Molecular cloning of cDNA species for rat and mouse liver α-methylacyl-CoA racemases

Werner SCHMITZ*, Heli M. HELANDER†, J. Kalervo HILTUNEN† and Ernst CONZELMANN*[1]

*Theodor-Boveri-Institut für Biowissenschaften (Biozentrum) der Universität Würzburg, Am Hubland, D-97074 Würzburg, Federal Republic of Germany, and †Biocenter Oulu, and Department of Biochemistry, University of Oulu, FIN-90570 Oulu, Finland

cDNA species coding for α-methylacyl-CoA racemase were cloned from rat and mouse liver cDNA libraries and characterized. The rat liver λgt11 cDNA expression library was screened with anti-racemase IgG [Schmitz, Albers, Fingerhut and Conzelmann (1995) Eur. J. Biochem. **231**, 815–822]. Several full-length clones were obtained that contained an open reading frame of 1083 bp, coding for a protein of 361 amino acid residues with a predicted molecular mass of 39679 Da. The sequences of three peptides that were isolated by HPLC from a tryptic digest of purified rat liver racemase fully matched the cDNA-derived amino acid sequence. The cDNA coding for mouse racemase was cloned from a mouse liver λZAP cDNA expression library and sequenced. The coding region of 1080 bp codes for a 360-residue protein (molecular mass 39558 Da) that shares 89.7% similarity with the rat protein. Expression of the rat racemase as a recombinant protein in *Escherichia coli* with the pTrcHisB-expression vector yielded enzymically active protein. The amino acid sequences of α-methylacyl-CoA racemases do not resemble any known sequence of β-oxidation or auxiliary enzymes, supporting the view of a highly diverse evolutionary origin of enzymes acting on fatty acyl-CoA *S*-esters.

INTRODUCTION

Various branched-chain fatty acids arise in the catabolism of isoprenoids. A prominent and quantitatively important example is phytanic acid (3,7,11,15-tetramethylhexadecanoic acid), which is derived from the chlorophyll component phytol. For degradation, it must, as a β-methyl fatty acid, first be shortened by one methylene unit, in a process called α-oxidation (reviewed in [1]), then the resulting pristanic acid (2,6,10,14-tetramethylpentadecanoic acid) can be further degraded by β-oxidation, yielding propionyl-CoA and acetyl-CoA in alternation. Some pristanic acid is also derived directly from dietary sources [2]. Interest in the catabolism of branched-chain fatty acids has been revived by the observation that phytanic acid [3] as well as pristanic acid [4] might accumulate in the sera of patients with deficient peroxisome assembly.

Phytanic acid occurs naturally as a mixture of the (3*R*) and (3*S*) diastereomers [5]. Because α-oxidation does not affect the hydrogen atom in the β-position [6,7], the resulting product, pristanic acid, is likewise a mixture of diastereomers [5]. The oxidases and dehydrogenases responsible for further β-oxidation, however, act only on (2*S*)-2-methylacyl-CoAs [8,9]. We have previously shown that in both rat liver [10] and human tissues [11], α-methyl-branched fatty acids are racemized as CoA *S*-esters by a specific α-methylacyl-CoA racemase (EC 5.1.99.4). The enzyme has a rather broad spectrum of substrates, which includes bile acid intermediates such as di- and trihydroxycoprostanoyl-CoA. Peroxisomal β-oxidation of the CoA *S*-esters of di- and trihydroxycoprostanoic acid was recently shown to be also stereospecific for the (25*S*) isomer [8,12]. Because mitochondrial oxidation of the cholesterol side chain specifically affords the (25*R*) enantiomer [13], racemization seems to be an essential step in bile acid formation. Apart from its biological function, α-methylacyl-CoA racemase might also be interesting from a mechanistic point of view because it seems to act without any cofactor or prosthetic group.

Here we describe the molecular cloning, characterization and sequencing of a cDNA encoding rat liver α-methylacyl-CoA racemase and the expression of an active recombinant protein in *Escherichia coli* cells. A homologous cDNA was also cloned from a mouse liver cDNA library and its sequence compared.

EXPERIMENTAL

Materials

Molecular biology reagents from the following companies were used: rat liver cDNA expression library in λgt11, Marathon-Ready® cDNA and Advantage® cDNA PCR Kit from Clontech Laboratories (Palo Alto, CA, U.S.A.); mouse liver cDNA library in λZAP from Stratagene (La Jolla, CA, U.S.A.); Xpress System® and TA Cloning Kit from Invitrogen (San Diego, CA, U.S.A.); QIAGEN Lambda Kit, QIAquick® from Qiagen (Hilden, Germany); T7 Sequencing® Kit, Sure Clone Ligation Kit and dNTPs from Pharmacia (Uppsala, Sweden); restriction endonucleases from Boehringer (Mannheim, Germany); *Taq* DNA polymerase ('DynaZyme) from Finnzyme (Espoo, Finland). Oligonucleotides were synthesized with an Applied Biosystems model 394 oligonucleotide synthesizer.

Abbreviations used: IPTG, isopropyl β-D-thiogalactoside; MALDI–TOF-MS, matrix-assisted laser desorption ionization–time-of-flight MS; RACE, rapid amplification of cDNA ends; UTR, untranslated region.

[1] To whom correspondence should be addressed.

The nucleotide sequence results reported here will appear in DDBJ, EMBL and GenBank Nucleotide Sequence Databases under the accession numbers U89905 (rat) and U89906 (mouse).

Table 1 Oligonucleotide probes used for the cloning and sequencing of rat and mouse α-methylacyl-CoA racemase cDNA species

Probe name	Nucleotide sequence	Basis of the oligonucleotide
R3F	5′-GAR-TGG-TGY-CAR-ATY/A-TTY-GA-3′	Rat racemase peptide A (bp 839–858)
R4R	5′-RTC-RAA-R/TAT-YTG-RCA-CCA-YTC-3′	Rat racemase peptide A (bp 858–839)
R28R	5′-ATC-TGT-CCC-GTC-AAA-GAT-CTG-GCA-C-3′	Rat racemase cDNA bp 868–844
L11F	5′-GGT-GGC-GAC-GAC-TCC-TGG-AGC-C-3′	Vector arm: λgt11 forward
L11R	5′-GAC-ACC-AGA-CCA-ACT-GGT-AAT-G-3′	Vector arm: λgt11 reverse
AP1	5′-CCA-TCC-TAA-TAC-GAC-TCA-CTA-TAG-GGC-3′	Adaptor primer 1 (Supplied with the Marathon-Ready® cDNA)

Other chemicals used were purchased from the following suppliers: Bacto tryptone, agar and Bacto yeast extract from Difco (Augsburg, Germany); gelatine, glycerol, Tris, glycine, methanol and chloroform from E. Merck (Darmstadt, Germany); Tween-20, H_2O_2 [30% (v/v) solution] and phenol from Aldrich (Steinheim, Germany); acetic acid and $TiO(SO_4)$ from Riedel-de Haen (Seelze, Germany); nitrocellulose sheets BA 85 from Schleicher & Schuell (Dassel, Germany); affinity-purified goat anti-(rabbit IgG) antibodies conjugated with alkaline phosphatase, Quantum Prep Plasmid Miniprep Kit and the prestained SDS/PAGE Standards (low range) from Bio-Rad (Munich, Germany); Coomassie Brilliant Blue G-250 from Roth (Karlsruhe, Germany); agarose (gene technology grade) and reverse-phase silica gel (RP-18) from ICN chemicals (Meckenheim, Germany); gel-loading solution, ethidium bromide, DNAse I, RNAse A, 5-bromo-4-chloro-3-indolyl phosphate, Nitro Blue Tetrazolium and Nycodenz from Sigma Chemie (Deisenhofen, Germany); molecular mass marker for SDS/PAGE, DEAE-cellulose and reagents for the phosphate determination from Serva (Heidelberg, Germany); dimethylformamide, tetramethylethylenediamine, acrylamide, bisacrylamide, ammonium peroxodisulphate and SDS from Fluka (Neu-Ulm, Germany). All other chemicals were of A.R. grade or of the highest purity available.

Antibodies

Purification of the enzyme α-methylacyl-CoA racemase from rat liver and preparation of a monospecific antiserum in rabbits have been described previously [10,11]. The IgG fraction was isolated from immune sera by ion-exchange chromatography on DEAE-cellulose ([14]).

Tryptic peptides

Purified racemase (100 μg) was treated with 10 μg of trypsin overnight; the peptides were separated by HPLC on a Vydac 218TP column (5 μm, 250 mm × 2.1 mm; RP18) with a trifluoroacetic acid/acetonitrile gradient. Three peptides (A, B and C) were selected and subjected to automated Edman degradation with an Applied Biosystems model 477A Protein Sequencer.

Molecular cloning

Screening of a λgt11 rat liver cDNA expression library and of the λZAP mouse liver cDNA library was performed with rabbit anti-racemase antibodies in accordance with the protocols provided by the suppliers. To identify potentially correct clones among those positive in immunoscreening of the rat liver library, PCR was performed applying hot-start and touch-down methods. The reactions were performed in a thermocycler, with 1 μM oligonucleotides R3F and L11R (for sequences see Table 1) as primers, 200 μM each of the dNTPs and 0.5 μl of template, in a total volume of 50 μl of DynaZyme-buffer. In PCR, the reaction conditions were as follows: after the mixture had been overlaid with 25 μl of mineral oil and incubated for 2 min at 80 °C, 0.5 μl of DynaZyme was added. The temperature programme was 40 s at 94 °C (denaturation), 40 s at different annealing temperatures (see below) and 2 min at 72 °C (extension). Annealing temperatures were, consecutively, 69 °C, 66 °C, 63 °C (2 cycles each), 56 °C (10 cycles) and 48 °C (15 cycles). Subsequently the phages from verified positive clones of the λgt11 library were isolated, the inserts were excised with *Eco*RI, size-fractionated by agarose gel electrophoresis, purified with the QIAquick® kit and ligated into the pUC18 vector supplied with the Pharmacia Sure Clone Kit. For λZAP clones, the Bluescript vector was excised *in vivo* from positive clones with the Exassist® Helper phage.

The 5′ end of the rat liver racemase cDNA was amplified by PCR on a Marathon-Ready® rat liver cDNA library, in accordance with the protocol of the manufacturer, with oligo R28R as gene-specific primer and the vector-specific primer (AP1) supplied by the manufacturer (Table 1). The product of 5′ rapid amplification of cDNA ends (RACE) PCR was cloned directly into the pCR 2.1 plasmid (TA cloning kit).

Cloning of α-methylacyl-CoA racemase from mouse liver

A λZAP cDNA expression library of mouse liver was screened by using the antiserum against the rat enzyme. From positive clones, Bluescript vectors were excised *in vivo*, inserts were isolated by digestion with *Eco*RI and their sizes were determined by agarose-gel electrophoresis.

Nucleic acid sequencing

This was performed by the dideoxynucleotide sequencing method [15].

Northern blot

A rat multiple-tissue Northern blot was purchased from Clontech. The nylon filter contained 2 μg of poly(A)$^+$ RNA per lane from various rat tissues. The filter was prehybridized at 42 °C for 16 h and hybridized with ^{32}P-labelled racemase cDNA from rat liver for 20 h and was washed in accordance with the instructions supplied with the kit. The radioactive signals were detected by exposing X-ray film for 72 h. The amount of poly(A)$^+$ RNA was standardized with the rat β-actin probe supplied with the kit.

Production of recombinant rat liver racemase

The insert of λ-clone 11 was excised with *Eco*RI, cloned into the pTrcHis expression vector (versions A, B and C) of the Xpress

system® protein expression kit and transfected into *E. coli* (strain TOP 10: mcrA, D[mrr$^-$ hsd, RMS$^-$ mrcBC] F80DlacZDM15, DlacX74, deoR, recA1, araD139, D[ara, leu] 7697, galU, galK, I$^-$, rpsL, endA1, nupG, F$^-$) and the cell culture was performed in accordance with the manufacturer's instructions in 3 ml of Luria–Bertani medium containing 100 μg/ml ampicillin. The expression of recombinant protein was induced with isopropyl β-D-thiogalactoside (IPTG), the cells were collected by centrifugation, resuspended in 100 μl of water and disrupted by sonication (Cup horn, 1 min, 100 % intensity, 50 % duty cycle). The debris was removed by centrifugation. The resulting supernatant was used for measurements of α-methylacyl-CoA racemase activity with [2-^{3}H]pristanoyl-CoA and trihydroxy-[24,25-^{3}H]-coprostanoyl-CoA as substrates [10]. As controls, cells transformed with pTrcHisB without insert were treated in the same manner.

Other methods

Subcellular fractionation of mouse liver on a Nycodenz gradient was done essentially as described elsewhere for rat liver [9]. SDS/PAGE and Western blotting was done as described elsewhere [11]. Protein content was determined by the dye-binding method of Bradford [16].

RESULTS AND DISCUSSION

Cloning and sequencing of α-methylacyl-CoA racemase from rat liver

When a total of 2×10^6 plaques of a rat liver λgt11 expression library were screened with anti-racemase IgG, 19 positive clones were detected.

Clone 11 was taken as a template for PCR reactions with primers R3F (constructed on the basis of the sequence information of peptide A) and L11R (λDNA-specific primer; for sequences of primers see Table 1). The amplified PCR fragment was found to encode part of peptide A, indicating that clone 11 represented α-methylacyl-CoA racemase. The insert of the clone 11 was excised from the λ phage and subcloned; then both strands were sequenced. Clone 11 had a length of 1504 bp, with a 5′ untranslated region (UTR) of 58 bp followed by an open reading frame of 1083 bp. In addition to peptide A, the deduced amino acid sequence also gave a full match with peptides B and C. The open reading frame encoded a polypeptide with 361 residues and a molecular mass of 39679.2 Da (Figure 1). The predicted molecular mass agreed with the mass of the purified rat liver racemase (39540 $\pm$ 881) as determined by matrix-assisted laser desorption ionization–time-of-flight (MALDI–TOF) MS (Kompakt Maldi III; Kratos Analytical, Manchester, U.K.). Direct N-terminal sequencing gave negative results, suggesting that the N-terminus is blocked. At the 3′ end, a UTR of 360 bp included the AATAAA polyadenylation signal (actually two overlapping signals at positions 1472 and 1476) followed by $(T)_9$ and terminating with a poly(A)$^+$ tail.

Clone 19 was likewise isolated and its 5′ and 3′ termini were sequenced. The 5′ terminus was identical with that of clone 11 but only $(T)_8$ was found at the 3′ end. Only the 5′ end of clone 18 was sequenced; it was identical with that of the other two.

The C-terminus of the protein is -KANL, which was shown to be a functional peroxisomal targeting signal in human cells [17]. With the two-hybrid system using the human PTS1 receptor PEX5p as a bait, even -ANL alone was shown to be a moderately strong signal (A. Hartig, personal communication). Starting from position 15, there is also a sequence [RL(X)$_7$HL], which closely resembles the consensus peroxisomal targeting signal

```
-58  GCTGCGTGGCGTCAGGGTTCTGGAGCTGGCAGGCCTGGCCCCAGGGCCGTTCTGCGGGATG
  1                                                           M

  4  ATCCTGGCGGACTTCGGCGCCGAGGTGGTGCTCGTGGACAGACTGGGCTCCGTGAACCAC
  2  I  L  A  D  F  G  A  E  V  V  L  V  D  R  L  G  S  V  N  H

 64  CCCAGTCACCTGGCCCGAGGCAAGCGCTCGCTGGCGCTGGACCTGAAGCGGTCTCCGGGA
 22  P  S  H  L  A  R  G  K  R  S  L  A  L  D  L  K  R  S  P  G

124  GCCGCGGTGTTGCGGCGCATGTGCGCACGCGCGGACGTGTTGCTGGAGCCCTTCCGTTGC
 42  A  A  V  L  R  R  M  C  A  R  A  D  V  L  L  E  P  F  R  C

184  GGTGTCATGGAGAAACTCCAGCTTGGGCCAGAGACTCTACGGCAGGACAATCCAAAGCTC
 62  G  V  M  E  K  L  Q  L  G  P  E  T  L  R  Q  D  N  P  K  L

244  ATCTATGCCAGGCTGAGTGGATTTGGCCAGTCGGGAATTTTCTCCAAAGTAGCTGGCCAT
 82  I  Y  A  R  L  S  G  F  G  Q  S  G  I  F  S  K  V  A  G  H

304  GACATCAACTATGTGGCTTTGTCAGGTGTCCTGTCAAAGATTGGCAGGAGCGGTGAGAAC
102  D  I  N  Y  V  A  L  S  G  V  L  S  K  I  G  R  S  G  E  N

364  CCATACCCTCCCCTGAACCTCCTGGCCGACTTTGGTGGCGGTGGCCTCATGTGCACATTG
122  P  Y  P  P  L  N  L  L  A  D  F  G  G  G  G  L  M  C  T  L

424  GGCATTTTGCTGGCTCTCTTCGAACGCACGCGGTCTGGCCTAGGGCAGGTCATTGATGCG
142  G  I  L  L  A  L  F  E  R  T  R  S  G  L  G  Q  V  I  D  A

484  AACATGGTGGAAGGAACGGCATACTTAAGTACTTTCCTGTGGAAAACTCAGGCCATGGGT
162  N  M  V  E  G  T  A  Y  L  S  T  F  L  W  K  T  Q  A  M  G

544  CTGTGGGCACAGCCTCGAGGGCAAAACCTGTTAGATGGCGGGGCACCTTTCTACACAACC
182  L  W  A  Q  P  R  G  Q  N  L  L  D  G  G  A  P  F  Y  T  T

604  TACAAGACCGCAGATGGGGAGTTCATGGCTGTAGGTGCAATAGAACCCCAGTTCTACACA
202  Y  K  T  A  D  G  E  F  M  A  V  G  A  I  E  P  Q  F  Y  T

664  CTGCTGCTTAAAGGACTTGGACTTGAGTCTGAGGAACTCCCCAGCCAGATGAGCATAGAA
222  L  L  L  K  G  L  G  L  E  S  E  E  L  P  S  Q  M  S  I  E

724  GATTGGCCAGAAATGAAGAAGAAATTTGCAGATGTGTTTGCAAGGAAGACTAAGGCAGAG
242  D  W  P  E  M  K  K  K  F  A  D  V  F  A  R  K  T  K  A  E

784  TGGTGCCAGATCTTTGACGGGACAGATGCATGTGTGACCCCAGTGCTGACTCTTGAGGAG
262  W  C  Q  I  F  D  G  T  D  A  C  V  T  P  V  L  T  L  E  E

844  GCCCTCCACCACCAGCACAACAGAGAACGGGGCTCCTTCATCACTGATGAGGAGCAGCAT
282  A  L  H  H  Q  H  N  R  E  R  G  S  F  I  T  D  E  E  Q  H

904  GCATGCCCCCGTCCTGCACCCCAGCTTTCCAGAACCCCTGCTGTTCCTTCTGCCAAAAGG
302  A  C  P  R  P  A  P  Q  L  S  R  T  P  A  V  P  S  A  K  R

964  GACCCTTCTGTGGGAGAGCACACTGTAGAGGTGCTTAAAGACTATGGATTCAGTCAGGAA
322  D  P  S  V  G  E  H  T  V  E  V  L  K  D  Y  G  F  S  Q  E

1024 GAGATCCATCAGCTGCACTCGGATAGAATCATTGAAAGTAATAAGCTAAAAGCCAACCTC
342  E  I  H  Q  L  H  S  D  R  I  I  E  S  N  K  L  K  A  N  L

1084 TGACTCAGGTTCACAGCTCAAGTGAATCTGAAGGCTGTATCTGTACTGGAGAAGGATGCC
     *

1144 CACCACTGTCCGTATGGAAATGTGAATGAACAGTAATGAAGTAATCCAAATATTCCAATC

1204 AAGACACAACGAAAGACTGATTACAGAGAAATGACTGTGCTCTCACACTGCTCATCCGAG

1264 CCTCTGATTGAGGAGTATTTTTGTGTGTGTACTGATATTAACTTGTGGCAGTTTTCTGCC

1324 TTTCAGCTTACTTGGTGAAGTGCATTCACTGATTAAAACCCTTTTGTAAATGCAACTCTG

1384 ATAATATATTAAATGAACTAATATAACTTTAATAAATAAAGCTTTTTTTTCCTTGAAAAA

1444 AAA
```

Figure 1 cDNA for rat α-methylacyl-CoA racemase and its deduced amino acid sequence

Nucleotides are numbered in the 5′ to 3′ direction, beginning with the first nucleotide encoding the initiation methionine residue, which is termed +1. The first nucleotide of the stop codon is in position 1084. The amino acid residues are numbered taking the initiation methionine residue as +1. The polyadenylation signal is underlined. The sequences of the three tryptic peptides, which appear in the order C, B, A, are doubly underlined.

RL(X)$_5$HL [18]. Because the β-oxidation of 2-methyl-branched fatty acids (e.g. pristanic acid) as well as of di- and trihydroxycoprostanoic acid takes place in peroxisomes [19,20], the racemase would also be expected to be located in peroxisomes. In human tissues, this was found to be so [11]. In mouse, the enzyme was found equally distributed between mitochondria and peroxisomes (Figure 2), whereas in rat the racemase activity was found almost exclusively in mitochondria [11]. However, the N-terminus of the translated amino acid sequence and its helical wheel do not agree with the known features of mitochondrial targeting signals [21]. In view of this, the fact that the mass determined by SDS/PAGE (45 kDa [10]) was higher than that calculated from the open reading frame of clone 11 (39679 Da) and determined by MALDI–TOF (39540 Da) raised the question of whether the 5′-terminus might still be incomplete. We therefore conducted RACE-PCR with a Marathon-Ready® rat liver cDNA library (Clontech), which is constructed to include a high percentage of full-length clones. With a combination of a gene-specific 25-mer primer for upstream amplification (R28R; Table 1) and the vector-specific primer (AP1; Table 1) supplied by the manu-

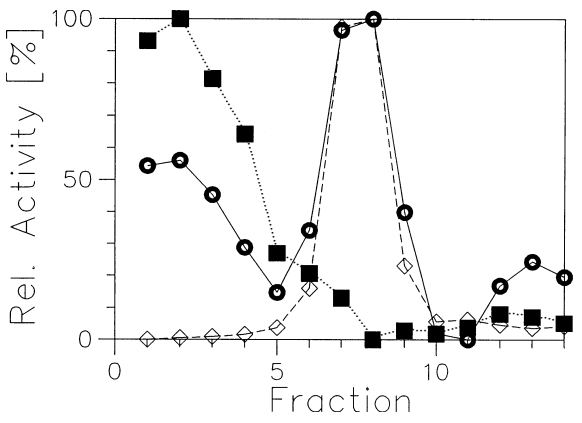

Figure 2 Subcellular distribution of α-methylacyl-CoA racemase in mouse liver

Mouse liver was fractionated on a Nycodenz/sucrose gradient, as described in the Experimental section. Aliquots (50 μl) of each fraction were assayed for α-methylacyl-CoA racemase activity (○) and for the marker enzymes catalase (■) and succinate dehydrogenase (◇).

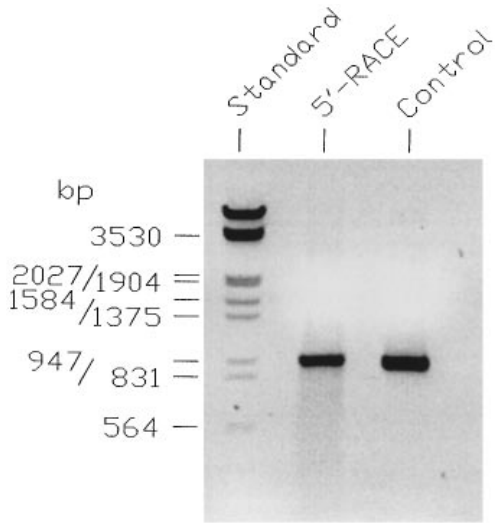

Figure 3 Agarose gel electrophoresis of the 5′-RACE PCR products

PCR was performed with oligonucleotides AP1 and either R28R (5′-RACE) or R3F (Control) (for oligonucleotides used, see Table 1) as primers and with a Marathon-Ready® rat liver cDNA library as template. After PCR, 2.5 μl aliquots were mixed with 2.5 μl of sample buffer (Sigma), applied to a 1% (w/v) agarose gel and subjected to electrophoresis at 75 V. The lane marked Standard is of size standards, λDNA digested with *Eco*RI and *Hin*dIII. For staining of DNA bands, ethidium bromide (0.5 μg/ml) was added to the gel solution.

facturer, only one product was obtained (Figure 3). The size of this product (approx. 950 bp) was very close to that expected from the cDNA sequence (921 bp). When the 5′ terminus of the PCR product was sequenced, the sequence was exactly the same as that of the λgt11 clones (11, 18 and 19). The other 16 λ clones were also analysed but did not give any additional information. One possible explanation for the lack of any mitochondrial targeting signal in our cDNA clones might be that some structural features, such as unusual secondary structures, invariably lead to the premature termination of reverse transcriptase action during first-strand synthesis of cDNA species. Examination of this possibility will have to await analysis of the genomic DNA.

Northern blots with mRNA from various rat tissues showed significant expression of the enzyme only in liver and kidney (Figure 4). In each of these two lanes, only one band was seen, corresponding to a transcript of approx. 1650 bp.

Expression of rat liver cDNA in *E. coli*

The insert of λgt11 clone 11 was excised with *Eco*RI and was cloned into the expression vector pTrcHis, versions A, B and C (for the three possible reading frames). After transformation of *E. coli* strain Top 10 with the constructs, induction of the transformed cells with IPTG revealed that reading frame B led to the production of a functional racemase (Figure 5). Expression of the recombinant enzymically active protein confirms that the cloned cDNA codes for α-methylacyl-CoA racemase.

In Western blotting of soluble extract from induced *E. coli*, the antibody against rat racemase recognized a protein band of 51 kDa (Figure 6), approx. 6 kDa larger than the native protein from rat liver. The difference is accounted for by the expression construct (which includes an additional 45 residues from the vector and 21 residues from the linker and the 5′UTR of the cDNA). The open reading frame of the construct translates to a polypeptide of 46742 Da. As with the native protein, the size determined by SDS/PAGE was approx. 10% larger than that calculated from the amino acid sequence, indicating a significantly deviant behaviour of this protein in SDS/PAGE. The reason for this discrepancy is not yet known.

Cloning and sequencing of α-methylacyl-CoA racemase from mouse liver

The cross-reaction of the antibody against the rat liver α-methylacyl-CoA racemase with the mouse liver enzyme (Figure 6) allowed the use of the antiserum against the rat enzyme to screen a λZAP cDNA expression library of mouse liver. Screening of 10^6 plaques yielded eight positive clones, all of which were purified. They all had inserts of a similar size, approx. 1500 bp. Two clones were fully sequenced in both directions with identical results. Of four other clones, only the 5′-termini were sequenced; they all proved to be identical with those of the first two clones. The complete mouse cDNA sequence is shown in Figure 7.

The total size (clone 3) was 1484 bp, an open reading frame starting at the second base. A second ATG in the same reading frame is at position 26. The latter corresponds to the start codon of the rat sequence. It is not yet clear which start codon is used for translation in the mouse. In Western blots, the mouse liver protein seemed to have the same size as the rat protein (Figure 6), which would argue for usage of the second start codon. Neither of the two possible N-termini constitutes a mitochondrial targeting sequence. As in rats, the C-terminus is -KANL, suggesting a peroxisomal localization. Depending on which start codon is used, the protein can be predicted to consist of 360 or 368 residues (molecular mass 39558 or 40319 Da). The 3′ UTR is 379 bases long and contains the same overlapping two polyadenylation signals (positions 1437 and 1441 respectively) as the rat cDNA and terminates with a poly(A)$^+$ tail.

The rat and mouse cDNA sequences are highly similar (86.5% identity). Interestingly, the mouse cDNA is not only shorter at the 5′ end but also lacks one complete triplet (bases 113–115 of the rat sequence). The amino acid sequences derived from the cDNA species share approx. 89.7% similarity (88.6% identity) (Figure 8).

Several data banks, both those of nucleotide (GenBank, EMBL, DDBJ, PDB) and protein [SwissProt, PDP, CDS translations, PIR, yeast (*Saccharomyces cerevisiae*)] sequences, were searched for related sequences with the BLAST algorithm [22].

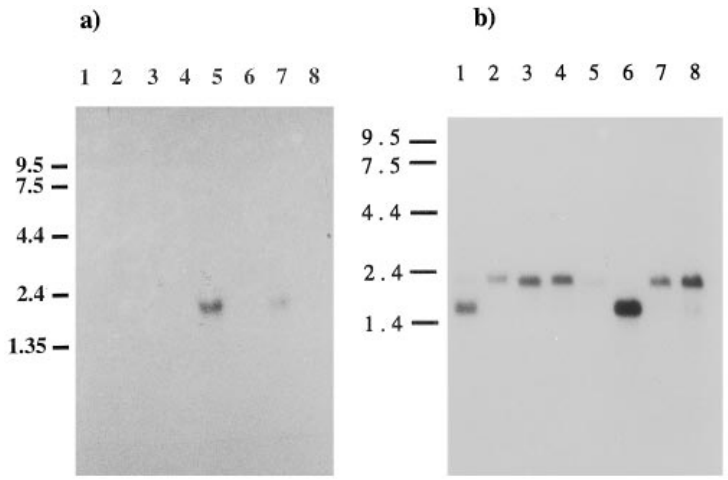

Figure 4 mRNA levels of α-methylacyl-CoA racemase in different rat tissues

Aliquots (2 μg) of mRNA isolated from various rat tissues were hybridized with cDNA for α-methylacyl-CoA racemase (**a**) and human β-actin cDNA control probe (**b**) to verify the amount of mRNA applied to lanes in the Northern blot. The size of the racemase transcript was approx. 1650 bp. The mRNA was isolated from: lane 1, heart; lane 2, brain; lane 3, spleen; lane 4, lung; lane 5, liver; lane 6, skeletal muscle; lane 7, kidney; lane 8, testis.

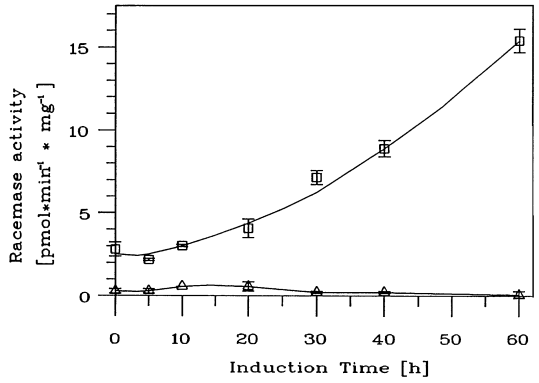

Figure 5 Expression of recombinant rat α-methylacyl-CoA racemase in *E. coli* as function of induction time

E. coli strain Top 10, transformed with pTrcHis vector variant B, containing the complete cDNA clone for the rat enzyme (□) or without insert (△) were induced by the addition of 1 mM IPTG to the culture medium. Samples were drawn at the indicated times, cells were immediately collected by centrifugation and homogenized, and their α-methylacyl-CoA racemase activity was determined with [2-^{3}H]pristanoyl-CoA as substrate. (For details see the Experimental section.)

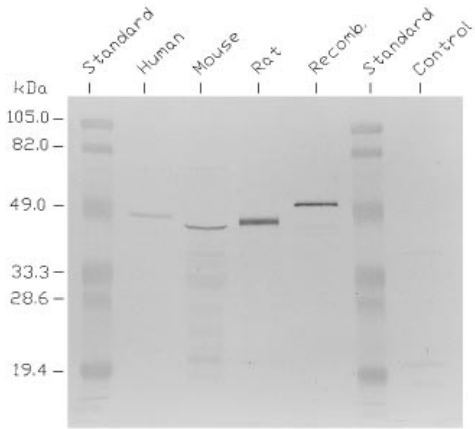

Figure 6 Immunoblotting of mouse, rat and human liver homogenates and of recombinant rat α-methylacyl-CoA racemase in *E. coli* with the antiserum against the rat liver enzyme

Lanes 1 and 6, molecular mass standards (their molecular masses are given at the right); lane 2, homogenate of human liver (15 μg of protein); lane 3, homogenate of mouse liver (7.5 μg of protein); lane 4, homogenate of rat liver (1.5 μg of protein); lane 5, homogenate of *E. coli* cells expressing recombinant rat enzyme (total protein 10 μg); lane 7, homogenate of control *E. coli* cells (total protein 10 μg). Rabbit anti-(rat liver α-methylacyl-CoA racemase) IgG was used as primary antibody.

An essentially identical sequence was found [23] and has been deposited by C. Reichel and colleagues for a rat liver enzyme termed 2-arylpropionyl-CoA epimerase (accession number Y08172). This enzyme had attracted attention because of its role in the pharmacology of ibuprofen and similar analgesic and anti-inflammatory compounds of the 2-arylpropionic acid type. Its identity with the 2-methylacyl-CoA racemase had been suspected before [10] and can now be regarded as established. Besides this, the best match found was a protein sequence from *Caenorhabditis elegans* (hypothetical 37.1 kDa protein ZK892.4 in chromosome III), which should have a size (335 residues) similar to that of the racemase and which in six regions between 20 and 88 residues in length (total of 225 residues) had 113 identical and another 42 structurally similar amino acids, resulting in 44.9 % identity with the rat racemase (Figure 8). Some similarity was also found with *E. coli* L-carnitine dehydratase (three regions, containing altogether 74 amino acids, of which 23 were identical and 14 similar). Identity with the rat racemase was 24.9 %. Because L-carnitine dehydratase might act via a similar mechanism (abstraction of a proton) as that proposed for the racemase [10], this could provide some evolutionary clue.

No similarities were found with other CoA-binding proteins. So far, the three-dimensional structures of approx. 15 such proteins are known. Except for the 3′-phosphate group pointing

```
-25   CATGGCCCCGGGGCCGTTCTGCGGAATGGTCCTGGCGGACTTCGGCGCCGAGGTGGTGCGC
  1                          M  V  L  A  D  F  G  A  E  V  V  R

 37   GTGAACCGGCTGGGCTCCACGGGCGAGAATTTTCTGGCCCGAGGCAAGCGCTCGCTAGCG
 13   V  N  R  L  G  S  T  G  E  N  F  L  A  R  G  K  R  S  L  A

 97   CTGGACCTGAAGCGCTCTCAGGGAGTCACGGTGTTGCGGCGCATGTGCGCACGCGCGGAC
 33   L  D  L  K  R  S  Q  G  V  T  V  L  R  R  M  C  A  R  A  D

157   GTGTTGCTGGAGCCCTTCCGCTGCGGTGTCATGGAGAAACTCCAGCTTGGGCCAGAGACT
 53   V  L  L  E  P  F  R  C  G  V  M  E  K  L  Q  L  G  P  E  T

217   CTACTGCAGGACAATCCAAAGCTCATCTATGCCAGGCTGAGCGGATTTGGCCAATCGGGA
 73   L  L  Q  D  N  P  K  L  I  Y  A  R  L  S  G  F  G  Q  S  G

277   ATTTTCTCCAAAGTAGCTGGCCATGACATCAACTATTTGGCTTTATCAGGCGTTCTGTCA
 93   I  F  S  K  V  A  G  H  D  I  N  Y  L  A  L  S  G  V  L  S

337   AAGATTGGCAGAAGCGGTGAGAACCCCTACCCACCGCTGAATCTCCTGGCTGACTTTGGC
113   K  I  G  R  S  G  E  N  P  Y  P  P  L  N  L  L  A  D  F  G

397   GGTGGAGGCCTCATGTGCACACTGGGCATTGTGCTGGCTCTCTTTGAACGCACACGCTCT
133   G  G  G  L  M  C  T  L  G  I  V  L  A  L  F  E  R  T  R  S

457   GGCCGAGGGCAGATCATCGATTCAAGCATGGTGGAAGGGACTGCATACTTAAGTTCTTTC
153   G  R  G  Q  I  I  D  S  S  M  V  E  G  T  A  Y  L  S  S  F

517   CTGTGGAAAACCCAGCCCATGGGTCTGTGGAAACAGCCTCGAGGACAAAACATCTTAGAT
173   L  W  K  T  Q  P  M  G  L  W  K  Q  P  R  G  Q  N  I  L  D

577   GGCGGTGCACCTTTCTACACAACCTACAAGACGGCAGACGGGGAGTTCATGGCTGTAGGT
193   G  G  A  P  F  Y  T  T  Y  K  T  A  D  G  E  F  M  A  V  G

637   GCCATAGAACCCCAGTTCTATGCACTGCTGCTTAAAGGACTTGGACTCGAGTCTGAGGAA
213   A  I  E  P  Q  F  Y  A  L  L  L  K  G  L  G  L  E  S  E  E

697   CTCCCCTCCCAGATGAGCTCAGCAGATTGGCCAGAGATGAAGAAGAAATTTGCAGATGTG
233   L  P  S  Q  M  S  S  A  D  W  P  E  M  K  K  K  F  A  D  V

757   TTTGCAAAGAAGACTAAGGCAGAATGGTGCCAGATCTTTGACGGGACAGATGCGTGTGTG
253   F  A  K  K  T  K  A  E  W  C  Q  I  F  D  G  T  D  A  C  V

817   ACCCCAGTGCTGACGTTTGAGGAGGCCCTCCACCACCAGCACAACAGAGAACGGGCCTCC
273   T  P  V  L  T  F  E  E  A  L  H  H  Q  H  N  R  E  R  A  S

877   TTCATCACTGATGGGGAGCAGCTCCCGAGCCCCGCCCTGCACCTCTGCTTTCCAGAACT
293   F  I  T  D  G  E  Q  L  P  S  P  R  P  A  P  L  L  S  R  T

937   CCTGCCGTCCCATCTGCCAAAAGGGACCCTTCTGTAGGGGAGCACACCGTAGAAGTGCTT
313   P  A  V  P  S  A  K  R  D  P  S  V  G  E  H  T  V  E  V  L

997   AGAGAGTATGGATTCAGTCAGGAAGAGATCCTTCAGCTGCACTCAGATAGAATCGTTGAA
333   R  E  Y  G  F  S  Q  E  E  I  L  Q  L  H  S  D  R  I  V  E

1057  AGTGATAAGCTAAAAGCCAATCTCTGACTCAGGCTTATAGCTCAAGAGAATCTGAAGGCT
353   S  D  K  L  K  A  N  L  *

1117  GCATCTCCACTGGGGAGGGATGCCCACAATTGTGTGTATGGAAATGTGGATGAACAGCAA

1177  TGAAGTCATCCAAATATCCCAATCACGATCCAACGAAAGGCTGGTTACAGGATAACGATT

1237  GCGCCCTTACGCTGCTTATCAGAACCTCTGATTGAGGAAAATTTGTGTGTGTGTACTGAT

1297  AGTAACTTGTGGCAGCTTTCTGCCTTTCAGTTCCCTTGGTGAAGTGTATCCATTTATTAA

1357  AACCCTTTTATAAATACAACTCTGATCATATATTAAATAAACTATTATAACTTTAATAAA

1417  TAAAACTCTTGTTTTCCTCCTCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

1477  AAAAAAAAAAAAA
```

Figure 7 cDNA for mouse α-methylacyl-CoA racemase and its deduced amino acid sequence

Nucleotides are numbered in the 5′ to 3′ direction, beginning with the first nucleotide encoding the (probable; see text) initiation methionine residue, which is termed +1. The first nucleotide of the stop codon is in position 1081. The amino acid residues are numbered taking the (probable) initiation methionine residue as +1. The polyadenylation signal is underlined.

```
RRACE   ------------------------------MLLADFGAEVVLVDRLGSVNHPS-------
MRACE   ------------------------------MVLADFGAEVVRVNRLGSTGEN--------
CAEL    -----MYRF--LSGIKVVEIAGLAPVPHCGMVLADFGADVTVIDKVRYINFISCWRTNNA
LCDHT   MDHLPMPKFGPLAGLRVVFSGIEIAGPFAGQMFAEWGAEVIWIENVAWADTIR-------

RRACE   ----------HLARGKRSLALDLKRSPGAAVLRRMCARADVLLEPFRCGVMEKLQLGPE
MRACE   ----------FLARGKRSLALDLKRSQGVTVLRRMCARADVLLEPFRCGVMEKLQLGPK
CAEL    VNFQKNPAIEQRLNRGKTMKQLDLKNPEDIKKVRDLCQTSDVLLDEYRPGTLEKMGLDPS
LCDHT   ----VQPNYPQLSPRNLHALSLNIFKDEGREAFLKLMETTDIFIEASKGPAFARRGITDE

RRACE   TLRQDNPKLIYARLSGFGQSGTFSKVAGHDINYVA--LSGVLSKIG-RSGENPYPPLNLL
MRACE   TLLQGQSKAHLCQAERIWPIGNFLQIAGHDINYLA--LSGVLSKIG-RSGENPYPPLNLL
CAEL    TLWNNNKGLILCKISGYGQTGRMSQETGHDINYVA--LSGVLPTFSGVNATRPWPPANML
LCDHT   VLWQHNPKLVIAHLSGFGQYGTEEYTNLPAYNTIAQAFSGYLIQNG--DVDQPMPAFPYT

RRACE   ADFGGGGLMCTLGILLALFERTRSG-LGQVIDANMVEGTAYLSTFLWKTQAMGLWAQPRG
MRACE   ADFGGGGLMCTLGIVLALFERTRSG-RGQIIDSSMVEGTAYLSSFLWKTQPMGLWKQPRG
CAEL    ADFAGGGLSAAFGILSAIYARSHNGGKGCLLDCSMTEGVAYLSSFVQHYYDQPNLFTDK-
LCDHT   ADYFSG-LTATTAALAALHKVRETG-KGESIDIAMYEVMLRMGQYFMMDYFNGGEMCPR-

RRACE   QNLLDGGAPFYT---TYKTADG----EFMAVGAIEPQFYTL-LLKGLGLESEELPSQMSI
MRACE   QNILDGGAPFYT---TYKTADG----EFMAVGAIEPQFYAL-LLKGLGLESEELPSQMSS
CAEL    YAIFSGEGPLYR---TYKTKDD----KFVAVGAVEPKFYQN-LFKILNVDGR--------
LCDHT   --MSKGKDPYYAGCGIYKGADGYIVMELVGITQTIEQGKDIGLAHLLGTPEIPEGTQLIH

RRACE   ----PDWPEMKKKFADVFARKTKAEWCQIFDGTDACVTPVLTLEEALHHQHNRERGSFIT
MRACE   ----ADWPEMKKKFADVFAKKTKAEWCQIFDGTDACVTPVLTLEEALHHQHNRERASFIT
CAEL    ---------------DLFVNP----------GQECCVTPVLDIHEVGSYGQHVDRNSETK
LCDHT   RIECPYGELVEEKLDAWIATHTIAEVKERFAELNIACAKVLTVPELESNPQYVARRSITQ

RRACE   DEEQHACPRPAP-QLSRTPAVPSAKRD--PSVGEHTVEVLKDYGFSQEEIHQLHSDRIIE
MRACE   DGEQLPSPRPAP-LLSRTPAVPSAKRD--PSVGEHTVEVIREYGFSQEEIHQLHSDRIVE
CAEL    TS------------SNWIANPSPR-----------------VWTQDELAALSSKK---
LCDHT   WQTMDGRTCKGFNIMPKFKNNFGQIWRGMPSHCMDTAAILKNICYSENDTQELVSKGLAK

RRACE   SNKLKANL
MRACE   SDKLKANL
CAEL    --------
LCDHT   VED-----
```

Figure 8 Alignments of rat and mouse racemase, *C. elegans* protein ZK 894.2 and *E. coli* L-carnitine dehydratase

The amino acid sequences of the rat and mouse racemases (RRACE and MRACE respectively) showed an overall identity of 88.6%. The RRACE had 44.9% and 24.9% identities with *C. elegans* hypothetical 37.1 kDa protein ZK 894.2 (CAEL) and *E. coli* L-carnitine dehydratase (LCDHT) respectively. White letters on black indicate amino acid residues identical with those in RRACE. Amino acid residues identical with those in MRACE are printed on a grey background. Multiple comparison of the amino acid sequences of the polypeptides was performed with the Clustal W program with default parameters.

out into the aqueous phase and the adenine ring facing the protein, there are no conserved features concerning the interaction between protein and coenzyme A [24]. The amino acid residues that participate in substrate binding are often distributed widely along the polypeptide chain and no conserved amino acid sequence motifs for CoA binding have emerged.

A great stereochemical variety of fatty acyl-CoA esters and their derivatives can undergo chain shortening by β-oxidation, although the purified enzymes of the pathway often show strict stereospecificity. The introduction of α-methylacyl-CoA racemase as an auxiliary enzyme of β-oxidation opens up the route for the oxidation of diastereomers of α-methyl-branched acyl-CoA esters. The molecular cloning of mammalian racemase(s) is an essential step in the future screening for potential inborn errors affecting this enzyme system or for studying more closely the role of racemase by methods involving transgenic animals. Interestingly, no significant homologies with racemase were found among polypeptides encoded by open reading frames of the *S. cerevisiae* genome, and this enzyme did not show any significant similarities with known amino acid sequences of enzymes of β-oxidation or its auxiliary enzymes.

We thank Monika Kempf for valuable help with cloning the mouse gene; Dr. J. Hoppe and Dr. V. Hoppe for protein sequencing; Wolfgang Hädelt for DNA sequencing; Eeva-Liisa Stefanius for synthesis of various oligonucleotides; and Dr. K.-H. Friedrich for invaluable practical advice. This work was financially supported by the Deutsche Forschungsgemeinschaft (grant Co 116/2-3), the Academy of Finland and the Sigrid Juselius Foundation.

REFERENCES

1 Steinberg, D. (1995) in The Metabolic Basis of Inherited Disease (Scriver, C. R., Beaudet, A. L., Sly, W. S. and Valle, D., eds.), 7th edn., pp. 2351–2369, McGraw-Hill, New York
2 Hansen, R. P. and Morrison, J. D. (1964) Biochem. J. **93**, 225–229
3 Poulos, A., Sharp, P. and Whiting, M. (1984) Clin. Genet. **25**, 579–586
4 Poulos, A., Sharp, P., Fellenberg, A. J. and Johnson, D. W. (1988) Eur. J. Pediatr. **147**, 143–147
5 Ackman, R. G. and Hansen, R. P. (1967) Lipids **2**, 357–362
6 Avigan, J., Steinberg, D., Gutman, A., Mize, C. E. and Milne, G. W. A. (1966) Biochem. Biophys. Res. Commun. **24**, 838–844
7 Fingerhut, R., Schmitz, W. and Conzelmann, E. (1993) J. Inher. Metab. Dis. **16**, 591–594
8 Van Veldhoven, P. P., Croes, K., Asselberghs, S., Herdewijn, P. and Mannaerts, G. P. (1996) FEBS Lett. **388**, 80–84
9 Schmitz, W. and Conzelmann, E. (1997) Eur. J. Biochem. **244**, 434–440
10 Schmitz, W., Fingerhut, R. and Conzelmann, E. (1994) Eur. J. Biochem. **222**, 313–323
11 Schmitz, W., Albers, C., Fingerhut, R. and Conzelmann, E. (1995) Eur. J. Biochem. **231**, 815–822
12 Pedersen, J. I., Veggan, T. and Björkhem, I. (1996) Biochem. Biophys. Res. Commun. **224**, 37–42

13 Shefer, S., Cheng, F. W., Batta, A. K., Dayal, B., Tint, G. S., Salen, G. and Mosbach, E. H. (1978) J. Biol. Chem. **253**, 6386–6392
14 Catty, D. (ed.) (1993) Antibodies, vol. 1, p. 72, IRL Press, Oxford
15 Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467
16 Bradford, M. M. (1976) Anal. Biochem. **72**, 248–254
17 Purdue, E. and Lazarow, P. B. (1996) J. Cell Biol. **134**, 849–862
18 De Hoop, M. J. and Ab, G. (1992) Biochem. J. **286**, 657–669
19 Pedersen, J. I. and Gustafsson, J. (1980) FEBS Lett. **121**, 345–348
20 Pedersen, J. I., Kase, B. F., Prydz, K. and Björkhem, I. (1987) in Peroxisomes in Biology and Medicine (Fahimi, H. D. and Sies, H., eds.), pp. 67–77, Springer-Verlag, Berlin
21 Roise, D., Horvath, S. J., Tomich, J. M., Richards, J. H. and Schatz, G. (1986) EMBO J. **5**, 1327–1334
22 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) J. Mol. Biol. **215**, 403–410
23 Reichel, C., Brugger, R., Bang, H., Geisslinger, G. and Brune, K. (1997) Mol. Pharmacol. **51**, 576–582
24 Engel, C. and Wierenga, R. (1996) Curr. Opin. Struct. Biol. **6**, 790–797

Received 27 February 1997/23 May 1997; accepted 3 June 1997