

# cDNA Cloning of the Human Peroxisomal Enoyl-CoA Hydratase: 3-Hydroxyacyl-CoA Dehydrogenase Bifunctional Enzyme and Localization to Chromosome 3q26.3-3q28: A Free Left *Alu* Arm Is Inserted in the 3' Noncoding Region

GERALD HOEFLER,\*<sup>1</sup> MICHAEL FORSTNER,\* MARTINA C. MCGUINNESS,† WOLFGANG HULLA,\*  
MICHAELA HIDEN,‡ PETER KRISPER,\* LUKAS KENNER,\* THOMAS RIED,§ CHRISTOPH LENGAUER,§  
RUDOLF ZECHNER,‡ HUGO W. MOSER,† AND GRACE L. CHEN†

\*Institute of Pathology and †Institute of Medical Biochemistry, University of Graz, Austria; †Kennedy Institute for Handicapped Children, Johns Hopkins School of Medicine, Baltimore, Maryland; and §Institute of Human Genetics, University of Heidelberg, Germany

Received April 5, 1993; revised September 7, 1993

**Enoyl-CoA hydratase:3-hydroxyacyl-CoA dehydrogenase bifunctional enzyme is one of the four enzymes of the peroxisomal  $\beta$ -oxidation pathway. Here, we report the full-length human cDNA sequence and the localization of the corresponding gene on chromosome 3q26.3-3q28. The cDNA sequence spans 3779 nucleotides with an open reading frame of 2169 nucleotides. The tripeptide SKL at the carboxy terminus, known to serve as a peroxisomal targeting signal, is present. DNA sequence comparison of the coding region showed an 80% homology between human and rat bifunctional enzyme cDNA. The 3' noncoding sequence contains 117 nucleotides homologous to an *Alu* repeat. Based on sequence comparison, we propose that these nucleotides are a free left *Alu* arm with 86% homology to the *Alu-J* family. RNA analysis shows one band with highest intensity in liver and kidney. This cDNA will allow in-depth studies of molecular defects in patients with defective peroxisomal bifunctional enzyme. Moreover, it will also provide a means for studying the regulation of peroxisomal  $\beta$ -oxidation in humans. © 1994 Academic Press, Inc.**

## INTRODUCTION

Peroxisomal  $\beta$ -oxidation of fatty acids, originally described by Lazarow and DeDuve (1976), is catalyzed by four enzymes, fatty acyl-CoA synthetase, acyl-CoA oxidase, enoyl-CoA hydratase:3-hydroxyacyl-CoA dehydrogenase bifunctional enzyme (referred to as the bifunctional enzyme in this paper), and 3-ketoacyl-CoA thiolase. In rat liver, these enzymes differ from their mitochondrial counterparts physically, catalytically, and

immunochemically (Hashimoto, 1982) and catalyze the chain shortening of a variety of different fatty acyl-CoAs (Lazarow, 1978).

The bifunctional enzyme as well as other peroxisomal proteins is synthesized on free polysomes, released into the cytosol, and then transported into preexisting peroxisomes (Miura *et al.*, 1984; Fujiki *et al.*, 1985). The peroxisomal  $\beta$ -oxidation system is active toward long-chain fatty acids (Singh *et al.*, 1984, 1987) and it has been demonstrated that impairment of this pathway leads to elevated long-chain fatty acid levels in patients with peroxisomal disorders (Singh *et al.*, 1984) such as Zellweger syndrome, neonatal adrenoleukodystrophy, and infantile Refsum's disease (for a review, see Lazarow and Moser, 1989). The bifunctional enzyme protein has been shown to be absent in patients affected with these disorders (Chen *et al.*, 1987). Additionally, patients with isolated defects of the peroxisomal bifunctional enzyme who either lack the protein (Watkins *et al.*, 1989; Guerroui *et al.*, 1989) or have an inactive form of the enzyme (Wanders *et al.*, 1990) have been reported.

To deal with questions concerning specific defects of the human peroxisomal bifunctional enzyme protein, we cloned and characterized a full-length human cDNA coding for the bifunctional enzyme using cDNA sequences from the rat (Osumi *et al.*, 1985). Additionally, we determined the chromosomal location of its corresponding gene. This information will allow the cloning of the gene for the human peroxisomal bifunctional enzyme of normal individuals and patients with defects of the bifunctional enzyme. Furthermore, it will provide the basis for studying tissue-specific expression of the enzyme and its regulation on the transcriptional and mRNA levels.

## MATERIALS AND METHODS

*cDNA library screening.* A 2.3-kb cDNA probe (Osumi *et al.*, 1985), derived from the coding region of the rat bifunctional enzyme

<sup>1</sup> To whom correspondence should be addressed at the Institute of Pathology, University of Graz Medical School, Auenbruggerplatz 25, A-8036 Graz, Austria, Telephone: 43 316 380 4429. Fax: 43 316 384329.

(kindly provided by T. Hashimoto), was used to screen a human fetal liver cDNA library according to standard procedures (Sambrook *et al.*, 1989). A single clone (pBC5) representing human bifunctional enzyme cDNA was isolated and used to screen two human adult liver cDNA libraries by conventional methods or by a modification of the PCR-facilitated cloning procedure (Trahey *et al.*, 1988; Hoefler *et al.*, 1991) using T3 sequencing primer (5'-ATTAACCCCTCACTAAAG-3') and bifunctional enzyme-specific primer H3 (5'-CAGCAAGATCAG-ACACTCTA-3').

To obtain cDNA from the 5' end, a specifically primed cDNA library, constructed in  $\lambda$ gt10 (Clontech Lab., Palo Alto, CA) using oligonucleotide primer P2 (5'-GATTCTAGAGGTTGATCTGA-3') and human liver RNA, was screened by the PCR-facilitated cloning procedure using pBC5, bifunctional enzyme-specific primer P3 (5'-TCT-GATTGCTTCTTCAACCG-3'), and a primer for the right (5'-TGA-GTATTCTTCCAGGGTA-3') or left (5'-CTTTTGAGCAAGTTC-AGCCT-3') lambda phage arm, respectively.

**DNA sequencing and primer extension analysis.** After verification by Southern blotting using pBC5 as a probe, either PCR products were cloned into pCRII vector using the TA-cloning kit (Invitrogen, San Diego, CA) or the respective phages were plaque purified and converted into plasmids using helper phage *in vitro* excision ( $\lambda$  ZAP, Stratagene).

cDNA inserts were sequenced using Sequenase II (U.S. Biochemical Corp., Cleveland, OH) or a Taquence kit (U.S. Biochemical Corp.) according to the manufacturer's recommendations. For some clones, unidirectional deletions were produced using an *ExoIII*/Mung Bean Deletion kit (Stratagene). The sequence was determined on both strands for several overlapping clones. Sequences were compared, assembled, and analyzed by the DNASIS software system (Hitachi Software Engineering Co., Ltd., Yokohama, Japan).

Primer extension analysis was performed according to standard protocol (Sambrook *et al.*, 1989) using 5  $\mu$ g of total human liver RNA, 10 pmol of a fluorescein-labeled primer corresponding to bases 14 to 43 of the cDNA of the human peroxisomal bifunctional enzyme and 50 units of either Moloney murine leukemia virus (M-MLV) or avian myeloblastosis virus (AMV) reverse transcriptase (Promega Co., Madison, WI). Products were analyzed by denaturing polyacrylamide gel electrophoresis (PAGE) on an A.L.F. DNA Sequencer (Pharmacia LKB, Uppsala, Sweden) and the fragment size was determined by comparison to a standard sequencing reaction.

**RNA blotting.** Total RNA was prepared from the human hepatoma cell line HepG2 using 4 M guanidinium isothiocyanate (Sambrook *et al.*, 1989). Northern blot analysis was performed after formaldehyde agarose gel electrophoresis of RNA samples (10  $\mu$ g each) and RNA size markers I (Boehringer Mannheim, Germany). Gels were washed, and the RNA was transferred to Hybond N<sup>+</sup> membrane (Amersham, Bucks., UK) and prehybridized according to standard techniques. RNA was probed with hybridization buffer containing randomly primed <sup>32</sup>P-labeled cDNA for human bifunctional enzyme (pBC93; see below).

**S1 nuclease protection assay.** Poly(A)<sup>+</sup> RNA was isolated from human liver, kidney, and brain tissues and cultured human hepatoma HepG2 cells using the Fast Tract mRNA isolation kit (Invitrogen). Three micrograms of Poly(A)<sup>+</sup> RNA was hybridized to <sup>32</sup>P-labeled antisense RNA probe (10<sup>9</sup> cpm/ $\mu$ g) generated by *in vitro* transcription using the TransProbe kit (Pharmacia LKB) from a cDNA clone (pBC19) containing the first 591 nucleotides of the human bifunctional enzyme. Hybridization was performed in 20  $\mu$ l of a buffer containing 60% formamide, 0.9 M NaCl, 6 mM EDTA, 60 mM Tris-HCl, pH 7.4, and 2.5 mg yeast tRNA/ml for 16 h at 68°C. Digestion of single-stranded nucleic acids was carried out using 540 U S1 nuclease (Pharmacia) in 300  $\mu$ l of a buffer containing 0.3 M NaCl, 30 mM sodium acetate, pH 4.8, and 3 mM ZnCl<sub>2</sub> for 1 h at 56°C. After phenol extraction and ethanol precipitation, the samples were separated on 5% PAGE, and the gels were dried and subjected to autoradiography.

**Chromosomal mapping by fluorescence *in situ* hybridization (FISH).** The chromosomal localization of the human bifunctional enzyme was determined by FISH. Human metaphase spreads were prepared from 5-bromodeoxyuridine-synchronized cultures of a

healthy female donor (46, XX) (Yunis, 1976) with the following modifications: cells were incubated with methotrexate (10<sup>-7</sup> M, final concentration) for 17 h at 37°C, collected by centrifugation, resuspended in fetal calf serum (FCS)-free medium, and then washed twice with FCS-free medium. The block was released by adding 5-bromodeoxyuridine (30  $\mu$ g/ml, final concentration) for 5 h at 37°C.

Insert preparations of two DNA fragments specific for the bifunctional enzyme (pBC93 and pBC19; see Results) were labeled in separate nick-translation reactions with biotin-11-dUTP. The insert of clone pBC93 was released from the *Alu* element at its 3' noncoding region by restriction enzyme digestion with *NdeI* and *EcoRI*. For FISH analysis, the probes were combined and hybridized to human metaphase chromosome spreads following standard protocols (Lengauer *et al.*, 1990). The biotinylated probe sequences were detected with avidin-DCS-fluorescein-isothiocyanate (Vector Lab., Burlingame, CA). The chromosome specimen was counterstained with 4,6-diamidino-2-phenylindole (DAPI), and the resulting G-banding pattern and the signal from the cDNA probes were imaged separately using a Zeiss Axiophot epifluorescence microscope coupled to a cooled CCD camera (Photometrics, Tucson, AZ). The sequentially recorded gray-scale images were pseudocolored and merged using an Apple Mac-Intosh IIX computer and custom computer software (for details see Ried *et al.*, 1992).

## RESULTS

### *Isolation of a cDNA for the Human Peroxisomal Bifunctional Enzyme*

1.5  $\times$  10<sup>5</sup> clones of a human fetal liver cDNA library were screened using a 2.0-kb cDNA probe (Osumi *et al.*, 1985) derived from the coding region of the rat bifunctional enzyme (kindly provided by T. Hashimoto). One clone, 2.3 kb in length, was obtained (pBC5). Comparison to the rat sequence revealed that its 5' end corresponded to nucleotide 1501. Extensive conventional and PCR-mediated screening of a total of 2  $\times$  10<sup>6</sup> clones from several different human liver cDNA libraries did not yield a clone extending to the transcription start site, with the longest one (pBC93) showing homology to the published rat sequence starting at nucleotide 462. Therefore, a specifically primed cDNA library was constructed in  $\lambda$ gt10 using a primer (P2) in a region of high homology between the human and the rat cDNA. Of 4  $\times$  10<sup>5</sup> clones screened, 34 hybridized with pBC93. Twelve of these yielded PCR products with the primer pairs described under Materials and Methods. The PCR products were cloned in pCRII vector (Invitrogen) and analyzed by Southern blotting, and the longest one (pBC19) was sequenced. This clone extended into the 5' noncoding sequence.

### *Sequence and Primer Extension Analysis of the Human Peroxisomal Bifunctional Enzyme*

The cDNA sequence is shown in Fig. 1. The composite nucleotide sequence derived from several overlapping cDNA clones spans 3779 nucleotides with an open reading frame of 2169 nucleotides coding for 723 amino acids. The molecular weight of the enzyme calculated from the deduced sequence is 79,335 Da.

Seven nucleotides of the 5' noncoding sequence were found. Primer extension analysis was performed with

```

-7
1  ATGGCCGAGT ATACGCGGCT GCACAACGCC TTGGCGCTAA TCCGCTCCG gggaaac
61  GTCACGCGGA TCAGTACGAC TTTACTCCGT GATATAAAG AAGGACTACA GAAAGCTGGA
121 AGAGACCATA CAATAAAAGC CATTGTGATT TGTGGAGCG AGGGCAAATT TCTGCAAGT
181 GTGTATATTC GTGGCTTCAG TGCTCCTAGG ACATTGGGCC TTATACTGGG ACATGTAGTA
241 GATGAAATAC AGAGAAATGA GAAGCCCGTG GTGGCAGCAA TCCAAGGCAT GGCTTTCGGA
301 GGGGGACTAG AGCTGGCCCT GGGCTGTAC TATAGGATTG CCCACGCAGA CGCTCAAGT
361 GGCTTACCAG AAGTTACACT TGGACTTCTC CCTGGTGCAA GAGGAACCCA GCTTCTCCCC
421 AGACTCACTG GAGTTCCTGC TGCACCTGAC TTAATTACCT CAGGAAGACG TATTTAGCA
481 GATGAAGCAC TCAAGCTGGG CATTCTAGAT AAAGTTGTAA ACTCAGACCC GGTGAAGAA
541 GCAATCAGAT TTGCTCAGAG AGTTTCAGAT CAACCTCTAG AATCCCGTAG ACTCTGCAAC
601 AAGCCAATTC AGAGCTTGCC CAACATGGAC AGCATTTTTA GTGAGGCCCT CTGAAGATG
661 CGGAGGCAGC ACCCTGGGTG TCTTGACAG GAGGCTTGTG TCCGTGCAGT CCAGGCTGCT
721 GTGCAGTATC CCTATGAAGT GGGCATCAAG AAGGAGGAGG AGCTGTTTCT ATATCTTTTG
781 CAATCAGGGC AGGCTAGAGC CCTGCAATAT GCTTCTTCG CTGAAAGGAA AGCAAATAAG
841 TGGTCAACTC CCTCCGGAGC ATCGTGGAAA ACAGCATCAG CGCGGCCCTG CTCTCAGTT
901 GGTGTTGTTG GCTGGGGAAC AATGGGCCGA GGCATTGTCA TTTCTTTTGC AAGGGCCAGG
961 ATTCCTGTGA TTGGTGTAGA CTCGGACAAA AACCAGCTAG CAACTGCAAA CAAGATGATA
1021 ACCCTGTCTT TGAAAAAGA AGCCTCCAAA ATGCAACAGA GCGGCCACCC TTGGTCAGGA
1081 CCAAACCCA GGTAACTTC ATCTGTGAAG GAGCTTGGTG GTGTAGATTT AGTCATTGAA
1141 GCAGTATTTG AGGAAATGAG CCTGAAGAAG CAGGTCTTTG CTGAACTCTC AGCTGTGTGC
1201 AAACCAGAAG CTTTTTGTG CACTAATACT TCAGCCCTGG ATGTTGATGA GATTGCTTCT
1261 TCCACTGATC GCTCCACTT ACCCACTTCT TTTGCCCAGC TCATGCTATG
1321 AAGTTGTTAG AGTTATTCC CAGCCAATAC TCTCCCCCA CTACCATTGC CACTGTTATG
1381 AACTTATCAA AAAAGATTAA AAAGATTGGA GTCGTTGTAG GCAACTGTTT TGGATTTGTG
1441 GGAATCGAA TGTGAATCC TTAACAATC CAGGCATATT TCTTGTAGA AGAAGGCAGC
1501 AAACCAGAGG AGGTAGATCA GGTGCTGGAA GAGTTTGGTT TAAAAATGGG ACCTTTTAGA
1561 GTGTCTGATC TTCTGGGTT GGATGTGGGC TGGAAATCTA GAAAGGGGCA AGGTCTTACT
1621 GGACCTACAT TGCTTCCAGG AACTCCTGCC CGAAAAGGG GTAATAGGAG GTACTGCCCA
1681 ATTCCTGATG TGCTCTGTGA ATTAGACGA TTTGGCCAGA AGACAGGTAA GGTGTGGTAT
1741 CAATATGACA AGCATTGGG TAGGATTCAC AAACCTGATC CCTGGCTTTC CACATTCCTA
1801 TCACGGTATA GAAAACCCCA TCACATTGAA CCACGTACCA TTAGCCAGGA TGAGATCCTT
1861 GAACGCTGCT TATATCACT TATCAATGAA GCATTCCGTA TCTTGGGAGA AGGGATAGCT
1921 GCTAGCCCAG AGCACATTGA TGTGTCTAT TTACATGGAT ATGGATGCGC AAGGCACAAG
1981 GCGGGGCCCA TGTTCTATGC TTCCACAGTT GGGTTGCCCA CAGTTCTAGA GAAATGCGAG
2041 AAATATTACA GGCAGAACC TGATATTCCC CAACTGGAGC CAAGTGACTA TCTAAAAAAA
2101 CTGGCTTCTC AGGGAACCC TCCCCTGAAA GAATGGCAAA CTTTGGCAGG CTCCCCTAGC
2161 AGTAAATTGT GAttcagtc tccagattat gcctcacatg ctagcatcag gtaatgctga
2221 ctgaatttca gtgaaattaa atcaaaaatc caaagtaaga ttgttctgaa atacaaagca
2281 aaataaataa tcattagaat cttctgtgta acgactctaa tggtaaatc tttaggaatg
2341 tgcttctctat gcctctgaat ctgtctttat cagataaatt caatgcata acttgtgtga
2401 atataatacc ataataagcta atgaaagagg ctcaggcata agttgagatt ctcaaagct
2461 tttatcattg gataaatgtg tcatcaatta ataaatgata aatgcagcta agtcatacat
2521 tcatcttgac tcctttcaat gtcacacaca tagtattgat cagaaatctt atgaatcata
2581 catacactca acaaacatta aagttgtagg aaaaagacag ttggaattg gtaagggaaac
2641 tgagtaactc aaaccagcac agggaaactta ggtagtgtg gcaagccttt cctctctg
2701 tctttctctc tctgtttatg gagaataaat agaaagtagt aagtcgttaa cttagtgtaa
2761 gaagggtctt agagaacatc taaccttcta ggatttccca attctgtgat agagtaatga
2821 caccagtttt cctgtcatga caagcctctg tgatgttaca tatggaatg gttgaatctt
2881 gaaaaatcta aaattgttgc aaaacatatt ttgtatgatt ttgtgtgaag agttcttctc
2941 ttttactttt ttgctttgtg tagttaaaaa ttaagggctt ggtcaataca aaaacttcta
3001 cacaaatatt tatagcagaa ttattcataa tggccaaaag ctgaatacaa cccaaatgtc
3061 tatgaaacta tgaatagata aaccaaatct ggtatatcca tacaatggac tatattatc
3121 agccataaaa agaataaag ggccagqcc agtggctcac acctgtaac ccagcacttt
3181 gggagqctga gqcaqqgqg ttgtttgagg ccagaaattt gagaccagcc tggqcaact
3241 agcaaaaacc tgtctctaca aaaaactt tccgtacatt agtggttacc taaggctaga
3301 attgggggtg atgaatggg agtacaggat tcatctaag ggtacaggat ttctttatgg
3361 ttcataaaaa tggcttaaac ttattgtggt aatggtcaca taacaatata ttgaaaaacc
3421 attgagttgt atattttaag tgggtgaatt atatggtatg tgaattatat ctcaataaag
3481 gtgttagtaa ataaatggc tgaagatgtt atccttcatt gtggtgtaaa tgaactttca
3541 caatattttc acctgtgaac ccaataaaa tgattaaagt tctgatggaa aaatctgtaa
3601 tggagattaa gtttccgtt gttaaaaacg acaaaaaaaa ccaaaaaaaa atcccagtg
3661 gccttgaatt gtacagagca caattattat gtttgaatg tgtactactt aattttatat
3721 aatttggtt gtgaaattaa agacatcaat aaaaatgatt cctgaaagta gt

```

FIG. 1. cDNA sequence of the human peroxisomal bifunctional enzyme (GenBank Accession No. L07077). The coding sequence is written in capital letters, and the TGA "stop codon" and a polyadenylation signal are underlined. One hundred seventeen nucleotides homologous to an *Alu* repeat are underlined and in bold letters.

total liver RNA and a fluorescein-labeled oligonucleotide primer (bases 14 to 43). Analysis of the extension mixture on denaturing PAGE yielded one major product of 57 nucleotides by comparison to a standard sequencing reaction (data not shown). This indicates that the actual 5' noncoding sequence has a length of 14 nucleotides.

The 3' noncoding sequence is 1603 nucleotides long, excluding the poly(A) tail. It contains a 117-nucleotide-long DNA stretch highly homologous to an *Alu* repeat. No short repetitive sequences, many times observed as insertion signals, were found flanking the *Alu* sequence. To classify this *Alu* type element, an extensive homology search was performed. Figure 2 shows a comparison of

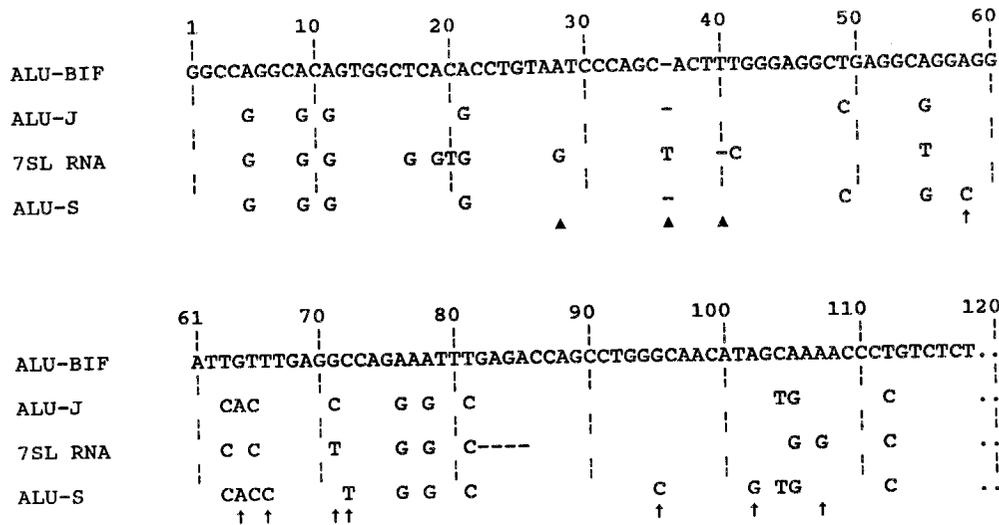


FIG. 2. Comparison of the FLA found in the peroxisomal bifunctional enzyme 3' untranslated region (*Alu*-BIF) to *Alu*-J, *Alu*-S, and 7SL RNA. Diagnostic bases for different *Alu* families according to Jurka and Zuckerkandl (1991) are denoted by ▲, and those according to Jurka and Smith (1988) by ↑.

the *Alu* sequence in the 3' untranslated region of the bifunctional enzyme to other *Alu* classes and 7SL RNA. The overall homology was highest (86%) to the *Alu*-J family (Jurka and Smith, 1988) followed by 83% for *Alu*-S and 78% for 7SL RNA. Comparison of the diagnostic bases (Jurka and Smith, 1988), however, showed a match of 6 of 8 bases both for 7SL RNA and *Alu*-J and only 2 of 8 for *Alu*-S. For bases used to distinguish 7SL RNA from both *Alu*-S and *Alu*-J (Jurka and Zuckerkandl, 1991), our sequence corresponds to the *Alu*-S/*Alu*-J families in all three positions (Fig. 2). Comparison of the dominant nucleotides in the diagnostic positions for classes I-IV (Britten *et al.*, 1988) revealed 11 of 14 diagnostic bases to be characteristic for class I and 7SL RNA (8/14 for class II, 4/14 for class II, and 3/14 for class IV).

*RNA Blotting and S1 Nuclease Protection Assay*

To estimate the size of the bifunctional enzyme mRNA, Northern blotting analysis was performed on isolated total RNA from cultured human hepatoma (HepG2) cells and human brain. For all samples, a single band with an approximate size of 3.7 kb (Fig. 3) was obtained. The intensity of the signal for brain RNA, however, was much weaker than that for HepG2 RNA.

A similar result concerning the amount of mRNA was obtained in an S1 nuclease protection assay (Fig. 3). Brain mRNA gave only a faint band compared to liver, kidney, and HepG2 mRNAs. The fact that only a single band was observed indicated that only a single transcript was detected by the antisense RNA.

*Chromosomal Localization*

The chromosomal mapping position of cDNA of the human bifunctional enzyme was determined by FISH (Fig. 4). To assign the position of the cDNA with respect to cytogenetically defined bands, chromosomes were

counterstained with DAPI, which revealed a G-banding pattern of human chromosomes. A total of 15 randomly selected metaphases were investigated. Nine showed two hybridization signals on both chromatids of 3q26.3-3q28. Three showed hybridization signals on only one homolog chromosome, again on 3q26.3-3q28. One metaphase plate revealed hybridization signals on only one chromatid of both homologous chromosomes, whereas two metaphase plates did not show any signal. Signals on different chromosomal mapping positions have not

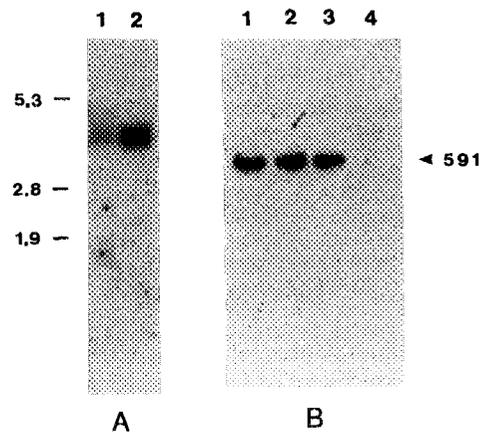
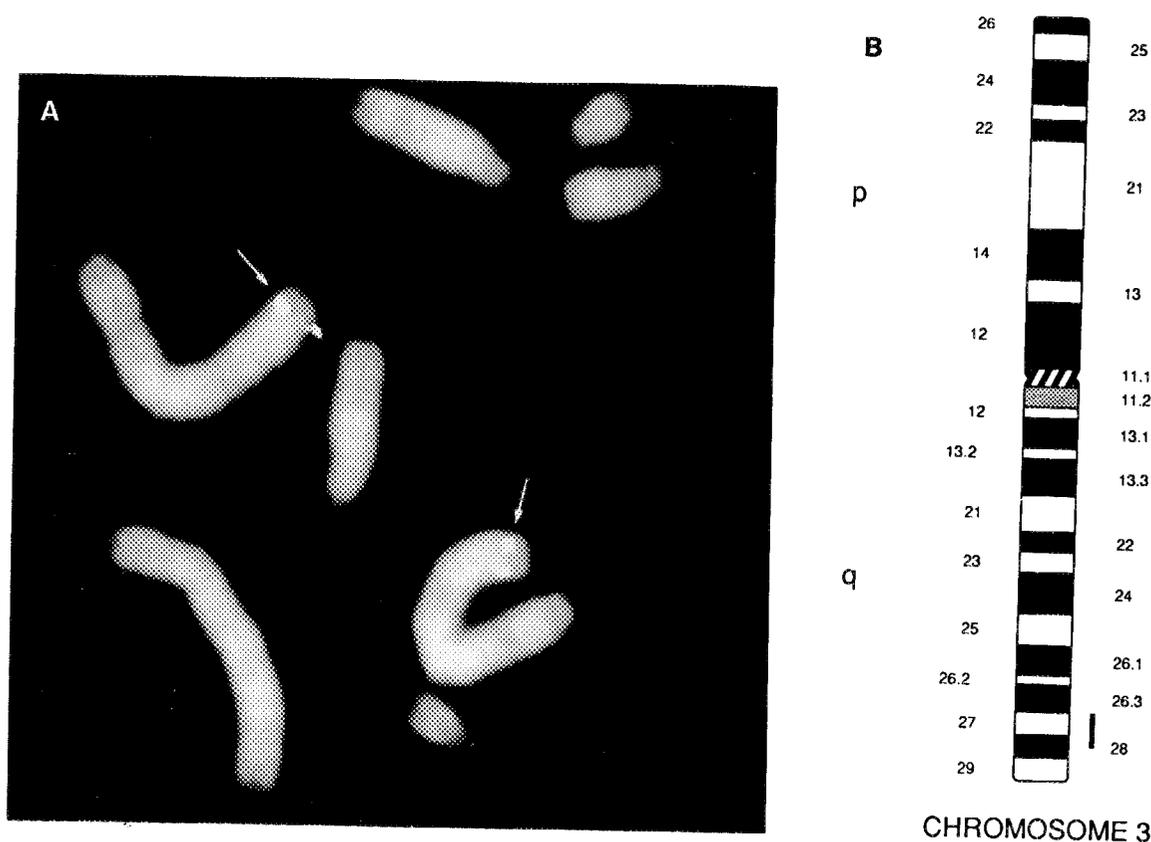


FIG. 3. Analysis of human peroxisomal bifunctional enzyme transcripts. (A) Northern blot. Total RNAs (10 µg) from human brain (lane 1) and human liver (lane 2) were subjected to formaldehyde agarose electrophoresis on 1% gels, blotted on positively charged Nylon paper, and probed with radioactively labeled cDNA for the human bifunctional enzyme (pBC93). Specific signals were visualized by autoradiography (for details, see Materials and Methods). Molecular weight markers (kb) are on the left. (B) S1 nuclease protection assay. Poly(A)<sup>+</sup> RNAs (3 µg) isolated from human liver (lane 1), cultured human hepatoma HepG2 cells (lane 2), human kidney (lane 3), and brain (lane 4) were hybridized to <sup>32</sup>P-labeled antisense RNA probe containing the first 591 nucleotides of the human bifunctional enzyme. After digestion with S1 nuclease, samples were run on 5% PAGE and visualized by autoradiography as described under Materials and Methods.



**FIG. 4.** Chromosomal mapping of cDNA of the human bifunctional enzyme by FISH. (A) The biotinylated cDNA fragments were hybridized to a normal human metaphase spread. Only a section of the metaphase spread is shown. The probe was detected with avidin-FITC (pseudocolored in yellow). The chromosomes were counterstained with DAPI (pseudocolored in blue), producing a G-banding pattern. The probes hybridize to chromosomal map positions 3q26.3-3q28 (arrows). (B) Idiogram of chromosome 3, displaying a G-banding pattern with a 400-band resolution. The bar denotes the chromosomal map position of the cDNA fragments from the human peroxisomal bifunctional enzyme.

been observed. Due to the limited resolution of FISH on metaphase chromosomes, the chromosomal map position could not be assigned with respect to a single chromosomal band. Figure 4A displays the signal on a partial metaphase spread. Figure 4B illustrates the mapping position on chromosome 3 on a 400-band resolution idiogram.

#### DISCUSSION

The peroxisomal bifunctional enzyme is important in the  $\beta$ -oxidation pathway of vertebrates. The human enzyme contains hydratase and dehydrogenase activities. Additionally, an isomerase activity has been reported for the enzyme from rat, which has therefore been termed a trifunctional enzyme (Palosaari and Hiltunen, 1990). Among the isolated defects of human peroxisomal  $\beta$ -oxidation, the loss of bifunctional enzyme activity is the most common (Watkins *et al.*, 1989; Guerroui *et al.*, 1989; Wanders *et al.*, 1990). Detailed analysis of the molecular defects in these patients might lead to a better understanding of the enzymatic activities associated with the putative domains. As a necessary first step in allowing the molecular defects to be identified and in studying tissue-specific expression and regulation of gene expression in humans, we have isolated a cDNA

spanning the complete coding region of the human peroxisomal bifunctional enzyme.

#### cDNA and Amino Acid Comparison

The full-length cDNA for the human peroxisomal bifunctional enzyme spans 3779 nucleotides. This is 696 nucleotides longer than the rat cDNA (Osumi *et al.*, 1985). The difference is mostly due to an extension in the 3' untranslated region. The open reading frame of 2169 nucleotides codes for 723 amino acids with a calculated molecular weight of 79,335 Da. This is in good agreement with the value determined by SDS-PAGE analysis (79 kDa; Chen *et al.*, 1987). A higher content of basic residues (Arg + Lys = 89 versus Glu + Asp = 71) is consistent with the observed *pI* of 9.8 of the purified enzyme (Reddy *et al.*, 1987).

Sequence comparison between the human and the rat bifunctional enzyme coding region shows an 80% homology at the nucleotide level and a 78% homology at the protein level, which becomes 96% if conservative amino acid changes are taken into account. The homology in the putative regions of 3-hydroxyacyl-CoA dehydrogenase (aa 296 to 527) and enoyl-CoA hydratase (aa 22 to 187) (Ishii *et al.*, 1987) does not differ significantly from the overall homology. Compared to the rat sequence, one

	10	20	30	40	50	60
HUMBIF	MAEYTRLHNALALIRLRNPPVNAISTLLRDIKEGLQKAGRDRHTIKAIVICGAEGKFSAG					
RATBIF	MAEYLRPLHSLAMIRLCNPPVNAVSPVIREVVRNGLQKAGSDHTVKAIVICGANGNFCAG					
	70	80	90	100	110	120
HUMBIF	ADIRGFSAPRTFGLILGHVVDEIQRNEKPVVAIQGMAGGGLELALGCHYRIAHADAQV					
RATBIF	ADIHGFSA-FTPGLALGSLVDEIQRYQKPVLAIQGVALGGLELALGCHYRIANAKARV					
	130	140	150	160	170	180
HUMBIF	GLPEVTLGLLPGARGTQLLPRLTGVPAAALDLITSGRRILADEALKGLDKVNSDPVEE					
RATBIF	GLPEVTLGILPGARGTQLLPVVGVPVALDLITSGKYLSADEALRLGILDVAVKSDPVEE					
	190	200	210	220	230	240
HUMBIF	AIRFAQRVSDQPLESRRLCNKPIQSLPNMDSIFSEALLKMRQHPGCLAQEACVRAVQAA					
RATBIF	AIKFAQKIIDKPIEPRRIFNKPVPSLPNMDSVFAEAIKVRKQYPGVLAPETCVRSIQAS					
	250	260	270	280	290	300
HUMBIF	VQYPYEVGIKKEEELFLYLLQSGQARALQYAFFAERKANKWSTPSGASWKTASARPVSSV					
RATBIF	VKHYPYEVGIKEEELFMYLRLASGQAKALQYAFFAEKSANKWSTPSGASWKTASAPVSSV					
	310	320	330	340	350	360
HUMBIF	GVVGLGTMGRGIVISFARARIPVIGVSDKNQLATANKMITSVLEKEASKMQSGHPWSG					
RATBIF	GVLGLGTMGRGIAISFARVGISVAVESDPKQLDAAKKIITFTLEKEASRAHQNGQASAK					
	370	380	390	400	410	420
HUMBIF	PKPRLTSSVKELGGVDLVI EAVFEEMSLKQVFAELSAVCKPEAFLCTNTSALDVDEIAS					
RATBIF	PKLRFSSSTKELSTVDLVVEAVFEDMNLKVKVFAELSAVCKPEAFLCTNTSALNVDDIAS					
	430	440	450	460	470	480
HUMBIF	STDRPHLVIGTHFFSPAHVMLLEVIPSYSSPTTIATVMNLSKKIKKIGVVVGNCFGFV					
RATBIF	STDRPQLVIGTHFFSPAHVMLLEVIPSYSSPTTIATVMNLSKKIKKIGVVVGNCFGFV					
	490	500	510	520	530	540
HUMBIF	GNRMLNPYYNQAYFLLEEGSKPEEVDQVLEEFQKMGFFRVSDLAGLDVGVKSRKQGGLT					
RATBIF	GNRMLAPYYNQGFLLLEEGSKPEDVDGVLEEFQKMGFFRVSDLAGLDVGVKIRKQGGLT					
	550	560	570	580	590	600
HUMBIF	GPTLLPGTTPARKRGNRRYCPIDVLCVLRGFRGQKTGKGYQYDKPLGRIHKPDPWLSLTF					
RATBIF	GPSLPPGTPVRKRGNSRYSPLGDMCEAGRFQKTGKGYQYDKPLGRIHKPDPWLSLTF					
	610	620	630	640	650	660
HUMBIF	SRYRKPHHIEPTISQDEILERCLYSLINEAFRILGEGIAASPEHIDVVYLHGYGCRHK					
RATBIF	SQYREVHHEIQTISKEEILERCLYSLINEAFRILEGMAARPEHIDVYLHGYGWRPHK					
	670	680	690	700	710	720
HUMBIF	GGPMFYASTVGLPTVLEKIQYYRQNPDIQLEPSDYLLKKLASQGNPPLKEWQSLAGSPS					
RATBIF	GGPMFYAASVGLPTVLEKIQYYRQNPDIQLEPSDYLRRLVAQGSPLKEWQSLAGPHG					
	660	670	680	690	700	710
HUMBIF	SKL					
	:::					
RATBIF	SKL					
	720					

FIG. 5. Primary structure of the human peroxisomal bifunctional enzyme (HUMBIF) compared to that of the rat (RATBIF; Osumi *et al.*, 1985). Homologous bases are denoted by double dots (:), and conservative amino acid changes by single dots (.). Thirty-one amino acids are nonconservatively changed.

additional amino acid (Pro) is found at amino acid 68. The sizes of both enzymes therefore differ by only one amino acid residue. The human peroxisomal bifunctional enzyme as well as its rat counterpart contains the peroxisomal import signal peptide Ser-Lys-Leu at the carboxy terminus (Fig. 5), like many other peroxisomal proteins (Gould *et al.*, 1987). The sequence of the car-

boxy terminus differs from the findings of Reddy *et al.* (1987), who reported the sequence of the purified enzyme to be Gly-Ser-Leu-Ile-COOH. However, we have been able to show that a protein, derived from the 3' end of our cDNA by *in vitro* translation, is imported into peroxisomes. This import did not occur when the last nine carboxy-terminal amino acids had been deleted

(Chen *et al.*, 1991) and provided additional evidence that the cDNA sequence indeed codes for a peroxisomal protein.

The 3' noncoding sequence, excluding the poly(A) tail, is 693 nucleotides longer than the rat sequence with little overall homology. No repeated sequences of the type found in the rat 3' noncoding region, i.e., (TGC)<sub>6</sub> and (TA)<sub>21</sub>, are present in the human sequence. Putative polyadenylation signals (AATAAA) are located after nucleotides 3489, 3563, and 3746. All cDNA clones isolated, however, extended to base 3772, suggesting that only one species of mRNA for the bifunctional enzyme is produced in humans. This is in agreement with the results of the Northern blot analysis, in which a single transcript was always detected.

#### *A Free Left Alu Arm (FLA) Is Inserted in the 3' Noncoding Region of the Bifunctional Enzyme*

A sequence highly homologous to human *Alu* was detected in the 3' untranslated region of the human peroxisomal bifunctional enzyme. Based on sequence homology, *Alu* sequences can be grouped into subfamilies. The oldest known primate *Alu* subfamily (*Alu-J*; Jurka and Smith, 1988) is estimated to be about 55 million years old (Labuda and Striker, 1989). The *Alu-J* family corresponds to class I (Britten *et al.*, 1988), which, in turn, is similar to the "divergent" subset of *Alu* repeats (Willard *et al.*, 1987). *Alu* elements organized as head-to-tail dimers approximately 300 bp in length are a common feature found in primates (Daniels and Deininger, 1983), whereas rodents have homologous sequences that are monomers. It is generally believed that an *Alu* monomer arose through deletions in the 7SL RNA. Later in evolution, these ancestral free left *Alu* arms dimerized to form the structure of the *Alu-J* family. Several examples of FLAs sharing several features with the gene for 7SL RNA have been described in the literature (Quentin, 1988; Jurka and Zuckerkandl, 1991). The *Alu* sequence we are describing, however, is equally homologous to 7SL RNA and *Alu-J* when the criteria of Britten *et al.* (1988) and Jurka and Smith (1988) are applied. In addition, at the bases used to discriminate 7SL RNA and FLAs from *Alu-J/Alu-S* families (Jurka and Zuckerkandl, 1991), this sequence shows the features of the latter. Overall homology to the *Alu-J* family is greatest. Our sequence differs, therefore, from other FLAs and might be interpreted as a FLA that has already acquired a number of mutations characteristic of the oldest dimeric *Alu* family (*Alu-J* or class I). It is therefore the evolutionarily youngest FLA detected thus far, indicating that FLAs also changed during evolution.

#### *RNA Blotting and S1 Nuclease Protection Assay*

Northern blot analysis (Fig. 2) of RNA from various tissues reveals a single band with an approximate size of 3.7 kb. This can be taken as evidence that only one species of mRNA is produced in these tissues. The length is in accordance with that of the cDNA sequence described

in this report. The amount of mRNA for the bifunctional enzyme in brain tissue is considerably lower than that in HepG2 cells. This correlates well with the lower number of peroxisomes and the lower amount of bifunctional enzyme protein in brain tissue (G. Hoefler, unpublished observation).

In S1 nuclease protection assays, mRNAs isolated from human liver, kidney, and brain as well as HepG2 cells protect an antisense RNA transcript corresponding to the first 177 amino acids. This includes the additional amino acid at position 68. Therefore, it can be concluded that this additional amino acid is present in all samples examined and that all mRNA for the bifunctional enzyme carries this insertion in humans. Again, as shown in Northern blot analysis, a lower amount of mRNA for the bifunctional enzyme in the brain sample is found by nuclease protection assays. We demonstrate the presence of similar amounts of mRNA for the bifunctional enzyme in HepG2 cells compared to human liver tissue. These cells have been shown to contain an active peroxisomal  $\beta$ -oxidation system (Watkins *et al.*, 1991).

The isolation of a full-length cDNA for the peroxisomal bifunctional enzyme and the cDNA sequence will allow in-depth studies of molecular defects in patients with deficient peroxisomal bifunctional enzyme protein or deficient catalytic activity. These studies might provide additional information on the structure and the different catalytic functions as well as the regulation of this enzyme in humans.

#### ACKNOWLEDGMENTS

The authors thank Professor T. Hashimoto for his generous gift of cDNA for the rat peroxisomal bifunctional enzyme. This work was supported by the Austrian FWF projects P 8586-MED to G.H. and S 4603 to R.Z. and by the National Institutes of Health Grants HD10981 and HD24061 to G.C.

#### REFERENCES

- Britten, R. J., Baron, W. F., Stout, D. B., and Davidson, E. H. (1988). Sources and evolution of human *Alu* repeated sequences. *Proc. Natl. Acad. Sci. USA* **85**: 4770-4774.
- Chen, W. W., Watkins, P. A., Hashimoto, T., and Moser, H. W. (1987). Peroxisomal beta-oxidation enzymes in adrenoleukodystrophy. *Proc. Natl. Acad. Sci. USA* **84**: 1425-1428.
- Chen, G. L., Balfe, A., Erwa, W., Hoefler, G., Gaertner, J., Aikawa, J., and Chen, W. W. (1991). Import of human bifunctional enzyme into peroxisomes of human hepatoma cells *in vitro*. *Biochem. Biophys. Res. Commun.* **178**: 1084-1091.
- Daniels, G. R., and Deininger, P. L. (1983). A second major class of *Alu* family repeated DNA sequences in a primate genome. *Nucleic Acids Res.* **11**: 7595-7610.
- Deininger, P. L., Jolly, D. J., Rubin, C. M., Friedmann, T., and Schmid, C. W. (1981). Base sequence studies of 300 nucleotide renatured repeated human DNA clones. *J. Mol. Biol.* **151**: 17-33.
- Fujiki, Y., Rachubinski, R. A., Mortensen, R. M., and Lazarow, P. B. (1985). Synthesis of 3-ketoacyl-CoA thiolase of rat liver peroxisomes on free polyribosomes as a larger precursor: Induction of thiolase mRNA activity by clofibrate. *Biochem. J.* **226**: 697-704.
- Gould, S. J., Keller, G. A., and Subramani, S. (1987). Identification of

- a peroxisomal targeting signal at the carboxy terminus of firefly luciferase. *J. Cell Biol.* **105**: 2923-2931.
- Guerroui, S., Aubourg, P., Chen, W. W., Hashimoto, T., and Scotto, J. (1989). Molecular analysis of peroxisomal  $\beta$ -oxidation enzymes in infants with peroxisomal disorders indicates heterogeneity of the primary defect. *Biochem. Biophys. Res. Commun.* **161**: 242-251.
- Hashimoto, T. (1982). Individual peroxisomal  $\beta$ -oxidation enzymes. *Ann. N.Y. Acad. Sci.* **386**: 5-12.
- Hoefler, G., Forstner, M., Hulla, W., and Chen, G. L. (1991). Cloning of cDNA for the human peroxisomal bifunctional enzyme: Comparison of conventional and PCR methods. In "Advances in Gene Technology: The Molecular Biology of Human Genetic Disease" (F. Ahmad *et al.*, Eds.), Vol. 1, p. 10, Miami Short Reports.
- Ishii, N., Hijikata, M., Osumi, T., and Hashimoto, T. (1987). Structural organization of the gene for rat enoyl-CoA hydratase/3-Hydroxyacyl-CoA dehydrogenase bifunctional enzyme. *J. Biol. Chem.* **262**: 8144-8150.
- Jurka, J., and Smith, T. (1988). A fundamental division in the Alu family of repeated sequences. *Proc. Natl. Acad. Sci. USA* **85**: 4775-4778.
- Jurka, J., and Zuckerkandl, E. (1991). Free left arms as precursor molecules in the evolution of Alu sequences. *J. Mol. Evol.* **33**: 49-56.
- Labuda, D., and Striker, G. (1989). Sequence conservation in Alu evolution. *Nucl. Acids Res.* **17**: 2477-2491.
- Lazarow, P. B. (1978). Rat liver peroxisomes catalyze the  $\beta$ -oxidation of fatty acids. *J. Biol. Chem.* **253**: 1522-1528.
- Lazarow, P. B., and Deduve, C. (1976). A fatty acyl-CoA oxidizing system in rat liver peroxisomes: Enhancement by clofibrate, a hypolipidemic drug. *Proc. Natl. Acad. Sci. USA* **73**: 2043-2046.
- Lazarow, P., and Moser, H. W. (1989). Disorders of peroxisome biogenesis. In "The Metabolic Basis of Inherited Disease" (C. R. Scriver, A. L. Beaudet, W. S. Sly, and D. Valle, Eds.), pp. 1479-1509, McGraw-Hill, New York.
- Lengauer, C., Riethman, H., and Cremer, T. (1990). Painting of human chromosomes with probes generated from hybrid cell lines by PCR with Alu and L1 primers. *Hum. Genet.* **86**: 1-6.
- Miura, S., Mori, M., Takiguchi, M., Tatibana, M., Furuta, S., Miyazawa, S., and Hashimoto, T. (1984). Biosynthesis and intracellular transport of enzymes of peroxisomal  $\beta$ -oxidation. *J. Biol. Chem.* **259**: 6397-6402.
- Osumi, T., Ishii, N., Hijikata, M., Kamijo, K., Ozasa, H., Furuta, S., Miyazawa, S., Kondo, K., Inoue, K., Kagamiyama, H., and Hashimoto, T. (1985). Molecular cloning and nucleotide sequence of the cDNA for rat peroxisomal enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase bifunctional enzyme. *J. Biol. Chem.* **260**: 8905-8910.
- Palosaari, P. M., and Hiltunen, J. K. (1990). Peroxisomal bifunctional protein from rat liver is a trifunctional enzyme possessing 2-enoyl-CoA hydratase, 3-hydroxyacyl-CoA dehydrogenase, and  $(\delta)^3$ ,  $(\delta)^2$ -enoyl-CoA isomerase activities. *J. Biol. Chem.* **265**: 2446-2449.
- Quentyn, Y. (1988). The Alu family developed through successive waves of fixation closely connected with primate lineage history. *J. Mol. Evol.* **27**: 194-202.
- Reddy, M. K., Usuda, N., Reddy, M. N., Kuczmariski, E. R., Rao, M. S., and Reddy, J. K. (1987). Purification, properties, and immunocytochemical localization of human liver peroxisomal enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase. *Proc. Natl. Acad. Sci. USA* **84**: 3214-3218.
- Ried, T., Baldini, A., Rand, T. C., and Ward, D. C. (1992). Simultaneous visualization of seven different DNA probes by in situ hybridization using combinatorial fluorescence and digital imaging microscopy. *Proc. Natl. Acad. Sci. USA* **89**: 1388-1392.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory Press, New York.
- Singh, I., Moser, A. E., Goldfischer, S., and Moser, H. W. (1984). Lignoceric acid in oxidized in the peroxisome: Implications for the Zellweger cerebro-hepato-renal syndrome and adrenoleukodystrophy. *Proc. Natl. Acad. Sci. USA* **81**: 4203-4207.
- Singh, H., Derwas, N., and Poulos, A. (1987). Very long-chain fatty acid  $\beta$ -oxidation by rat liver mitochondria and peroxisomes. *Arch. Biochem. Biophys.* **259**: 382-390.
- Trahey, M., Wong, G., Halenbeck, R., Rubenfeld, B., Martin, G. A., Ladner, M., Long, C. M., Crosier, W. J., Watt, K., Koths, K., and McCormick, F. (1988). Molecular cloning of two types of GAP complementary DNA from human placenta. *Science* **242**: 1697-1700.
- Wanders, R. J., Van Roermund, C. W., Schelen, A., Schutgens, R. B., Tager, J. M., Stephenson, J. B., and Clayton, P. T. (1990). A bifunctional protein with deficient enzymatic activity: Identification of a new peroxisomal disorder using novel methods to measure the peroxisomal  $\beta$ -oxidation enzyme activities. *J. Inherited Metab. Dis.* **13**: 375-379.
- Watkins, P. A., Chen, W. W., Harris, C. J., Hoefler, G., Hoefler, S., Blake, D. C., Jr., Balfe, A., Kelley, R. I., Moser, A. B., Beard, M. E., and Moser, H. W. (1989). Peroxisomal bifunctional enzyme deficiency. *J. Clin. Invest.* **83**: 771-777.
- Watkins, P. A., Ferrell, E. V., Pedersen, J. I., and Hoefler, G. (1991). Peroxisomal fatty acid  $\beta$ -oxidation in HepG2 cells. *Arch. Biochem. Biophys.* **289**: 329-336.
- Willard, C., Nguyen, H. T., and Schmid, C. W. (1987). Existence of at least three distinct Alu subfamilies. *J. Mol. Evol.* **26**: 180-186.
- Yunis, J. J. (1976). High resolution mapping of human chromosomes. *Science* **191**: 1268-1270.