrow Val$, and $Arg^{541} \rightarrow Lys$. These differences may be simply due to strain-related genetic polymorphism (129/sv *versus C57/Bl*). In addition, sequencing of *AOH1* exon 31 led to the identification of an error between nucleotides 3649 and 3662 of the reported AOH1 cDNA (12). This results in the following amino acid changes: Arg^{1151}-Lys^{1152}-Val^{1153}-Thr^{1154}-Phe^{1155} to Glu^{1151}-Gly^{1152}-Asp^{1153}-Ile^{1154}. Thus, AOH1 consists of 1335 (see also Fig. 8) and not 1336 amino acids, as originally reported (12). Nine of the 15 nucleotide differences observed in the *AOH2* gene relative to the corresponding cDNA are responsible for amino acid substitutions ($Arg^{106} \rightarrow Gln$, $Gly^{326} \rightarrow Glu$, $Arg^{336} \rightarrow His$,

TABLE III

Exon–intron organization of the mouse AOH1 gene

Exon sequences are shown in *uppercase letters*, and intron sequences are in *lowercase letters*. The positions of nucleotides close to the 5′ and 3′ ends of the exons are numbered below the DNA sequence. The numbering is the same as that of the AOH1 cDNA (12), up to nucleotide 3650. From this nucleotide on, the numbering changes to account for the observed single-codon difference between the coding sequence of the gene and the reported sequence of the cDNA (see text). The sizes of exons 1bis and 1 are calculated on the basis of the published cDNA sequence, although their lengths, as determined by RACE experiments, are also presented in *parentheses*. Amino acids bordering the splice junctions are shown above the nucleotides. Putative mRNA cleavage/polyadenylation signals at the 3′-end of exon 35 are *underlined*. The unusual splicing site between exons 1bis and 1 is also *underlined*. The length of exon 1bis is calculated on the basis of the major transcription initiation site in AOH1 type I transcripts, whereas that of exon 1 is calculated according to the position of the junction between exon 1bis and exon 1 in the same type of transcripts.

Exon	Exon size (bp)	5′Splice site	Intron size (Kb)	3′Splice site
1bis	76 (90)	..cctttgctaggatagagagca..	7.0	..aggcaaagccacagagcagat..
		66		77
		..Gly Lys Lys		Val Thr Glu..
1	177 (135)	..GGA AAA AAGgtaagttctc..	1.6	..ctctctccagGTC ACC GAG..
		245		254
		..Lys Val I		le Arg Leu..
2	58	..AAA GTC Agtatccttttt..	3.6	..acccttccagTC CGA CTC..
		305		312
		..Arg Ile Se		r His Phe..
3	97	..AGG ATC AGgtacctgccc..	0.8	..aactttgaagT CAT TTC..
		401		409
		..Pro Val Gln		Glu Arg Ile..
4	109	..CCT GTC CAGgtaagagcgt..	1.4	..ttcattttagGAA AGG ATT..
		509		510
		..Leu Gly G		ly Asn Leu..
5	127	..CTG GGC Ggtaagttgat..	4.3	..cctctctcagGG AAT CTA..
		638		645
		..Phe Cys Pro		Ser Ser Thr..
6	62	..TTC TGC CCGgtgagtatga..	2.5	..ttttccacagAGT TCA ACT..
		698		707
		..Lys Asn Ser		Val Cys Thr..
7	81	..AAA AAC AGTgtgagtttcc..	15.0	..tcttcttcagGTT TGC ACC..
		779		788
		..Glu Leu Met		Arg Met Ala..
8	81	..GAG CTG ATGgtttgttctg..	3.0	..aatattgcagAGG ATG GCC..
		860		869
		..Tyr Leu G		ly Leu His..
9	145	..TAT CTG Ggtgggtagcg..	1.1	..tcctatctagGG CTT CAT..
		1007		1014
		..Lys Gln G		ly Leu Thr..
10	93	..AAG CAA Ggtatgttctt..	3.5	..tcactctcagGG CTG ACA..
		1100		1107
		..Asn Val Ala		Ser Leu Gly..
11	152	..AAT GTG GCCgtaggttctc..	0.7	..tttcctttagTCC TTA GGT..
		1250		1259
		..Ser Thr G		lu Gly Ile..
12	94	..TCA ACA Ggtaagtgact..	3.0	..cccttcacagAA GGA ATA..
		1346		1353
		..Ser Ser Lys		Trp Glu Phe..
13	110	..TCC AGC AAGgtatgatgtc..	2.3	..cattttccagTGG GAG TTT..
		1454		1463
		..Ile Gly Ar		g Cys Trp..
14	185	..ATT GGA AGgtgacttaca..	0.5	..ctgtgcttagG TGC TGG..
		1640		1648
		..Lys Thr Arg		Asp Pro His..
15	163	..AAG ACG AGGgtaagtttgt..	0.4	..ttcattccagGAC CCC CAC..
		1802		1811
		..Ser Phe Gln		Asp Val Asp..
16	93	..TCA TTT CAGgtgagtccat..	4.5	..tttaacacagGAT GTA GAC..
		1895		1904
		..Lys Ile Il		e Ser Leu..
17	170	..AAA ATC ATgtaagttatg..	0.6	..tctttcctagC TCT CTT..
		2066		2074
		..Gln Asp Glu		Val Ile Cys..
18	115	..CAG GAT GAGgtagtgtgat..	0.8	..ttcttcccagGTG ATC TGC..
		2180		2189

Ser^{417} → Phe, Arg^{570} → Gln, Glu^{597} → Gly, Ser^{1165} → Ala, Ala^{1166} → Cys, and Pro^{1167} → Ser). Except for Ser^{417} → Phe, however, all these amino acid differences are the consequence of base misincorporations caused by the *rTth* DNA polymerase used for the original amplification of part of the AOH2 cDNA (12). In fact, sequencing of the cDNA following amplification with *Pfu* DNA polymerase results in the same nucleotide sequence determined for the *AOH2* gene. None of the observed amino acid differences fall within highly conserved domains of the AOH1 and AOH2 proteins (see also Fig. 8).

The sequence of *AOH1* exon 35 and of its 3′-flanking region does not contain any canonical polyadenylation signal. However, as already noticed for the corresponding AOH1 cDNA, 20 and 32 nucleotides upstream of the end of exon 35, there are two sequences that may be used as polyadenylation signals. By contrast, a canonical consensus sequence for the addition of a polyadenylated tail to the corresponding mRNA is present in exon 35 of the *AOH2* gene.

Determination of the Transcription Start Sites of the AOH1 and AOH2 Genes, Nucleotide Sequencing, and Characterization of the Corresponding 5′-Flanking Regions—To determine the transcription start site of the *AOH1* gene, a number of nested RACE-PCR experiments using different specific amplimers were performed on RNA extracted from adult and embryonic liver. Fig. 5*A* shows one representative study on the AOH1 transcript(s). Following amplification with AP1 (Anchor

TABLE III—*Continued*

Exon	Exon size (bp)	5′Splice site	Intron size (Kb)	3′Splice site
		..Thr Val Gln		Asp Ala Leu..
19	123	..**ACA GTG CAG**gtaagttctg..	3.6	..accttcccag**GAT GCA CTG**..
		2303		2312
		..Leu Glu G		ly Glu Val..
20	97	..**CTC GAA G**gtaaattctt..	2.5	..tatccaaaag**GG GAG GTG**..
		2402		2409
		..Phe Thr Gln		Glu Met Val..
21	125	..**TTT ACA CAG**gtaggttcca..	1.7	..ctgctcatag**GAA ATG GTG**..
		2525		2534
		..Ala Gln Ly		s Thr Gly..
22	134	..**GCA CAG AA**gtgagtggat..	2.3	..tcttttctag**G ACC GGC**..
		2660		2668
		..Lys Tyr Lys		Ile Gly Phe..
23	88	..**AAA TAC AAG**gtgagtgggt..	0.7	..ctgttcacag**ATT GGC TTC**..
		2747		2756
		..Ser Glu Leu		Val Ile Glu..
24	87	..**TCT GAA CTG**gtaagtgact..	1.2	..ttttttttag**GTG ATA GAA**..
		2834		2843
		..Pro Glu Lys		Val Arg Glu..
25	192	..**CCA GAG AAG**gtaagaggag..	0.3	..catccaccag**GTT CGA GAG**..
		3026		3035
		..Tyr Tyr Gln		Ala Ala Ala..
26	228	..**TAT TAT CAG**gtgattattg..	4.2	..cttgtttcag**GCT GCC GCT**..
		3254		3263
		..Met Ile Gln		Val Ala Ser..
27	96	..**ATG ATA CAG**gtgagagagg..	1.0	..ctccggacag**GTG GCC AGC**..
		3350		3359
		..Ala Val Gln		Asn Ala Cys
28	129	..**GCT GTT CAG**gtgataggct..	2.0	..catcccacag**AAT GCC TGT**..
		3479		3488
		..Glu Glu Trp		Val Lys Glu..
29	75	..**GAA GAA TGG**gtgagtacct..	1.2	..tcttttttcag**GTT AAG GAA**..
		3554		3563
		..Tyr Phe Ar		g Gly Tyr..
30	53	..**TAT TTT AG**gtaagaggtg..	2.8	..gtcttctcag**G GGT TAC**..
		3608		3616
		..Ala His Lys		Asn Ile Arg..
31	115	..**GCT CAC AAG**gtaagtggga..	2.7	..ttctccacag**AAC ATT AGA**..
		3722		3731
		..Ile Gly Gln		Ile Glu Gly..
32	66	..**ATA GGC CAG**gtaggcaatgc..	2.0	..tgtatcccag**ATT GAA GGG**..
		3788		3797
		..Ser Ser Lys		Gly Leu Gly..
33	189	..**TCT TCT AAG**gtaagtatat..	6.1	..acatccccag**GGC CTT GGC**..
		3977		3986
		..Thr Asn Leu		Val Pro Gln..
34	168	..**ACG AAC CTG**gtatgggatc..	5.0	..ttctttttag**GTT CCA CAA**..
		4145		4154
35	276	..AAAACTTCCACAGGTCACAAATATAAATATCTATATCTAAAATATGTAGAAAATTAGA GGCAATTCTTAAATTTTCATATTTACAAAATATTGTAATCTTTAATCACATGCACTATAA GCAACGATGTTTACGGAGCTTCAGCCAAAGGTTATGTTTAATTCAATATATTAAACGGAA TTTTCCACT....		

Primer 1) and SP1 (Specific Primer 1; complementary to nucleotides 388–409 of the AOH1 cDNA), a nested PCR step was performed with the use of AP2 and either NP1 (Nested Primer 1; complementary to nucleotides 51–76 of the AOH1 cDNA) or NP2 (complementary to nucleotides 205–230 of the AOH1 cDNA). The use of NP1 allows the amplification of a single band of ~90 bp, whereas NP2 leads to the appearance of a fragment of ~250 bp from both adult and embryonic liver RNA. Sequencing of these two DNA products in numerous derived cDNA clones resulted in the identification of a major transcription initiation site located 213 nucleotides upstream of the first ATG codon (type I transcript), *i.e.* 14 nucleotides upstream of the published AOH1 cDNA 5′-end (12). The first 90 nucleotides are located in a region of the *AOH1* gene, which is separated from the ATG-containing exon 1. This defines a leader exon, which we named exon 1bis (see Table III). When the nested amplification step is carried out with NP2, a second and minor band of ~110 bp is observed in adult and to a lesser extent also in embryonic liver. Following subcloning and sequence analysis of this DNA, we identified a minor transcription initiation site located 81 bases upstream of the first ATG codon (type II transcript). Type I and II transcripts represent ~90 and 10%, respectively, of all the cDNAs amplified by RACE. Surprisingly, Southern blot analysis and long range PCR experiments, using DNA obtained from PAC1 and other overlapping clones, demonstrate that exon 1bis is located downstream of exon 26. In addition, the direction of transcription of exon 1bis is opposite to that of all the other exons in the gene, as schematized in Fig. 5*A*. The molecular mechanisms underlying the generation of type I transcripts are presently unknown, although they may be post-transcriptional in nature. A recombination between precursor RNA molecules or a *trans*-splicing event (the use of a splicing donor site from one RNA by the acceptor site from a second) (32–34) may explain the phenomenon. This would involve two precursor mRNAs, which originate from the activity of distinct promoters regulating transcription of exon 1bis and exon 1.

Transcription of the *AOH2* gene is less complicated than that of its *AOH1* counterpart. Both nested RACE-PCR with different primers (SP2, complementary to nucleotides 380–401 of the AOH2 cDNA; NP3, complementary to nucleotides 76–105; NP4, complementary to nucleotides 189–210) (Fig. 5*B*) and

TABLE IV

Exon–intron organization of the mouse AOH2 gene

Exon sequences are shown in *uppercase letters*, and intron sequences are in *lowercase letters*. The positions of nucleotides close to the 5′ and 3′ ends of the exons are numbered below the DNA sequence. Nucleotide numbering is the same as that of the reported AOH2 cDNA sequence. The size of exon 1 is calculated on the basis of the published cDNA sequence, although its length, as determined by RACE experiments, is also presented in *parentheses*. Amino acids bordering the splice junctions are shown above the nucleotides. Putative mRNA cleavage/polyadenylation signals at the 3′-end of exon 35 are *underlined*.

Exon	Exon size (bp)	5′Splice site	Intron size (Kb)	3′Splice site
		..Gly Lys Lys		Val Ile Glu..
1	159 (178)	..GGA AAA AAGgtaagtttct..	2.1	..cgcaatatagGTC ATA GAG..
		151		160
		..Lys Val L		eu Asn Leu..
2	58	..AAA GTT Cgtatccttta..	3.5	..tcccttccagTC AAT CTC..
		211		218
		..Lys Ile Hi		s His Tyr..
3	97	..AAG ATC CAgtatcctttg..	1.5	..cccatcacagC CAC TAC..
		307		315
		..Pro Val Arg		Glu Arg Leu..
4	109	..CCA GTC AGGgtaagagcct..	1.9	..tgtctctcagGAA CGA CTT..
		415		424
		..Leu Gly G		ly Asn Leu..
5	127	..CTT GGA Ggtgaggcttt..	2.5	..tattaatcagGG AAT TTG..
		544		551
		..Phe Ser Gln		Lys Ser Thr..
6	62	..TTC TCC CAGgtaagttaaa..	3.5	..acttcccaagAAG TCT ACT..
		604		613
		..Glu Lys Lys		Met Cys Thr..
7	84	..GAG AAA AAGgtaagtctct..	0.7	..ttttcttaagATG TGT ACT..
		688		697
		..Glu Leu Ile		Arg Met Ala..
8	81	..GAA CTG ATTgtaagttgct..	2.2	..ttctctgcagAGA ATG GCA..
		769		778
		..Thr Val G		ly Pro Gly..
9	145	..ACA GTG Ggtatgtgtta..	0.3	..tctctttaagGA CCC GGT..
		916		923
		..Asn Asn G		ly Val Thr..
10	93	..AAC AAT Ggtaagcatct..	2.3	..ctgctttcagGG GTA ACA..
		1009		1016
		..Asn Met Ala		Thr Leu Gly..
11	152	..AAT ATG GCTgtatgttcta..	1.2	..tttttattagACT TTA GGA..
		1159		1168
		..Ser Arg G		lu Gly Lys..
12	94	..TCT AGA Ggtaagagatc..	1.5	..tctctaacagAA GGA AAA..
		1255		1262
		..Thr Ala Gln		Trp Gln Phe..
13	113	..ACT GCT CAGgtgaggttcc..	3.0	..ctatctctagTGG CAG TTT..
		1366		1375
		..Ile Gly Ar		..g Gln Trp..
14	185	..ATT GGG AGgtaggtccca..	1.6	..gtcctgccagG CAA TGG..
		1552		1560
		..Asn Glu Met		Asp Pro Gln..
15	163	..AAT GAA ATGgtaaatagct..	3.5	..ttttctccagGAT CCT CAG..
		1714		1723
		..Met Phe Gln (Arg)		Cys Val Asp..
16	93	..ATG TTC CAGgtgggtagag..	0.8	..ttcttaacagTGT GTG GAT..
		1807 G		1816
		..Lys Ile Th		r Ser Leu..
17	170	..AAG ATC ACgtaaggaaat..	0.5	..ttcctcccagA TCA CTT..
		1978		1986
		..Gln Ser Glu		Val Ile Cys..
18	115	..CAG AGT GAGgtgtgtggtt..	1.0	..ttttctacagGTC ATT TGC..
		2092		2101

primer-extension analysis (data not shown) concur in defining a single major transcription initiation site for the AOH2 mRNA synthesized in mouse skin as well as stomach (data not shown). The transcription start site lays 124 nucleotides upstream of the first ATG codon, *i.e.* 19 nucleotides upstream of the 5′-end of the published AOH2 cDNA (12).

The data obtained indicate the presence of potential promoter elements in the 5′-flanking regions of exon 1bis (5FR1/1bis) and exon 1 (5FR1/1) of the *AOH1* gene. The sequences of 5FR1/1bis and 5FR1/1 (Fig. 6, *A* and *B*) are characterized by the presence of possible binding sites for numerous general and tissue-specific transcription factors. Consistent with the high level of expression of the *AOH1* gene in liver, both sequences contain multiple copies of cEBP binding sites. 5FR1/1bis shows a typical TATA box located 30 nucleotides upstream of the transcription initiation site, whereas an almost canonical consensus sequence for an SP1 binding site lays 55 nucleotides upstream of the transcriptional start site of exon 1. Interestingly, both DNA regions have binding sites for aryl hydrocarbon receptor nuclear translocator, a transcription factor involved in the regulation of genes by polycyclic aromatic hydrocarbons (35).

Fig. 6*C* shows the nucleotide sequences of the 5′-flanking region of the *AOH2* gene (5FR2), which indicates the presence of a canonical TATA box located 17 nucleotides upstream of the transcription initiation site. Consistent with the completely different pattern of tissue- and cell-specific expression, the 5′-flanking region of *AOH2* contains binding sites for transcription factors different from those observed in the corresponding regions of *AOH1*. Of particular notice, is the presence of multiple AP1 and nuclear factor of activated T-cells sites as well as that of SREBP1 (sterol regulatory element binding protein 1), one of the transcriptional factors that regulate the activity of genes involved in the metabolism of cholesterol (36).

TABLE IV—*Continued*

Exon	Exon size (bp)	5′Splice site	Intron size (Kb)	3′Splice site
		..Thr Ile Glu		Glu Ala Leu..
19	123	..ACC ATA GAGgtagatcact..	1.6	..tctactgcagGAA GCC CTA..
		2215		2224
		..Val Glu G		ly Glu Ile..
20	97	..GTT GAA Ggtatgtacag..	1.4	..tgctttgcagGT GAG ATC..
		2314		2321
		..His Val Gln		Glu Phe Val..
21	125	..CAC GTG CAGgtaacggggg..	2.0	..tctcttgtagGAA TTT GTG..
		2437		2446
		..Ala Asn Ly		s Thr Gly..
22	134	..GCC AAC AAgtaagtactgt..	1.7	..tttcccacagG ACT GGC..
		2572		2580
		..Lys Tyr Lys		Ile Gly Phe..
23	88	..AAA TAC AAAgtgagtgaga..	0.6	..cactttgcagATT GGG TTC..
		2659		2668
		..Ser Glu Leu		Val Ile Glu..
24	87	..TCT GAG CTGgtaaatatac..	0.4	..tttaattcagGTG ATA GAA..
		2746		2755
		..Pro Glu Glu		Val Arg Glu..
25	192	..CCT GAA GAGgtgagcagtt..	1.5	..caccttt cagGTT AGA GAA..
		2938		2947
		..Tyr Asn Gln		Ala Ala Ala..
26	228	..TAC AAT CAGgtgaggtgga..	2.0	..tttcccacagGCT GCT GCT..
		3166		3175
		..Met Ile Gln		Val Ala Ser..
27	96	..ATG ATT CAGgtgagaccag..	2.1	..catccggtagGTA GCC AGT..
		3262		3271
		..Ala Val Gln		Asn Ala Cys..
28	129	..GCT GTA CAGgtgacttgtc..	3.5	..tgcttttcagAAT GCC TGC..
		3391		3400
		..Glu Glu Trp		Ile Lys Met..
29	75	..GAG GAA TGGgtgagtgttc..	1.0	..tctttttcagATT AAA ATG..
		3466		3475
		..Tyr Phe Ly		s Gly Tyr..
30	53	..TAT TTC AAgtaaatgagt..	0.5	..ctcttctcagA GGC TAC..
		3520		3528
		..Ala His Lys		Leu Leu Arg..
31	115	..GCT CAC AAGgtagagtgat..	0.5	..ctctgtgcagCTC CTG AGG..
		3634		3643
		..Ile Gly Gln		Val Glu Gly..
32	66	..ATC GGG CAGgtatgtaatg..	0.8	..gtgcccacagGTT GAG GGA..
		3700		3709
		..Ser Ser Lys		Gly Leu Gly..
33	186	..TCA TCC AAGgtaagaattt..	1.7	..ttttcctcagGGA TTG GGT..
		3886		3895
		..Thr Glu Met		Ile Pro Arg..
34	168	..ACT GAG ATGgtaagggaag ..	1.0	..ctttttttagATT CCC AGA..
		4054		4063
35	834	..AGTCCATCTACTGGGACAGAAAGGAGGTCCTCCACCATGGGCAGTCTTGTGAACTTGCC TCTAAACTATGAACCAAAATAAAACTCTTTTCTTCAC.....		

The 5′-flanking regions of *AOH1* and *AOH2* contain functional promoters, as assessed by transient transfection experiments performed with appropriate constructs containing the luciferase reporter gene under the control of 5FR1/1bis (*p5FR1bis-Luc*), 5FR1/1 (*p5FR1-Luc*), or 5FR2 (*p5FR2-Luc*). Transfection of HEK293 cells with *p5FR1bis-Luc* causes a 1.98 ± 0.16-fold induction of luciferase activity (mean ± S.D. of three experiments) relative to what is observed in extracts of cells programmed with a promoter-less luciferase construct (pLuc-basic; 1.00 ± 0.07). Following transfection of *p5FR1-Luc* and *p5FR2-Luc*, luciferase activity is stimulated 11.50 ± 0.83 (pLuc-basic 1.00 ± 0.08) and 3.85 ± 0.11 (pLuc-basic; 1.00 ± 0.01), respectively.

Comparison of the Structure and the Position of Exon/Intron Junctions of AOH1 and AOH2 with Those of Other Members of the Molybdo-flavoprotein Family—GenBank™ contains a number of putative or characterized XOR and AO protein sequences derived from various animal and plant species. Putative proteins, whose primary structure was deduced from the available nucleotide sequence of assembled genomes, can be classified as XORs according to the presence of a conserved NAD binding site (28) and two invariant amino acid residues in the substrate binding pocket (37). At present all the other putative molybdo-proteins can be classified as AOs. On the basis of these criteria, the genome of *Arabidopsis thaliana* is predicted to encode two XOR isoenzymes and four different genes coding for AO isoenzymatic forms (38). The genome of *Drosophila melanogaster* codes for four different putative isoenzymatic forms of AO besides XOR, which is the product of the well known *rosy* locus (39). Furthermore, the *Bombyx mori* genome codes for two isoenzymatic forms of XOR (40). Finally, analysis of the *Caenorhabditis elegans* genome predicts the presence of one XOR and one AO protein (41). The entire amino acid sequence of mouse AOH1 and AOH2 can be readily aligned with that of all the AO and XOR currently known, allowing the generation of the dendrogram shown in Fig. 7*A*. This analysis indicates that AOH1 and AOH2 proteins represent a subgroup of mammalian AOs. Remarkably, AOH1 and AOH2 are more similar to plant, fungus, nematode, and insect *XOR* genes than to the corresponding *AO* genes. A comparison of the exon structure of *AOH1* and *AOH2* with that of all the other known or predicted eukaryotic *XOR* or *AO* genes is shown in Fig. 7*B*. This scheme indicates that *AOH1* and *AOH2* have the same number of exons as mouse and human *AOs* and one exon less than mouse and human *XORs*. Furthermore, the number of exons of both *AO* and *XOR* genes in mammals is much higher than in any other animal or plant species.

As indicated in Fig. 8, the position and type of all the coding-exon boundaries of the *AOH1* and *AOH2* genes are conserved with those of mouse *AO*. 32 of 35 junctions are identical in

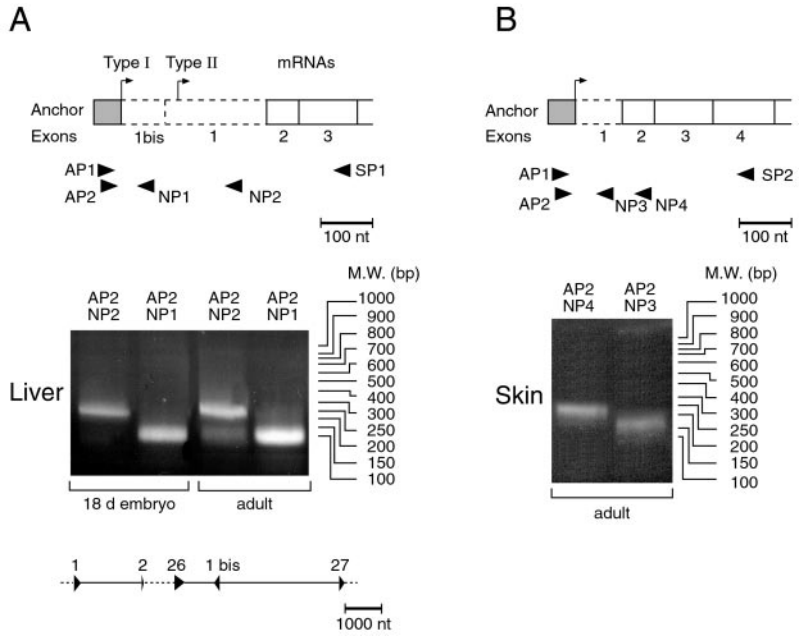

FIG. 5. **Mapping of the 5′-end of the mouse AOH1 mRNA and primary structure of the 5′-flanking region of the corresponding genes.** Poly(A^+) RNA isolated from 18-day-old embryo and adult mouse liver (*A*) or skin (*B*) was subjected to a first round of RACE using the specific AOH1 and AOH2 downstream amplimers SP1 (*A*) and SP2 (*B*), respectively, and the anchor-amplimer AP1. The second round of RACE was performed with the AOH1 or AOH2 specific amplimers, NP1 or NP2 (*A*), and NP3 or NP4 (*B*), respectively, and the anchor-amplimer, AP2. A scheme showing the position of each primer relative to the relevant exon junctions of the AOH1 and AOH2 transcripts is indicated (*top panel*). The *dashed lines* indicate the 5′ portion of the AOH1 transcripts whose length and sequence were determined by RACE. The *middle portion of each panel* shows an ethidium bromide staining of the RACE bands obtained following the second round of amplification with the indicated amplimers along with the position and the size of DNA molecular weight markers (*right*). A scheme indicating the position of exon 1bis and exon 1 in the *AOH1* gene is presented at the *bottom of panel A*. Only relevant exons are represented, and the *dotted lines* indicate the portions of the *AOH1* gene containing the 5′-flanking region, exons 3–25 and exons 28–35.

AOH1, *AOH2*, *AO*, and *XOR*. The first two positional discrepancies are observed at the boundaries between exons 7 and 8, and exons 15 and 16. These fall within regions of relatively low amino acid identities between *XOR* and the other three genes. The fusion of exons 26 and 27 of the *XOR* gene observed in *AOH1* and *AOH2* has already been reported in the case of AO (14). Interestingly, whenever a junction of type 0, I or II is determined in *AO*, an identical junction is also observed in *AOH1* and *AOH2*, or *XOR*. This striking conservation of the exon/intron junctions represents convincing evidence that the four loci coding for mouse molybdo-flavoproteins have a common genetic origin and evolved through one or more rounds of duplication from the same ancestral precursor. It is of note that the exon distribution of mouse *AOH1*, *AOH2*, *AO*, and *XOR* is conserved also in human *AO* and *XOR* (42).

To determine whether homologs of the *AOH1* and *AOH2* genes are present in animal species other than the mouse, we performed the Southern blot analysis shown in Fig. 9, using the full-length AOH1 and AOH2 cDNAs as probes. The high stringency hybridization pattern was compared with that observed following probing of the same genomic DNAs with the mouse AO and XOR cDNAs. AOH1 cross-hybridization bands are evident in human, bovine, and rabbit, and possibly also in lizard, whereas AOH2-specific signals are present only in the first three species. This suggests the existence of *AOH1* and *AOH2* homologs distinct from the *AO* and *XOR* genes in the genomes of the three animal species.

DISCUSSION

In this report, we present the structural and biochemical characterization of the native AOH1 protein, a recently identified aldehyde oxidase isoenzymatic form, which we purified from mouse liver. In addition, we report the isolation of the genetic loci coding for AOH1 and AOH2,[2] a second structurally related and less abundant molybdo-flavoenzyme.

The existence of proteins encoded by novel cDNAs needs to be verified by chemico-physical methods. In addition, it is important to define the structural and biochemical characteristics of protein products that are considered to require cofactors or post-translational modifications in their holoenzymatic form. This is especially true for purported molybdo-flavoenzymes, such as AOH1 and AOH2, which are difficult to express as a catalytically active recombinant protein in heterologous systems. For these reasons, we devised a strategy for the purification of AOH1 from mouse liver. The purification procedure is rapid and relatively efficient, resulting in the recovery of ~0.5–0.8 mg of pure and catalytically active protein from 40 g of mouse liver. Following purification, we determined the exact molecular mass of the AOH1 monomeric subunit with MALDI-TOF, the mass of numerous tryptic peptides with the same methodology and the sequence of five of them by ESI-MS/MS. These data confirm and expand the structural results indirectly obtained from the cDNA cloning experiments (12). Furthermore, we unequivocally demonstrate that AOH1 is a molybdenum- and iron-containing flavoprotein with an absorption spectrum very similar to that of the prototypical molybdo-flavoenzymes XOR and AO. The metal content of AOH1 is consistent with the presence of one atom of molybdenum and four atoms of iron per enzyme subunit. This is identical to what was reported in the case of XOR and AO, which contain 1 mol of molybdenum cofactor and 4 mol of iron–sulfur per enzyme subunit (1). As in other members of the molybdo-flavoprotein

[2] The unassembled and incomplete mouse genomic working draft sequence contained in GenBank™ (GI 8705112) contains most of the primary structure of the *AO*, *AOH1*, and *AOH2* genes. We have reordered this sequence, which can be obtained upon request.

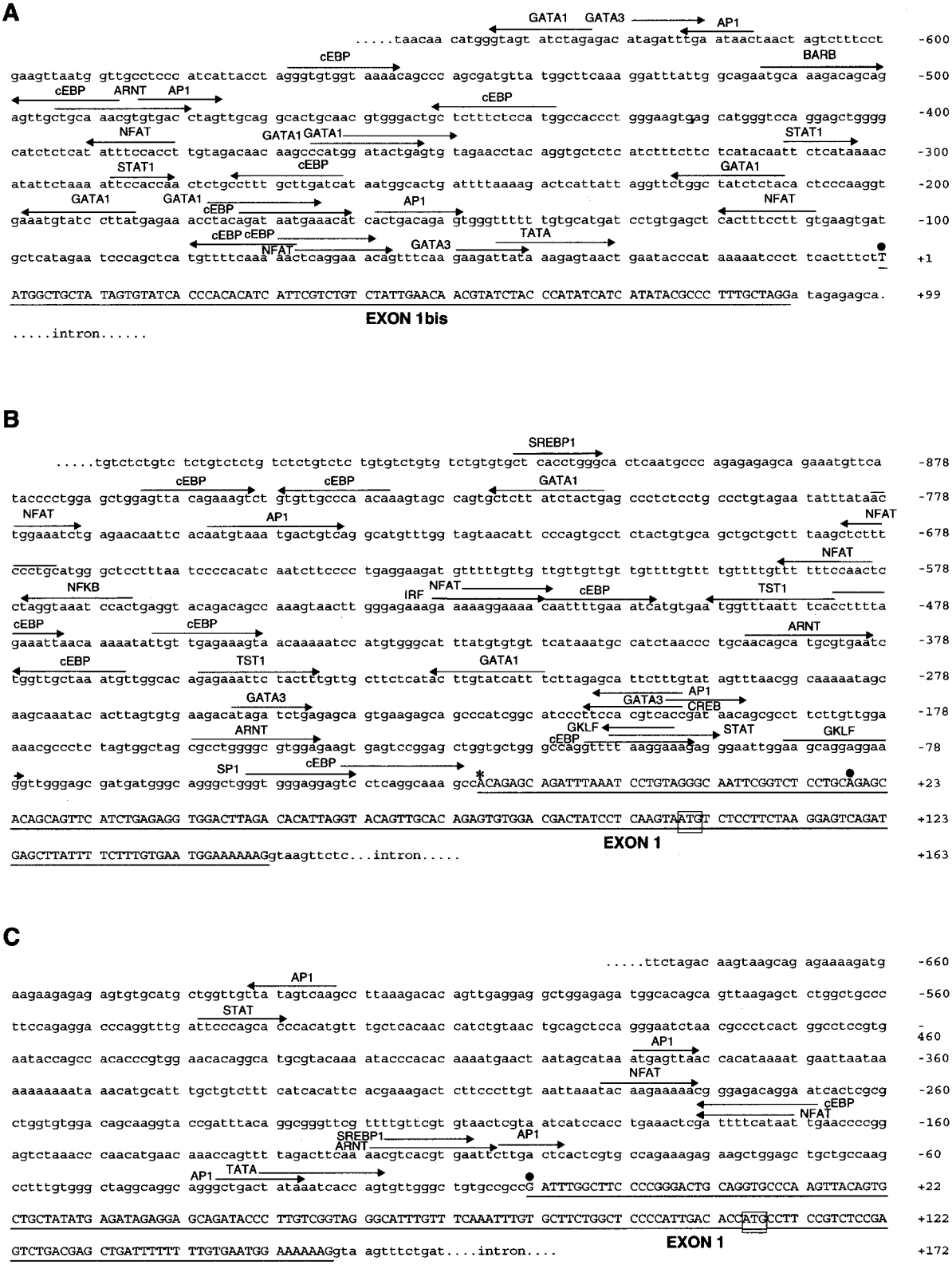

FIG. 6. **Primary structure of the 5′-flanking region of the corresponding genes.** The nucleotide sequence of the 5′-flanking regions of the *AOH1* exon 1bis (*A*), exon 1 (*B*), and *AOH2* exon 1 (*C*) are shown. In each case the nucleotides corresponding to the determined transcription start sites (*dots above the sequences*) are numbered as +1. Upstream sequences are indicated by *negative numbers*, and exon sequences are *underlined*. The first methionine codons are *boxed*. The *asterisk above the sequence shown in B* indicates the splicing junction between exon 1bis and 1 observed in type I transcripts. Consensus sequences for the binding of known transcription factors are indicated by *arrows above the sequence*.

family (1), the flavin cofactor necessary for the catalytic activity of the enzyme is FAD. *Finally*, we establish that AOH1 oxidizes phthalazine with a K_m in the low micromolar range. Considering that phthalazine is a specific substrate of mammalian AO (14), the data are consistent with our proposal that AOH1 is an isoenzymatic form of this protein. Similar to AO and XOR, AOH1 enzymatic activity is irreversibly inhibited by CN^-, suggesting the presence of cyanolyzable sulfur in the catalytic

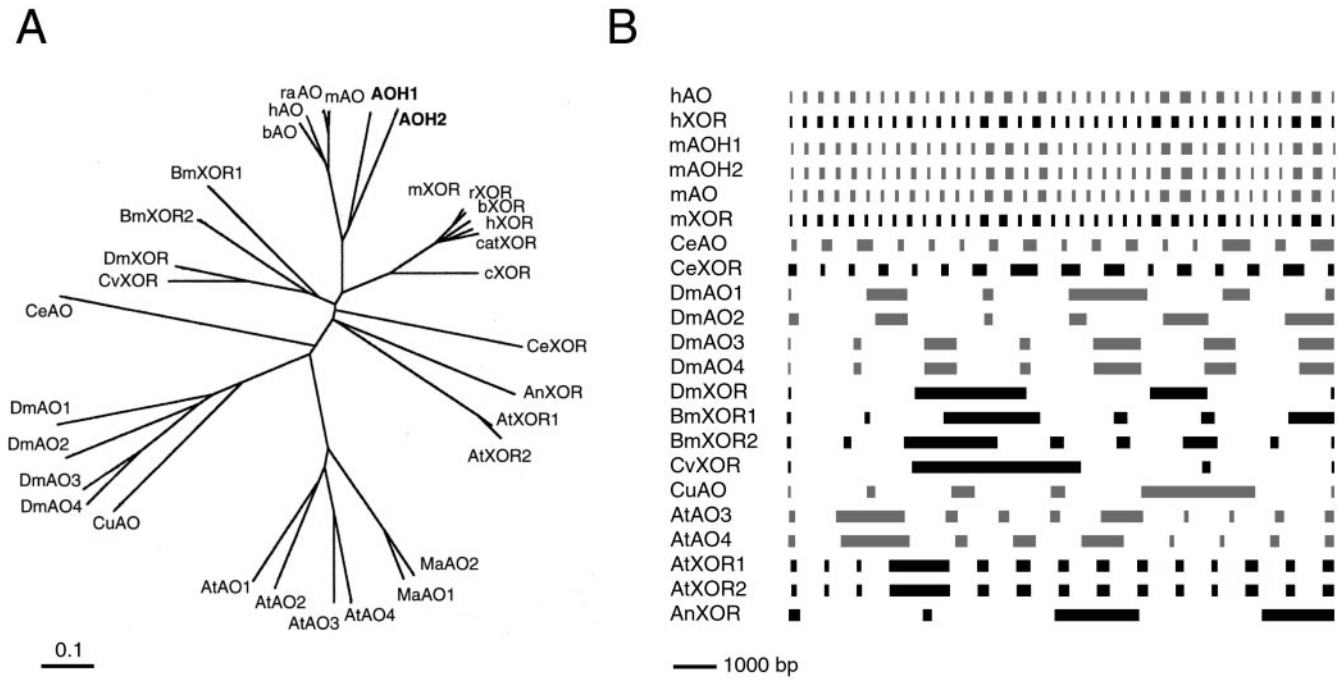

FIG. 7. **Phylogeny of eucaryotic molybdo-flavoenzymes.** *A*, an unrooted dendrogram obtained by a ClustalW computer-aided alignment of the indicated proteins and subsequent application of the Phylodendron software package are shown. AOH1 and AOH2 are *highlighted in boldface*. *B*, the coding exon structure of *AO* (*gray boxes*) and *XOR* genes (*black boxes*) of different animal and plant species is shown. *A. nidulans* xanthine oxidoreductase = AnXOR (GenBank™ accession number X82827; other GenBank™ numbers are shown in parens below); *A. thaliana* aldehyde oxidases-1, -2, -3, -4, and xanthine oxidoreductase-1 and -2 = AtAO1, AtAO2, AtAO3, AtAO4 (AB005804, AB005805, AB016622, and AB037271), AtXOR1 and AtXOR2 (AL161586 and AL161591); *B. mori* xanthine oxidoreductase-1 and -2 = BmXOR1 and BmXOR2 (D38159 and D43965); bovine aldehyde oxidase and xanthine oxidoreductase = bAO and bXOR (X87251 and X83508 or X98491); chicken xanthine oxidireductase = cXOR (D13221); cat xanthine oxidoreductase = catXOR (AF286379); *C. elegans* aldehyde oxidase and xanthine oxidoreductase = CeAO and CeXOR (gi 3877697 and AAB92058); *Calliphora vicina* xanthine oxidoreductase = CvXOR (X07323); *Culex pipiens quinquefasciatus* aldehyde oxidase = CuAO (AF202953); *D. melanogaster* aldehyde oxidase-1, -2, -3, -4 and xanthine oxidoreductase = DmAO1, DmAO2, DmAO3, DmAO4 and DmXOR (GenBank™ accession number AE003709; protein identification numbers: AAF55207.1, AAF55208.2, AAF55209.1 and AAF55210.1; Acc. no. Y00308); human aldehyde oxidase and xanthine oxidoreductase = hAO and hXOR (XM_002522 and NM_000379); mouse aldehyde oxidase and mouse xanthine oxidoreductase = mAO and mXOR (NM009676 and X62932); *Zea mays* aldehyde oxidase-1 and -2 = MaAO1 and MaAO2 (D88451 and D88452); rabbit aldehyde oxidase = raAO (AB009345); rat xanthine oxidoreductase = rXOR (NM_017154).

domain. Purification of mouse AOH1, free of AO and XOR contaminations, is a first and important step toward the crystallization of the enzyme and the characterization of its substrate specificity, two aims that we are currently pursuing.

As to the second aspect of our work, we establish that the *AOH1* and *AOH2* genes map to chromosome 1 in close proximity to the *AO* locus with which they form a molybdo-flavoenzyme gene cluster. Both genes consist of 35 coding exons and give rise to transcripts, which encode single protein products. In fact, reverse transcription-PCR experiments with a large number of appropriate amplimer couples rule out the existence of alternatively spliced forms of the coding portion of the *AOH1* and *AOH2* genes in the tissues and cell types where the two corresponding proteins are synthesized.[3] However, in the liver, the *AOH1* gene is transcribed in two distinct mRNA types, which differ from each other for the presence of a non-coding leader exon (exon 1bis). The predominant mRNA form (type I transcript) contains the transcription product of exon 1bis fused to exon 1 through the use of a rare splicing junction. Interestingly, *AOH1* exon 1bis is located within intron 26 and is transcribed in the opposite direction to that of all the other exons. The unorthodox position of exon 1bis in the *AOH1* gene is demonstrated by long-range PCR and sequencing experiments, following subcloning of the corresponding genomic fragment (data not shown). In addition, the presence of identical DNA sequences in the context of the molybdo-flavoprotein gene cluster on chromosome 1 or somewhere else in the mouse genome is excluded by Southern blot experiments and by a BLAST search of the draft mouse genomic sequence available in GenBank™. Type I mRNA is not an artifact, because it is abundant and was isolated and identified with different types of techniques and in many independent experiments. The molecular mechanism underlying the generation of type I transcripts is unusual and as yet unknown. Recently, the mRNA of the *Drosophila mod* (*mdg4*) gene has been shown to consist of coding exons transcribed from both strands of the DNA (32). Recombination between precursor RNA molecules and *trans*-splicing (32–34) have been postulated as possible mechanisms underlying the phenomenon. Similar mechanisms may be operative also in the case of AOH1 type I transcripts. Nevertheless, our data make it likely that the transcription of *AOH1* is controlled by the activity of two separate promoter elements. Indeed, the 5′-flanking regions of both exon 1bis and exon 1 of the *AOH1* gene contain functional, albeit weak, constitutive promoter elements, as verified by transient transfection experiments with appropriate reporter constructs. The low promoter activity of our constructs may be related to the cell type used in our experiments and needs to be further studied in other cellular contexts.

The structural organization of the *AOH1* and *AOH2* genes is virtually identical to that of mouse *AO* (25) and extremely similar to that of the corresponding *XOR* locus (30). This represents compelling evidence that the four genes are derived from a single ancestral precursor through one or more duplication events. The level of nucleotide identity among the coding regions of mouse *AOH1*, *AOH2*, *AO*, and *XOR* also indicates

[3] M. Terao and E. Garattini, unpublished results.

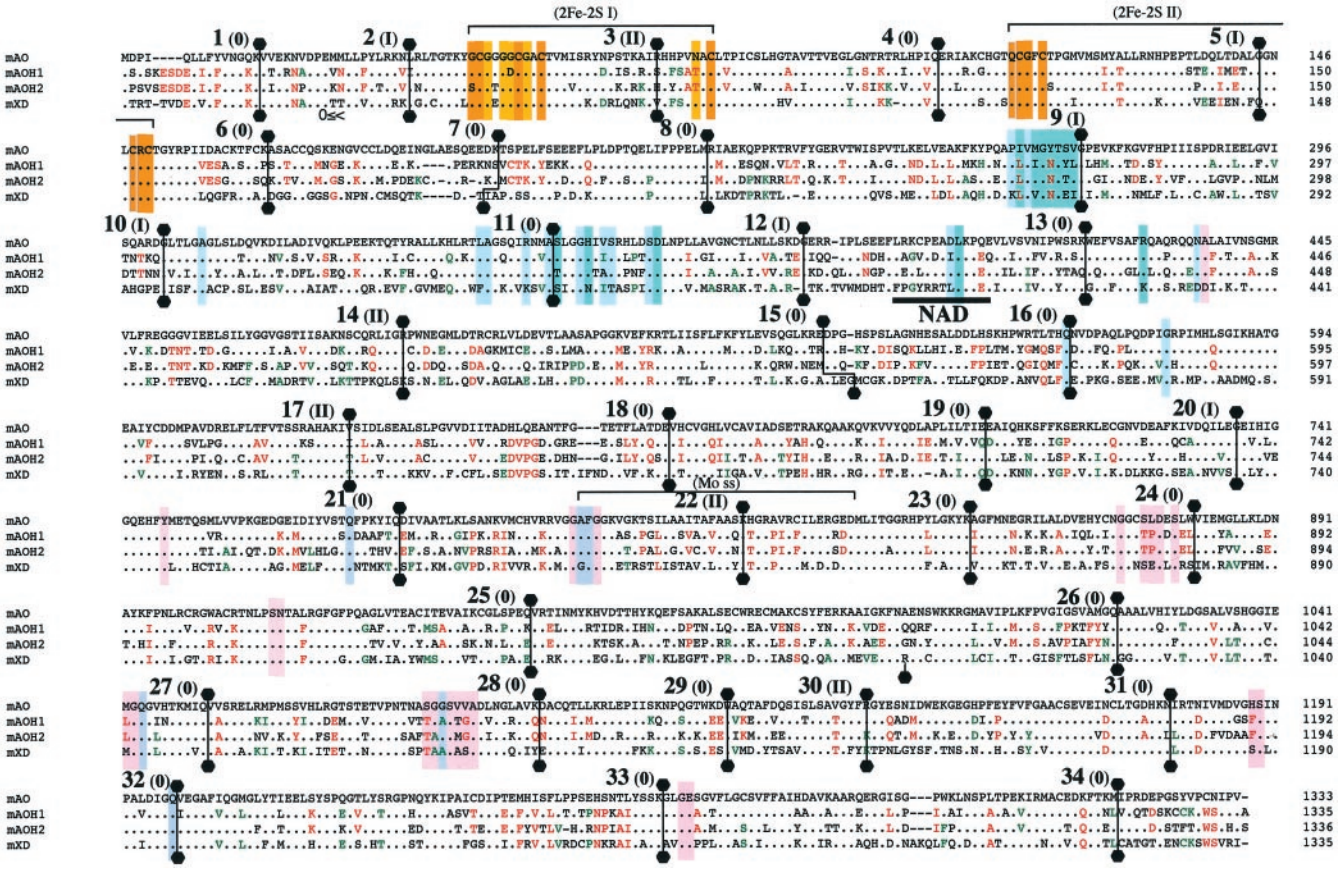

FIG. 8. **Comparison between exon/intron boundaries of the mouse *AO* and *XOR* genes relative to the amino acid sequence of the corresponding proteins.** The amino acid sequence of the mouse AOH1 and AOH2 proteins deduced from the nucleotide sequence of the corresponding genes are aligned with the mouse *AO* (25) and *XOR* (30). Amino acid residues are numbered from the N terminus to the C terminus from the putative first methionine of each sequence. Residues identical to AO are indicated by *dots*. *Hyphens* represent gaps introduced to obtain the best alignment among the four sequences. Amino acids that are identical in AOH1 and AOH2, but not in AO, are *red*. Amino acids that are identical in XOR and AOH1 or AOH2, but not in AO, are *green*. The position of the exon/intron boundaries is indicated by *solid hexagons connected with lines*. On the *left side of each solid hexagon*, only the numbers of the *AOH1*, *AOH2*, and *AO* gene exons are indicated along with the type of junction. Whenever an exon/intron junction is placed after the first (type I), second (type II), or third nucleotide (type 0), this is indicated. The amino acid residues reported to be involved in the formation of the two iron–sulfur centers (2Fe-2S), and the fingerprint sequence observed in molybdenum (*Mo ss*)-containing proteins are indicated. By analogy with the crystal structure of bovine milk XOR (37), amino acids potentially involved in the ionic and hydrophobic interactions with the 2Fe-2S prosthetic groups (*brown*, ionic; *yellow*, hydrophobic) the FAD cofactor (*green*, ionic; *blue*, hydrophobic) and the molybdenum cofactor (*purple*, ionic; *pink*, hydrophobic) are indicated by *colored boxes*. The amino acid sequence reported to be responsible for the binding of NAD in XORs are *underlined* with a *thick line*.

that the last gene is the most divergent and is probably the oldest one. In terms of deduced amino acid sequence, AOH1 is more similar to XOR than AOH2 or AO. This would be compatible with a first duplication of the *XOR* gene into *AOH1* and the subsequent generation of *AO* and/or *AOH2*. The remarkable resemblance of the *AOH1*, *AOH2*, *AO*, and *XOR* genes both in terms of nucleotide sequence and exon/intron structure suggests that the duplication event(s) giving rise to the four genes is (are) relatively recent. Support *for* this theory comes from a comparison of all the known or predicted genetic loci of non-human origin coding for proteins that could be aligned along their entire sequence with mouse AOH1, AOH2, AO, and XOR. In *A. thaliana*, *C. elegans*, or *D. melanogaster*, AO homologs show a significant (of the order of 28–30%) resemblance to the XOR counterparts; however, this is accompanied by limited concordance in the position of the exon/intron junctions of the relative genes. In *A. thaliana*, the genes coding for the two isoenzymes AO3 and AO4 have almost identical intron-exon boundaries; however, they share only one exon/intron junction with *XOR1* and *XOR2*. A similar situation is observed in *D. melanogaster* where only the junction between exons 1 and 2 of *AO1*, *AO3*, and *AO4* is common to XOR. In *C. elegans*, a slightly higher level of convergence is observed, where four junctions are conserved in the *AO* and *XOR* genes. These observations indicate that the *XOR* and *AO* loci in the aforementioned plant and animal species have also a common ancestor and are likely to be the result of duplication events. However, they suggest that the duplications of *XOR* into the *AO* genes observed in plants, nematodes, and fruit flies are events more ancient and independent from those observed in mouse. The original duplication event must have been followed by similar events leading to the appearance of a variable number of *AO* genes in different plant and animal species. Thus, it is likely that the mouse *AOH1*, *AOH2*, and *AO* genes are just structural homologs and not true orthologs of the *AOs* present in more primitive animal species and plants. By contrast, despite a vastly different number of exons, *XOR* genes show a remarkable degree of concordance in the position of many of the exon/intron junctions across the various organisms considered. Mouse *XOR* shows conservation in the position and type of exon/intron boundaries with insects (3/3 with *D. melanogaster*, 3/5 with *B. mori* XOR1, 5/7 with *B. mori* XOR2, and 3/3 with *Calliphora vicina*), nematodes (7/15 with *C. elegans*), plants (7/13 with *A. thaliana*), and even mycetes (1/3 with *Aspergillus nidulans*). Thus, it is possible that the persistence of a minority of identical exon/intron boundaries in mouse *AOH1*, *AOH2*, and *AO* relative to *A.*

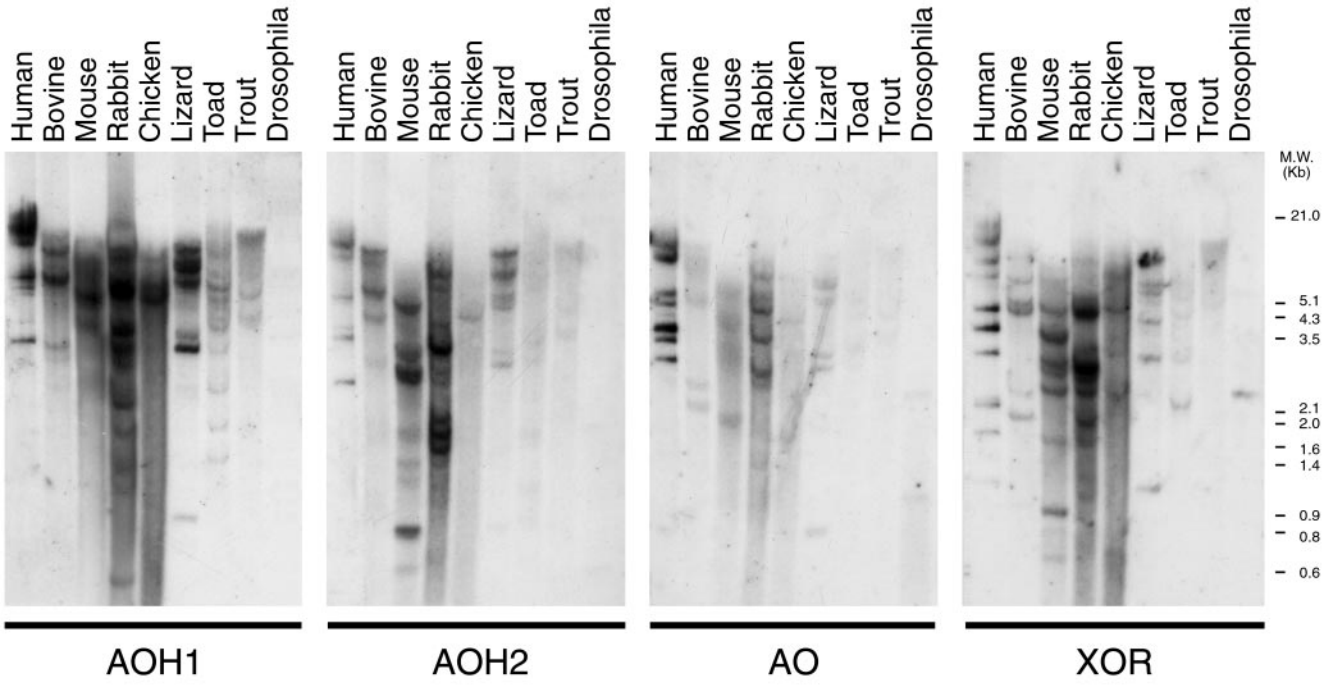

FIG. 9. **Southern blot analysis of genomic DNA from various animal species.** Genomic DNA (20 μg) from the indicated species was cleaved with *Bam*HI endonuclease and subjected to Southern blot analysis with radiolabeled AOH1, AOH2, AO, and XOR full-length cDNAs. The position of DNA molecular weight markers is indicated on the *right*.

thaliana, *C. elegans*, or *D. melanogaster AOs* is the result of a common origin from the corresponding *XORs*.

A key question arising from the identification and characterization of the *AOH1* and *AOH2* genes in the mouse is whether orthologs exist and are expressed in humans and other mammalian or vertebrate species. Our Southern blot experiments indicate the presence of human, rabbit, and bovine genes with structural similarity to *AOH1* and *AOH2* (42, 43–46). In humans, a BLAST computer search in GenBank™ demonstrates the presence of putative exon sequences showing high similarity with portions of the mouse AOH1 and AOH2 cDNAs. Interestingly, these sequences are located at a relatively short distance from the *AO* locus, suggesting the presence, on human chromosome 2, of a molybdo-flavoprotein gene cluster similar to that observed on mouse chromosome 1. At present, it is impossible to reconstruct the entire exon structure of the two human genes, probably as the result of a defective assembly of this genomic region. However, the presence of expressed sequence tags coding for the 3′ portion of these putative human orthologs of AOH1 and AOH2 indicates that they are transcribed and unlikely to be pseudogenes. Given their chromosomal location and clustering with the *AO* locus, the two human genes could be considered as potential candidate genes for the rare motor-neuron disease known as the recessive form of familial amyotrophic lateral sclerosis (47, 48).

In conclusion, besides its importance from a phylogenetic point of view, the cloning of the loci coding for AOH1 and AOH2 described in this report is the first step toward the generation of knockout animals, which are likely to give insight into the functional significance of the two enzymes.

Acknowledgments—We thank Prof. Bruno Curti (University of Milano) for providing hospitality and free access to instrumentation in his laboratory. We are grateful to Prof. Claudio Minoia and Dr. Anna Ronchi (Fondazione S. Maugeri IRCCS, Pavia) for the measurement of molybdenum and iron content in AOH1. We are thankful to Dr. Metodej Kolek and Dr. Ruth Vila-Pont for their skillful help in the project and Prof. Silvio Garattini for critical reading of the manuscript.

REFERENCES

1. Hille, R. (1996) *Chem. Rev.* **96,** 2757–2816
2. Rebelo, J., Macieira, S., Dias, J. M., Huber, R., Ascenso, C. S., Rusnak, F., Moura, J. J., Moura, I., and Romao M. J. (2000) *J. Mol. Biol.* **297,** 135–146
3. Glatigny, A., and Scazzocchio, C. (1995) *J. Biol. Chem.* **270,** 3534–3550
4. Sekimoto, H., Seo, M., Dohmae, N., Takio, K., Kamiya, Y., and Koshiba, T. (1997) *J. Biol. Chem.* **272,** 15280–15285
5. Ori, N., Eshed, Y., Pinto, P., Paran, I., Zamir, D., and Fluhr, R. (1997) *J. Biol. Chem.* **272,** 1019–1025
6. Wright, R. M., Clayton, D. A., Riley, M. G., McManaman, J. L., and Repine, J. E. (1999) *J. Biol. Chem.* **274,** 3878–3886
7. Kurosaki, M., Zanotta, S., Li Calzi, M., Garattini E., and Terao, M. (1996) *Biochem. J.* **319,** 801–810
8. Kurosaki, M., Li Calzi, M., Scanziani, E., Garattini E., and Terao, M. (1995) *Biochem. J.* **306,** 225–234
9. Fabre, G., Seither, R., and Goldman, I. D. (1986) *Biochem. Pharmacol.* **35,** 1325–1330
10. Cates, L. A., Jones, G. S., Jr., Good, D. J., Tsai, H. Y., Li, V. S., Caron, N., Tu, S. C., and Kimball, A. P. (1980) *J. Med. Chem.* **23,** 300–304
11. Krenitsky, T. A., Neil, S. M., Elion, G. B., and Hitchings, G. H. (1972) *Arch. Biochem. Biophys.* **150,** 585–599
12. Terao, M., Kurosaki, M., Saltini, G., Demontis, S., Marini, M., Salmona, M., and Garattini, E. (2000) *J. Biol. Chem.* **275,** 30690–30700
13. Holmes, R. S. (1978) *Comp. Biochem. Physiol.* **61B,** 339
14. Beedham, C., Bruce, S. E., Critchley, D. J., and Rance, D. J. (1990) *Biochem. Pharmacol.* **39,** 1213–1221
15. Shambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
16. Rosenfeld, J., Capdevielle, J., Guillemot, J. C., and Ferrara, P. (1992) *Anal. Biochem.* **203,** 173–179
17. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) *Anal. Chem.* **68,** 850–858
18. Perkins, D. N., Pappin, D. J,. Creasy, D. M., and Cottrell, J. S. (1999) *Electrophoresis* **20,** 3551–3567
19. Clauser, K. R., Baker, P. R., and Burlingame, A. L. (1999) *Anal. Chem.* **71,** 2871–2882
20. Triplett, E. W., Blevins, D. G., and Randall, D. D. (1982) *Arch. Biochem. Biophys.* **219,** 39–46
21. Aliverti, A., Curti, B., and Vanoni, M. A. (1999) *Methods Mol. Biol.* **131,** 9–23
22. Segel, I. H. (1975) in *Enzyme Kinetics*, John Wiley & Sons, New York
23. Saccone, S., Cacciò, S., Perani, P., Andreozzi, L., Rapisarda, A., Motta, S., and Bernardi, G. (1997) *Chromosome Res.* **5,** 293–300
24. Wingender, E., Kel, A. E., Kel, O. V., Karas, H., Heinemeyer, T., Dietze, P., Knuppel, R., Romaschenko, A. G., and Kolchanov, N. A. (1997) *Nucleic Acids Res.* **25,** 265–268
25. Demontis, S., Kurosaki, M., Saccone, S., Motta, S., Garattini, E., and Terao, M. (1999) *Biochim. Biophys. Acta* **1489,** 207–222
26. Hille, R. (1992) In: *Chemistry and Biochemistry of Flavoenzymes* (Mueller, F., ed) Vol. III, pp. 21–68, CRC Press, Boca Raton, FL
27. Komai, H., Massey, V., and Palmer, G. (1969) *J. Biol. Chem.* **244,** 1692–1700

28. Nishino, T., and Nishino, T. (1989) *J. Biol. Chem.* **264,** 5468–5473
29. Sato, A., Nishino, T., Noda, K., Amaya, Y., and Nishino, T. (1995) *J. Biol. Chem.* **270,** 2818–2826
30. Cazzaniga, G., Terao, M., Lo Schiavo, P., Galbiati, F., Segalla, F., Seldin, M. F., and Garattini, E. (1994) *Genomics* **23,** 390–402
31. Burge, C., and Sharp, P. A. (1997) *Cell* **91,** 875–879
32. Labrador, M., Mongelard, F., Plata-Rengifo, P., Baxter, E. M., Corces, V. G., and Gerasimova, T. I. (2001) *Nature* **409,** 1000
33. Stover, N. A., and Steele, R. E. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 5693–5698
34. Vandenberghe, A. E., Meedel, T. H., and Hastings, K. E. (2001) *Genes Dev.* **15,** 294–303
35. Sadek, C. M., Jalaguier, S., Feeney, E. P., Aitola, M., Damdimopoulos, A. E., Pelto-Huikko, M., and Gustafsson, J. (2000) *Mech. Dev.* **97,** 13–26
36. Yokoyama, C., Wang, X., Briggs, M. R., Admon, A., Wu, J., Hua, X., Goldstein, J. L., and Brown, M. S. (1993) *Cell* **75,** 187–197
37. Enroth, C., Eger, B. T., Okamoto, K., Nishino, T., Nishino, T., and Pai, E. F. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 10723–10728
38. The Arabidopsis Genome Initiative. (2000) *Nature* **408,** 796–815
39. Hilliker, A. J., Duyf, B., Evans, D., and Phillips, J. P. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 4343–4347
40. Yasukochi, Y., Kanda, T., and Tamura, T. (1998) *Genet. Res.* **71,** 11–19
41. The Washington University Genome Sequencing Center. (1998) *Science* **282,** 2012–2018
42. Terao, M., Kurosaki, M., Demontis, S., Zanotta, S., and Garattini, E. (1998) *Biochem. J.* **332,** 383–393
43. Xu, P., Huecksteadt, T. P., and Hoidal, J. R. (1996) *Genomics* **34,** 173–180
44. Huang, D. Y., Furukawa, A., and Ichikawa, Y. (1999) *Arch. Biochem. Biophys.* **364,** 264–272
45. Li Calzi, M., Ghibaudi, E., Salmona, M., Cazzaniga, G., Terao, M., and Garattini, E. (1995) *J. Biol. Chem.* **270,** 31037–31045
46. Terao, M., Kurosaki, M., Zanotta, S., and Garattini, E. (1997) *Biochem. Soc. Trans.* **25,** 791–796
47. Wright, R. M., Weigel, L. K., Varella-Garcia, M., Vaitatis, G., and Repine, J. E. (1997) *Redox Rep.* **3,** 135–144
48. Berger, R., Mezey, E., Clancy, K. P., Harta, G., Wright, R. M., Repine, J. E., Brown, R. H., Brownstein, M., and Patterson, D. (1995) *Somat. Cell Mol. Genet.* **21,** 121–131