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

From the §Laboratory of Molecular Biology, Centro Calullo 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 ¶Department of Scienze Chemiche, Fisiche e Matematiche, Università dell’Insubria, via Valleggio 11, Como 22100, the ¶Department of 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 oxidase and aldehyde oxidase. 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 the same 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-
Purification of AOH1 and Cloning of the AOH1 and AOH2 Genes

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 mice weighing 20-30 g (CD-1 strain, Charles River Laboratories, Wilmington, MA) were anesthetized with diethyl ether and sacrificed by cervical dislocation. The livers were excised, rinsed briefly in cold saline, and homogenized in 10 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 and once with 20 ml of 100 mM Tris-glycine buffer, pH 9.0. Following elution with 5 ml of 1 M benzamidine (Sigma Chemical Co., St Louis, MO), the eluate was incubated at 0.5 ml/min with a linear gradient (30 ml) from 0 to 1 M ammonium bicarbonate at 56 °C for 2 h. The resin was washed five additional times with 100 mM Tris-glycine buffer, pH 9.0. The eluate was then dialyzed against 100 mM sodium phosphate buffer, pH 7.5, and frozen as aliquots at −80 °C.

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 trypsin digestion (16, 17). Briefly, proteins or Coomassie Blue-stained gel slices were incubated with 10 mg/ml 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 eluted twice with 50% acetonitrile/methanol and evaporated to dryness. 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, Waltham, MA). 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.

Identification of the flavin cofactor bound to AOH1 was carried out fluorometrically as described previously (21). Briefly, aliquots of AOH1 protein were incubated at 37 °C with 20 μM FAD for 30 min. Following removal of the denatured protein by centrifugation, the centrifugation, the yield 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.

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 trypsin digestion (16, 17). Briefly, proteins or Coomassie Blue-stained gel slices were incubated with 10 mg/ml 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 eluted twice with 50% acetonitrile/methanol and evaporated to dryness. 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, Waltham, MA). 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.

Identification of the flavin cofactor bound to AOH1 was carried out fluorometrically as described previously (21). Briefly, aliquots of AOH1 protein were incubated at 37 °C with 20 μM FAD for 30 min. Following removal of the denatured protein by centrifugation, the centrifugation, the yield 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.

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 trypsin digestion (16, 17). Briefly, proteins or Coomassie Blue-stained gel slices were incubated with 10 mg/ml 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 eluted twice with 50% acetonitrile/methanol and evaporated to dryness. 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).
TABLE I

Purification scheme of mouse liver AOH1

| Step       | Volume | Protein | Phthalazine-oxidizing activity | AOH1 immunoreactive protein |
|------------|--------|---------|--------------------------------|-----------------------------|
|            | ml     | mg      | Total | Specific activity | Purification factor | Yield | Total | Specific activity | Purification factor | Yield |
| Extract    | 25     | 1041    | 6250 | 6.3             | 1                   | 100   | 1090 | 1.0            | 100             |        |
| 55 °C      | 20     | 704     | 6380 | 9.1             | 1                   | 102   | 941  | 1.3            | 39              | 72    |
| (NH₄)₂SO₄ | 20     | 420     | 4090 | 9.7             | 1.5                | 65    | 812  | 1.9            | 74              |        |
| Benzamidine-Sepharose | 1     | 0.54    | 520  | 963.0           | 8.3                | 39    | 722  | 7.2            | 3.6             |        |
| Mono Q     | 0.5    | 0.11    | 1100 | 155.7           | 1.8                | 19    | 1727 | 1.7           | 1.7             |        |

Ten grams of mouse liver was isolated, homogenized in 30 ml of buffer, and ultracentrifuged to obtain a cytosolic extract (Extract), which was processed as described. Enzymatic activity was measured as the ability of the various purification fractions to oxidize phthalaline and concomitantly reduce potassium ferricyanide. The reduction of potassium ferricyanide was monitored at 420 nm. One unit of enzymatic activity is defined as the amount 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.

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 Sv129 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 32P-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 amplifiers.

**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 (38). The purified PCR products were subcloned in pBluescript, and multiple clones were sequenced, to determine the 5’-flanking regions of exons 1bs and 1 of the AOH1 and exon 1 of the AOH2 genes, we constructed the plasmids: p5FR1/1bs-Luc, p5FR1/1bs-Luc, and p5FR2-Luc. To obtain p5FR1/1bs-Luc, a 0.9-kb XbaI-Pro fragment, containing 583 bp of the 5’-flanking region of AOH1 exon 1bs, exon 1 and 221 bp of the corresponding intron, was subcloned in pBluescript (p5FR1/1bs-blue). The insert was released from p5FR1bs-blue by XbaI digestion and inserted into the pGL-Basic Vector (CLONTECH) pre-digested with the same endonucleases. To construct p5FR1-Luc, a 1.7-kb EcoRI-HincII 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 EcoRI, blunted, and digested with Xhol. The resulting fragment was inserted into the pGL-Basic Vector pre-digested with Xhol and Xbal. To prepare p5FR2-Luc, a 1.4-kb XbaI-Apal 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 NotI-Apal, blunted, and inserted into the pGL-Basic Vector, which was pre-digested with Xhol 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 phthalaline-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 phthalaline-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. 1A demonstrates that all the phthalaline-oxidizing activity elutes in correspondence with anti-AOH1 immunoreactivity (WB inset) and a single peak of proteins (see Coomassie staining). As documented in Fig. 1B (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. 1B). In fact, upon cellulose acetate electrophoresis of the Mono Q-active fractions, only the benzaldehyde-oxidizing band a (corresponding to AOH1) (12) is visible, whereas the presence of band b (corresponding to AO) (12) is evident.
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. 1C, 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-
Identification of AOH1 tryptic peptides by MALDI-MS and sequencing of selected peptides by ESI-MS/MS

TABLE II

| m/z     | Charge state | Sequence                      | Identification     |
|---------|--------------|-------------------------------|--------------------|
| 1151.4  | 2+           | EFPQLDFTQLIFFPELMLR (206–224) | TIIHQEFDPTNLQCWECVENSSYYNR |
| 948.0   | 2+           | GLSPWAINSPATAEVI9 (1291–1308) | 958–985            |
| 893.2   | 2+           | TTIAGTLNDLLELK (241–256)      | 265–289            |
| 612.6   | 2+           | TTIAGTLNDLLELK (241–256)      | 958–985            |
| 528.4   | 3+           | TTIAGTLNDLLELK (241–256)      | 1041.7–1042.2      |

The second column number of the first and last residue in the sequence of the identified AOH1 peptides (numbering as in Fig. 8), whereas 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.

Purification of AOH1 and Cloning of the AOH1 and AOH2 Genes

AOH1 protein was trypsinized and subjected to MALDI-MS analysis. MALDI-MS was used to identify tryptic peptides of 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.

| Table II | MALDI-MS M + H+ obs | M + H+ cal | Identification                     | Sequence          |
|----------|----------------------|------------|------------------------------------|-------------------|
| 3376.1   | 3376.6               | 958–985    | TIIHQEFDPTNLQCWECVENSSYYNR         | EFPQLDFTQLIFFPELMLR (206–224) |
| 2430.4   | 2429.8               | 205–224    | EFQPLDFTQLIFFPELMLR                | GLSPWAINSPATAEVI9 (1291–1308) |
| 2301.7   | 2301.7               | 206–224    | KEFPQPLDFTQLIFFPELMLR              | TTIAGTLNDLLELK (241–256) |
| 1949.0   | 1949.3               | 259–276    | HSPAILVGNGLLHMK                    | 958–985           |
| 1896.3   | 1896.2               | 1291–1308  | GLSPWAINSPATAEVI9 (1291–1308)      | 1041.7–1042.2     |
| 1869.5   | 1869.0               | 225–240    | MAEESQNTVLTFGER2                  | TTIAGTLNDLLELK (241–256) |
| 1579.4   | 1579.8               | 277–290    | FTDVSPHIISPAR                      | 958–985           |
| 1568.2   | 1568.7               | 988–999    | AVDFENQIQRFWK                      | 23–35             |
| 1564.4   | 1564.8               | 23–35      | ESDELIFVNYK                        | 1041.7–1042.2     |
| 1526.3   | 1526.7               | 225–237    | MAEESQNTVLTFGER2                  | TTIAGTLNDLLELK (241–256) |
| 1461.2   | 1461.6               | 1112–1123  | QNPSTGWEWVK                        | 958–985           |
| 1358.3   | 1358.6               | 824–834    | TGRPIRFLER4                        | 1041.7–1042.2     |
| 1344.0   | 1344.5               | 419–429    | SSKWEPVSAFR                        | 958–985           |
| 1223.8   | 1224.4               | 351–363    | NVASLGGHSSR                        | 1041.7–1042.2     |
| 1164.7   | 1185.3               | 1218–1227  | YSPFQVLTYR                         | 958–985           |
| 1120.9   | 1121.2               | 988–996    | TNLPMNTAFR                         | 1041.7–1042.2     |
| 1060.8   | 1072.7               | 988–996    | AVDFENQQR                          | TTIAGTLNDLLELK (241–256) |
| 1041.7   | 1042.2               | 422–429    | WEFVSAFR                           | 958–985           |
| 887.0    | 887.0                | 343–350    | TLAGQQR                            | 1041.7–1042.2     |
| 825.66   | 825.9                | 948–953    | ELNMYR                             | 1041.7–1042.2     |
| 805.0    | 805.0                | 1051–1057  | MIQVAASR                           | TTIAGTLNDLLELK (241–256) |

The use of genomic probes labeled with two different fluorescent labels allowed the chromosome 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 chromosomes G1 and G2. Three signals map to the same region, where the AOH gene was previously localized. (25). The use of genomic probes containing the AOH1 (Fig. 3C) and AOH2 (Fig. 3G) 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-
rescent tags were performed. Fig. 3 demonstrates that the AOH1 and AO hybridization signals are coincident. As shown in Fig. 3H, 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. 4A, 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. 4B, 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. 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.](http://www.jbc.org/)

![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.](http://www.jbc.org/)
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: Ser261 → Asn, Leu530 → Val, and Arg541 → 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: Arg1151-Lys1152-Val1153-Thr1154-Phe1155 to Glu1151-Gly1152-Asp1153-Ile1154. 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 (Arg106 → Gln, Gly326 → Glu, Arg336 → His,

![Physical map of the AOH1 and AOH2 loci.](http://www.jbc.org/Downloaded from http://www.jbc.org/)

**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.
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.

### Table III

Exon–intron organization of the mouse AOH1 gene

| Exon | Exon size (bp) | 5′Splice site | Intron size (Kb) | 3′Splice site |
|------|---------------|---------------|-----------------|---------------|
| 1bis | 76 (90)       | ...cccttgctagggtagctagcag.. | 7.0 | ...agggcagaggcagagagcat.. |
| 1   | 177 (135)     | ...Gly Lys Lys | 1.6 | ...ctctctcagcagacc GAG.. |
| 2   | 58            | ...Glu Val I | 3.6 | ...acccctcagcagcagat CGA CTC.. |
| 3   | 97            | ...Arg Ile Se | 312 | ...His Phe.. |
| 4   | 109           | ...TTT GTC CAGgtagccgt.. | 1.4 | ...ttttcatttggGAA AAG ATT.. |
| 5   | 127           | ...CTG GCC Ggtatggtgat.. | 4.3 | ...cttcctcctcAGG AAT CTA.. |
| 6   | 62            | ...TTT GTC CGGTgctgtgat.. | 2.5 | ...ttttcccacagag TCA ACT.. |
| 7   | 81            | ...Glu Asn Ser | 789 | ...Val Cys Thr.. |
| 8   | 81            | ...GAG CTG ATGgtgtgtctg.. | 3.0 | ...aatattgcatGAG ATG GCC.. |
| 9   | 145           | ...Lys Ala Gln | 1.1 | ...ttctctctcctGAG CTT CAT.. |
| 10  | 93            | ...AAG CAA Ggtatgcttctt.. | 3.5 | ...taagctcctcagag CTG ACA.. |
| 11  | 152           | ...Mak Val Ala | 1107 | ...Ser Leu Gly.. |
| 12  | 94            | ...TTA ACA Ggtatgct.. | 7.0 | ...tttttttgcagttc CTT GTA.. |
| 13  | 110           | ...Ser Ser Lys | 1346 | ...lu Gly Ile.. |
| 14  | 185           | ...TCC ACC AAAGtactgatc.. | 2.3 | ...cattttcgctcag AGA GAA.. |
| 15  | 163           | ...Ile Gly Ar | 1454 | ...Trp Glu Phe.. |
| 16  | 1802          | ...ATT GGA AGGgacttaca.. | 0.5 | ...cttgtctctg TGC TGG.. |
| 17  | 1895          | ...Glu Thr Thr | 1640 | ...Asg Pro His.. |
| 18  | 2066          | ...AAA ATC ATYtacttat.. | 0.6 | ...tttttttcagt CTT CT.. |
| 19  | 2180          | ...CAG GAT GAGgtgtgct.. | 0.8 | ...ttttcccacagag TGG CTC.. |

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 amplifiers were performed on RNA extracted from adult and embryonic liver. Fig. 5A shows one representative study on the AOH1 transcript(s). Following amplification with AP1 (Anchor
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. 5A. 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. 5B) and
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. 6C 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).
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 four different putative isoenzymatic forms of AO besides XOR, which is the product of the well known rosy locus (39). Additionally, 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. 7A. 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. 7B. This scheme indicates that AOH1 and AOH2 proteins represent a subgroup of mammalian AOs. 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
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 directly 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 GenBankTM (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.
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$^-$/H$_2$O$_2$, suggesting the presence of cyanolyzable sulfur in the catalytic

---

**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 (data 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.
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 substructure 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 genes map to chromosome 1 in close proximity to the AO locus (30). This representation of 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 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 transsplicing (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 1a 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

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 (grey 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 oxidase-1, -2, -3, -4 and xanthine oxidoreductase-1 and -2 (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, -2, -3, -4 and xanthine oxidoreductase = DmAO1, DmAO2, DmA03, DmA04 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); 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 = DmA01, DmA02, DmA03, DmA04 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).
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 XOR (25) and AOH1 (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 (blue, 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 nonhuman 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, they are not as completely conserved in the position of the exon/intron junctions of the relative genes. In A. thaliana, the genes coding for the two isoenzymes A03 and A04 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, A03, and A04 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 intron/exon 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 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.
**Purification of AOH1 and Cloning of the AOH1 and AOH2 Genes**

**Fig. 9. Southern blot analysis of genomic DNA from various animal species.** Genomic DNA (20 μg) from the indicated species was cleaved with BamH1 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* AO-*s* is the result of a common origin from the corresponding XOR-*s*.

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 Minoa 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. Glattigny, A., and Scazzocchio, C. (1995) *J. Biol. Chem.* 270, 3554–3556
4. Sekimoto, H., Soo, M., Dohmae, N., Takiz, K., Kamiya, Y., and Koshiha, 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., Sarautia, S., Li Cali, M., Garattini E., and Terao, M. (1996) *Biochem. J.* 319, 811–810
8. Kurosaki, M., Li Cali, M., Scanzani, 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, 1235–1330
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. (1986) *J. Med. Chem.* 29, 300–304
11. Krebszky, 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.* 59, 1213–1221
14. Aliverti, A., Curti, B., and Vanoni, M. A. (1999) *Methods Mol. Biol.* 131, 9–23
15. Segel, I. H. (1975) in *Enzyme Kinetics*, John Wiley & Sons, New York
16. Aliverti, A., Curti, B., and Vanoni, M. A. (1999) *Methods Mol. Biol.* 131, 9–23
17. Segel, I. H. (1975) in *Enzyme Kinetics*, John Wiley & Sons, New York
18. Perkins, D. N., Pappin, D. J., Creasy, D. M., and Cottrell, J. S. (1999) *Electrophoresis* 20, 3551–3567
19. Clauer, 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., Acieo, C., Piacentini, F., Andreozzi, L., Rapisarda, A., Motta, S., and Bernardi, G. (1997) *Chromosoma 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 Kelchhauer, N. A. (1997) *Nucleic Acids Res.* 25, 265–268
25. Demontis, S., Kurosaki, M., Sacco, S., Motta, S., Garattini, E., and Terao, M. (1999) *Biochim. Biophys. Acts* 1490, 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
Purification of AOH1 and Cloning of the AOH1 and AOH2 Genes

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, 390–402
33. Stover, N. A., and Steele, R. E. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 5693–5698
34. Vandenberghhe, 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. (1998) Cell 75, 187–197
37. Erneth, 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., Ghiaudi, 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., Varela-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
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 MOLYBDENUM-FLAVOPROTEIN GENE CLUSTER ON MOUSE CHROMOSOME 1

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

J. Biol. Chem. 2001, 276:46347-46363.
doi: 10.1074/jbc.M105744200 originally published online September 18, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M105744200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 46 references, 18 of which can be accessed free at http://www.jbc.org/content/276/49/46347.full.html#ref-list-1