Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 12185–12190, October 1998
Biochemistry

Natural trans-splicing in carnitine octanoyltransferase pre-mRNAs in rat liver

CONCHA CAUDEVILLA*, DOLORS SERRA*, ANGEL MILIAR*, CARLES CODONY†, GUILLERMINA ASINS*, MONTSERRAT BACH†, AND FAUSTO G. HEGARDT*‡

*Department of Biochemistry, School of Pharmacy, University of Barcelona, 08028 Barcelona, Spain; and †Consejo Superior de Investigaciones Científicas, Jordi Girona, 18-26, 08034 Barcelona, Spain

Edited by Thomas R. Cech, University of Colorado, Boulder, CO, and approved August 17, 1998 (received for review May 5, 1998)

ABSTRACT **Carnitine octanoyltransferase (COT) transports medium-chain fatty acids through the peroxisome. During isolation of a COT clone from a rat liver library, a cDNA in which exon 2 was repeated, was characterized. Reverse transcription-PCR amplifications of total RNAs from rat liver showed a three-band pattern. Sequencing of the fragments revealed that, in addition to the canonical exon organization, previously reported [Choi, S. J. *et al.* (1995) *Biochim. Biophys. Acta* 1264, 215–222], there were two other forms in which exon 2 or exons 2 and 3 were repeated. The possibility of this exonic repetition in the COT gene was ruled out by genomic Southern blot. To study the gene expression, we analyzed RNA transcripts by Northern blot after RNase H digestion of total RNA. Three different transcripts were observed. Splicing experiments also were carried out *in vitro* with different constructs that contain exon 2 plus the 5′ or the 3′ adjacent intron sequences. Our results indicate that accurate joining of two exons 2 occurs by a trans-splicing mechanism, confirming the potential of these structures for this process in nature. The trans-splicing can be explained by the presence of three exon-enhancer sequences in exon 2. Analysis by Western blot of the COT proteins by using specific antibodies showed that two proteins corresponding to the expected M_r are present in rat peroxisomes. This is the first time that a natural trans-splicing reaction has been demonstrated in mammalian cells.**

In trans-splicing, two pre-mRNAs are processed to produce a mature transcript that contains exons from both precursors. This process is believed to proceed through two trans-esterification steps that result in the linking of the two exons by a normal 3′-5′ phosphodiester bond. This process has been described mostly in trypanosoma, nematodes, plant/algal chloroplasts, and plant mitochondria (1).

Trans-splicing of artificial pre-mRNAs in mammalian cells *in vitro* has been reported but with some limitations (2–5). In addition, spliced leader RNAs from nematodes or from Simian virus 40 can be accurately trans-spliced in transfected COS cells, which reveals functional conservation in the splicing machinery between lower eukaryotes and mammals and demonstrated the potential for trans-splicing in mammalian cells (6). Studies *in vitro* also have shown that a synthetic pre-mRNA substrate containing an exon and a 5′ donor splice site can be efficiently trans-spliced to another synthetic pre-mRNA (3′ trans-splicing substrate) if this contains either exonic enhancers or a downstream 5′ splice site (7–8). Several examples of possible natural trans-splicing in mammalian cells have been reported (9–12), but none of these trans-splicing have been demonstrated *in vitro*.

During our current investigation on the carnitine octanoyltransferase (COT) gene, that encodes for an enzyme, which transports medium-chain fatty acids through the peroxisome, we isolated a cDNA COT clone, which had exon 2 repeated. We report here that the pre-mRNA of COT from rat liver produces three different mature transcripts: one transcript produced by cis-splicing and in addition, two transcripts produced by a trans-splicing mechanism, in which exon 2 or exons 2 and 3 are repeated. Western blot analysis of peroxisomal proteins using specific antibodies showed the occurrence of two translated proteins with molecular masses corresponding to the cis-spliced mRNA as well as the trans-spliced mRNA in which exon 2 and 3 are repeated. To our knowledge, natural trans-splicing in which two identical molecules of pre-mRNA produce a trans-spliced mature mRNA has not been described in genes involved in the metabolism of vertebrates.

MATERIALS AND METHODS

Isolation and Characterization of cDNA and Genomic Clones. To isolate cDNA clones corresponding to COT, a λgt10 rat liver cDNA library 5′-stretch from adult male Sprague–Dawley rats (CLONTECH) was used. To isolate genomic clones corresponding to COT cDNA, a λFIX II rat genomic library Vector (Stratagene) was used. The screenings were probed with a fragment of 1,190 bp (*Xba*I–*Eco*RI) of plasmid λ2Ax COT (13) supplied by B. Chatterjee (University of Texas, San Antonio), corresponding to the coding sequence from 510 to 1,700 bp of the cDNA reported by Choi *et al.* (14). Hybridization was performed in 800 mM NaCl, 20 mM Pipes (pH 6.5), 50% deionized formamide, 0.5% SDS, and denatured salmon sperm DNA (100 μg/ml). Washes was carried out in 0.1% SSC and 0,1% SDS at 65°C. The λ DNAs obtained by screening from genomic and cDNA libraries were analyzed by Southern blot. Appropriate DNA fragments were subcloned into pBluescript for sequencing.

DNA Blot Hybridization. Total DNA from male Sprague–Dawley rat liver was purified as described (15). Restriction enzyme digestions, standard saline citrate transfers to Hybond-C membrane (Amersham), hybridizations with random primed ^{32}P-labeled probes, and filter washes were carried out according to the manufacturer's recommendations. Filters were autoradiographed at −70°C in contact with Kodak x-ray film with an intensifying screen.

RNase H Digestion of RNA. Total RNA from rat liver was extracted with guanidine isothiocyanate and purified by centrifugation through a CsCl cushion (16). Samples of 20 μg of total RNA were incubated separately with oligonucleotides, E4r, E4r-2, E5r, and E7r (Table 1) for 30 min at 37°C in a 20-μl mixture containing 20 mM Hepes-KOH (pH 8.0), 50 mM KCl, 70 mM $MgCl_2$, 2 mM DTT, and 40 units of RNasin (GIBCO/BRL). 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.

This paper was submitted directly (Track II) to the *Proceedings* office.
Abbreviations: COT, Carnitine octanoyltransferase. RT-PCR, reverse transcription–PCR. ESE, exonic-splicing enhancer. PL, plasmid polylinker region.
Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF056298 and AF056299).
‡To whom reprint requests should be addressed. e-mail: hegardt@farmacia.far.ub.es.

Table 1. Oligonucleotides used in this study

Position	Primers	Orientation	Sequence	Expt(s)
Exon 1	E1f	Sense, 22-mer	5′GAG TGC AGA GAG CCA AGC CGG G3′	RT-PCR
Exon 2	E2f	Sense, 25-mer	5′TTT TCT TAC TGT GAC TAT ACC ATG G3′	RT-PCR and trans-splicing
Exon 2	E2r	Antisense, 22-mer	5′GGA ATG TTC GTT CTT CAA TTG A3′	RT-PCR
Exon 3	E3r	Antisense, 20-mer	5′TGA TGC AAT GTC TTG CCA AC3′	RT-PCR
Exon 4	E4r	Antisense, 20-mer	5′GGC GAC ATT GAG CCA CCA CT3′	RT-PCR and RNase H protection
Exon 5	E5r	Antisense, 18-mer	5′GGT TCA TGT CTA GAG GAG3′	RT-PCR and RNase H protection
Exon 7	E7r	Antisense, 21-mer	5′GGC CCA ACA GGT TCA TTC CAG3′	RT-PCR and RNase H protection
Exon 4	E4r-2	Antisense, 21-mer	5′CTT CTT AGC AGC TGC CAG TAG3′	RNase H protection
Intron 2	I2r	Antisense, 25-mer	5′CTA CAC TAG AAG ATT ATG AAC ATA C3′	Trans-splicing
Intron 1	I1f	Sense, 20-mer	5′CAT GGG GTG CTA AAT CCA CA3′	Trans-splicing
Exon 2	E2r-2	Antisense, 22-mer	5′CTG ACT CAA GGT ACT TCT TCA G3′	Trans-splicing

samples were incubated for 1 h at 30°C after addition of 0.8 units of RNase H (Boehringer Mannheim) and then ethanol precipitated. RNA was loaded onto 1.6% agarose formaldehyde gels and transferred to NY13 nytran filters (Schleider & Shuell). Filters were hybridized with a random primed ^{32}P-labeled cDNA probe, in ExpressHyb hybridization solution (CLONTECH) according to the manufacturer's recommendations and then autoradiographed.

Reverse Transcription (RT)–PCR. The RT reaction was performed in 200 μl of a mixture containing 10 μg of DNA-free total RNA (in diethyl pyrocarbonate-treated water). Either oligo dTs or random hexamers were used as primers except for the *in vitro*-splicing assay, in which primers E2r-2 and I2r (see Table 1) were used. The reaction was performed with Moloney murine leukemia virus reverse transcriptase (GIBCO/BRL) according to the supplier's conditions. The incubations were at 37°C for 1 h and 95°C for 5 min. The mixtures were then quickly chilled on ice and stored at −20°C. PCR was performed in a total volume of 50 μl with 5 μl of cDNAs, 1 μM of each primer, 0.2 mM dNTPs, and 1.25 mM $MgCl_2$. The conditions for PCR were 94°C for 30 s; 55°C for 30 s; and 72°C for 1 min for 30 cycles. The amplified fragments were separated in agarose gels or acrylamide gels. PCR products were sequenced with an automatic sequencer ABI prism by using nested primers. Sequence alignments were performed to check the identities by using the Wisconsin Package Version 9.1, Genetics Computer Group, Madison, WI.

Pre-mRNA Synthesis and Trans-Splicing Reaction *in Vitro*. The A and B constructs were obtained by PCR amplification by using the DNA obtained from the screening of the genomic library as template. The purified 18-kb DNA insert (λgr5), comprises 4.7 kb upstream from the transcription site and 13.7 kb downstream to exon 6. The nucleotides of intron 1 and intron 2, next to exon 2, were sequenced. Using *Pfu* polymerase and primers E2f and I2r or I1f and E2r-2, two fragments were obtained. The fragments were subcloned in the *Eco*RV restriction site of pBluescript SK+ to obtain constructs A and B. Construct A was then linearized with Eco*RI*. Construct B was shortened by incubating with *Sac*I and *Eco*RI, making the extremes by blunt ending with T4 DNA polymerase and religated (Fig. 5). Two micrograms of the plasmid A was linearized with *Hin*dIII and transcribed in the following conditions: 40 mM Tris·HCl pH 7.9, 6 mM $MgCl_2$, 10 mM NaCl, 2 mM spermidine, 10 mM DTT, 2 units/μl RNase out (GIBCO/BRL), 0.6 mM cap dinucleotide (GpppG), 0.5 mM each ATP and UTP, 0.05 mM each CTP and GTP, 40 μCi [^{32}P]CTP (800 Ci/mmol), and 2 units/μl T7 RNA polymerase. The incubation was performed for 60 min at 37°C. B pre-mRNA was transcribed unlabeled in the same conditions with the following modifications: 30 mM $MgCl_2$, 2 units/μl T3 RNA polymerase, and 5 mM of each NTP. All the pre-mRNAs were purified from a 5% (30:1) acrylamide-bisacrylamide gel (16).

In vitro trans-splicing assays (25 μl) contained 40% of HeLa nuclear extract (17) 0.5 mM ATP, 2 mM $MgCl_2$, 20 mM creatine-phosphate, ≈5 ng (65 fmol) of radiolabeled A pre-mRNA and increasing amounts, 10–200 ng (0.11–2.22 pmol), of B pre-mRNA, (see Fig. 5 legend). The trans-splicing mix was incubated for 2 h at 30°C without preincubation. Reactions were arrested by addition of 2 μl of proteinase K (20 mg/ml), 10 μl 10% SDS, 63 μl H_2O, and incubated for 30 min at 37°C. Then RNAs were extracted with phenol/chloroform/isoamyl alcohol and precipitated with ethanol. RNAs were loaded on a denaturing 8% polyacrylamide/7 M urea gel. RNAs were eluted from the gel and used in RT-PCR experiments.

Isolation of Rat Liver Peroxisomes. Rat liver peroxisomes were isolated in a Nycodenz cushion as described in (18). The peroxisomes were dispersed in 250 mM sucrose, 10 mM Tris/HCl pH 7.4 and 1 mM EDTA and used in Western blot experiments.

Generation of Anti-Carnitine Octanoyltransferase Antibodies and Western Blot Analysis. A peptide corresponding to the N terminus of COT protein (sequence 43–54, ANEDEYKKTEEI) (14) was synthesized by the solid-phase method developed by Marglin and Merrifield (19). The peptide showed little identity to other carnitine transferases. A cysteine residue was added to the N-terminal end. The peptides were coupled to keyhole-limped haemocyanin with maleimidobenzoyl-*N*-hydroxysuccinimide ester. Two male New Zealand white rabbits were each injected s.c. on days 0, 21, and 42 with 200 μg of the peptide coupled to haemocyanin. The synthetic peptide-haemocyanin conjugates were emulsified with Freund's complete adjuvant (day 0) or incomplete adjuvant (days 21 and 42) in a total volume of 1 ml. Rabbits were bled from an ear vein 12 days after each booster injection. γ-Globulins from preimmune and immune sera were purified by $(NH_4)_2SO_4$ fractionation (16). The γ-globulins from peptide 43–54 were called A43. Purified peroxisomes were subjected to 10% SDS/PAGE. Electroblotting to nitrocellulose sheets was carried out for 2 h at 120 mA. Immunodetection was performed by using anti-COT antiserum A43 and the blots were developed by using the ECL Western blotting system from Amersham Pharmacia Biotech.

RESULTS

Cloning and Characterization of Liver COT cDNAs. To study different properties of the COT gene, we isolated a cDNA clone from a rat liver cDNA library by using as a probe the *Xba*I--*Eco*RI fragment of the reported cDNA (13). The exonic organization of the isolated cDNA did not correspond to that reported (14); instead, exon 2 was repeated. To rule out the possibility that the abnormal exon 2-exon 2-exon 3, etc., organization of the isolated clone was a product of an artifactual rearrangement and to determine whether a trans-splicing reaction occurred naturally in this gene, RT-PCR with different primers were carried out. The templates for these amplifications were the cDNAs produced after the rat liver mRNAs were reacted with reverse transcriptase by using either oligo dTs or random hexamers as primers. Fig. 1*A*

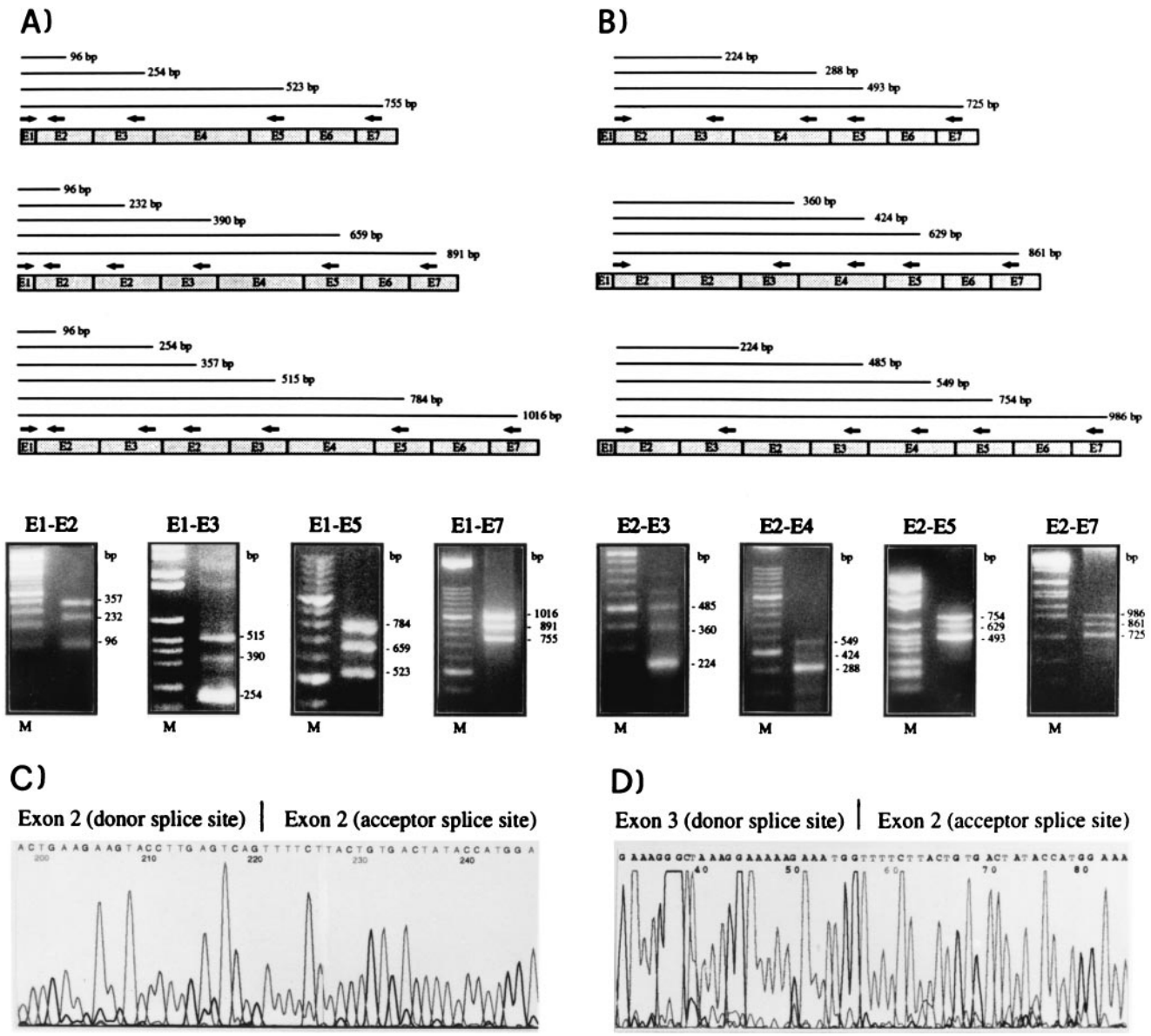

FIG. 1. Three-band pattern obtained by RT-PCR amplification of the mRNAs of rat liver COT. PCR products were obtained by using four sets of primers, E1f-E2r, E1f-E3r, E1f-E5r, and E1f-E7r (*A*) and E2f-E3r, E2f-E4r, E2f-E5r, and E2f-E7r (*B*). Schematic representation of the possible origin of PCR products is shown at the top. Exons are indicated by boxes and capital letters; arrows indicate the annealing regions of the PCR primers, lines above denote expected size for the three possible exonic organizations. Representative agarose gels of PCR products obtained with the four sets of primers are shown at the foot of the figure. M indicates the molecular DNA markers. The size and the sequence analysis of PCR products correspond to the exonic representation. The chromatograms show the accurate junction exon 2-exon 2 (*C*) and exon 3-exon 2 (*D*).

shows that when primers located in exon 1 and exon 3 were used (Table 1), three different bands were visible. One of the bands was of the expected size, but the other two were of sizes corresponding to the inclusions of exon 2 and exons 2 and 3. The sequencing of these bands unequivocally showed that they were formed by (*i*) exon 1-exon 2-exon 2-exon 3 (390 bp) (Fig. 1*C*), (*ii*) exon 1-exon 2-exon 3-exon 2-exon 3 (515 bp) (Fig. 1*D*), in addition to the canonical cis-spliced band, and (*iii*) exon 1-exon 2-exon 3 (254 bp). To confirm these data, PCR amplifications were carried out with forward primers for either exon 1 or 2 and reverse primers complementary to exons 2, 3, 4, 5, and 7. As also shown in Fig. 1*A* and *B*, the same three-band pattern was obtained in all cases. Sequencing of all these bands showed that the repetition affected only exons 2 and 3.

A possible explanation of the duplication of exons 2 and 3 in the mature transcripts was that the genomic DNA contained duplication of exons 2 and 3. However, genomic Southern blot showed that this is a single-copy gene and that exons 2 (Fig. 2) and 3 (not shown) are present only once. Moreover, Southern blot and restriction mapping of three different clones obtained from a rat genomic library confirmed these data.

Kinetic analysis of rat mRNA levels in adult rats was performed by RT-PCR with a variable number of cycles. The proportion of the various bands in the PCR products starting with exon 1 was ≈80%, 4%, and 16%, respectively, for the sequence 1–2–3–4–5, 1–2–2–3–4–5, and 1–2–3–2–3–4–5 in one experiment (Fig. 3). We studied the reproducibility of these results and found variability in the quantitative results: values ranging from 4% to 30% for the two bands with high molecular mass was observed in different rats.

Detection of Three Different Size Transcripts in Rat Liver COT. The occurrence of three different transcripts could not be detected by Northern blot analysis of the mRNAs because the difference in size between the three transcripts was so small that they could not been seen separately. To improve the sensitivity of the method and determine whether three differ-

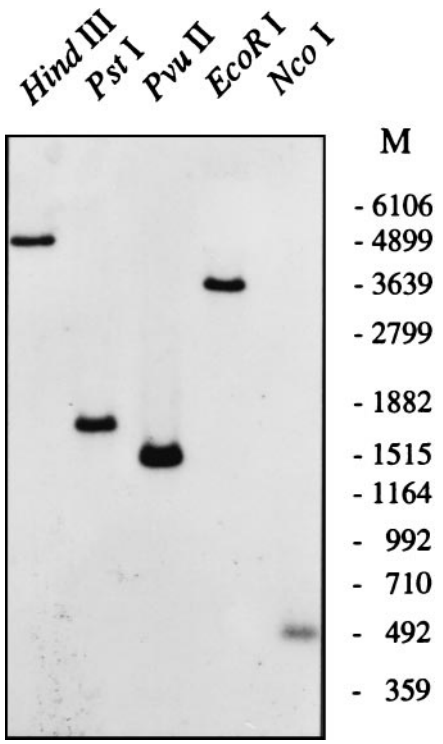

FIG. 2. Southern blot analysis of genomic rat DNA. Total rat genomic DNA, 15 μg per lane, was digested with *Hin*dIII, *Pst*I, *Pvu*II, *Eco*RI, and *Nco*I. Fragments were fractionated electrophoretically, bound on a Hybond-C membrane, and hybridized with a ^{32}P-labeled cDNA probe corresponding to exon 2, with positions 31–166 bp in the cDNA sequence of rat COT (14). DNA molecular size markers are indicated at the right.

ent transcripts of COT were present as independent molecules, total RNA was hybridized with specific oligonucleotides within different exons of the COT mRNA and digested with RNase H, which digests double-stranded regions of RNA–DNA hybrids. If the trans-splicing transcripts were present, then the treatment would result in the appearance of four different bands detectable on a Northern blot (Fig. 4*A*). Using the primers described in the legend to Fig. 4, the band of higher molecular mass would be the most 3′ and the other three of lower molecular mass would be at the most 5′ regions of the transcripts. Fig. 4*B* shows a Northern blot by using the whole cDNA COT as a probe. The results observed fit with these predictions. Bands corresponding to the 3′ fragment after RNase H digestion (lanes 2–5) ran with a slightly faster mobility (2,250–2,500 bp) than the uncut transcript (lane 1) (3,000 bp). Moreover, three bands with faster mobility corresponding to the 5′ fragment after RNase H digestion also were seen. The results of these analyses are consistent with the presence of 3 transcripts of COT. The size of each fragment corresponds to those expected.

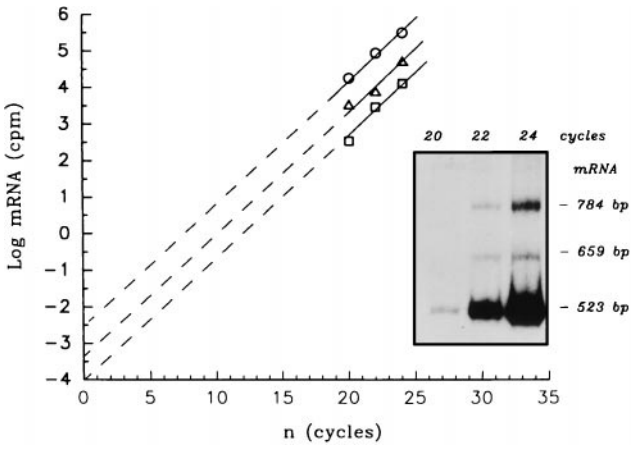

FIG. 3. Quantification of the RT-PCR products by PCR kinetic analysis cDNAs were obtained from total adult rat liver RNA as described in *Materials and Methods*. PCR amplifications were performed with E1f and E5r primers (Table 1) in the presence of 2.5 μCi (0.25 μl) of [α-^{32}P]dCTP (300 Ci/mmol). After 20 amplification cycles, a small portion of the PCR mix (10%), was removed every two cycles and the products were resolved on 8% acrylamide gels. The amount of radioactivity in each band was determined by scintillation counting. For kinetic analysis, values of log counts per minute were plotted against the number of cycles. The antilog of the respective intercepts shows the proportion of each mRNA. The upper straight line denotes the 523 pb mRNA (○), and the lower lines correspond to the 659 bp (□) and 784 bp mRNA (△). The autoradiography is shown in the *Inset*.

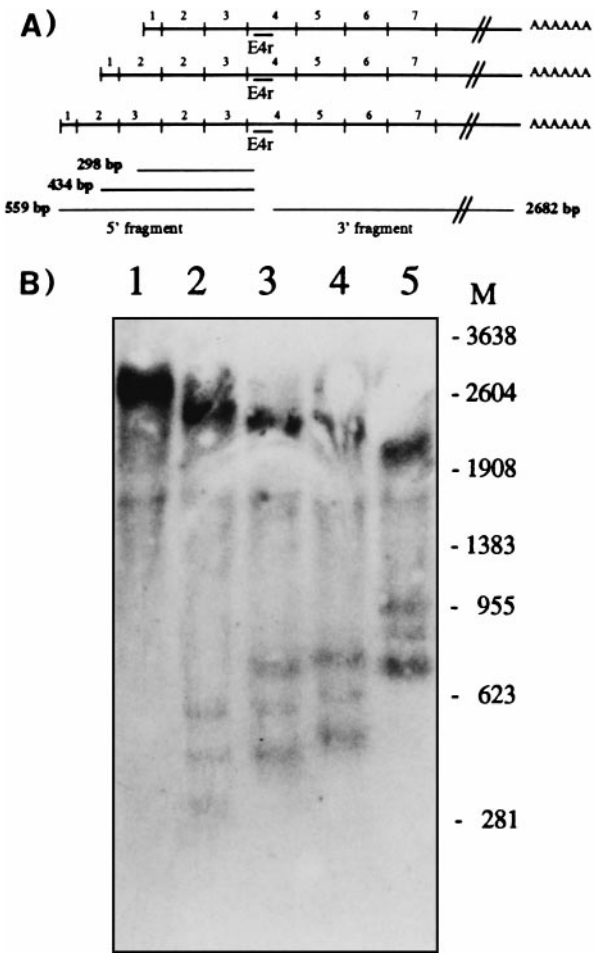

FIG. 4. RNase H Digestion of RNA. (*A*) The diagram shows the products formed after binding the primer E4r to RNA and digesting with RNase H. An analogous diagram can be formulated with the other three primers used. (*B*) Total RNA from liver of adult rat was bound in different tubes with primers E4r (exon 4), E4r-2 (exon 4), E5r (exon 5), and E7r (exon 7) and then digested with RNase H and analyzed on a 1.6% agarose gel with formaldehyde as described in *Materials and Methods* (lanes 2–5). Lane 1 denotes control nondigested RNA. The filter containing the transferred RNA was hybridized to a ^{32}P-labeled cDNA probe corresponding to positions 31–1,700 bp in the cDNA sequence from COT (14). RNA molecular size markers (M) are indicated to the right.

The COT Gene Contains Natural Sequences That Were Trans-Spliced in a Trans-Splicing Assay *in Vitro* with Human Nuclear Extracts. To gain further insight into the possible role of the sequences around exon 2 of the COT gene, two truncated pre-mRNAs were prepared: A donor pre-mRNA, containing

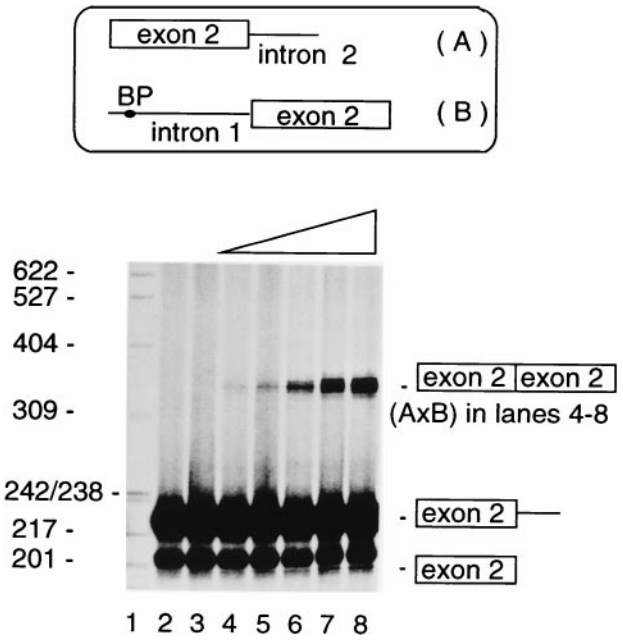

FIG. 5. Trans-splicing of COT pre-mRNAs *in vitro*. Transcription plasmids (Bluescript SK^+) with the corresponding PCR fragments were obtained as described in *Materials and Methods*. A Pre-mRNA was 220 nucleotides (nt) long: 55 nt from the plasmid polylinker region (PL) + exon 2 (136 nt) + 25 nt of the 5′ end of intron 2 + 4 nt PL; B Pre-mRNA was 278 long: 23 nt PL + 115 nt of the 3′ end of intron 1 + exon 2 (136 nt) + 4 nt PL. BP indicates branch-point. The ^{32}P-labeled A pre-mRNA was mixed with increasing concentrations of unlabeled B pre-mRNA and incubated in trans-splicing conditions. Lane 2 shows incubation reaction of labeled A pre-mRNA with 50 ng of B pre-mRNA without ATP. Lane 3 shows the incubation reaction with A pre-mRNA but without construct B. A pre-mRNA was incubated with increasing amounts of B pre-mRNA (10 ng, 25 ng, 50 ng, 100 ng, and 200 ng). The trans-splicing product (AxB) (lanes 4–8) was confirmed by RT-PCR and sequencing. The exon 2 band might correspond to the intermediate exon 2 without the attached intron.

only exon 2 and the 5′ splice site of intron 2 (A pre-mRNA), and an acceptor pre-mRNA, containing the branch point region and the 3′ splice site of intron 1 and exon 2 (B pre-mRNA) (Fig. 5).

Fig. 5 also shows typical trans-splicing assays with the two pre-mRNAs in which the donor, A pre-mRNA, was ^{32}P-labeled during transcription and the acceptor, B pre-mRNA, was added unlabeled to the trans-splicing reaction. The combination of A donor with B acceptor (Fig. 5, lanes 4–8) showed that there is a band above the pre-mRNA that is dependent on ATP (cf. lane 2 to lanes 4–8). This band was cut, assayed by RT-PCR, and sequenced. This band contained a correct ligation of exon 2-exon 2, which rules out the possibility of an alternative splice site. It turned out to be a trans-splicing product between the A donor and the B acceptor. This product was not seen when the B acceptor was not added to the reaction (lane 3). The comparison of lane 2 and lanes 4–8 also shows the occurrence of another band (exon2), which is assumed to be one of the two intermediates of the first-step trans-splicing reaction. This assumption is based on its size (191 nt) and on the fact that there is a correlation between the amount of this band and the amount of the trans-splicing product exon 2-exon 2. Lanes 4–8 also show that when the amount of receptor was increased more product was obtained. The comparison of our results with other trans-splicing reactions with artificial pre-mRNAs leads us to conclude that exon 2 and the splice site regions around it are strong trans-splicing signals and that they may be suitable for the study of the trans-splicing mechanisms with other genes *in vitro*.

Western Blot Analysis of the COT Proteins in Rat Liver Peroxisomes. To test whether the three mature mRNAs were translated *in vivo*, Western blots of peroxisomal proteins were carried out. As observed in Fig. 6 two proteins corresponding to 69 and 79 kDa were immunolocalized. The theoretical M_r of the translated proteins of the cis-spliced mRNA is 70,302 Da and that of the trans-spliced mRNA correponding to the sequence exons 1–2–3–2–3–4 . . . , etc. is 80,572 Da, consistent with the two proteins observed. The third trans-spliced mRNA whose sequence is represented by 1–2–2–3–4 . . . , etc. is expected to be translated like the cis-spliced mRNA because the first in-frame AUG is located in the second exon 2.

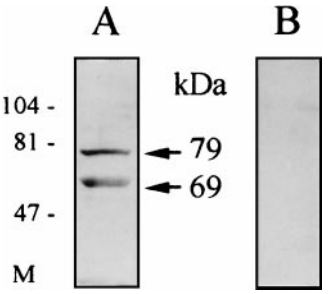

FIG. 6. Immunolocalization of peroxisomal COT. Ten micrograms of rat liver peroxisomal extracts was separated by SDS/PAGE and subjected to immunoblotting by using specific A43 antibodies (*A*) for carnitine octanoyltransferase or preimmune sera (*B*). Two bands corresponding approximately to M_r of 69 and 79 kDa are observed. The markers (M) were used to determine the approximate molecular weights of the species indicated.

DISCUSSION

In this study, we have demonstrated that the rat-liver carnitine octanoyltransferase is a single-copy gene, which is able to be processed in several mature transcripts involving cis- and trans-splicing events. This has been demonstrated by different experimental approaches such as screening of a cDNA library, RT-PCR analysis, RNase H digestion, and DNA sequencing. We also have demonstrated the ability of this mRNA to generate trans-splicing in experiments *in vitro*.

The isolation of a cDNA clone containing a repeat of exon 2, could be interpreted as an artifact generated in the construction of the library. However the accurate junction between the two exons 2 indicates a natural phenomenon. Moreover, clear evidence for the occurrence of such organization came after the sequencing of the fragments obtained by RT-PCR from total adult rat liver RNA. Not only the transcript composed by exon 1-exon 2-exon 2-exon 3, etc., whose sequence was the same as the cDNA clone isolated was found, but also the canonical exon 1-exon 2-exon 3, etc. and a third that corresponds to exon 1-exon 2-exon 3-exon 2-exon 3-exon 4, etc. The trans-splicing episode was always produced within exon 2 and 3 because the RT-PCR amplifications with primers complementary to sequences of exon 4, exon 5, exon 6, exon 7, and exon 15 always produced the same three-band pattern corresponding to an identical organization to that described above. The possibility of the duplication of exons 2 and 3 in rat genomic DNA was ruled out after a genomic Southern blot experiment, which unequivocally showed that there was only one exon 2 and one exon 3 in the genome.

Experiments with RNase H by using different annealing primers clearly confirmed the presence of the three different transcripts and that they were not attributable to artifacts, suggesting the occurrence of trans-spliced mRNA molecules in the rat liver. Splicing experiments *in vitro* confirmed that COT exon 2 functions as an acceptor and as a donor in trans-splicing reactions.

The hypothesis that trans-splicing is a regular event in mammalian cells was first suggested by Dandekar *et al.* (20). Later, the capability of mammalian cells to perform the trans-splicing reaction with appropriate foreign RNAs (21) also was demon-

strated *in vitro*. The results presented here confirm this hypothesis.

The mechanism responsible for the trans-splicing in lower eukaryotic cells or mammalians cells is not well known. Solnick (2, 4) had reported that trans-splicing in mammals *in vitro* can occur only if the 5′ and 3′ substrate RNAs form base pairs via a segment of complementarity in their introns. However, Konarska *et al.* (3) clearly show that trans-splicing does not need complementarity in the introns. Chiara and Reed (7) agreed with this view. Recently, reports have been published describing mammalian (22) and yeast (23) trans-splicing systems by using short RNA substrates. Bruzik and Maniatis (8) showed that the occurrence of an exonic-splicing enhancer (ESE) in the downstream exon is necessary for trans-splicing to occur. In addition, SR proteins have been shown to mediate alternative splicing (24–27) and commit pre-mRNAs to the splicing pathway (28, 29). Splicing enhancers facilitate the assembly of protein complexes on mRNAs containing a 3′ splice site, and these complexes are sufficiently stable to interact with 5′ splice sites located on separate mRNAs. When the sequence of exon 2 of the COT gene is carefully examined (Fig. 7) three purine rich ESEs are detected at positions 49 (GAAGAACGAA), 106 (GAAGAA), and 117 (GAAGAAG), which are similar to those previously described (27). The need of these ESE to produce trans-splicing in the COT gene remains to be demonstrated.

Conrad *et al.* (30, 31) demonstrated that a cis-splice acceptor can act as a trans-splice acceptor if it is moved into a proper context, which these authors refer to as an outron, that is, an intron without a well defined 5′ donor splice site. These studies carried out in *Caenorhabditis elegans* show that when an exon of 27 nt was placed upstream of an outron it was cis-spliced at a very low level. The trans-splicing was dominant over the cis-splicing with a poorly defined short untranslated exon even though two molecules are necessary for trans-splicing to occur. This is apparently the case shown here. The short, untranslated, poorly defined exon 1 with 30 nt (32) may not be long enough to define the intron 1, and it becomes outron-like, which allows trans-splicing and cis-splicing to take place, simultaneously.

From this study, it is clear that a trans-splicing occurs in the rat liver mRNAs when the processing of the pre-mRNAs takes place. Trans-splicing in the COT gene may occur as a result of the following: (*i*) the presence of several ESEs in exon 2, (*ii*) the strength of the 3′ acceptor splice site, and (*iii*) the poorly defined exon 1 with only 30 nt.

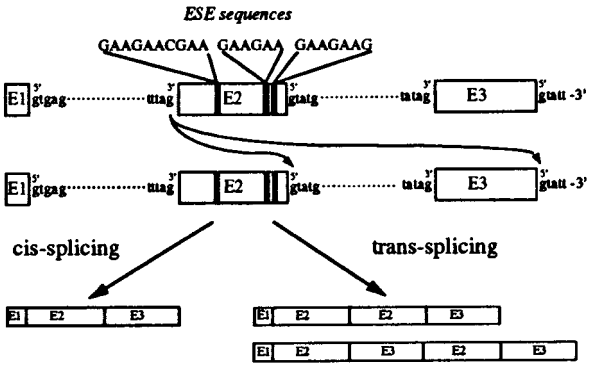

FIG. 7. Model for cis- and trans-splicing in COT pre-mRNA. The exons are represented by boxes and the introns by dotted lines. Sequences of the 5′- and 3′-intronic splice sites are indicated in lowercase. The ESE are represented in exon 2 in capital letters. Simultaneous cis- and trans-splicing of hepatic mRNA COT are produced. Trans-splicing can be promoted by ESE sequences. The 3′ acceptor site of intron 1 reacts with the 5′ donor sites of introns 2 or 3 of other molecules of pre-mRNA (trans-splicing). The organization of the mRNAs is shown at the foot.

The Western blot experiment revealed the presence in peroxisomes of two proteins translated from the mRNAs generated by cis- and trans-splicing. In theory only two proteins should be expected to appear. One, of 70,302 Da would correspond to the cis-spliced exon 1-exon 2-exon 3 . . . , etc. mRNA, which is indistinguishable from the protein translated from the trans-spliced form exon 1-exon 2-exon 2-exon 3 . . . , etc. mRNA, as exon 1 is before the first AUG codon and the first exon 2 is out of frame with respect to the second exon 2 and the following exons. The second protein (M_r 80, 572 Da), translated from the trans-spliced mRNA composed by exons 1–2–3–2–3–4 . . . , etc. should also appear. As seen in Fig. 6, two proteins with molecular mass of 69 and 79 kDa were clearly seen in the immunoblot of peroxisomal proteins when we used specific antibodies against a selected hydrophilic peptide of COT. The occurrence of the two forms of the COT enzyme in peroxisomes may be a direct consequence of this trans-splicing mechanism in the COT gene.

We thank Bandana Chatterjee for the COT probe. We thank other members of our group, in particular Diego Haro, Pedro F. Marrero, Carme Caelles, and Carles J. Ciudad for critical reading of this manuscript. The editorial help of Robin Rycroft is greatly acknowledged. This work was supported by Grant PB95-0012 from the Dirección General de Investigaci-n Científica y Técnica, Spain. C.C. and A.M. are recipient of fellowships from the Dirección General de Investigación Científica y Técnica and the Instituto de Cooperación Iberoamericana, Spain, respectively.

1. Bonen, L. (1993) *FASEB J.* **7,** 40–46.
2. Solnick, D. (1985) *Cell* **42,** 157–164.
3. Konarska, M. M., Padgett, R. A. & Sharp, P. A. (1985) *Cell* **42,** 165–171.
4. Solnick, D. (1986) *Cell* **44,** 211.
5. Sharp, P. A. & Konarska, M. M. (1986) *Cell* **44,** 211.
6. Bruzik, J. P. & Maniatis, T. (1992) *Nature (London)* **360,** 692–695.
7. Chiara, M. D. & Reed, R. (1995) *Nature (London)* **375,** 510–513.
8. Bruzik, J. P. & Maniatis, T. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 7056–7059.
9. Sullivan, P. M., Petrusz, P., Szpirer, C. & Joseph, D. R. (1991) *J. Biol. Chem.* **266,** 143–154.
10. Shimizu, A. & Honjo, T. (1993) *FASEB J.* **7,** 149–154.
11. Vellard, M., Sureau, A., Soret, J., Martinerie, C. & Perbal, B. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 2511–2515.
12. Breen, M. A. & Ashcroft, S. J. H. (1997) *FEBS Lett.* **409,** 375–379.
13. Chatterjee, B., Song, C. S., Kim, J. M. & Roy, A. K. (1988) *Biochemistry* **27,** 9000–9006.
14. Choi, S. J., Oh, D. H., Song, C. S., Roy, A. K. & Chatterjee, B. (1995) *Biochim. Biophys. Acta* **1264,** 215–222.
15. de Wet, J. R., Wood, K. V., DeLuca, M., Helinski, D. R. & Subramani, S. (1987) *Mol. Cell. Biol.* **7,** 725–737.
16. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
17. Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. (1983) *Nucleic Acids Res.* **11,** 1475–1489.
18. Derrick, J. P. & Ramsay, R. R. (1989) *Biochem. J.* **262,** 801–806.
19. Marglin, A. & Merrifield, R. B. (1970) *Annu. Rev. Biochem.* **39,** 395–420.
20. Dandekar, T. & Sibbald, P. (1990) *Nucleic Acids Res.* **18,** 4719–4725.
21. Eul, J., Graessmann, M. & Graessmann, A. (1995) *EMBO J.* **13,** 3226–3235.
22. Konforti, B. B. & Konarska, M. M. (1995) *RNA* **1,** 815–827.
23. Ghetti, A. & Abelson, J. N. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 11461–11464.
24. Horowitz, D. S. & Krainer, A. R. (1993) *Trends Genet.* **10,** 100–106.
25. Watakabe, A., Tanaka, K. & Shimura, Y. (1993) *Genes Dev.* **7,** 407–418.
26. Lavigueur, A., La Branche, H., Kornblihtt, A. R. & Chabot, B. (1993) *Genes Dev.* **7,** 2405–2417.
27. Sun, Q., Mayeda, A., Hampson, R. K., Krainer, A. R. & Rottman, F. M. (1993) *Genes Dev.* **7,** 2598–2608.
28. Fu, X. D. (1993) *Nature (London)* **365,** 82–85.
29. Bruzik, J. P. (1996) *Microb. Pathog.* **21,** 149–155.
30. Conrad, R., Thomas, J., Spieth, J. & Blumenthal, T. (1991) *Mol. Cell. Biol.* **11,** 1921–1926.
31. Conrad, R., Liou, R. F. & Blumenthal, T. (1993) *EMBO J.* **12,** 1249–1255.
32. Berget, S. M. (1995) *J. Biol. Chem.* **270,** 2411–2414.