J. Biochem. 119, 926-933 (1996)

Sterol 14-Demethylase P450 (P45014DM*) Is One of the Most Ancient and Conserved P450 Species[1]

Yuri Aoyama,* Mitsuhide Noshiro,† Osamu Gotoh,‡ Susumu Imaoka,§ Yoshihiko Funae,§ Norio Kurosawa,* Tadao Horiuchi,* and Yuzo Yoshida||,[2]

*Department of Bioengineering, Faculty of Engineering, Soka University, Hachioji, Tokyo 192; †Department of Biochemistry, Hiroshima University School of Dentistry, Hiroshima 734; ‡Department of Biochemistry, Saitama Cancer Center Research Institute, Saitama 362; §Department of Chemistry, Osaka City University School of Medicine, Osaka 545; and ||School of Pharmaceutical Sciences, Mukogawa Women's University, 11-68 Koshien-Kyubancho, Nishinomiya 663

Received for publication, December 25, 1995

To determine the orthology of sterol 14-demethylase (P45014DM), the only known P450 enzyme distributed widely in eukaryotes with a conserved metabolic role, the full-length amino acid sequences of rat and human P45014DMs were determined from the cloned cDNA sequences, and compared with those of the corresponding fungal proteins (CYP51). The amino acid identity value between given pairs of P45014DMs ranged from 93% (human/rat) to 39% (human or rat/*Saccharomyces cerevisiae*). All the P45014DMs formed a single cluster in a phylogenetic tree constructed from representative P450 protein sequences currently available. The nearest neighbors to the P45014DM cluster in the phylogenetic tree were CYP7 (cholesterol 7α-hydroxylase) and CYP8 (prostacyclin synthase), and the divergence point of fungal and mammalian P45014DMs was clearly more recent than that of P45014DM and CYP7/CYP8. These lines of evidence show that fungal and mammalian P45014DMs are really orthologous. This is the first example of orthologous P450s occurring in distinct kingdoms. P45014DM may be an ancient P450 which arose before the divergence of major eukaryotic branches and has been conserved throughout evolution. The amino acid identity value (93%) between human and rat P45014DMs was comparable to those observed for some housekeeping enzymes. In addition, a processed pseudogene of P45014DM was found in a rat genomic DNA library, suggesting the expression of P45014DM in germ line cells. These facts suggest that P45014DM may be a housekeeping enzyme essential for the viability of mammals.

Key words: molecular evolution, orthologous gene, P450, processed pseudogene, sterol 14-demethylase.

Monooxygenase P450 comprises a gigantic gene superfamily of several hundred or more molecular species distributed widely in most eukaryotes and a few prokaryotes (*1*). Through comparison of encoded amino acid sequences, it has been revealed that P450 has undergone extensive divergence since an early stage of the evolution, with the formation of a wide variety of monooxygenases showing considerably different substrate specificities (*1-3*). Such extensive divergence producing various functionally different monooxygenases is considered to be the most outstanding characteristic of the P450 superfamily (*4*). However, such characteristics have made it difficult to identify the orthologous nature of P450s occurring in different organisms, and there is no confirmative information about the orthology of P450s occurring in distinct kingdoms.

Sterol 14-demethylase P450 (P45014DM or CYP51 for the fungal enzyme) (*5, 6*) occurs in different kingdoms, such as fungi, higher plants, and animals, with the same metabolic role, *i.e.*, removal of the 14-methyl group of sterol precursors such as lanosterol (*4, 7, 8*), and this is the only known P450 distributed widely in eukaryotes with essentially the same metabolic role. Therefore, P45014DM is the most interesting subject for examining the orthology of P450s. However, there is another possibility that P45014DMs occurring in different kingdoms originated independently and acquired similar catalytic properties through convergence. The most orthodox way for estimating the orthology of genes occurring in different organisms is to compare the primary structures of their products, together with related protein sequences. The primary structures of three yeast CYP51s were revealed some years

[1] Sterol 14-demethylase P450 of fungi was designated as CYP51 by the P450 Nomenclature Committee organized by Drs. D.W. Nebert and D.R. Nelson [Nelson, D.R. *et al.* (1993) *DNA Cell Biol.*, **12**, 1-51], but CYP50s are allocated as fungal P450s. Therefore, in this paper, "CYP51" is used only for the fungal enzyme, and the trivial name, "P45014DM," is used for the mammalian enzyme and also as the general abbreviation of this enzyme. The sequence data reported in this paper have been entered in the DDBJ/EMBL/GenBank databases under the following accession numbers: D55653 for human P45014DM; D29962 and D55681 for rat P45014DM; D78370 for rat P45014DM pseudogene.

[2] To whom correspondence should be addressed. Email: met10540@miet.mukogawa-u.ac.jp

Abbreviations: ORF, open reading frame; P45014DM, sterol 14-demethylase P450.

ago (*9-12*). In the preceding communication (*13*), we presented preliminary evidence suggesting the occurrence of a P450 showing high amino acid sequence similarity to CYP51 in rat liver, although its identification as P45014DM was incomplete. This paper presents lines of confirmative evidence indicating the orthologous nature of mammalian P45014DM and fungal CYP51, obtained by comparing the complete amino acid sequences of rat and human P45014DMs with those of fungal CYP51s. This is the first report describing confirmative evidence for the existence of orthologous P450 in as widely divergent species as unicellular eukaryotes and mammals. This paper also presents a preliminary evidence showing the occurrence of a processed pseudogene of P45014DM in the rat genome, suggesting the housekeeping-enzyme-like nature of P45014DM in mammals.

MATERIALS AND METHODS

Purification of P45014DM from Rat Liver Microsomes—The starting material was a stored preparation (3.2 nmol P450/mg protein) of a pass-through fraction obtained on the DEAE-5PW (TOSOH) HPLC (*14*). This material was subjected to preparative cation-exchange HPLC (*14*), and the fractions showing lanosterol 14-demethylase activity were collected. To remove CYP2C11 (P450UT-2), the pooled eluate was treated with anti-CYP2C11 antibodies according to the previously described method (*16*). Briefly, rabbit anti-CYP2C11 antibodies (*15*) conjugated with Protein A-Sepharose CL-4B (Pharmacia) were added to the eluate and the resultant suspension was allowed to stand overnight at 4°C with gentle shaking. The supernatant showing lanosterol 14-demethylase activity was further purified by HPLC on a hydroxylapatite column (KB-column; Koken) according to the previously described method (*14*). The fractions showing high lanosterol 14-demethylase activity were collected and Emulgen 911 was removed with a small column of BioGel HT (Bio-Rad) (*14*).

Molecular Cloning and Sequencing of the Full-Length cDNAs of Rat and Human P45014DM—The methods for immunoscreening of a rat liver cDNA expression library in λgt11 (*17*) and cloning of a cDNA (pRT-9) encoding part of a protein similar to yeast P45014DM were described in the preceding communication (*13*). The same library as above was screened with the *Eco*RI-*Pvu*II fragment of the pRT-9 cDNA. Hybridization and detection were carried out using the ECL direct nucleic acid labeling and detection system (Amersham). The cDNA inserts of two positively reacting clones (pRT-10 and pRT-11) were subcloned into the *Eco*RI site of the pBluescript SK(−) plasmid, and then subjected to restriction enzyme mapping and nucleotide sequencing.

A λgt11 cDNA library of human liver (*18*) was screened with ^{32}P-labeled pRT-11 cDNA. The cDNA inserts of five positively reacting clones were subcloned into the *Eco*RI site of the pBluescript SK(−) plasmid, and then subjected to restriction enzyme mapping. One of them, having the longest insert (3.1 kbp), was subjected to nucleotide sequencing.

Double-strand DNA sequencing was carried out by the dideoxy chain termination method using a *Bca*BEST™ dideoxy sequencing kit (Takara) with FITC-labeled primers. The nucleotide sequence was analyzed with a HITACHI SQ-3000 DNA sequencer.

Expression of a Putative P45014DM cDNA in COS7 Cells—Expression of a putative rat P45014DM cDNA in COS7 cells was performed by the method described previously (*18*). Briefly, the *Xho*I-*Xba*I fragment (2.0 kbp) of pRT-10 was ligated to the pSVL vector. The recombinant plasmids were transfected into COS7 cells by electroporation. The cells were maintained on Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum and harvested 3 days after the transfection. The cells were washed once with phosphate-buffered saline (pH 7.4), suspended in 0.25 M sucrose-containing 50 mM potassium phosphate buffer (pH 7.4), 1 mM EDTA, and 1 μg/ml each of leupeptin and pepstatin A, and then homogenized by ultrasonication, followed by Potter homogenizer treatment. The microsomal fraction of COS7 cells was prepared by the usual differential centrifugation and microsomes were suspended in the medium used for the homogenization. Lanosterol 14-demethylase activity of COS7 microsomes was assayed by the previously described method (*19*) except that purified rat liver NADPH-P450 reductase was added to the reaction mixture. The lanosterol demethylase activity due to liver microsomes slightly contaminating the cytosolic fraction was corrected for.

Identification of a Processed Pseudogene of P45014DM in a Rat Genomic DNA Library—A rat genomic DNA library (*20*) was screened with pRT-11 cDNA. One of the 11 positive clones, having 17-kbp DNA insert (RG3), was digested with *Sal*I plus *Eco*RI. The resulting fragments were separated on an agarose gel, transferred to Nytran™ membranes, and then subjected to Southern hybridization with *Eco*RI-*Pst*I (140 bp, 5′-end) and *Pst*I-*Eco*RI (600 bp, 3′-end) fragments of pRT-11 cDNA. Only one fragment of 2.3 kbp hybridized with both of these probes. This *Eco*RI-*Eco*RI fragment (2.3 kbp) was subcloned into the pBluescript SK(−) plasmid and sequenced. Other experimental conditions were given in the previous paper (*20*).

Amino Acid Sequencing of the Cyanogen-Bromide-Digested Fragments of Rat P45014DM—An aliquot (30 μg protein) of the purified preparation of rat liver P45014DM described above was dialyzed overnight against pure water. The dialyzed specimen was dried up and then digested overnight at room temperature with 200 μl of a cyanogen bromide solution (10 mg/ml of 80% formic acid). The digested preparation was dried up and the resulting polypeptide fragments were separated by HPLC on a Beckman Ultrasphere column (2.0×150 mm). The elution was conducted with a linear gradient from 0.1% trifluoroacetic acid to 70% acetonitrile containing 0.1% trifluoroacetic acid over 30 min at the flow rate of 0.5 ml/min. Five distinct polypeptides were obtained. Their amino acid sequences were analyzed with an Applied Biosystems gas-phase protein sequencer.

Amino Acid Sequence Alignment and Construction of a Phylogenetic Tree—Amino acid sequence alignment of rat and human P45014DMs together with three yeast CYP51s (*10-12*) was performed with a program adopting the randomized iterative refinement strategy, that optimizes the weighted sum-of-pairs scores (*21*). To construct a phylogenetic tree showing the evolutionary relationship of P45014DM to other P450 families and subfamilies, single representative sequences were arbitrarily selected from each of the 63 families or subfamilies of E-class (eukaryote-class) P450s (*3*), and were aligned together with

P45014DM sequences using the same program as above. All sequence data, except for those for human and rat P45014DMs described here, were taken from publicly available nucleotide or protein sequence databases. A distance matrix was calculated from the multiple sequence alignment based on the PAM model (*22*), and a phylogenetic tree was constructed by the neighbor-joining method (*23*).

RESULTS AND DISCUSSION

Purification of Rat Liver P45014DM—The breakthrough in the purification of P45014DM from rat liver described here was the removal of CYP2C11 (P450UT-2) by immunoabsorption. A method for the purification of P45014DM from rat liver microsomes has already been reported (*24*). However, it was hard to obtain a P45014DM preparation showing a turnover number of over 4 nmol/min/nmol P450. It was found on Western blotting with antibodies against various P450s (*15*) that such P45014DM preparations were contaminated by CYP2C11, one of the major P450 components of untreated male rats. Efforts to remove contaminating CYP2C11 from P45014DM preparations by HPLC under various conditions were unsuccessful. The major use of the purified P45014DM in this work was as an antigen for raising anti-P45014DM antibodies necessary for the immunoscreening of the rat liver cDNA expression library, and CYP2C11 is one of the most highly expressed P450s in untreated male rat liver. Therefore, contamination by CYP2C11 of an antigen for raising antiserum for immunoscreening must strictly be avoided.

The purification technique was essentially the same as that used for the simultaneous purification of P450s in rat liver microsomes (*14*), except that immunoabsorption with anti-CYP2C11 antibodies was added. The purified P45014DM preparation obtained by this method showed lanosterol 14-demethylase activity of 9.8 nmol/min/nmol P450 upon reconstitution with rat liver NADPH-P450 reductase. No protein band assignable as an other P450 was detected on SDS-PAGE. No protein band reacted with anti-CYP2C11 antibodies upon Western blotting (data not shown). Thus, the P45014DM preparation obtained here was suitable as the antigen. This P45014DM preparation also served as the material for cyanogen bromide digestion to determine the partial amino acid sequence of P45014DM.

Cloning and Sequencing of Putative Rat P45014DM cDNA—The results of the molecular cloning of a cDNA (pRT-9) containing an incomplete ORF encoding a protein having an amino acid sequence similar to that of yeast P45014DM were described in the preceding communication (*13*). To obtain full-length cDNA of this protein, the same cDNA library (*17*) was screened with the *Eco*RI-*Pvu*II fragment of pRT-9 cDNA, and two positive clones (pRT-10 and pRT-11) were isolated. The nucleotide sequence of the pRT-11 cDNA was determined (2,270 bp; sequence data, DDBJ D55681, restriction enzyme map, see Fig. 4), and a complete ORF (1,512 bp) between ATG (118-120) and TGA (1627-1629) was identified. This ORF encodes a protein consisting of 503 amino acids (see the first line of Fig. 2), and the sequence of its 430 amino acids (from alanine 74 to lysine 503) is identical to that of the pRT-9 protein (*13*). The same ORF was also identified in the pRT-10 cDNA, though the 64 nucleotides from the 5′-end of pRT-11 were absent in pRT-10 cDNA (data not shown).

Identification of pRT-10 and pRT-11 as P45014DM cDNA—The *Xho*I-*Xba*I fragment (2.0 kbp) of pRT-10 cDNA was ligated to the pSVL vector and then expressed in COS7 cells as described under "MATERIALS AND METHODS." The microsomal fractions of recombinant-pSVL transfected cells, control cells transfected only with the vector, and non-transfected COS7 cells were isolated, and their lanosterol 14-demethylase activities were determined. Conversion of lanosterol to its 14-demethylated product, 4,4-dimethylcholesta-8,14,24-trienol (*5, 6*), by the microsomes of the recombinant-pSVL transfected cells was confirmed by gas-chromatographic analysis of sterols extracted from the reaction mixture (Fig. 1), indicating expression of P45014DM in the transfected cells. The apparent lanosterol 14-demethylase activity calculated from the chromatogram (Fig. 1) was 10.9 pmol/min/mg protein. The lanosterol demethylase activity of the microsomes from non-transfected or pSVL-transfected cells was significantly lower than the above value. The highest activity obtained with these microsomes was 4.2 pmol/min/mg protein, and in some experiments no demethylated product was detected (data not shown). These observations indicated that the lanosterol 14-demethylase activity of COS7 cells was apparently elevated by the transfection of pRT-10/pSVL. This fact strongly indicated that the product of the pRT-10 ORF was catalytically active P45014DM.

Five polypeptide fragments were obtained on cyanogen bromide digestion of the purified rat liver P45014DM, and the amino acid sequences of their N-terminal parts were analyzed as described under "MATERIALS AND METHODS." The N-terminal amino acid sequences of these five polypeptides were detected downstream of methionines in the

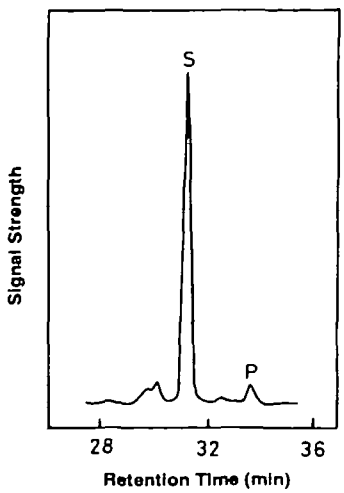

Fig. 1. **Gas-chromatographic detection of the 14-demethylated metabolite of lanosterol formed on incubation with microsomes of pRT-10 transfected COS7 cells.** Lanosterol (23.5 nmol) was incubated aerobically at 37°C for 60 min with microsomes of pRT-10 transfected COS7 cells (1.56 mg protein) in the presence of an NADPH-generating system (*6*), purified rat NADPH-P450 reductase (0.5 unit) and rat liver 105,000×*g* supernatant (3.0 mg protein). Sterols in the reaction mixture were extracted after saponification (*6*) and analyzed by GLC as described previously (*6*). Peak P was identified as the demethylated product, 4,4-dimethylcholesta-8,14,24-trienol (*6*), from its relative retention time as to the lanosterol peak (peak S).

amino acid sequence deduced from the ORF of pRT-11 cDNA (overlines 1 through 5 in Fig. 2). This was strong evidence that the pRT-11 insert was P45014DM cDNA. However, efforts to determine the N-terminal amino acid sequence of the native P45014DM protein were unsuccessful. This suggested that the N-terminal of the purified P45014DM preparation might be blocked.

Amino Acid Sequence of Human P45014DM—Using pRT-11 cDNA as a probe, a cDNA clone having a 3.1 kbp insert was isolated from a human liver cDNA library (*18*), and its nucleotide sequence was determined (sequence data, DDBJ D55653). The nucleotide sequence of the ORF

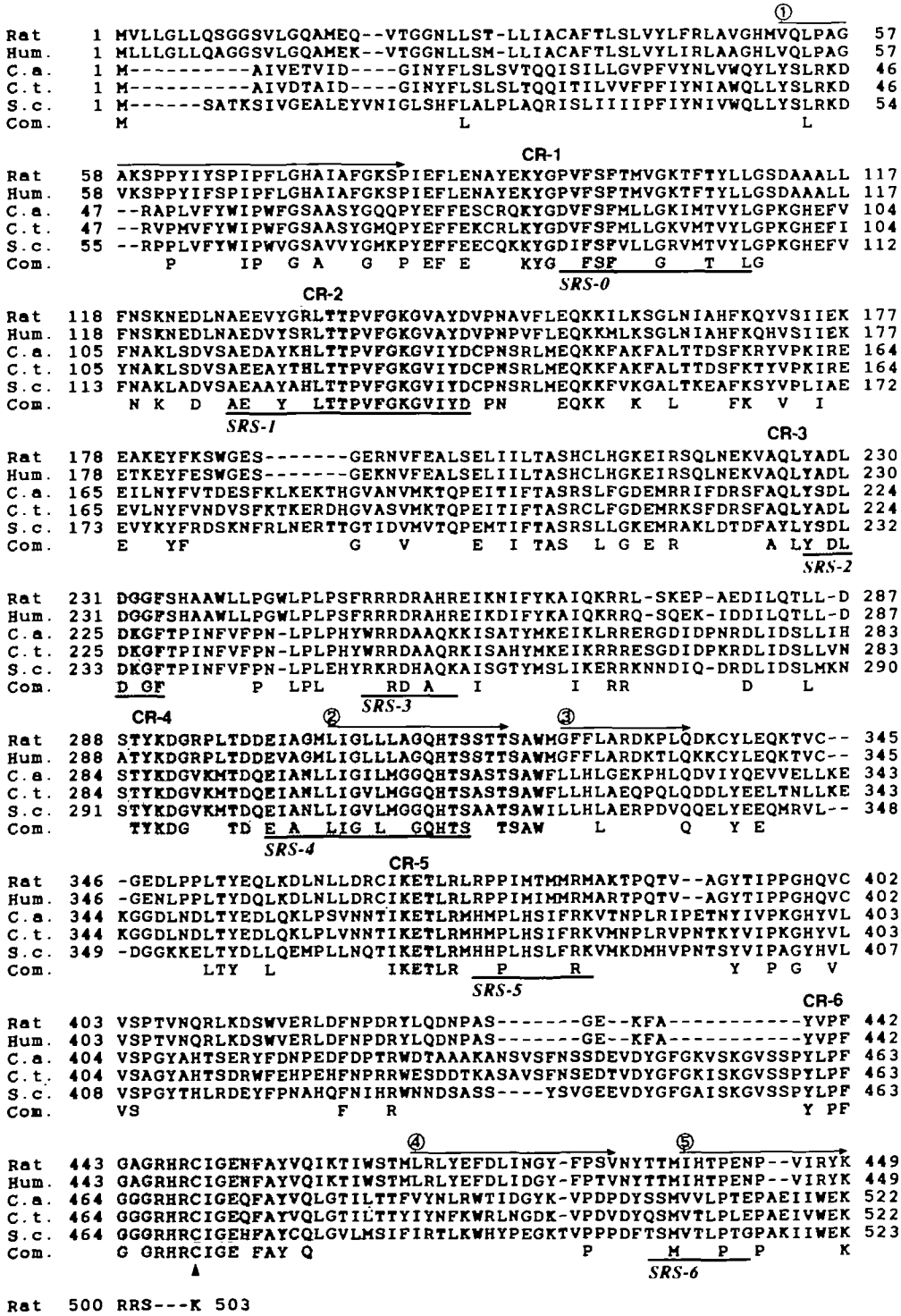

Fig. 2. **Amino acid sequences of rat and human P45014DMs, and their alignment with those of three yeast CYP51s.** The amino acid sequences of rat (*Rat*) and human (*Hum.*) P45014DMs were deduced from the nucleotide sequences of their cDNAs. These sequences were aligned together with those of three yeast CYP51s (*10–12*) by means of a program adopting the randomized iterative refinement strategy that optimizes the weighted sum-of-pairs scores (*21*). *S.c.*, *C.t.*, and *C.a.* denote *Saccharomyces cerevisiae*, *Candida tropicalis*, and *Candida albicans*, respectively. The bottom line (*Com.*) shows the amino acid residues common in all sequences. Lines 1 through 5 over the rat sequence indicate the amino acid sequences corresponding to the N-terminal parts of the five polypeptides obtained on cyanogen bromide-digestion of the purified rat liver P45014DM. Shaded regions (CR-1 through 6) are highly conserved regions. Underlines (SRS-0 through 6) indicate the regions corresponding to the putative substrate recognition sites (*27*). The arrowhead in CR-6 denotes the cysteine providing the thiolate heme ligand.

(1,530 bp) of this cDNA showed 87% identity to that (1,512 bp) of pRT-11, indicating its identity to human P45014DM cDNA. The ORF of human P45014DM cDNA encodes a protein consisting of 509 amino acids. However, another ATG was found at 18 bp downstream of the formal initiation ATG. If the downstream ATG is assumed to be the translation initiation point, the number of amino acid residues of the encoded protein becomes 503. This shorter sequence of human P45014DM was completely aligned with the sequence of rat P45014DM, as shown in the first and second lines of Fig. 2, the amino acid identity between them being found to be 93%. Thus, we tentatively employed the shorter sequence as the primary structure of human P45014DM. If human P45014DM has the longer sequence (509 amino acids), a 6 amino acid fragment (MAAAAG) that cannot be aligned with the rat sequence extends from the N-terminal of the sequence shown in the second line of Fig. 2.

Amino Acid Sequence Identity between Given Pairs of P45014DMs—Table I summarizes the amino acid sequence identities between given pairs among two mammalian P45014DM and five fungal CYP51s (*10, 11*; EMBL Z48164 and EMBL Z49750). The identity between mammalian P45014DM and fungal CYP51s (39-42%) is comparable to those observed for corresponding mammalian and fungal housekeeping enzymes, such as glucose-6-phosphate dehydrogenase (46.5%, human/yeast), δ-aminolevulinic acid synthase (40.2%, human/mold), and HMG-CoA reductase (38.6%, human/yeast). Therefore, P45014DM is considered to be strongly conserved in distinct kingdoms such as fungi and mammals.

The degree of sequence conservation between rat and human P45014DMs (93%) is also noteworthy since it is the highest among those observed for all known pairs of rodent and human P450s. A NBRF-PIR protein sequence database search suggested that most of the enzymes showing such high similarity values (>90%) between rodents and primates were orthologous pairs of housekeeping enzymes (Table II). It can thus be concluded that P45014DM is conserved as strongly as housekeeping enzymes in mammals as well as between mammals and fungi.

Evolutionary Relationship of P45014DM and Other P450 Families—Single representative sequences were selected from each of the 63 families and subfamilies of E-class P450s (*3*), and aligned together with eight known P45014DM sequences. A phylogenetic tree (Fig. 3) consisting of these 71 P450s constructed by the neighbor joining method (*23*) clearly indicates that mammalian P45014DMs and fungal CYP51s form a single cluster, and the divergence point of mammalian P45014DMs and fungal CYP51s was clearly more recent than the divergence point of P45014DM and its closest neighbors, CYP7 (cholesterol 7α-hydroxylase) and CYP8 (prostacyclin synthase). This evidently indicates that P45014DM was established as a distinct family before the divergence of primordial animals and fungi, and has been conserved throughout evolution, supporting the idea that mammalian P45014DMs and fungal CYP51s are true orthologues. This may be a unique evolutionary characteristic of P45014DM, since most other P450 species occur in limited classes of organisms and participate in various distinct metabolic processes characteristic of individual classes of organisms.

Functional Regions in Mammalian and Yeast P45014DMs—Figure 2 shows the amino acid sequence alignment of two mammalian P45014DMs together with three yeast CYP51s. As indicated in the bottom line, 156 amino acid residues are common. The conserved amino acids seem to form six clusters (CR-1 through 6 in Fig. 2), where more than 65% amino acids are common. Although the exact role of each region has yet to be clarified experimentally, the following discussion can be made regarding their possible functions.

CR-1, CR-2, CR-3, and CR-4 overlap with four of the seven putative substrate-recognition sites (underlines SRS-0 through -6 in Fig. 2) inferred by Gotoh (*25*). Among these regions, the extents of overlapping of CR-2 and CR-3 with SRS-1 and SRS-2 are prominent, and the degree of sequence conservation in these regions is remarkably high. As pointed out by Gotoh (*3, 25*), the ratio of nonsynonymous to synonymous substitution observed in the genes of P450 family 2 is significantly high in SRS-1 and SRS-2, and this excessive substitution is considered to be partly responsible for the different substrate specificities of many family-2 P450s. Inversely, the strong conservation of these regions among P45014DMs indicates their importance in the recognition of their common substrates, 14-methylsterols. CR-4 overlaps SRS-4 located in the putative distal helix (helix I) (*10, 13*). Since the distal helix constitutes a part of the active site (*25, 26*) and a mutation of this region converts *Saccharomyces cerevisiae* CYP51 into an inactive form called P450SG1 (*10*), conservation of this region may also be related to the conserved catalytic function of this P450.

CR-5 and CR-6 correspond to the so called K-helix and heme binding regions, respectively, that are generally conserved in all P450 species (*3*). The sequence conserva-

TABLE I. **Amino acid identities between P45014DM sequences of different origins.** Identity values were calculated with DNASIS software, and values are expressed in percentage of amino acids identical between two molecules. Sequence data for yeast and filamentous fungal CYP51s were taken from Refs. *10* and *11*, and EMBL Z48164 and Z49750, respectively.

	Hum.	Rat	S.c.	C.t.	P.i.	U.m.
Human	—	93	39	41	39	41
Rat		—	39	42	39	40
S. cerevisiae			—	65	46	47
C. tropicalis					46	47
P. italicum					—	45
U. maydis						—

TABLE II. **Examples of enzymes showing high amino acid sequence identity between primates and rodents.**

Enzyme name	Identity (%)	Species
Alcohol dehydrogenase	94.2	Human/rat
Hexokinase	94.1	Human/rat
Deoxycytidine kinase	93.7	Human/mouse
Glucose-6-P dehydrogenase	93.6	Human/mouse
C_1-Tetrahydrofolate synthase	93.2	Human/rat
Hydroxymethylglutaryl-CoA reductase	93.1	Human/hamster
Phosphoglycerate mutase	92.3	Human/rat
Phenylalanine hydroxylase	92.1	Human/rat
Ornithine transcarbamylase	91.5	Human/mouse
Glutathione peroxidase	91.5	Human/rat
Tryptophan hydroxylase	91.2	Human/rat
δ-Aminolevulinic acid synthase	90.2	Human/rat

tion among the heme binding regions of five P45014DMs (16/20 amino acids are identical) is considerably high. The heme binding site is situated opposite the substrate binding site with respect to the heme plane, and is considered to contribute little to the substrate recognition. Therefore, it is difficult to consider that the high structural conservation of this region is due to a selective force to maintain substrate specificity, and the reason for this high conservation of the heme-binding region has yet to be examined. However, such high structural conservation of this region may be taken as a strong evidence for the close evolutionary relationship between mammalian P45014DM and fungal CYP51.

Three putative substrate-recognition sites, SRS-3, SRS-5, and SRS-6, do not overlap with the conserved regions. However, the amino acid identities of these sites within the respective classes are considerably high. Although the contribution of these regions to the substrate recognition by P45014DM has yet to be examined, these facts may be related to the slightly different substrate specificities

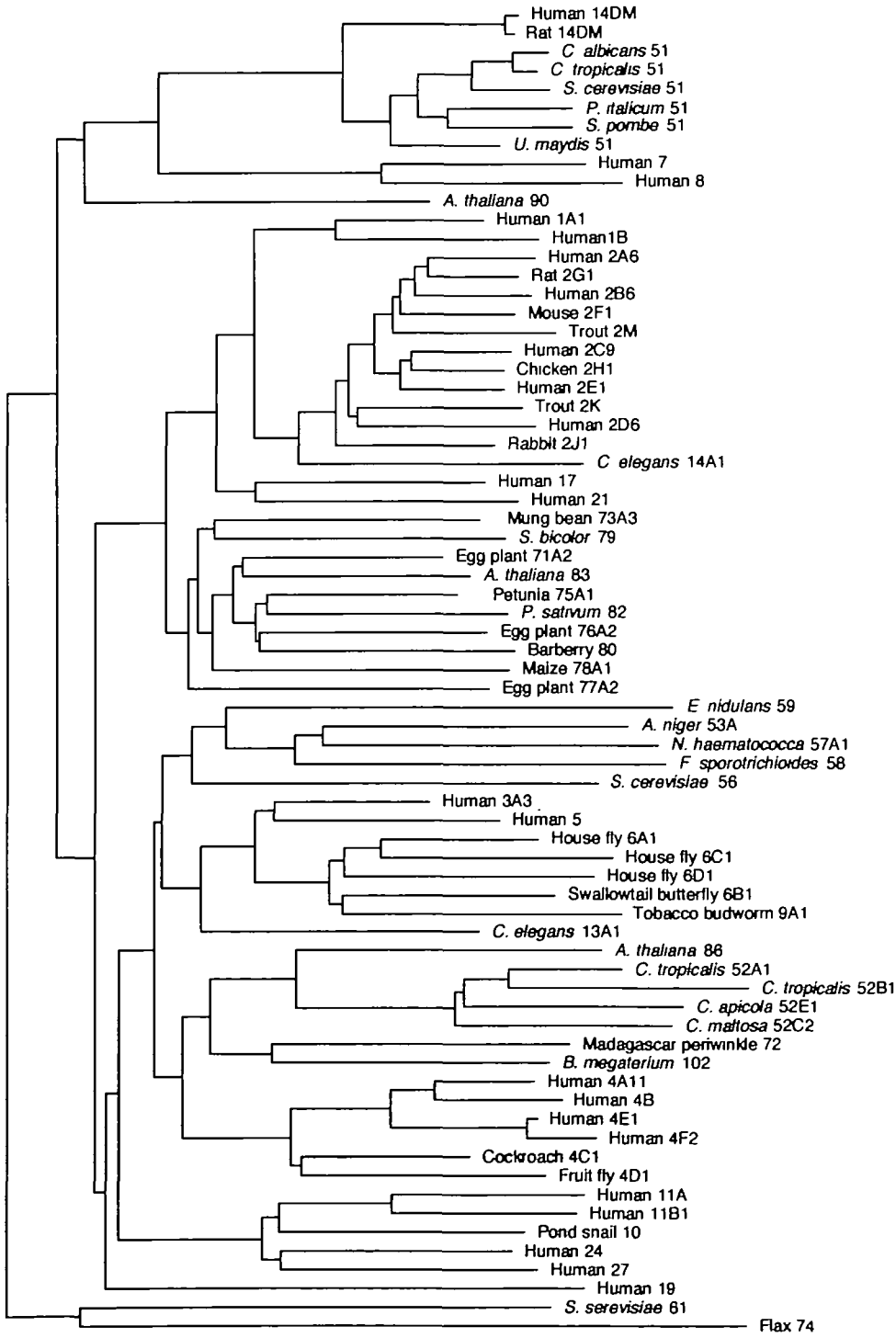

Fig. 3. **A phylogenetic tree showing the evolutionary relationship of P45014DM and other P450 families and subfamilies.** Single representative amino acid sequences were arbitrary selected from each of the 63 families or subfamilies of E-class P450s (*3*), and were aligned together with two mammalian P45014DM and six fungal CYP51 sequences as described under "MATERIALS AND METHODS." All sequence data, except for the two mammalian P45014DM sequences described here, were taken from publicly available nucleotide or protein sequence databases.

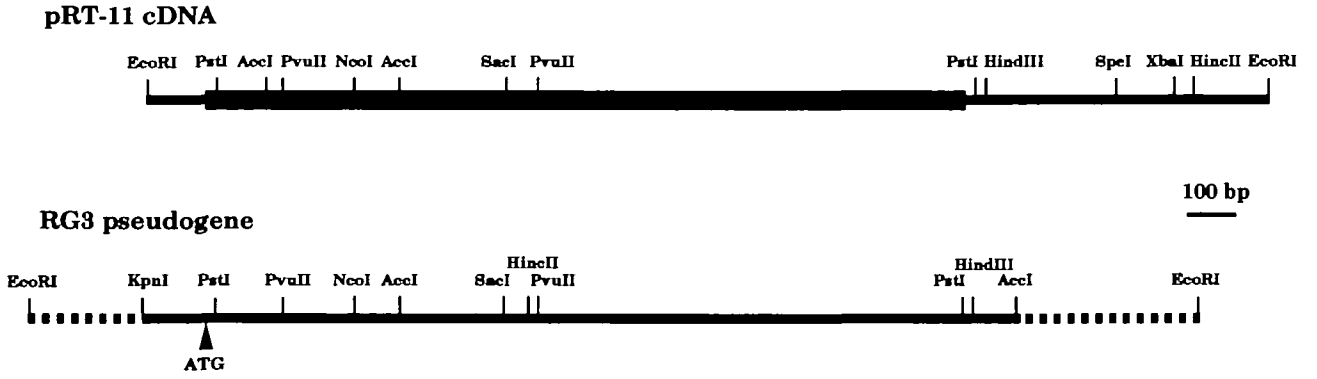

Fig. 4. **Comparison of the restriction enzyme maps of pRT-11 (rat P45014DM) cDNA and the RG3 pseudogene.** Restriction enzyme maps of pRT-11 cDNA and the 2.3-kbp DNA fragment isolated from a rat genomic clone, RG3, are illustrated based on their sequences and the results of digestion with several restriction enzymes. The open reading frame (118-1629 nt) of the cDNA is indicated by a thick bar. The region of the RG3 DNA fragment highly identical to that of the cDNA in the nucleotide sequence is indicated by a solid line. The ATG codon of the RG3 DNA fragment corresponding to the initiation codon of the cDNA is indicated in the figure, its position having been adjusted to that of the cDNA.

observed between fungal CYP51 and mammalian P45014DM (*4, 7, 8*).

Occurrence of a Processed Pseudogene in the Rat Genome—A 2.3-kbp DNA fragment was subcloned from a rat genomic clone, RG3, as described under "MATERIALS AND METHODS," and subjected to restriction enzyme mapping. As shown in Fig. 4, the restriction enzyme map of the RG3 DNA fragment was similar to that of the P45014DM cDNA (pRT-11), although some restriction enzyme sites were different from each other. Moreover, the RG3 DNA fragment hybridized with both the *Eco*RI-*Pst*I fragment (140 bp) and the *Pst*I-*Eco*RI fragment (600 bp), that were obtained respectively from the 5′-end and 3′-end regions of pRT-11 cDNA (Fig. 4). These facts suggested the possibility that the RG3 DNA fragment was a genomic DNA covering the entire region of the P45014DM gene. However, the length of this DNA fragment was 2.3 kbp, which is comparable to that of P45014DM cDNA. Therefore, if the above-mentioned assumption is correct, no intron must be included in this gene. Then, the RG3 DNA fragment was sequenced as described under "MATERIALS AND METHODS." The nucleotide sequence of this fragment (DDBJ D78370) showed 95.2% identity to that of the rat P45014DM cDNA within 1,800 aligned nucleotides, which is shown by the solid line in the restriction enzyme map of the RG3 DNA (Fig. 4). No intron was detected in the sequence, and the reading frame corresponding to the ORF of P45014DM cDNA was disturbed by several nucleotide substitutions and deletions. Since several DNA fragments that must be derived from the functional intron-containing P45014DM gene have been identified in the same genomic library (Noshiro, M., Aoyama, Y., Gotoh, O., and Yoshida, Y., preliminary observation), the gene contained in the RG3 DNA fragment was identified as a processed pseudogene of P45014DM. Processed pseudogenes of P45014DM have also been found in a human genomic DNA library (Rozman, D. and Waterman, M.R., personal communication). P45014DM is the first and only known example of a P450 having processed pseudogenes. Processed pseudogenes are generally formed through the insertion of a cDNA produced by reverse transcription of mRNA expressed in germ line cells (*27*). Accordingly, it is highly likely that P45014DM is expressed in germ line cells of mammals as a housekeeping enzyme.

General Consideration and Conclusion—Taking all the observations together, there is no doubt of the orthologous nature of fungal CYP51 and mammalian P45014DM. As shown in Fig. 3, P45014DM was established as a distinct family at an early stage of eukaryotic P450 evolution and has been conserved throughout evolution. Apparently, P45014DM is one of the most ancient and most conserved P450 species. Sterol 14-demethylation occurs in all organisms exhibiting *de novo* sterol biosynthesis, and this reaction is known to be also catalyzed by P450 monooxygenase in higher plants (*4, 28*). Since human and fungal (one of the most evolutionary distant pairs of eukaryotes) P45014DMs are orthologous, it is highly likely that the P450 monooxygenases catalyzing sterol 14-demethylation in higher plants are also orthologues of P45014DM. Comparison of the amino acid sequences of a group of orthologous P450s occurring in different kingdoms, phyla, classes, and species will provide important information about the evolution of P450. P45014DM is a suitable material for such investigation, and currently such comparative studies can only be performed with this P450. Cloning and sequencing of P45014DMs of other organisms, especially higher plants and non-mammalian vertebrates, are expected.

The amino acid sequence identity between rat and human P45014DM (93%) is outstanding and comparable to those of housekeeping enzymes. Moreover, the occurrence of processed pseudogenes suggests the possibility that P45014DM is expressed in germ line cells of mammals. Accordingly, P45014DM must have some fundamental and essential role in mammals. The 14-demethylation of lanosterol catalyzed by P45014DM is an essential step in cholesterol biosynthesis, and cholesterol is an indispensable component of animal cells. This may be the primary reason for the housekeeping nature of P45014DM. Recently, it was reported that some 4,4-dimethylsterols, such as 4,4-dimethylcholesta-8,14,24-trienol and 4,4-dimethylzymosterol, trigger the resumption of meiosis of cultured mammalian oocytes (*29*). These sterols are formed through P45014DM-mediated lanosterol 14-demethylation (*6, 24*). Accordingly, there is a possibility that P45014DM partici-

pates not only in sterol biogenesis but also in the production of the biosignal substance regulating the meiosis of mammalian oocytes. This may be another principal and interesting function of P45014DM in mammals.

Ms. C. Yamashita of Mukogawa Women's University and Ms. T. Hiroi of Osaka City University are acknowledged for their technical assistance in the lanosterol 14-demethylase assay and purification of P45014DM, respectively.

REFERENCES

1. Nelson, D.R., Kamataki, T., Waxman D.J., Guengerich, F.P., Estabrook, R.W., Fryereisen, R., Gonzalez, F.J., Coon, M.J., Gunsalus, I.C., Gotoh, O., Okuda, K., and Nebert, D.W. (1993) The P450 superfamily: Update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature. *DNA Cell Biol.* **12**, 1-51
2. Nebert, D.W., Nelson, D.R., Coon, M.J., Estabrook, R.W., Feyereisen, R., Fujii-Kuriyama, Y., Gonzalez, F.J., Guengerich, F.P., Gunsalus, I.C., Johnson, E.F., Loper, J.C., Sato, R., Waterman, M.R., and Waxman, D.J. (1991) The P450 superfamily: Update on new sequences, gene mapping, and recommended nomenclature. *DNA Cell Biol.* **10**, 1-14
3. Gotoh, O. (1993) Evolution and differentiation of P-450 genes in *Cytochrome P-450,* 2nd ed. (Omura, T., Ishimura, Y., and Fujii-Kuriyama, Y., eds.) pp. 255-272, Kodansha, Tokyo and VCH, Weinheim
4. Yoshida, Y. and Aoyama, Y. (1994) The P450 superfamily: A group of versatile hemoproteins contributing to the oxidation of various small molecules in *Regulation of Heme Protein Synthesis* (Fujita, H., ed.) pp. 75-88, AlphaMed Press, Dayton
5. Yoshida, Y. and Aoyama, Y. (1984) Yeast cytochrome P-450 catalyzing lanosterol 14-demethylation: I. Purification and spectral properties. *J. Biol. Chem.* **259**, 1655-1660
6. Aoyama, Y., Yoshida, Y., and Sato, R. (1984) Yeast cytochrome P-450 catalyzing lanosterol 14-demethylation: II. Lanosterol metabolism by purified P45014DM and by intact microsomes. *J. Biol. Chem.* **259**, 1661-1666
7. Yoshida, Y. (1993) Sterol biosynthesis in *Cytochrome P-450,* 2nd ed. (Omura, T., Ishimura, Y., and Fujii-Kuriyama, Y., eds.) pp. 93-101, Kodansha, Tokyo and VCH, Weinheim
8. Yoshida, Y. (1993) Lanosterol 14α-demethylase (cytochrome P45014DM) in *Cytochrome P450* (Schenkman, J.B. and Grein, H., eds.) pp. 627-639, Springer-Verlag, Berlin
9. Kalb, V.F., Woods, C.W., Turi, T.G., Dey, C.R., Sutter, T.R., and Loper, J.C. (1987) Primary structure of the P-450 lanosterol demethylase gene from *Saccharomyces cerevisiae. DNA* **6**, 529-537
10. Ishida, N., Aoyama, Y., Hatanaka, R., Oyama, Y., Imajo, S., Ishiguro, M., Oshima, T., Nakazato, H., Noguchi, T., Maitra, U.S., Mohan, V.P., Sprinson, D.B., and Yoshida, Y. (1988) A single amino acid substitution converts cytochrome P-450/14DM to an inactive form cytochrome P-450/SG1: Complete primary structures deduced from cloned cDNAs. *Biochem. Biophys. Res. Commun.* **155**, 317-323
11. Chien, C., Kalb, V.F., Turi, T.G., and Loper, J.C. (1988) Primary structure of the cytochrome P-450 lanosterol 14α-demethylase gene from *Candida tropicalis. DNA* **7**, 617-626
12. Lai, M.H. and Kirsch, D.R. (1989) Nucleotide sequence of cytochrome P-450 LIA1 (lanosterol 14α-demethylase) from *Candida albicans. Nucleic Acids Res.* **17**, 804
13. Aoyama, Y., Funae, Y., Noshiro, M., Horiuchi, T., and Yoshida, Y. (1994) Occurrence of a P450 showing high homology to yeast lanosterol 14-demethylase (P45014DM) in the rat liver. *Biochem. Biophys. Res. Commun.* **201**, 1320-1326
14. Funae, Y. and Imaoka, S. (1987) Purification and characterization of liver microsomal cytochrome P-450 from untreated male rats. *Biochim. Biophys. Acta* **926**, 349-358
15. Imaoka, S., Terano, Y., and Funae, Y. (1990) Changes in the amounts of cytochrome P450s in rat hepatic microsomes with starvation. *Arch. Biochem. Biophys.* **278**, 168-178
16. Ohishi, N., Imaoka, S., Suzuki, T., and Funae, Y. (1993) Characterization of two P-450 isozymes placed in the rat CYP2D subfamily. *Biochim. Biophys. Acta* **1158**, 227-236
17. Noshiro, M., Nishimoto, M., Morohashi, K., and Okuda, K. (1989) Molecular cloning of cDNA for cholesterol 7α-hydroxylase from rat liver microsomes. *FEBS Lett.* **257**, 97-100
18. Noshiro, M. and Okuda, K. (1990) Molecular cloning and sequence analysis of cDNA encoding human cholesterol 7α-hydroxylase. *FEBS Lett.* **268**, 137-140
19. Aoyama, Y. and Yoshida, Y. (1991) Different substrate specificities of lanosterol 14-demethylase (P-45014DM) of *Saccharomyces cerevisiae* and rat liver for 24-methylene-24,25-dihydrolanosterol and 24,25-dihydrolanosterol. *Biochem. Biophys. Res. Commun.* **178**, 1064-1071
20. Nishimoto, M., Gotoh, O., Okuda, K., and Noshiro, M. (1991) Structural analysis of the gene encoding rat cholesterol 7α-hydroxylase, the key enzyme for bile acid biosynthesis. *J. Biol. Chem.* **266**, 6467-6471
21. Gotoh, O. (1995) A weighting system and algorithm for aligning many phylogenetically related sequences. *Comput. Appl. Biosci.* **11**, 543-551
22. Dayhoff, M.O., Schwartz, R.M., and Orcutt, B.C. (1978) A model of evolutionary change in proteins in *Atlas of Protein Sequence and Structure,* Vol. 5, Suppl. 3 (Dayhoff, M.O., ed.) pp. 345-352, National Biomedical Research Foundation, Washington DC
23. Saitou, N. and Nei, M. (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406-425
24. Trzaskos, J., Kawata, S., and Gaylor, J.L. (1986) Microsomal enzymes of cholesterol biosynthesis. Purification of lanosterol 14α-methyl demethylase cytochrome P450 from hepatic microsomes. *J. Biol. Chem.* **261**, 14651-14657
25. Gotoh, O. (1992) Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. *J. Biol. Chem.* **267**, 83-90
26. Boscott, P.E. and Grant, G.H. (1994) Modeling cytochrome P450 14α-demethylase (*Candida albicans*) from P450cam. *J. Mol. Graph.* **12**, 185-192
27. D'Eustachio, P. and Ruddle, F.H. (1983) Somatic cell genetics and gene families. *Science* **220**, 919-924
28. Taton, M. and Rahier, A. (1991) Properties and structural requirements for substrate specificity of cytochrome P450-dependent obtusifoliol 14α-demethylase from maize (*Zea mays*) seedlings. *Biochem. J.* **277**, 483-492
29. Byskov, A.G., Andersen, C.Y., Nordholm, L., Thogersen, H., Guoliang, X., Wassmann, O., Andersen, J.V., Guddal, E., and Roed, T. (1995) Chemical structure of sterols that activate oocyte meiosis. *Nature* **374**, 559-562