

# Structural organization of mouse peroxisome proliferator-activated receptor $\gamma$ (mPPAR $\gamma$ ) gene: Alternative promoter use and different splicing yield two mPPAR $\gamma$ isoforms

(peroxisome proliferation/nuclear receptor superfamily/fatty acid  $\beta$ -oxidation)

YIJUN ZHU\*, CHAO QI\*, JULIE R. KORENBERG<sup>†</sup>, XIAO-NING CHEN<sup>†</sup>, DAVID NOYA<sup>†</sup>, M. SAMBASIVA RAO\*, AND JANARDAN K. REDDY<sup>‡</sup>

\*Department of Pathology, Northwestern University Medical School, Chicago, IL 60611; and <sup>†</sup>Department of Pediatrics, Division of Genetics, Cedars-Sinai Research Institute, University of California, Los Angeles, CA 90048

Communicated by N. Edward Tolbert, Michigan State University, East Lansing, MI, April 28, 1995 (received for review November 22, 1994)

**ABSTRACT** To gain insight into the regulation of expression of peroxisome proliferator-activated receptor (PPAR) isoforms, we have determined the structural organization of the mouse PPAR  $\gamma$  (mPPAR $\gamma$ ) gene. This gene extends >105 kb and gives rise to two mRNAs (mPPAR $\gamma$ 1 and mPPAR $\gamma$ 2) that differ at their 5' ends. The mPPAR $\gamma$ 2 cDNA encodes an additional 30 amino acids N-terminal to the first ATG codon of mPPAR $\gamma$ 1 and reveals a different 5' untranslated sequence. We show that mPPAR $\gamma$ 1 mRNA is encoded by eight exons, whereas the mPPAR $\gamma$ 2 mRNA is encoded by seven exons. Most of the 5' untranslated sequence of mPPAR $\gamma$ 1 mRNA is encoded by two exons, whereas the 5' untranslated sequence and the extra 30 N-terminal amino acids of mPPAR $\gamma$ 2 are encoded by one exon, which is located between the second and third exons coding for mPPAR $\gamma$ 1. The last six exons of mPPAR $\gamma$  gene code for identical sequences in mPPAR $\gamma$ 1 and mPPAR $\gamma$ 2 isoforms. The mPPAR $\gamma$ 1 and mPPAR $\gamma$ 2 isoforms are transcribed from different promoters. The mPPAR $\gamma$  gene has been mapped to chromosome 6 E3-F1 by *in situ* hybridization using a biotin-labeled probe. These results establish that at least one of the PPAR genes yields more than one protein product, similar to that encountered with retinoid X receptor and retinoic acid receptor genes. The existence of multiple PPAR isoforms transcribed from different promoters could increase the diversity of ligand and tissue-specific transcriptional responses.

Peroxisome proliferators are a structurally diverse group of compounds that include certain phthalate ester plasticizers, herbicides, hypolipidemic drugs, and leukotriene D4 inhibitors (1). These agents, when administered to rodents and nonrodents including primates, cause profound proliferation of peroxisomes in hepatic parenchymal cells, as well as marked increases in the activities of enzymes required for the peroxisomal  $\beta$ -oxidation of fatty acids—namely, fatty acyl-CoA oxidase, enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase bifunctional enzyme, and 3-ketoacyl-CoA thiolase (2–4). Despite the structural diversity and differences in their potencies, these agents induce qualitatively predictable pleiotropic responses characterized by hepatomegaly, an increase in the number of peroxisomes in liver cells, and rapid induction of the mRNA levels of all three enzymes of the peroxisomal fatty acid  $\beta$ -oxidation system due to coordinate increase in the rates of transcriptional activation of the nuclear genes encoding these enzymes (4, 5). Continued administration of peroxisome proliferators leads to the development of hepatocellular carcinomas in rats and mice despite the inability of these compounds to interact with and damage DNA directly (6). In view of the

non-genotoxic nature of these agents, the hepatocarcinogenesis by peroxisome proliferators is postulated to involve the secondary events associated with an enormous increase in the activity of the H<sub>2</sub>O<sub>2</sub>-generating peroxisomal fatty acyl-CoA oxidase and downstream events related to the overproduction of H<sub>2</sub>O<sub>2</sub> and possibly other reactive oxygen intermediates (1).

Given the tissue-specific induction of peroxisome proliferation, the rapidity of response, the demonstration of a reversible peroxisome proliferator-binding moiety in the liver cytosol, and the transcriptional activation of  $\beta$ -oxidation enzyme system genes, it was proposed that peroxisome proliferators exert their pleiotropic responses via a receptor-mediated signal-transduction mechanism (4). Support for this concept was provided by the identification of a peroxisome proliferator-activated receptor (PPAR), a member of the nuclear hormone receptor superfamily in mouse liver (7). Three isoforms of PPAR have been described in *Xenopus* (xPPAR $\alpha$ , xPPAR $\beta$ , and xPPAR $\gamma$ ) and in mouse (mPPAR $\alpha$ , mPPAR $\gamma$ , and mPPAR $\delta$ ) (8–10). After the isolation of mPPAR $\gamma$  cDNA from mouse liver (9), an isoform of this receptor, designated mPPAR $\gamma$ 2, which exhibited adipocyte specificity in mouse, has recently been described (11). The mPPAR $\gamma$ 2 cDNA encodes an additional 30 amino acids, N-terminal to the first ATG codon of mPPAR $\gamma$ 1 (previously mPPAR $\gamma$ ), and also differs in the 5' untranslated sequence. The restricted tissue distribution of xPPAR $\gamma$ , mPPAR $\gamma$ 1, and mPPAR $\gamma$ 2, as well as the identification of other isoforms, raises some important questions regarding the physiological role and the structural basis of the diversity of this PPAR subfamily. Elucidation of the structural organization of mammalian PPAR genes should provide a starting point to unravel the molecular mechanisms controlling tissue specificity as well as the evolution of diversity of PPAR subfamily. Here, we report the structural organization of the mPPAR $\gamma$  gene and show that this mouse gene gives rise to two distinct mRNAs (mPPAR $\gamma$ 1 and mPPAR $\gamma$ 2) due to alternative promoter use and tissue-specific differences in splicing patterns. This mPPAR $\gamma$  gene has been mapped to the E3-F1 region of mouse chromosome 6, by *in situ* hybridization, using a biotinylated probe.

## MATERIALS AND METHODS

**Isolation of Genomic Clones.** Three pairs of oligonucleotide primers (5'-CAGTTTCGATCCGTAAGAAGCCGT-3' and 5'-CTCAAACCTTAGGCTCCATA-3'; 5'-ATTGGGTGCGC-GCGACC-3' and 5'-AGAACGTGATTCTCAGCC-3'; and 5'-CCAGTGTGAATTACAGCAAAATCTCTGTTT-TATGCTG-3' and 5'-TTGTGAAGTGCTCATAGGCAAGTG) were used to screen a mouse P1 bacteriophage library

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.

Abbreviations: PPAR, peroxisome proliferator-activated receptor; mPPAR, mouse PPAR; xPPAR, *Xenopus* PPAR.  
<sup>‡</sup>To whom reprint requests should be addressed.

(Genome Systems, St. Louis) by PCR method (12). Six clones that contained inserts ranging from 60 to 80 kb in length were obtained.

**Restriction Mapping, Determination of Exon/Intron Boundaries, and DNA Sequencing.** P1 clone DNA was prepared by the alkaline lysis procedure, digested with *Sac* I, *Apa* I, and *Xba* I, respectively. The restriction fragments were purified and ligated into the corresponding sites of pBluescript SK vector (Stratagene) using standard laboratory procedures (13). The ligated product was used to transform the *E. coli* DH5 $\alpha$ , and plasmid DNA was isolated. Positive subclones were identified by Southern blot analysis of the plasmid DNA using mPPAR $\gamma$ 1 cDNA as probe (9), isolated, and subjected to various restriction enzyme digestions to map the mPPAR $\gamma$  gene. Double-stranded sequencing of denatured plasmid DNA was done to determine the intron/exon boundaries and also to obtain sequence information of promoters by using modified T7 DNA polymerase and the dideoxynucleotide chain-termination method (13). The exon/intron boundaries were identified by using mPPAR $\gamma$ 1 and mPPAR $\gamma$ 2 cDNA sequence information (9, 11). Introns that were not included in these plasmids were amplified from the P1 clones using extra long (XL) PCR (14).

**Determination of Transcription Initiation Site.** Transcription initiation sites for mPPAR $\gamma$ 1 and mPPAR $\gamma$ 2 were determined by a ribonuclease protection assay and/or primer-extension analysis. Primer extension for mPPAR $\gamma$ 1 was done according to a described method (13). For the ribonuclease protection assay, a fragment from -272 to +145 in the promoter region for mPPAR $\gamma$ 1 was amplified with primers (5'-TATCAGGATCCATATATATGATATATAAAT-3' and 5'-CGGGTCTCGAGCGTCCGGTCCGGCCGACCC-3') including *Bam*HI and *Xho* I sites and then subcloned into the *Bam*HI/*Xho* I site of pBluescript SK. This plasmid was linearized with *Xba* I and used to synthesize the RNA probe by *in vitro* transcription using [ $\alpha$ - $^{32}$ P]CTP (10 mCi/ml, 800 Ci/mmol; 1 Ci = 37 GBq) and T7 RNA polymerase (13). Probe (1  $\times$  10 $^5$  cpm) was added to 10  $\mu$ g of mouse liver RNA and hybridized overnight at 42°C. The reactions were digested with 5  $\mu$ g of ribonuclease A per ml and 100 units of ribonuclease T1 per ml and analyzed on 5% polyacrylamide/8 M urea gel. Control reactions included 10  $\mu$ g of yeast tRNA with or without ribonuclease digestion.

**Analysis of Promoter Activity.** To test the promoter activity of the  $\gamma$ 1 promoter region (mPPAR $\gamma$ 1 promoter), a blunted *Xba* I-*Apa* I fragment and a *Hind*III-*Apa* I fragment from the 5'-flanking region of mPPAR $\gamma$ 1 gene were cloned into blunted *Hind*III site of pGL2-basic vector (Promega), which contains luciferase gene as a reporter. Luciferase construct (6  $\mu$ g of DNA) and  $\beta$ -galactosidase expression plasmid pCH110 (4  $\mu$ g DNA) were used to transfect the H4IIEC3 cells using the calcium phosphate precipitation technique (13). The culture medium was replaced by fresh medium 23 hr later, and cell extracts were prepared 48 hr after transfection by three cycles of freezing and thawing. The homogenates were assayed for luciferase and  $\beta$ -galactosidase activities (13). To ascertain the promoter activity of  $\gamma$ 2 promoter (promoter for mPPAR $\gamma$ 2), a fragment from position -611 to +62 was amplified with primers (5'-TCACCGCTAGCAGCATAAAAACAGAGATT-TGCT-3' and 5'-CTTTTCTAGAATTTGGATAGCAGT-3') including *Nhe* I and *Xba* I sites and subcloned into the *Nhe* I site of pGL2-basic vector (Promega). 3T3-L1 cells were induced to differentiate into adipocytes (15) and then transfected with the luciferase construct (6  $\mu$ g) and  $\beta$ -galactosidase expression plasmid pCH110 (4  $\mu$ g) by the calcium phosphate precipitation technique (13). The cells were incubated with the DNA-calcium phosphate precipitate for 16 hr and washed and refed. Cells were collected after 36 hr and were assayed for luciferase and  $\beta$ -galactosidase activities (13).

**Chromosomal Localization.** The pBluescript SK plasmid with an insert of a 19.8-kb *Sac* I fragment from mPPAR $\gamma$  gene containing exons 2 and 3 was labeled with biotin-14-dATP (GIBCO/BRL) using nick translation and was hybridized to female mouse metaphase chromosomes. Fluorescence *in situ* hybridization (FISH) was done essentially according to a method described by Korenberg and Chen (16). The image was captured by a Photometric cooled-charged coupled device camera and Oncor image analysis system.

## RESULTS

**Structural Organization of mPPAR $\gamma$  Gene.** To isolate the mPPAR $\gamma$  gene a mouse P1 bacteriophage library was first screened by PCR using a pair of oligonucleotide primers (5'-CAGTTTCGATCCGTTAGTAGAAGCCGT-3' and 5'-CTCAAACCTTAGGCTCCATA-3'), which turned out to cover the same exon (no. 5) of mPPAR $\gamma$  gene. Two clones were isolated, the authenticity of one of which (clone 646) was confirmed by Southern blotting using mPPAR $\gamma$ 1 cDNA as probe. This clone was chosen for detailed mapping and analysis. DNA was digested with *Sac* I and *Xba* I and then subcloned into pBluescript KS by a shot-gun procedure. Plasmids containing inserts were selected by Southern blotting with mPPAR $\gamma$ 1 cDNA. Mapping and sequencing indicated that this clone (clone 646) contained all exons covering the translating region, but it did not include most of the 5' untranslated region of mPPAR $\gamma$ . A second primer pair (5'-ATTGGGTCGCGCGCAGCC-3' and 5'-AGAACGTGATT-TCTCAGCC-3'), which covered a region close to the 5'-end of the mPPAR $\gamma$ 1 cDNA was then used to screen the library. Two positive clones (nos. 1222 and 1223) were isolated; they shared a similar pattern after digestion with different restriction endonucleases, but Southern blot analysis revealed that clone 1223 had one more positive band with some restriction endonucleases than the other clone, indicating that it contained an additional exon. This clone, clone 1223, was subcloned into pBluescript SK and mapped. Upon sequencing, it revealed the presence of the 5' untranslated sequence of mPPAR $\gamma$ 1 and the promoter region, but it did not overlap with clone 646, which contained the entire translating region, suggesting the presence of a large intron. Because none of these clones contained the coding sequence for the extra 30 N-terminal amino acids of adipocyte-specific mPPAR $\gamma$ 2 cDNA, a third primer pair (5'-CCAGTGTGAATTACAGCAAATCTCTGTTTTATGCTG-3' and 5'-TTGTGAAGT-GCTCATAGGCAGTG-3') was used to screen the P1 library, and two clones (nos. 2991 and 2992) were obtained. These primers were derived from the region encoding the first 30 N-terminal amino acids of mPPAR $\gamma$ 2. Southern blot analysis indicated that both of these clones overlapped clone 646 and contained mPPAR $\gamma$ 2 extra sequence but did not overlap with clone 1223, which contained the 5'-noncoding sequences and the promoter of mPPAR $\gamma$ 1. Clone 2991 was subcloned and mapped.

The location of exon sequences within the genomic clones is depicted in Fig. 1. The mPPAR $\gamma$  gene spans >105 kb, and it contains exons that encode for both mPPAR $\gamma$ 1 and mPPAR $\gamma$ 2 mRNAs (Fig. 1). The mPPAR $\gamma$ 1 mRNA species is transcribed by eight exons with most of the 5' untranslated region derived from two exons designated  $\gamma$ A1 and  $\gamma$ A2. The protein-coding region of this mRNA is derived from six exons designated 1-6. These six downstream exons are also shared by mPPAR $\gamma$ 2 mRNA (Fig. 1). The extra 30 N-terminal amino acids of mPPAR $\gamma$ 2, as well as its 5'-untranslated sequence, are encoded by one exon designated  $\gamma$ B1 (Fig. 1). This exon is located between the second exon (exon  $\gamma$ A2) and the third exon (exon 1) of mPPAR $\gamma$ 1.

Each of the two zinc fingers of the DNA-binding domain of mPPAR $\gamma$  is encoded by a separate exon (exons 2 and 3,

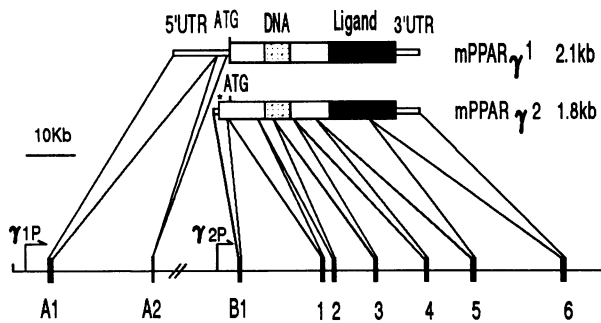


FIG. 1. Organization of mPPAR $\gamma$  gene. The eight exons (A1, A2, and 1–6) encoding the mPPAR $\gamma$ 1 (solid boxes) as well as the seven exons (B1 and 1–6) encoding mPPAR $\gamma$ 2 are shown in the genomic DNA.  $\gamma$ P1 and  $\gamma$ P2 represent the promoter of mPPAR $\gamma$ 1 and mPPAR $\gamma$ 2, respectively. The corresponding positions of the exons with respect to mPPAR $\gamma$ 1 and mPPAR $\gamma$ 2 cDNAs are shown above. Exons numbered 1–6 code for identical sequences in both isoforms starting at the ATG. The mPPAR $\gamma$ 2 cDNA encodes an additional 30 amino acids N-terminal to the first ATG of mPPAR $\gamma$ 1 from an upstream ATG (indicated by \*) from exon B1 shown on the genomic DNA. At the top, the location of 5' and 3' untranslated regions (UTR) and the DNA- and ligand-binding domains are shown.

respectively). The ligand-binding domain is encoded by two exons (exons 5 and 6). All the splice acceptor and donor sequences conformed to the GT-AG consensus rule for splicing (Table 1). The splice junctions between exons  $\gamma$ 2 B1/1, 1/2, 3/4, and 5/6 occurred after the first nucleotide of the codon (type 1), and the rest of the intron–exon boundaries occurred between codons (type 0). These values are consistent with the ratios observed for vertebrate genes.

**Transcription Initiation Site of mPPAR $\gamma$ 1 and mPPAR $\gamma$ 2.** The transcription initiation site of mPPAR $\gamma$ 1 was determined by primer-extension analysis using mouse liver RNA. An oligonucleotide (5'-CCGCGCGGACCGAGGAG-3') that is located in the 5' untranslated region was used as a primer. A primer-extension product, 74 bp in size, was identified (data not shown). This product indicated that the transcription start site is the cytosine residue at 495 bp upstream from the translation start site ATG of the mPPAR $\gamma$ 1 mRNA (Fig. 2). This site was further verified as the transcription initiation site by ribonuclease protection assay, which produced a 145-bp protected fragment (data not shown) with a probe from -272 to +145 of the promoter region,  $\gamma$ 1 promoter, for mPPAR $\gamma$ 1 (Fig. 3). The assignment of tran-

scription initiation site for mPPAR $\gamma$ 2 is based on the primer-extension assay using adipocyte RNA. An oligonucleotide primer (5'-AACAGGAGAATCTCCAGAGTTTCA-3') located in the exon B1 yielded a 94-bp extension product (data not illustrated), which indicated that the transcription start site is an adenine residue, 64 bp upstream from the translation start site ATG for the mPPAR $\gamma$ 2 (Fig. 3). This assignment is in agreement with the primer-extension results obtained by Tontonoz *et al.* (11).

**5' Upstream Regions of mPPAR $\gamma$ 1 and mPPAR $\gamma$ 2.** The promoter region of mPPAR $\gamma$ 1 (see  $\gamma$ 1P in Fig. 1) was sequenced (Fig. 2). A 550-bp sequence of 5' untranslated region and the promoter region (-147 to +404) is extremely G+C rich (75%). Although precise GC box motifs were not found in this promoter sequence, there are three GGGCGC motifs in this sequence. There is a 200-bp (A+T)-rich sequence upstream of the (G+C)-rich region (-397 to -198). The first TATA-like sequence is located 210 bp upstream from the transcription start site. The mPPAR $\gamma$ 1 promoter also contains consensus binding sequences of transcription factors Ap2, C/EBP, HNF5, and HNF3. Liver is rich in these transcription factors, and the presence of consensus binding identity of the  $\gamma$ 1 promoter is consistent with the expression of mPPAR $\gamma$ 1 transcript in liver. The sequence of the 5' region of mPPAR $\gamma$ 2 is shown in Fig. 3. The TATA-like element (TATTA) is located 49 bp upstream of the transcription start site. This promoter region also contains consensus binding sequences of transcription factors HNF3, C/EBP, and Ap1. Transfection assay confirmed the identity of the region as the promoter of mPPAR $\gamma$ 2. Identity of the  $\gamma$ 1 promoter as promoter was confirmed by transfection assay. The 930-bp *HindIII*-*Apa* I fragment was inserted into the *HindIII* site of the luciferase expression vector pGL2-basic in both forward and reverse orientations. Transfection of the forward expression construct into H4IIEC3 rat hepatoma cells showed  $\approx$ 2-fold increase in luciferase activity as compared with the control pGL2 basic vector (Fig. 4A). Cloning of the fragment in the reverse orientation demonstrated no increase in luciferase activity. A 2.7-kb *Xba* I-*Apa* I fragment, which includes a 1.8-kb upstream sequence, and the *HindIII*-*Apa* I sequence, which includes a 1.8-kb upstream sequence and the 930-bp *HindIII*-*Apa* I sequence, when transfected into H4IIEC3 cells showed 6-fold increase in luciferase activity, suggesting that this 1.8-kb fragment possibly serves as an enhancer. To assay the promoter activity of  $\gamma$ 2 promoter, a 673-bp fragment from -611 to +62 (see Fig. 3) was inserted into the *Nhe* I site of pGL2-basic in both forward and reverse directions. The for-

Table 1. Mouse PPAR $\gamma$  gene: Sequence of intron/exon boundaries for mPPAR $\gamma$ 1 and mPPAR $\gamma$ 2 isoforms

Exon size, bp	Intron		Exon
	Donor	Acceptor	
$\gamma$ 1A1 (417)	ACAAG gtgagggggg .....	(21.9 kb).....aattattcag ATTTG	A2
$\gamma$ 1A2 (65)	TACTG gtatgtacat.....	(>33.3 kb).....tccacttcag AAATT	1
$\gamma$ 2B1 (121)	A, CAA, G gtaaagtcca .....	(18.2 kb).....tccacttcag AA, ATT	1
	S Q	E I	
1 (228)	C, CAA, A gtatgtttac .....	(1.5 kb).....ccatcgacag GT, GCG	2
	Y Q	S A	
2 (170)	GC, AAG gtaaaaatta .....	(9.4 kb).....ccttttgcag GGT, TT	3
	C K	G F	
3 (139)	C, AAT, G gtaagtggat.....	(10.9 kb).....tcctctgtag CC, ATC	4
	H N	A I	
4 (200)	AA, TCA gttagtctc .....	(9.6 kb).....tcatttgcag CCA, TT	5
	K S	P F	
5 (451)	T, GGA, G gtaagattca.....	(20.1 kb).....attctcccag AC, CGC	6
	S G	D R	
6 (445)	CATTA taaaattgttt		

Exons 1–6 code for identical sequences in both isoforms.  $\gamma$ 1A1 and  $\gamma$ 1A2 are two exons encoding the 5' untranslated region of mPPAR $\gamma$ 1 cDNA, and  $\gamma$ 2B1 is the exon that encodes 30 amino acids that are additional at the N terminus of mPPAR $\gamma$ 2.

-847 **AAGCTTTCTGAATGTGAATCTTAGGACAGATCATGACTGTGAGAGCTGAA**  
 HindIII  
 -797 **AGAGTGAAGAAAAGGAACGTCTATCATGTGGGCTTCAGGCTCCCGAGG**  
 -747 **CAAGAAGCTCCAATAGTCTAACTTAAAAACAAGAGATTCTCAGGCCCTCT**  
 -697 **CCACCCTATGTGTAAGGTCTCCTATGTGAAGAAATGGTCTAAAGAAATTT**  
 -647 **AAAAAGCCAGTGAACAAGGTTAA**CAATA**GATTGAAGGCAGGAGAGATTG**  
 HNF5 C/EBP  
 -597 **CTCAGCGGTTAGGAACACTTTGTGTTCTCCAGAGGACCCCTTGTTCCGGTT**  
 -547 **CCCAGCACCACCCCTGTAACCTCCAGCTCCAGGG**CGGAGCC**CACACCTCTC**  
 Ap2  
 -497 **CCTCCGTT**CCCTGC**CATACACACAATTTGTTTTTAAAGGAAATTTAA**  
 Myc  
 -447 **TGATTAGCACTAAAGTCT**CTGTTAAG**CGGTTCCGCCTTCATAACATTTCT**  
 HNF3  
 -397 **GAGATTAATAATAATAAACTGAGATTAATAATAATAAAAGAAAAAAG**  
 -347 **AAAAAAGGGAGATCCTAAATTCACCTTAATTCATATAAATATATATATGGG**  
 -297 **TGTGTATCTATTATATATCATATATATATATATATATATATATATATG**  
 -247 **TTATGTATGATATATAAATATATATATG**TATATA**TATAGAGAGAGATGA**  
 -197 **AAAGCACATCTAGAAAAAACAACCTCTCCATGACAGACATGGACATG**  
 -147 **GACATCGGCTCTGAGGGACACGGGACCTTCCTGGGTCGCCTCCACGGCGCT**  
 -97 **GTGAGGAGCAAGGCGGCCAGGTAACACGCGCACGGCGCGGGCGGGGAG**  
 -47 **GGCCACGCGGCTCTG**GGGCGC**CGGCTCCGGGTGCGGGCGCGGGGCT**  
 +4 **CCGGCTCACCTGGGCACGGCCCCGGTCCGGCCGGGTGGCTCGGCTCG**  
 +54 **GCTCCTCCGGTCCGGCGGGCCAGGCGGGCCCGTCCGACTCAGAGCGG**  
 Ap1  
 +104 **CAGCCGCTGGGGCGCTCGGGTCCGGTCCGGCCACCGCACAGCACC**  
 +154 **CGCAAGCCGCGCTCAGGTGAGAGTCCGCCCCGGCCACGCGGCCGCGC**  
 +204 **GGAGGGACGCGGAAGAAGAGACTGGGGCGTGCCTGGGGTATTGGGTCG**  
 +254 **CGCGCAGCTGAGGGACCGAGTGTGACGACAAGGTGACGGGGCTGAGGG**  
 +304 **GACGGGCTGAGGAGAAGTCACTCTGACAGGAGCCTGTGAGACCAACAG**  
 +354 **CCTGACGGGTCTCGGTTGAGGGGACGGGCTGAGAAATCACGTTCTGACA**  
 +404 **GGACTGTGTGACAGACAAG/gt...intron 1..ag/ATTGAAAGAAG**  
 +435 **CGGTGAACCACTGATATTCAGGACATTTTTAAAAACAAGACTACCCTTTA**  
 +485 **CTG/gt...intron 2..ag/AAATTACCA**ATG****

FIG. 2. The promoter region of mPPAR $\gamma$ 1. The mPPAR $\gamma$ 1 5' untranslated and promoter region from the initiation ATG methionine (boldface type and underlined). The transcription start site, indicated by a star (nt C) is designated as +1. The imperfect GC box motifs are underlined. The TATA box-like sequence is double underlined. The HNF5, C/EBP, Ap2, myc, and HNF3 motifs are shown as boldface italic letters.

ward expression vector, when transfected into adipocyte cells, showed  $\approx$ 7-fold increase in luciferase activity as compared with the pGL2-basic control (Fig. 4B).

**Chromosomal Localization.** The biotinylated plasmid probe, which contained a 19.8-kb fragment from mPPAR $\gamma$  gene, was mapped to mouse chromosome band MMU6E3-F1. Three independent experiments were done, and >100 metaphase cells were evaluated. Signals were clearly seen on two chromatids of at least one mouse chromosome MMU6E3-F1 in 85% of cells (Fig. 5). There were no other hybridization signals seen on other chromosomes in more than one single chromosomal band.

## DISCUSSION

In this paper we report the genomic structure and chromosomal location of the PPAR $\gamma$  gene. This gene, which spans >105 kb, is at the E3-F1 region of mouse chromosome 6 and gives rise to two mRNA species, mPPAR $\gamma$ 1 and mPPAR $\gamma$ 2, with distinct 5' exons that are spliced onto common downstream sequences. Furthermore, the data we present here clearly indicate that the two mRNAs arise as a result of alternative promoter use. The mPPAR $\gamma$ 1 is encoded by eight exons, whereas the mPPAR $\gamma$ 2 is encoded by seven exons. The last six exons, which encode a 475-amino acid protein, are identical for both mPPAR $\gamma$ 1 and mPPAR $\gamma$ 2. Although these

-615 **TT**TATAGAA**TTTGGATAGCAGTAACATTTTGGACCTTTTATAGACTTGT**  
 HNF3  
 -565 **GAATAAATCACCTTAGAAATGTATTTTGTACCTGTGTGGGTAAACAAAT**  
 -515 **CTAAAATAAGAATGTGTATATGTTTGGAGTACAAGAATATTTCTCAGATG**  
 -465 **GTGATTAGGAGTTTCAACCAAAGATAAATACTTAAGAAAACTTTGGCCA**  
 -415 **AATACGTTTATCTGGTGTTCATAACTTAGAGATTAAGGTTTTCTATTTT**  
 -365 **AAAAGCCACTGGTGTGATTTTACT**CAAT**TTTAAAG**CAAT**CAATATT**  
 C/EBP C/EBP  
 -315 **GAACAATCTCTGCTCTGGTAATTCAACTACTGTACAGTTACGCGCCCTC**  
 -265 **ACAGAACAGTGAATGTGTGGGCTACTGGCGAGACAATGTAGCAACGTTTT**  
 -215 **CCTTGTAATGTACCAAGTCTTGCCAAAGCAGCAGCAGCATTATGACACA**  
 -165 **CCATTTTGTACAACTGGCT**TCAGTCAG**GACAGTCCAGCCAATTCAGG**  
 Ap1  
 -115 **CCTGATCTTTCTGTGTTTATTTCCACCTCTCCAAATATTTGAAAAGT**  
 -65 **GTGCTTGACTT**TATT**AGCATATTCATAAGCTCGATGACCATAAGCCTTTT**  
 -15 **TCCTTAAACCA**CAAT**CTTTTGAAGACATAGACAAAACACCAAGTGTGA**  
 +36 **ATTACAGCAAATCTCTGTTTTATGCTGTT**ATG****

FIG. 3. The promoter region of mPPAR $\gamma$ 2. The mPPAR $\gamma$ 2 5' untranslated and promoter region from the initiation ATG methionine (boldface type and underlined). The transcription start site, indicated by a star (nt A) is designated as +1. A putative TATA box (TATTA) is underlined. HNF3, C/EBP, and Ap1 motifs are shown as boldface italic letters.

six exons encode the entire mPPAR $\gamma$ 1 protein and also account for the last 475 amino acids of the mPPAR $\gamma$ 2 protein, the extra 30 N-terminal amino acids of mPPAR $\gamma$ 2 are encoded by a different exon ( $\gamma$ B1) located in the second intron (between exon  $\gamma$ A2 and exon 1) of mPPAR $\gamma$ 1. Exons  $\gamma$ A1 and  $\gamma$ A2 of this mPPAR $\gamma$  gene encode almost the entire portion of the N-terminal untranslated region of mPPAR $\gamma$ 1, whereas exon  $\gamma$ B1, as mentioned above, encodes the 5' untranslated region and the N-terminal 30 amino acids that are not present

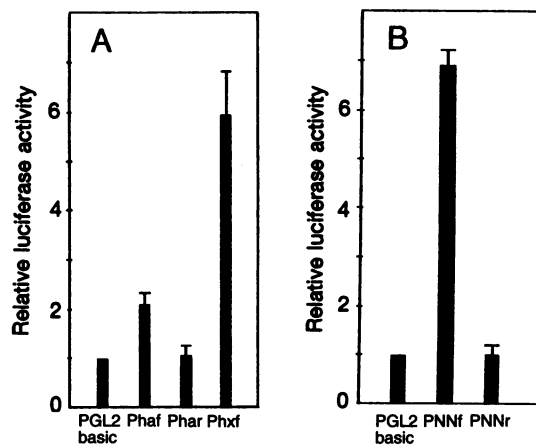


FIG. 4. Promoter activity of  $\gamma$ 1 promoter of mPPAR $\gamma$ 1 and  $\gamma$ 2 promoter of mPPAR $\gamma$ 2. (A) Transfection assay of the 5' region of mPPAR $\gamma$ 1. A 930-bp fragment and a 2.7-kb fragment were subcloned into a luciferase expression vector (pGL2-basic) and transiently transfected into H4IIEC3 cells. The relative levels of luciferase expression are shown for the promoterless vector (pGL2-basic), the forward 930-bp promoter construct (Phaf), the reverse 930-bp promoter construct (Phar), and the forward 2.7-kb promoter construct (Phxf). Values are the means of three independent experiments normalized to a  $\beta$ -galactosidase transfection controls. (B) Reporter plasmid expression driven by mPPAR $\gamma$ 2 promoter. A 673-bp fragment, from position -611 to +62, was subcloned into the luciferase reporter plasmid (pGL2-basic) and transiently transfected into adipocyte cells. The relative luciferase activities were shown for the promoterless vector (pGL2-basic), the forward 673-bp promoter construct (PNNf), and the reverse 673-bp promoter construct (PNNr). Results are the means of three independent transfections normalized to the internal controls of  $\beta$ -galactosidase expression.

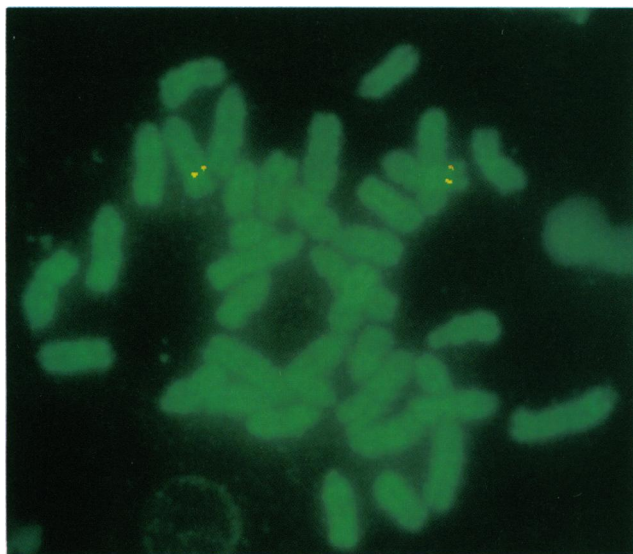


FIG. 5. Fluorescence *in situ* hybridization mapping of mPPAR $\gamma$ . Mouse chromosome preparation was hybridized with a plasmid probe containing a 19.8-kb fragment from mPPAR $\gamma$  gene which was labeled with biotin-14-dATP. The fluorescein isothiocyanate signals were clearly shown on chromomycin/distamycin reverse-banded mouse chromosome 6 at the region of E3-F1. A mouse chromosome 6 ideogram (not illustrated) showed the location of mPPAR $\gamma$  gene at the region of MMU6E3-F1.

in mPPAR $\gamma$ 1. Delineation of the structure of the mPPAR $\gamma$  gene constitutes the only description of the genomic organization of a member of the mammalian PPAR gene subfamily.

Comparison of the mPPAR $\gamma$  gene with other nuclear receptor genes of the steroid-hormone receptor subfamily reveals some interesting facets. The size of mPPAR $\gamma$  gene (>105 kb) is in the same range as that of the human estrogen receptor, which is >140 kb in length (17). The two zinc fingers of mPPAR $\gamma$  are separated by one intron, which is located one amino acid after the last cysteine residue of the first zinc finger, a splicing pattern shared by the TR/RAR/*ear 1* subfamily (thyroid hormone receptor  $\alpha$  gene, retinoic acid receptor gene, and *ear 1* gene) (1, 18). In contrast, this intron is located in a different position between the two zinc fingers in the steroid-hormone receptor gene and the *nur77* gene (19). There is only one intron in the ligand-binding domain of mPPAR $\gamma$ , whereas the *ear 1* gene, which is considered closest to the mPPAR $\gamma$  gene, possesses two or more introns in this domain (18). It is pertinent to note that only a single intron is found in the ligand-binding domain of *Xenopus* PPAR $\beta$  (xPPAR $\beta$ ) (20). The presence of a single intron in the mouse (mPPAR $\gamma$ ) and the *Xenopus* (xPPAR $\beta$ ) genes characterized thus far supports the concept that PPAR and *ear 1* genes are derived from a common ancestral gene but does not provide any clues as to whether the presence of two introns in *ear 1* represents a loss or gain of an intron in the evolutionary divergence of *ear 1* and PPAR genes.

We have also determined the nucleotide sequences of the mPPAR $\gamma$ 1 and mPPAR $\gamma$ 2 promoter regions and confirmed that they can function as promoters (Figs. 2 and 3). mPPAR $\gamma$ 1 promoter contains a 550-bp (C+G)-rich sequence, which is lacking in the mPPAR $\gamma$ 2 promoter. This CpG island is relatively small in comparison with the typical 1- to 2-kb CpG island observed near the 5' end of genes (21). The CpG islands, though frequent in housekeeping genes, are also known to occur in some genes that show tissue-specific expression. Methylation of a CpG island inhibits expression of the associated gene. Thus, methylation of a CpG island could dictate the tissue specificity of mPPAR $\gamma$  gene expression. Methylation of this CpG island 5' to mPPAR $\gamma$ 1 may possibly account

for the expression of only mPPAR $\gamma$ 2 in adipocytes. A transfection assay confirmed the promoter activity of this mPPAR $\gamma$ 1 promoter region. A 930-bp fragment possessed a weak promoter activity, whereas inclusion of an additional 1.8-kb upstream sequence significantly enhanced the functional activity.

The nucleotide sequence of the proximal region of mPPAR $\gamma$ 2 reveals two C/EBP sites. The C/EBP family of transcription factors, which consists of four members, plays a role in adipocyte differentiation and initiation of 3T3-L1 adipocyte differentiation (11). It appears that mPPAR $\gamma$ 2 is expressed earlier than C/EBP $\alpha$  during the course of adipocyte differentiation (11). The C/EBP sites in the mPPAR $\gamma$ 2 promoter may at first bind to other members of C/EBP family and then switch to C/EBP $\alpha$  in the course of adipocyte differentiation. Our results demonstrate the functional activity of the  $\gamma$ 2 promoter (mPPAR $\gamma$ 2 promoter) by transfection assay. The reporter construct revealed promoter activity of this region in adipocyte cells. Additional studies are needed to dissect the role of the two promoters in the mPPAR $\gamma$  gene in determining the transcriptional regulation and tissue specificity of expression. It is important to note that mPPAR $\gamma$ 2 has recently been shown to stimulate adipogenesis in fibroblasts, suggesting that mPPAR $\gamma$ 2 regulates the development of adipose tissue (22). The existence of multiple PPAR protein isoforms increases the range of potential retinoid X receptor-PPAR heterodimerizations that can specifically transcribe the responsive genes essential not only for differentiation and development but also in xenobiotic-inducible pleiotropic responses.

This work is supported, in part, by National Institutes of Health Grant R37GM23750 (J.K.R.), R29 HG00037 (J.R.K.), the Department of Energy (DE-FG03-94ER-61402) to J.R.K., and The Joseph L. Mayberry, Sr., Endowment Fund (J.K.R.). J.R.K. holds the Brawerman Endowed Chair in Molecular Genetics.

- Reddy, J. K. & Mannaerts, G. P. (1994) *Annu. Rev. Nutr.* **14**, 343-370.
- Lazarow, P. B. & de Duve, C. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 2043-2046.
- Tolbert, N. E. (1981) *Annu. Rev. Biochem.* **50**, 133-157.
- Reddy, J. K. & Lalwani, N. D. (1983) *Crit. Rev. Toxicol.* **12**, 1-58.
- Reddy, J. K., Goel, S. K., Nemali, M. R., Carrino, J. J., Laffler, T. G., Reddy, M. K., Sperbeck, S. J., Osumi, T., Hashimoto, T., Lalwani, N. D. & Rao, M. K. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 1747-1751.
- Reddy, J. K., Azarnoff, D. L. & Hignite, C. E. (1980) *Nature (London)* **283**, 397-398.
- Issemann, I. & Green, S. (1990) *Nature (London)* **374**, 645-650.
- Dreyer, C., Krey, G., Keller, H., Givel, F., Helftenbein, G. & Wahli, W. (1992) *Cell* **68**, 879-887.
- Zhu, Y., Alvares, K., Huang, Q., Rao, M. S. & Reddy, J. K. (1993) *J. Biol. Chem.* **268**, 26817-26820.
- Kliwer, S. A., Forman, B. M., Blumberg, B., Ong, E. S., Borgmeyer, U., Mangelsdorf, D. J., Umesono, K. & Evans, R. M. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 7355-7359.
- Tontonoz, P., Hu, E., Graves, R. A., Budavari, A. I. & Spiegelman, B. M. (1994) *Genes Dev.* **8**, 1224-1234.
- Pierce, J. C., Sternberg, N. & Sauer, B. (1992) *Mamm. Genome* **3**, 550-558.
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K., eds. (1993) *Current Protocols in Molecular Biology* (Wiley, New York).
- Cheng, S., Fockler, C., Barnes, W. M. & Higuchi, R. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 5695-5699.
- Student, A. K., Hsu, R. Y. & Lane, M. D. (1980) *J. Biol. Chem.* **255**, 4745-4750.
- Korenberg, J. R. & Chen, X. N. (1995) *Cytogenet. Cell Genet.* **69**, 196-200, 1995.
- Ponglikitmongkol, M., Green, S. & Chambon, P. (1988) *EMBO J.* **7**, 3385-3388.
- Miyajima, N., Horiuchi, R., Shibuya, Y., Fukushima, S. I., Matsubara, K., Toyoshima, K. & Yamamoto, T. (1989) *Cell* **57**, 31-39.
- Laudet, V., Hanni, C., Coll, J., Catzeflis, F. & Stehelin, D. (1992) *EMBO J.* **11**, 1003-1013.
- Frey, G., Keller, H., Mahfoldi, A., Medin, J., Ozato, K., Dreyer, C. & Wahli, W. (1993) *J. Steroid Biochem. Mol. Biol.* **47**, 65-73.
- Bird, A. P. (1986) *Nature (London)* **321**, 209-213.
- Tontonoz, P., Hu, E. & Spiegelman, B. M. (1994) *Cell* **79**, 1147-1156.