

Purification, molecular cloning, and expression of 2-hydroxyphytanoyl-CoA lyase, a peroxisomal thiamine pyrophosphate-dependent enzyme that catalyzes the carbon–carbon bond cleavage during α -oxidation of 3-methyl-branched fatty acids

(phytanic acid/ Peroxisomes/ Refsum's disease/ oxalyl-CoA decarboxylase/ protein targeting)

VEERLE FOULON*, VASILY D. ANTONENKOV*, KATHLEEN CROES*, ETIENNE WAELEKENS†, GUY P. MANNAERTS*, PAUL P. VAN VELDHoven*‡, AND MINNE CASTEELS*‡

Divisions of *Pharmacology and †Biochemistry, Department of Molecular Cell Biology, Katholieke Universiteit Leuven, Campus Gasthuisberg B-3000 Leuven, Belgium

Edited by Philip W. Majerus, Washington University School of Medicine, St. Louis, MO, and approved July 13, 1999 (received for review April 1, 1999)

ABSTRACT In the third step of the α -oxidation of 3-methyl-branched fatty acids such as phytanic acid, a 2-hydroxy-3-methylacyl-CoA is cleaved into formyl-CoA and a 2-methyl-branched fatty aldehyde. The cleavage enzyme was purified from the matrix protein fraction of rat liver peroxisomes and identified as a protein made up of four identical subunits of 63 kDa. Its activity proved to depend on Mg^{2+} and thiamine pyrophosphate, a hitherto unrecognized cofactor of α -oxidation. Formyl-CoA and 2-methylpentadecanal were identified as reaction products when the purified enzyme was incubated with 2-hydroxy-3-methylhexadecanoyl-CoA as the substrate. Hence the enzyme catalyzes a carbon–carbon cleavage, and we propose calling it 2-hydroxyphytanoyl-CoA lyase. Sequences derived from tryptic peptides of the purified rat protein were used as queries to recover human expressed sequence tags from the databases. The composite cDNA sequence of the human lyase contained an ORF of 1,734 bases that encodes a polypeptide with a calculated molecular mass of 63,732 Da. Recombinant human protein, expressed in mammalian cells, exhibited lyase activity. The lyase displayed homology to a putative *Caenorhabditis elegans* protein that resembles bacterial oxalyl-CoA decarboxylases. Similarly to the decarboxylases, a thiamine pyrophosphate-binding consensus domain was present in the C-terminal part of the lyase. Although no peroxisome targeting signal, neither 1 nor 2, was apparent, transfection experiments with constructs encoding green fluorescent protein fused to the full-length lyase or its C-terminal pentapeptide indicated that the C terminus of the lyase represents a peroxisome targeting signal 1 variant.

3-Methyl-branched fatty acids cannot be degraded by β -oxidation because of the presence of a methyl group in the 3-position, which prevents the third reaction of the pathway, a dehydrogenation of the 3-hydroxyl group. For this reason, 3-methyl-branched fatty acids first undergo α -oxidation, in which they are shortened by one carbon atom (1). The α -oxidation pathway has been elucidated only recently (see Fig. 1). Watkins *et al.* (2) and Croes *et al.* (3) could demonstrate that α -oxidation in rat liver starts with an activation reaction requiring ATP, CoA, and Mg^{2+} , leading to the formation of a 3-methylacyl-CoA. The activated branched fatty acid is then 2-hydroxylated in a reaction that depends on O_2 , 2-oxoglutarate, iron, and ascorbate. Finally, the resulting 2-hydroxy-3-

methylacyl-CoA is cleaved into formyl-CoA (4) and a 2-methyl fatty aldehyde (5, 6). Formyl-CoA is enzymatically hydrolyzed to formate and subsequently converted to CO_2 (4), whereas the 2-methyl fatty aldehyde is dehydrogenated in an NAD^+ -dependent reaction to the corresponding acid (5). The 2-methyl-branched fatty acid is then degraded via β -oxidation, mainly in peroxisomes (7–10).

In the rat, the entire α -oxidation pathway appears to be peroxisomal (3, 6) (except for the conversion of formate to CO_2 , which occurs mainly in the cytosol). Also in humans, the α -oxidation pathway (activation, 2-hydroxylation, cleavage) occurs in peroxisomes (11), but the dehydrogenation of the 2-methyl-branched aldehyde to the corresponding acid may be catalyzed in the endoplasmic reticulum (12). Of the enzymes involved in α -oxidation, the 2-hydroxylase (phytanoyl-CoA hydroxylase; see below) (13, 14) and possibly the activation enzyme (15) have been purified and cloned. The hydroxylase is targeted to the peroxisome by a cleavable N-terminal peroxisome targeting signal (PTS2).

Quantitatively the most important 3-methyl-substituted fatty acid in humans is phytanic (3,7,11,15-tetramethylhexadecanoic) acid. This isoprenoid-derived fatty acid of vegetal origin is taken up with the diet in dairy products and ruminant fat (1). An accumulation of phytanic acid occurs in generalized peroxisome biogenesis disorders (e.g., Zellweger syndrome), which confirms the peroxisomal involvement in α -oxidation, and in rhizomelic chondrodysplasia punctata (RCDP) and Refsum's disease (1). Although in generalized peroxisome biogenesis disorders the whole α -oxidation sequence is most probably malfunctioning, in RCDP and in Refsum's disease the defect seems to be confined to the 2-hydroxylase (16). The hydroxylase deficiency in RCDP is explained by a defective peroxisomal import of proteins containing a PTS2, because of mutations in the gene encoding the peroxisomal PTS2 import receptor (Pex7p) (17–19), whereas in Refsum's disease the hydroxylase gene itself is mutated (13, 14).

In this paper, we describe the purification, functional and molecular characterization, and eukaryotic expression of the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

This paper was submitted directly (Track II) to the *Proceedings* office. Abbreviations: EST, expressed sequence tag; GFP, green fluorescent protein; 2-HPCL, 2-hydroxyphytanoyl-CoA lyase; PTS, peroxisome targeting signal; TPP, thiamine pyrophosphate.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. HSA131753).

‡To whom reprint requests should be addressed. E-mail: paul.vanveldhoven@med.kuleuven.ac.be or minne.casteels@med.kuleuven.ac.be.

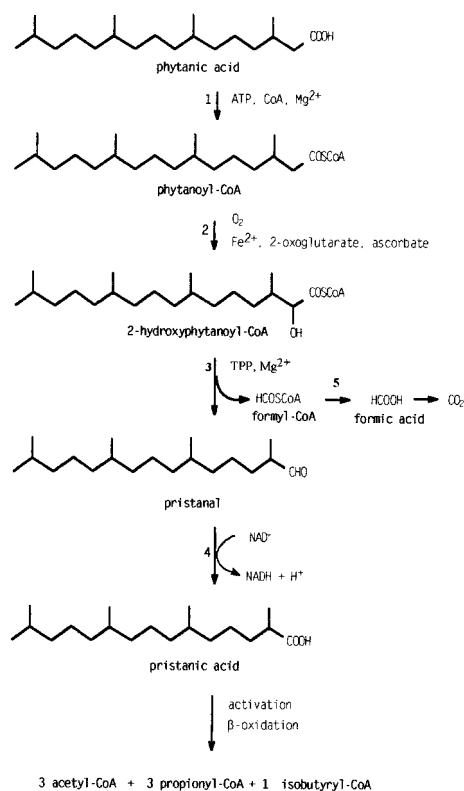


FIG. 1. α -Oxidation pathway. The figure represents the recently revised α -oxidation pathway for phytanic acid, including the results of this paper. The enzymes involved are: 1, phytanoyl-CoA synthetase; 2, phytanoyl-CoA hydroxylase; 3, 2-HPCL; 4, branched fatty aldehyde dehydrogenase; and 5, formyl-CoA hydrolase.

cleavage enzyme of α -oxidation, which we propose to term 2-hydroxyphytanoyl-CoA lyase (2-HPCL).

METHODS

Preparation of Homogenates and Subcellular Fractions.

Livers from male Wistar rats weighing approximately 200 g and fasted overnight were homogenized in 0.25 M sucrose/5 mM Mops, pH 7.2/0.1% (vol/vol) ethanol and fractionated by differential centrifugation (20). Determination of marker enzymes and protein was done as described (20).

Enzyme Purification. Peroxisomes were purified from 70 g of rat liver, homogenized in 0.25 M sucrose/10 mM Mops, pH 7.4/0.1% (vol/vol) ethanol/1 mM EDTA/1 mM EGTA/2 mM DTT/1 μ g/ml leupeptin, by combination of differential and isopycnic centrifugation, as described (21). The matrix proteins were released by sonication (21) and subjected to five successive low-pressure chromatographic steps at 4°C. After dialysis against 20 mM Tris, pH 8.0/10 mM K-phosphate/1 mM EDTA/2 mM DTT/20% (vol/vol) glycerol/0.2% (wt/vol) Tween-20, the enzyme preparation was first applied to a DEAE-Fast Flow Sepharose column (Amersham Pharmacia; 2.6 \times 10 cm). After the column was washed for 90 min, bound proteins were eluted with KCl (0 to 0.4 M over 90 min) at a flow rate of 1.47 ml/min. The enzymatically active fractions, eluted between 20 and 100 mM salt, were combined, concentrated by ultrafiltration (Centriprep-10, Amicon), and dialyzed against 20 mM K-phosphate, pH 7.0/1 mM EDTA/2 mM DTT/20% (vol/vol) glycerol/0.2% (wt/vol) Tween-20. This preparation was subjected to phosphocellulose chromatography (Whatman P11; 2.5 \times 8 cm) and eluted with a linear KCl gradient at a flow rate of 1.18 ml/min (0 to 0.6 M KCl over 86 min). Enzymatically active fractions, eluted between 40 and

100 mM salt, were combined and concentrated as described above and stored at -20°C for further experiments or dialyzed against 20 mM Tris, pH 7.5/10 mM K-phosphate/2 mM DTT/20% (vol/vol) glycerol/0.2% (wt/vol) Tween-20. The dialyzed sample was passed through two Econo-Pac HTP hydroxyapatite cartridges (Bio-Rad; 5 ml) connected in series, and the proteins were eluted with a KCl gradient (0 to 0.5 M over 120 min) at a flow rate of 0.5 ml/min. After concentration, the active fractions, eluted from the hydroxyapatite column between 160 and 420 mM salt, were dialyzed against 20 mM Tris, pH 8.0/10 mM K-phosphate/2 mM DTT/20% (vol/vol) glycerol/0.2% (wt/vol) Tween-20 and applied on an anion-exchanger (Econo-Pac Q cartridges, Bio-Rad, 2 \times 5 ml). Proteins were eluted with a KCl gradient (0 to 0.6 M over 120 min) at a flow rate of 1 ml/min. Final purification was obtained by subjecting the combined and concentrated active fractions of the anion exchange column (eluted between 25 and 50 mM salt) to gel filtration (Ultrogel AcA-34, 1.6 \times 90 cm) in 20 mM Tris, pH 8.0/0.1 M KCl/2 mM DTT/20% (vol/vol) glycerol/0.2% (wt/vol) Tween-20 at a flow rate of 0.2 ml/min. The column was calibrated as described previously (21).

Fractions from the different chromatographic steps were examined by SDS/PAGE (10–20% polyacrylamide gradient gels), and protein bands were visualized by silver or Coomassie blue staining.

Measurement of Lyase Activity. The production of [¹⁴C]formate, measured as ¹⁴CO₂ (22), was used to quantify lyase activity. In this assay, [¹⁴C]formyl-CoA, the primary α -oxidation product that is readily hydrolyzed to [¹⁴C]formate, is also measured as ¹⁴CO₂ (4). Incubations (37°C) were started by adding 50 μ l of the enzyme source to 200 μ l of reaction medium. Final assay concentrations were 50 mM Tris buffer, pH 7.5, 6.6 μ M BSA, 0.8 mM MgCl₂, 20 μ M thiamine pyrophosphate (TPP), and 40 μ M 2-hydroxy-3-methyl[1-¹⁴C]hexadecanoyl-CoA (4). Reactions were terminated after 10 min with 125 μ l 6% (wt/vol) HClO₄, and the acidified mixtures were used for the measurement of radioactive formate (22).

Other incubations were terminated by adding 50 μ l of 1 M K-phosphate, pH 2, and [¹⁴C]formate and [1-¹⁴C]formyl-CoA were measured separately by HPLC as described (4). For the analysis of lipid-soluble reaction products, similar incubations with unlabeled 2-hydroxy-3-methylhexadecanoyl-CoA were quenched by adding 0.5 ml of methanol/glacial acetic acid (98/2, vol/vol) and extracted with hexane. Identification of 2-methylpentadecanoic acid and 2-methylpentadecanal was done by GC as described (5), with heptadecanoic acid and heptadecanal as internal standards. Heptadecanal was synthesized by oxidation of heptadecanol (Aldrich) with tetra-*n*-propylammoniumperruthenate (23) and purified by silica column chromatography in dichloromethane. Purity was checked by TLC and GC (yield 65%, purity >95%).

Amino Acid Sequencing. Internal amino acid sequencing was carried out after SDS/PAGE of the combined active fractions from the gel filtration column and in gel digestion of the band of interest (24). Peptide fragments were separated on a C₂/C₁₈ SC 2.1/10 column (SMART-system, Amersham Pharmacia) at a flow rate of 80 μ l/min by using an 83-min linear gradient from 0 to 70% (vol/vol) acetonitrile in 0.1% (vol/vol) trifluoroacetic acid in water. Manually collected fractions with a high level of purity, as monitored at 215 nm, were collected on a Biobrene coated glass fiber membrane (Perkin-Elmer) and subjected to automated Edman degradation on a Procise 492 sequenator (Applied Biosystems).

Cloning and Molecular Characterization of 2-HPCL. To obtain the cDNA sequence coding for the human 2-HPCL, we probed, using the BLAST algorithm (advanced options: expect 1,000, filter none) (25), the Expressed Sequence Tag (EST) database (National Center for Biotechnology Information) with the peptide sequences derived from the purified rat protein.

Highest scoring ESTs were always assigned as being homologous to oxalyl-CoA decarboxylases. Reprobating the *HsEST* database with a *Caenorhabditis elegans* sequence (Z66519; see *Results*) yielded 12 partially overlapping ESTs. Two of the ESTs (accession nos. T77417 and R17722) appeared to contain the full-length human 2-HPCL cDNA and were obtained from the I.M.A.G.E. Consortium (Human Genome Mapping Project Resource Center, Hinxton, England). The inserts were sequenced (ALF-sequencer, Amersham Pharmacia) with fluorescent vector primers [M_{13} forward and reverse; cycle sequencing kit (Amersham Pharmacia)] and gene-specific primers [forward primers F_1 : 5'-CCAGGAGTCTGCCTTGTG, F_2 : 5'-CCAGGTTGATACTGTGCAG, and F_3 : 5'-GCAATGGATCATCTGTGTGG; reverse primers R_1 : 5'-CTGCACAGATATCAACCTGG and R_2 : 5'-CCACACAGATGATCCAATGC; AutoRead sequencing kit (Amersham Pharmacia)]. Comparison of these sequences to other ESTs revealed that in EST T77417, a stretch of 246 bases (from 310 to 555) was missing, whereas the first 66 nucleotides of EST R17722 did not match other ESTs similar to this cDNA. To determine the correct 5'-end, human liver cDNA was analyzed by PCR with primers corresponding to the 5'-end of both sequences [forward primer F_4 for EST T77417 5'-GGGGTACCCCGAAGATGCCGACAGTAAC (a *KpnI* site (underlined) was introduced to allow subcloning); F_5 for EST R17722 5'-CGGGG-TACCCCGAAGGATGTGGCGATATGG; reverse primer R_3 5'-GGCATGGATGAGACCTGG (bases 256–273)] (denaturation at 94° for 30 s, annealing at a primer-dependent temperature for 30 s and extension at 72° for 45 s; 30 cycles). Only in the presence of the forward primer for the sequence of EST T77417 was an amplicon generated. In another experiment, the human liver cDNA was amplified by PCR with gene-specific primers (forward primer from base 229 to 247, reverse primer from base 909 to 925), encompassing the gap in EST T77417. Only one amplicon of 696 bp could be detected.

A 1.8-kb fragment containing the ORF was obtained by amplifying human liver cDNA with forward primer F_4 and reverse primer R_4 [5'-TGCTCTAGAGCATTACATATTAGAGCGGGTC; an *XbaI* site (underlined) was introduced to allow subcloning]. This fragment was labeled by using the Strip-EZ DNA kit (Ambion, Austin, TX) and used to probe a human multiple tissue Northern blot (MTN, CLONTECH).

Transfection Studies of the Human 2-HPCL. To express the 2-HPCL, the 1.8-kb *KpnI-XbaI* restricted PCR product (see above) was cloned into the pcDNA3.1/Myc-His B vector (Invitrogen). Human embryonic kidney (HEK 293T) cells were transiently transfected with the resultant recombinant plasmid by using polyethyleneimine (26). As a control for transfection efficiency and cell viability, parallel cultures were transfected with the pcDNA3.1/Myc-His/*lacZ* plasmid.

Forty-eight hours after transfection, cell monolayers were harvested by scraping into phosphate buffered saline and centrifuged at $2,000 \times g$ for 5 min. Pelleted cells were resuspended in 0.25 M Tris buffer, pH 8.0, and sonicated. Cell lysates were analyzed for 2-HPCL activity and protein content. β -Galactosidase activity was measured according to the Invitrogen protocol.

To generate a green fluorescent protein (GFP) fusion product of the *Hs* 2-HPCL, the 5' overhang of the 1.8-kb *KpnI-XbaI* restricted PCR product (see above) was filled in with T_4 DNA-polymerase and ligated between the *KpnI-SmaI* sites of the EGFP-C1 vector (CLONTECH). To generate a GFP fusion product containing the five C-terminal amino acids of the *Hs* 2-HPCL, two 24-mers (GTACATTACCCGCTCTAATATGTA and GGCCTACATATTAGAGCGGGTAAT) were allowed to hybridize to generate a DNA insert with the appropriate overhangs and were cloned between the *BsrGI-NotI* sites of the EGFP-N1 vector, according to a strategy described before (27). Transfection of fibroblasts from PEX5^{+/-} mice and PEX5 knockout mice (28) with the EGFP-

fusion constructs and investigation of colocalization with peroxisomal thiolase were performed as described (27).

RESULTS AND DISCUSSION

Subcellular Distribution of the Lyase in Rat Liver. Previously we reported that rat liver peroxisomes can convert 2-hydroxy-3-methylacyl-CoAs into formyl-CoA (4) and a 2-methyl-branched aldehyde (5). Apparently, no specific cofactors were required for this cleavage reaction. The subcellular distribution of the lyase activity was now investigated in more detail in rat liver (Fig. 2). A clearly peroxisomal distribution pattern was seen, which correlates with the earlier observed distribution of the whole α -oxidation pathway in rat liver (3). As is the case for catalase, the cytosolic lyase activity is presumably caused by leakage of matrix enzymes from the peroxisomes during tissue homogenization. In agreement, sonication of purified peroxisomes resulted in the release of activity to a similar extent as that of catalase, confirming a matrix localization (V.F., K.C., G.P.M., P.P.V.V. and M.C., unpublished data).

Purification of the Cleavage Enzyme; Dependence on TPP. In our first efforts to purify the lyase, we noted an important loss of activity during purification and/or storage of the active

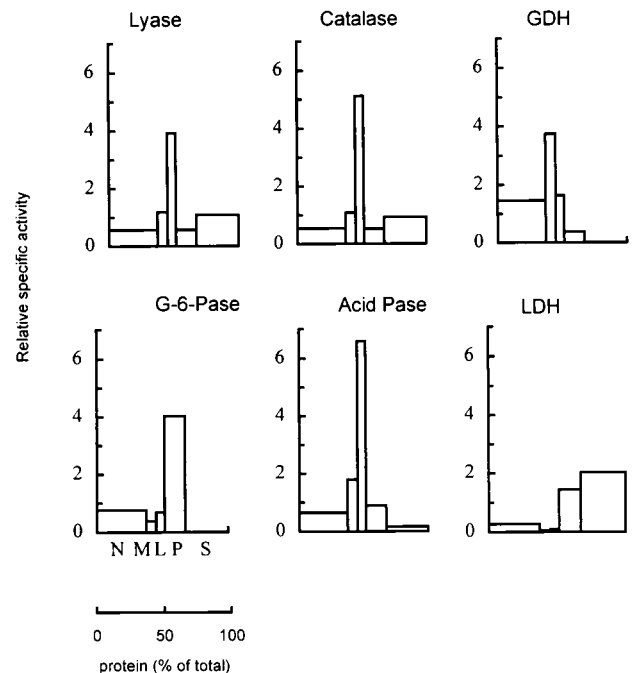


FIG. 2. Subcellular distribution of 2-HPCL activity in rat liver. A fresh rat liver homogenate was fractionated by differential centrifugation into a nuclear (N), a heavy mitochondrial (M), a light mitochondrial (L), a microsomal (P), and a cytosolic (S) fraction. The fractions were incubated with 2-hydroxy-3-methyl-[1-¹⁴C]hexadecanoyl-CoA, and labeled formate was measured as reaction product. In these experiments, no TPP or Mg^{2+} was present in the assay mixtures. Marker enzymes and protein were determined in each fraction: catalase (peroxisomal matrix), glutamate dehydrogenase (GDH, mitochondria), glucose-6-phosphatase (G-6-Pase, endoplasmic reticulum), acid phosphatase (Acid Pase, lysosomes), and lactate dehydrogenase (LDH, cytosol). Relative specific activities are presented vs. cumulative percentage of total protein (relative specific activity is defined as the percentage of total recovered activity in a given fraction vs. the percentage of total recovered protein in that fraction). Overall recovery for lyase was 92%; overall recoveries for marker enzymes were between 99 and 113%. Total lyase activity was 112.8 mU (nmol/min) per gram of liver. A similar distribution pattern was found in the presence of TPP and Mg^{2+} (V.F., K.C., G.P.M., P.P.V.V. and M.C., unpublished data).

fractions and this despite various attempts to stabilize the enzyme. The amino acid sequences of tryptic peptides of the purified and virtually inactive enzyme (see below) suggested that the cleavage enzyme was related to a putative *C. elegans* protein that displays homology to bacterial oxalyl-CoA decarboxylases. These enzymes, which have hitherto been described only in bacteria, catalyze the TPP-dependent decarboxylation of oxalyl-CoA to formyl-CoA and CO₂ (29, 30). This TPP-dependence suggested that 2-HPCL might also require TPP, an as yet unrecognized cofactor for α -oxidation, and that enzyme-bound TPP had gradually been lost from the enzyme during purification and storage. In accordance with this contention, enzyme activity could be restored by the addition of 20 μ M TPP and 0.8 mM MgCl₂ (see below), which from then on were routinely included in the assay mixtures.

The lyase could be purified 67-fold as compared with the peroxisomal matrix protein fraction with a yield of 3%. The specific activity of the purified enzyme amounted to 558 milliunits (mU)/mg protein. Fig. 3 shows the SDS/PAGE analysis of fractions from the gel filtration column and the activity measured on these fractions. The results show that the enzyme was purified to near homogeneity and that it consists of a polypeptide with a molecular mass of 63 kDa. Because during gel filtration the activity was eluted as a protein of 250 kDa (Fig. 3), the native enzyme appears to be a homotetramer.

Functional Characterization of the Lyase. As mentioned above, addition of TPP and Mg²⁺ stimulated the activity of the enzyme during purification. In the presence of 0.8 mM Mg²⁺, optimum activity was reached at 20 μ M TPP (V.F., K.C., G.P.M., P.P.V.V. and M.C., unpublished data). Under these conditions, the activity of the partially purified preparation after phosphocellulose chromatography was enhanced 4-fold, whereas addition of TPP to whole homogenates or crude peroxisomal fractions stimulated the lyase activity only 1.3-

fold. In the final purified preparation, no activity at all could be measured in the absence of TPP/Mg²⁺. Fig. 4 shows the substrate dependency as studied on the partially purified lyase. An apparent K_m of 15 μ M for 2-hydroxy-3-methylhexadecanoyl-CoA was calculated. When assayed in different overlapping buffer systems (Mops-NaOH, pH 5.5–6.5; Hepes-NaOH, pH 6.5–7.5; Tris-HCl, pH 7.5–8.5; glycine-NaOH, pH 8.5–9), the lyase had a pH optimum between 7.5 and 8 in Tris buffer (V.F., K.C., G.P.M., P.P.V.V. and M.C., unpublished data).

GC analysis of the hexane extract, obtained from an incubation of the partially purified enzyme with 2-hydroxy-3-methylhexadecanoyl-CoA, revealed a peak coinciding with 2-methylpentadecanal. The amount of aldehyde formed was of the same magnitude as the amount of formate produced (velocities: 5.8 mU/ml and 5.26 mU/ml, respectively). Incubation in the presence of NAD⁺ [a cofactor for fatty aldehyde dehydrogenation (5)] did not alter the amount of formate and 2-methylpentadecanal formed. Methylation of the hexane extracts and GC analysis for the presence of fatty acid methyl esters also revealed no peak at the position of 2-methylpentadecanoic acid, indicating that the conversion of the aldehyde to a fatty acid is performed by a separate enzyme.

In another series of experiments, partially purified enzyme from a different batch was incubated with 2-hydroxy-3-methyl-[1-¹⁴C]hexadecanoyl-CoA. HPLC analysis of the ultrafiltrate of the incubation mixture revealed two labeled peaks, which coeluted with formate and formyl-CoA, respectively. When the HPLC fractions were collected and analyzed for [1-¹⁴C]-formate, the summated counts of both peaks grossly equaled the amount of [1-¹⁴C]-formate, measured in parallel samples of the unfractionated incubation mixtures (velocities: 2.36 mU/ml vs. 2.03 mU/ml). These results confirm that 2-HPCL forms formyl-CoA and a 2-methyl-branched fatty aldehyde in equal amounts and that the amount of measured [1-¹⁴C]-formate fully corresponds to the amount of formyl-CoA produced.

Given the presence of 3-hydroxy-3-methylglutaryl-CoA lyase in peroxisomes (31), it was important to exclude a possible contribution of this enzyme to our measurements. No 3-hydroxy-3-methylglutaryl-CoA lyase activity could be detected with the purified enzyme preparation (V.F., K.C., G.P.M., P.P.V.V. and M.C., unpublished data).

Molecular Cloning and Expression of 2-HPCL. N-terminal amino acid sequencing of tryptic peptides of the purified lyase

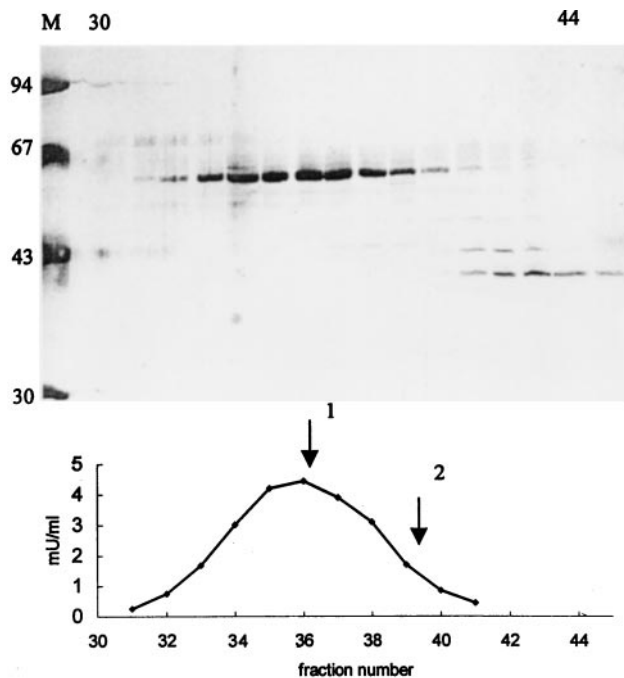


FIG. 3. Purification of 2-HPCL. The lyase was purified from a peroxisomal matrix protein fraction by successive chromatographic steps as described in *Methods*. (Upper) The polypeptide pattern of fractions 30–45 from the final gel filtration column as analyzed by SDS/PAGE (100 μ l fraction; 10–20% gradient gel; silver staining). The migration of molecular mass (expressed in kDa) markers is represented in lane M. (Lower) The elution of lyase activity, expressed in mU/ml, from the same gel filtration column. Elution of calibration proteins is indicated by arrows: 1, catalase (240 kDa); 2, lactate dehydrogenase (140 kDa).

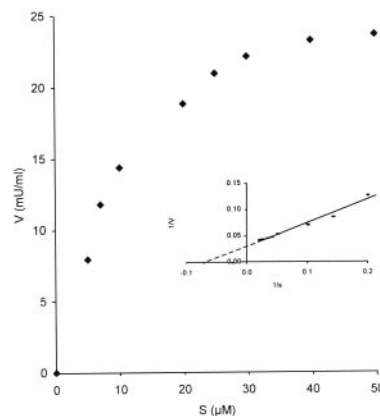


FIG. 4. Substrate dependency of 2-HPCL. Lyase activity was measured at increasing concentrations of 2-hydroxy-3-methylhexadecanoyl-CoA on the partially purified enzyme preparation after phosphocellulose chromatography. The velocity (V , expressed in mU/ml) vs. substrate concentration (S , expressed in μ M) curve was linearly transferred according to the method of Lineweaver-Burk as shown (Inset). An apparent K_m of 15 μ M could be calculated.

from rat liver yielded three useful sequences: (K/R)TQD-VEYMFVGVGIPVT, (K/R)GAAYSHAEDSIR, and (K/R)NPWQYPTDSK. None of these peptides displayed full homology to known mammalian proteins. The first peptide very closely matched (81% identity; 87% similarity) an internal sequence of a putative *C. elegans* protein, displaying similarity to a number of bacterial oxalyl-CoA decarboxylases. By probing (25) the human EST database with the rat peptides as well as the *C. elegans* protein sequence, 12 partially overlapping ESTs were obtained that allowed the reconstruction of a human cDNA sequence of 1,974 bases (Fig. 5). This sequence contains an ORF of 1,734 bases, starting with a translation initiation codon surrounded by a Kozak motif (32), and encodes a protein of 578 amino acids with a calculated molecular mass of 63,732 Da, which corresponds to the size of the subunit of the rat enzyme. The other two peptides of the

```

-99  ggtaccagaggtggcagtgctgctgcaacttcgctgttgccttgcctggtgatctccgcttgt
-39  ttgccgctgctgagtgctgtagctcttcctcggtggaagatccggcagcagtaacttcgca
      M P D S N F A
22   gagcgcagcaggagcaggtgcttggctaaagtcacgctcaggccctgaaacgcgaa
      E R S E E Q V S G A K V I A Q A L K T Q
82   gatgtggagtcacatatttggcatcgtagctccagcagcgaatccgcaattgctgccc
      D V E Y I F G I V G I P V T E I A I A A
142  cagcagctagcctacagtcacatcgggatgaggaatgagcaagcggctgttattgctgccc
      Q Q L G I K Y I G M R N E Q A A C Y A A
202  tccgcgattggatatacgaacagcaggcaggagctgctgctgttgccttgcggccaggt
      S A I G Y L T S R P G V C L V V S G F G
262  ctcatccatgcttggcggatggaatgcaaacatgcaactgctggcccttgccttgtg
      L I H A L G G M A N A N M N C W F L L V
322  attggtggttccctctgaaagaacaaagaacaaatggagcttccaggaatttccctcag
      I G G S S E R N Q E T M G A F Q E F P Q
382  gttgaagcttggatattataccacagttctctgcccgcgcaagcagcatagaagctatt
      V E A C R L Y T T S R P G V C L V V S G F G
442  cctttgttattgaaagcagtgagaagcagctatctatgctgctccaggtgcttgcctat
      P F V I E K A V R S S I Y G R P G A C Y
502  gttgacataccagcagatatttgggaacctcaggtgaatgtaattctataaagtacatg
      V D I P A D F V N L Q V N V N S I K Y M
562  gaacgctgcatgcaactcttattgcaatgagcagaacacctgctgctgagcagcggcct
      E R C M S P P I S M A E T S A V C T A A
622  tctgttattaggaatgcaaacaccttcttctcctggaagctgctgctgctgctgctgctg
      S V I R N A K Q P L L I I G K G A A Y A
682  catgcagaagagatcaagaatggtgagcaataaactgcaacttctgcccacc
      H A E E S I K K L V E Q Y K L P F L P T
742  cctatgggaaagggtgctgctcagcaaccctcactatgctgagtgagcagcagctc
      P M G K G V V P D N H P Y C V G A A R S
802  agggcttgcgaatttgcagtgtaattggttattggtgccaagctaaattggaattta
      R A L Q F A D V I V L F G A R L N W I L
862  cattttgagctgctcccaagatcagcagcagtggaatttatccaggtgatactctg
      H F G L P P R Y Q P D V K F I Q V D I C
922  gcagaagaattggggaataatgaaagccctgcttacttgcaggaacacatgct
      A E E L G N N V K P A V T L L G N I H A
982  gtcaactaacgctttagagaaacttgaaacaccatggcagctatccctcagagagc
      V T K Q L L E E L D K T P W Q Y P P E S
1042 aagtgggtgaaactctgagaaaaatgaaagcaatgaagctgcatcacaaggaacta
      E W K T I R E K M K S N E A A S K E L
1102 gcttcaaaaaatccctgctatgaattattacacagctatctaccatgctcaagaaca
      A S K K S L P M N Y Y T V F Y H V Q E Q
1162 ctacctagagactgttctggtgaagtgaaagagcaaatactatggacattggcggact
      L P R D C F V V S E G A N T M D I G R T
1222 gtgcttcagaactccttccctgctcagagcttgcctgctgcttctcggacaatggga
      V L Q N Y L P R H R L D A G T F G T M G
1282 gttggttgggattgctattgcaagctgctgctgctgctgctgctgctgctgctgctg
      V G L G F A I A A A V V A K D R S P G Q
1342 ttgatcattctggtgaaagagcagctgcttgggttcttctgctgagaggtgagaacc
      W I I C V E G D S A F G F S G M E V E T
1402 atctcaggttacaacttgcacatcactggtgtagtgaataacaatggaatttaccaa
      I C R Y N L P I I L L V V N N N G I Y Q
1462 ggttttgatcacagacttggaaagaaatgtaaaattcaagatgctactgagctgctg
      G F D T D T W K E M L K F Q D A T A V V
1522 cctccaatggttctgctcacaattcacattatgagcaagctgactgcttggagcc
      P P M C L L P N S H Y E Q V M T A F G G
1582 aaaggttatttgcacaacaccagaagaactcacaacacctcctgagcagagcctagca
      K G Y F V Q T P E E L Q K S L R Q S L A
1642 gacacaactaaccttcttcaacatcatgattgagccacaagccacacggaagcc
      D T T K P S L I N I M I E P Q A T R K A
1702 caggatttcttctgctgcccctcctaatgtaaaataagaccagctggtggtcct
      Q D F H W L T R S N M *
1762 gagtttctcttctgcaagatgaaatttttttccacagcaaatctactactggt
1822 aaaattggtcacaataaaataaacattttaaataagcattttacagtaaaaaaa
    
```

FIG. 5. cDNA and deduced amino acid sequence of *Hs* 2-HPCL. The cDNA shown is based on overlapping sequences, determined at least once in both directions of the *Hind*III/*Not*I inserts of the Lafmid BA vector containing either T77417 or R17722 and of the 1.8-kb amplicon. Accession numbers of other ESTs overlapping this sequence are AA306411, AA156133, W74012, AA056397, AA298588, F08168, W21374, R58040, H10428, AA150650, AA643224, A1380322, and Z41242. The start codon and the corresponding methionine are indicated in bold. The position of the stop codon is marked by an asterisk. Sequences corresponding to the tryptic peptides, obtained from rat enzyme, are underlined. The C-terminal peroxisome targeting sequence [R]SNM is printed in bold, and the TPP-binding region is underlined twice.

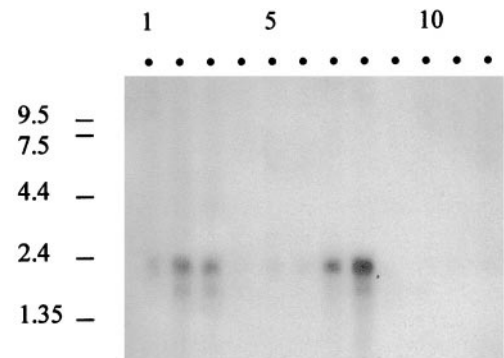


FIG. 6. Northern blot analysis of the *Hs* 2-HPCL. A human 12-lane multiple tissue Northern blot was analyzed with the 1.8-kb amplicon, derived from the *Hs* 2-HPCL cDNA, as a probe. Lanes: 1, brain; 2, heart; 3, skeletal muscle; 4, colon; 5, thymus; 6, spleen; 7, kidney; 8, liver; 9, small intestine; 10, placenta; 11, lung; 12, peripheral blood leukocytes. RNA size markers are indicated in kb.

rat 2-HPCL could also be located in the deduced human lyase amino acid sequence (see Fig. 5), confirming the identity of the cloned cDNA.

Attempts to express the human 2-HPCL in *Escherichia coli* yielded an inactive enzyme (V.F., K.C., G.P.M., P.P.V.V. and M.C., unpublished data). Expression of the 2-HPCL in human embryonic kidney cells, however, resulted in a significant increase in lyase activity: 0.771 ± 0.095 mU/mg protein in lysate of transfected cells vs. 0.063 ± 0.012 mU/mg protein in lysate of cells transfected with the β -galactosidase plasmid (mean values \pm SEM, $n = 4$). These data confirm that the obtained cDNA is derived from the 2-HPCL mRNA.

Northern blot analysis on mRNAs from multiple human tissues revealed that the abundance of the 2-HPCL mRNA transcript varied significantly among different tissues, with highest levels in liver and kidney (Fig. 6), corresponding with the presence of enzymatic activity in both human tissues (V.F., K.C., G.P.M., P.P.V.V. and M.C., unpublished data). The size of the 2-HPCL transcript was estimated to be 2.4 kb.

The human 2-HPCL amino acid sequence was used as query to recover homologous proteins (25). In addition to the *C. elegans* decarboxylase (Z66519), two fungal proteins with

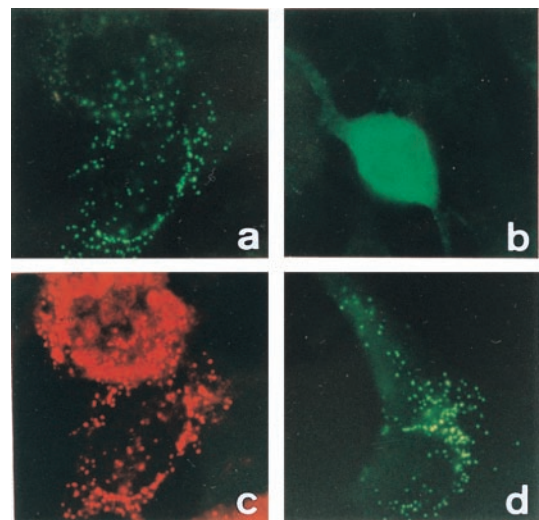


FIG. 7. Subcellular localization of GFP fusion proteins in cultured mouse fibroblasts. Immortalized mouse PEX5^{+/-} fibroblasts (a, c, and d) and PEX5^{-/-} fibroblasts (b) were transfected with the construct coding for GFP fused to *Hs* 2-HPCL (a, b, and c) or to the C-terminal pentapeptide of the human lyase (d). After 26 h, cells were analyzed by fluorescence microscopy, either for direct fluorescence of GFP (a, b, and d) or, after immunostaining, for peroxisomal thiolase (c).

unknown function (AL034352, *Schizosaccharomyces pombe*; P39994, *Saccharomyces cerevisiae*) and several bacterial oxalyl-CoA decarboxylases (e.g., A55219, *Oxalobacter formigenes*) were found. In the C-terminal part of the human lyase, a TPP-binding consensus domain is present (see Fig. 5). The consensus sequence of this domain, which is also located in the C terminus of a number of TPP-dependent decarboxylases (pyruvate decarboxylase, indolepyruvate decarboxylase, and benzoylformate decarboxylase), is now described as [LIVMF]-[GSA]-X-(5)-P-X-(4)-[LIVMFYW]-X-[LIVMF]-X-G-D-[GSA]-[GSAC] (Prosite PDOC00166). The corresponding peptide sequences in the lyase comply exactly with this consensus domain, except for the fact that only four instead of five amino acids precede the proline residue.

At first glance, the *Hs* 2-HPCL sequence did not contain a C-terminal or N-terminal PTS. Because the *C. elegans* ortholog ends in a putative PTS1 (SKM) and because PRL, the C-terminal tripeptide of the *S. cerevisiae* ortholog (see above), has been shown to bind to the human PTS1 import receptor (33), the C-terminal sequence SNM, which is also conserved in the mouse counterpart (V.F., K.C., G.P.M., P.P.V.V. and M.C., unpublished data), could have a targeting function. Transfection studies with constructs coding for 2-HPCL fused to GFP revealed that the fluorescence localized to peroxisomes in fibroblasts from PEX5^{+/-} mice and to the cytosol in fibroblasts from PEX5^{-/-} mice, lacking the PTS1 import receptor (see Fig. 7). Because a GFP construct containing only the last five amino acids of 2-HPCL also localized to peroxisomes (Fig. 7), we can conclude that targeting information is present within this pentapeptide and that SNM, preceded by a positive charge (27, 33), is a hitherto unrecognized PTS1.

CONCLUSION

We have described the purification, functional and molecular characterization, and eukaryotic expression of 2-HPCL, a peroxisomal enzyme that catalyzes the cleavage of the carbon-carbon bond during the α -oxidation of 3-methyl-branched fatty acids. The lyase is targeted to peroxisomes by a hitherto unknown PTS1, [R]SNM, and is homologous to bacterial oxalyl-CoA decarboxylases at the level of the primary and of the tertiary structure, being a tetrameric protein with a subunit size of approximately 63 kDa. In addition, the reaction mechanisms of the lyase and the decarboxylases are similar, generating formyl-CoA from 2-hydroxyacyl-CoAs in a TPP-dependent manner.

This work was supported by grants from the Geconcerteerde onderzoeksacties van de Vlaamse Gemeenschap (GOA 94/98-12 and GOA 99/03-09) and from the Fonds voor Wetenschappelijk Onderzoek-Vlaanderen (G-0239-98 and G.0164.96N). V.F. and K.C. were supported by fellowships from the Fonds voor Wetenschappelijk Onderzoek-Vlaanderen. We are grateful to Leen Amery and Dr. Myriam Baes for help with transfection studies and to Dr. Mark Franssen and Dr. Ivo De Baere for help with molecular cloning. We thank Tine Wylm for providing us with human liver cDNA and Luc Govaert, Els Meyhi, Kristien De Greef, Vanessa Brys, and Stanny Asselberghs for expert technical assistance. The secretarial help of Mouche Bareau is greatly appreciated.

- Steinberg, D. (1995) in *The Metabolic and Molecular Bases of Inherited Disease*, eds Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D. (McGraw-Hill, New York), pp. 2351-2369.
- Mihalik, S. J., Rainville, A. M. & Watkins, P. A. (1995) *Eur. J. Biochem.* **232**, 545-551.

- Croes, K., Casteels, M., de Hoffmann, E., Mannaerts, G. P. & Van Veldhoven, P. P. (1996) *Eur. J. Biochem.* **240**, 674-683.
- Croes, K., Van Veldhoven, P. P., Mannaerts, G. P. & Casteels, M. (1997) *FEBS Lett.* **407**, 197-200.
- Croes, K., Casteels, M., Asselberghs, S., Herdewijn, P., Mannaerts, G. P. & Van Veldhoven, P. P. (1997) *FEBS Lett.* **412**, 643-645.
- Verhoeven, N. M., Schor, D. S. M., ten Brink, H. J., Wanders, R. J. A. & Jakobs, C. (1997) *Biochem. Biophys. Res. Commun.* **237**, 33-36.
- Vanhove, G., Van Veldhoven, P. P., Vanhoutte, F., Parmentier, G., Eyssen, H. J. & Mannaerts, G. P. (1991) *J. Biol. Chem.* **266**, 24670-24675.
- Van Veldhoven, P. P., Huang, S., Eyssen, H. J. & Mannaerts, G. P. (1993) *J. Inherited Metab. Dis.* **16**, 381-391.
- Singh, H., Beckman, K. & Poulos, A. (1994) *J. Biol. Chem.* **269**, 9514-9520.
- Jakobs, B. S., van den Bogert, C., Dacremont, G. & Wanders, R. J. A. (1994) *Biochim. Biophys. Acta* **1211**, 37-43.
- Casteels, M., Croes, K., Van Veldhoven, P. P. & Mannaerts, G. P. (1997) *J. Inherited Metab. Dis.* **20**, 665-673.
- Verhoeven, N. M., Jakobs, C., Carneuse, S., Somers, M. P., Wanders, R. J. A. & Rizzo, W. B. (1998) *FEBS Lett.* **429**, 225-228.
- Mihalik, S. J., Morrell, J. C., Kim, D., Sacksteder, K. A., Watkins, P. A. & Gould, S. J. (1997) *Nat. Genet.* **17**, 185-189.
- Jansen, G. A., Ofman, R., Ferdinandusse, S., Ijlst, L., Muijsers, A. O., Skjeldal, O. H., Stokke, O., Jakobs, C., Besley, G. T. N., Wraith, J. E., *et al.* (1997) *Nat. Genet.* **17**, 190-193.
- Watkins, P. A., Howard, A. E., Gould, S. J., Avigan, J. & Mihalik, S. J. (1996) *J. Lipid Res.* **37**, 2288-2295.
- Pahan, K., Khan, M. & Singh, I. (1996) *J. Lipid Res.* **37**, 1137-1143.
- Braverman, N., Steel, G., Obie, C., Moser, A., Moser, H., Gould, S. J. & Valle, D. (1997) *Nat. Genet.* **15**, 369-376.
- Motley, A. M., Hetteema, E. H., Hogenhout, E. M., Brites, P., ten Asbroek, A. L. M. A., Wijburg, F. A., Baas, F., Heijmans, H. S., Tabak, H. F., Wanders, R. J. A., *et al.* (1997) *Nat. Genet.* **15**, 377-380.
- Purdue, P. E., Zhang, J. W., Skoneczny, M. & Lazarow, P. B. (1997) *Nat. Genet.* **15**, 381-384.
- Declercq, P. E., Haagsman, H. P., Van Veldhoven, P. P., Debeer, L. J., Van Golde, L. M. G. & Mannaerts, G. P. (1984) *J. Biol. Chem.* **259**, 9064-9075.
- Antonov, V. D., Van Veldhoven, P. P., Waelkens, E. & Mannaerts, G. P. (1997) *J. Biol. Chem.* **272**, 26023-26031.
- Casteels, M., Croes, K., Van Veldhoven, P. P. & Mannaerts, G. P. (1994) *Biochem. Pharmacol.* **48**, 1973-1975.
- Griffith, W. P. & Ley, S. V. (1990) *Aldrichimica Acta* **23**, 13-19.
- Rosenfeld, J., Capdevielle, J., Guillemot, J. & Ferrara, P. (1992) *Anal. Biochem.* **203**, 173-179.
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997) *Nucleic Acids Res.* **25**, 3389-3402.
- Boussif, O., Lezoualc'h, F., Zanta, M. A., Mergny, M., Sherman, D., Demeneix, B. & Behr, J.-P. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 7297-7303.
- Amery, L., Brees, C., Baes, M., Setoyama, C., Miura, R., Mannaerts, G. P. & Van Veldhoven, P. P. (1998) *Biochem. J.* **336**, 367-371.
- Baes, M., Gressens, P., Baumgart, E., Carmeliet, P., Casteels, M., Franssen, M., Evrard, P., Fahimi, D., Declercq, P. E., Collen, D., *et al.* (1997) *Nat. Genet.* **17**, 49-56.
- Baetz, A. L. & Allison, M. J. (1989) *J. Bacteriol.* **171**, 2605-2608.
- Lung, H. Y., Baetz, A. L., Peck, A. B. (1994) *J. Bacteriol.* **176**, 2468-2472.
- Ashmarina, L. I., Robert, M. F., Elsliger, M. A. & Mitchell, G. A. (1996) *Biochem. J.* **315**, 71-75.
- Kozak, M. (1984) *Nucleic Acids Res.* **12**, 857-872.
- Lametschwandtner, G. L., Brocard, C., Franssen, M., Van Veldhoven, P. P., Berger, J. & Hartig, A. (1998) *J. Biol. Chem.* **273**, 33635-33643.