Vol. 199, No. 1, 1994 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
February 28, 1994 Pages 255-263

STRUCTURE OF THE MOUSE PEROXISOME PROLIFERATOR ACTIVATED RECEPTOR α GENE[1]

K. L. Gearing,* A. Crickmore and J.-Å. Gustafsson **

Centre for Biotechnology and Department of Medical Nutrition, Karolinska Institute, Novum, Huddinge University Hospital, S-141 86 Huddinge, Sweden

Received January 10, 1994

We have isolated and characterised genomic clones spanning the mouse peroxisome proliferator activated receptor alpha gene (PPARα). The gene contains eight exons spanning at least 30 kilobases. The DNA binding domain of PPARa consists of two exons, each encoding one of the zinc fingers. Three exons encode the ligand binding domain and 3' untranslated regions of PPARα. There are two 5' untranslated exons and one exon encoding the N-terminal domain. The 5' flanking region is GC rich and contains several putative SP1 binding sites. The structure of the PPARα gene is unique but shares similar features with related members of the nuclear hormone receptor family.

The peroxisome proliferator activated receptor (PPAR) is a member of the steroid receptor gene superfamily. It mediates transcription of responsive genes which include those encoding peroxisomal enzymes and members of the cytochrome P450 family of drug metabolising enzymes (1-8).

The first PPAR to be cloned and characterised was obtained from mouse (9). It was found to resemble other members of the nuclear hormone receptor superfamily, especially in the region responsible for DNA binding. In particular, PPAR is most closely related to members of one of the three subfamilies of receptor genes that includes the thyroid hormone and retinoic acid receptors (10).

The PPAR was initially shown to be activated by compounds that cause peroxisome proliferation (9). More recently, fatty acids have been identified as receptor activators (11, 12) and non-metabolisable fatty acids were found to be potent activators (3, 13). Fatty acids have also been demonstrated to be the component of plasma which can activate PPAR,

[1]EMBL accession nos.: X75287, X75288, X75289, X75290, X75291, X75292, X75293, X75294.

*Current address: Glaxo Group Research, Greenford Road, Greenford, Middlesex UB6 0HE England.

** Corresponding author: Fax: 46-8-711 66 59.

0006-291X/94 $5.00

confirming their physiological relevance to this receptor (14). None of the receptor activators described above has yet been shown to be a ligand for the PPAR nor has it been demonstrated whether the receptor can be activated by other means, such as by covalent modification, rather than by binding of a ligand.

Studies of the promoters of several peroxisome proliferator responsive genes have shown that they contain specific binding sites for PPAR (1, 2, 6, 7, 15). These sites, termed peroxisome proliferator response elements (PPREs), resemble the recognition sequences of other nuclear receptors and are composed of direct repeats of a hexanucleotide sequence separated by a single nucleotide. Furthermore, binding of PPAR to a PPRE requires additional cofactors, such as the retinoid X receptor, in a similar fashion to its closest relatives within this receptor family (7, 12, 16, 17).

The homologues of mouse PPAR have been cloned from rat (11), Xenopus (18) and human (19), with two related but distinct PPARs also isolated from Xenopus (18). It is likely that other forms of PPAR exist in other species since another, more distantly related receptor (NUC 1), has been cloned from man and is reported to be expressed in both mice and rats (20). We have investigated the genomic organisation of one of the PPAR genes, the mouse PPARα gene. The structure of this gene and its 5' flanking region are discussed in relation to other nuclear receptor genes.

MATERIALS AND METHODS

Isolation and characterisation of mouse PPARa genomic clones

A mouse 129SV genomic library in lambda FIX (Stratagene) was screened with probes prepared from the rat PPAR cDNA (11). Probes labelled with ^{32}P-dCTP were prepared using a Megaprime DNA labelling system (Amersham) and hybridisations were performed under high stringency conditions. DNA was prepared from purified l clones using standard methods (21) and restriction enzyme mapping was performed using Not I, Pst I and Bgl II. Restriction fragments which hybridised to the rat cDNA were subcloned in pBluescript and were sequenced using PPAR specific primers. Alternatively, purified lambda DNA was sequenced directly. DNA sequencing was carried out using a Taq DyeDeoxy terminator cycle sequencing kit (Applied Biosystems) and samples were analysed on an Applied Biosystems Model 373A DNA sequencing system according to the manufacturers instructions. Oligonucleotide primers were synthesised using an ABI 308B DNA synthesiser.

RNase protection analysis

For mapping of the transcription start site, a Not I - Pvu II restriction fragment (nucleotides -949 - +106, fig. 2) encompassing a 70 nucleotide region homologous to the 5' end of the rat cDNA and approximately 1 kb of upstream sequence was subcloned in pBluescript. After linearisation with Not I, a 1.1 kb radiolabeled transcript was synthesised from this plasmid using T3 RNA polymerase and ^{32}P-UTP (Amersham). For analysis of differential splicing in exon 1, a 380 nucleotide Hae III fragment (beginning at nucleotide + 34, fig. 2) which encompasses the 3' ends of both exon 1a and 1b was used to prepare a probe. This probe contained approximately 100 nucleotides of sequence transcribed from the plasmid vector in addition to the PPAR genomic sequences. PolyA$^+$ RNA (1μg) isolated from adult male kidney or liver (Clontech) was precipitated together with 16 μg yeast tRNA and 10^5 cpm RNA probe. Control samples contained only tRNA. The RNAs were denatured at 85°C for 10 min and annealed for 16 h at 40°C in 80% formamide, 40mM PIPES (pH6.4), 1mM EDTA, 0.4M NaCl. The RNAs were digested with RNase (40μg/ml RNase A, 2μg/ml RNase T1 in 0.3M NaCl, 10mM Tris.HCl (pH 7.4), 5mM EDTA) for 45 min at 37°C, before

treatment with proteinase K, phenol extraction and precipitation with ethanol. The products were run on 5% denaturing polyacrylamide gels which were exposed to film overnight. End labelled DNA markers used to estimate the size of the products were φX174 digested with Hinf I and end labelled with ^{32}P-γATP (Amersham). The sizes of products were confirmed using RNA markers transcribed from the polylinker of pBluescript (data not shown).

RESULTS AND DISCUSSION

Structural organisation of the mouse PPARa gene

A l library containing mouse genomic DNA was screened with probes generated from the rat PPARa cDNA. Nine positive clones were isolated and of these, three distinct clones (λ2b, λ10 and λ18b) were chosen for further study. The exons present in these clones were mapped by hybridisation to the cDNA. There are eight exons in the mouse PPARα gene and it is at least 30 kb in length. Although the three clones characterised contain all of the PPAR exon sequences, they do not overlap each other. The two gaps in the gene occur within introns 2 and 6. Figure 1 shows restriction maps of clones λ2b, λ10 and λ18b and the position of the PPAR exons within them.

The nucleotide sequences of the PPAR exons and flanking introns were determined and these genomic sequences are shown in figure 2. Exon 3 contains the initiating methionine and encodes the first 69 amino acids. The third intron lies between the regions encoding the N-terminal domain and the DNA binding domain (DBD) of PPAR, while the fourth divides the DBD and is located between the two zinc fingers. The position of this intron varies between groups of receptor genes, but in PPAR it is in a similar position to introns in the thyroid hormone receptor (TR) and retinoic acid receptor (RAR) genes (22, 23). In PPAR, however, this intron occurs 1 nucleotide earlier. The position of the fifth intron separating the DBD from the putative ligand binding domain (LBD) is conserved amongst all nuclear receptor genes studied to date. Presumably this was the site of an intron in the primitive nuclear receptor gene since it is even conserved in COUP-TF, which has only two introns (24).

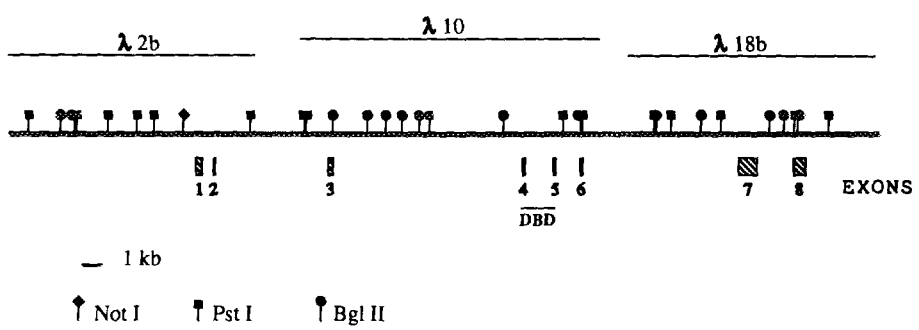

Figure 1.
Restriction map of lambda clones 2b, 10 and 18b. The sites for Not I, Pst I and Bgl II are shown together with the positions of PPAR exons 1a, 1b and 2-8. Introns 1-7 are estimated to be approximately 0.8 kb, at least 5 kb, 11 kb, 1.7 kb, 1.7 kb, at least 5 kb and 1.8 kb, respectively.

Figure 2.

Sequence of the mouse PPARα exons and flanking intron sequences. Intron sequences are shown in lower case letters. The major transcription start sites are underlined and are designated nucleotides +1 and +2. The 5' flanking sequence is shown beginning at nucleotide -951 and GC boxes are in bold type and underlined. The first A of the polyA tail of the cDNA (ref. 9) is also indicated.

Interestingly, the C-terminal domain of PPAR is contained within just three exons. The sixth intron is located at a position conserved between PPAR and TR. The seventh intron is in a position conserved between receptors and divides the region thought to mediate receptor dimerisation (25).

A single difference in the deduced amino acid sequence was found (at amino acid 75, fig. 2) between the mouse cDNA and genomic sequences. In the mouse cDNA there is an arginine at this position (9), whereas the genomic sequence and also the rat, Xenopus and human PPARα cDNAs all encode alanine at this position (11, 18, 19).

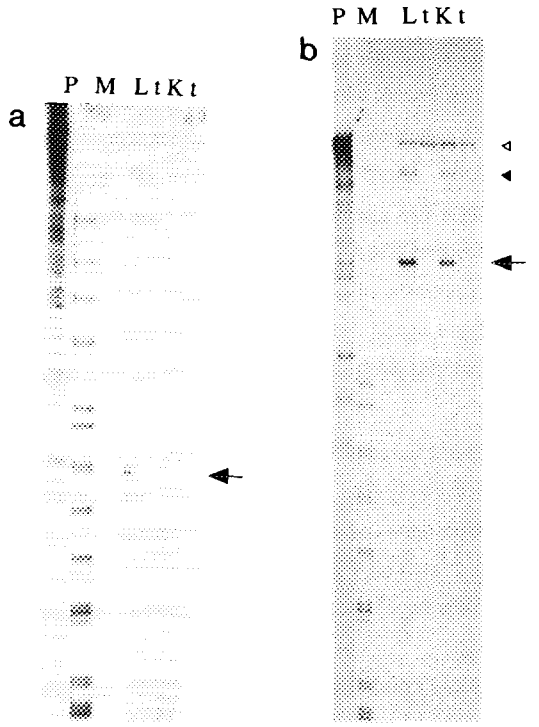

Figure 3.
RNase protection analysis of the mouse PPARα mRNA. L, liver RNA, K, kidney RNA, t, tRNA control, P, undigested RNA probe, M, DNA markers (40, 42, 48, 66, 82, 100, 118, 140, 151, 200, 249, 311, 413, 417, 427, 500, 553, 713 and 726 base pairs).
a) Mapping the transcription start site. The major protected RNA products of 105 and 106 nucleotides are indicated by an arrow.
b) Mapping introns in the 5' untranslated region. The protected products of 255 nucleotides (arrow) and 380 nucleotides (closed triangle) are indicated. Undigested probe is also marked (open triangle).

Analysis of the 3' end of the PPARα gene

There are no recognisable AATAAA elements in the 3' end of the cDNA or the sequences immediately 3' to this in the genomic DNA clones (see fig. 2). It is possible that the mPPAR cDNA (9) was produced by mispriming through an internal polyA tract in the mRNA (beginning at nucleotide +2523 in the genomic DNA sequence, fig. 2). This is quite likely since the cDNA, which was cloned using an oligo dT primer, is several kb shorter than the PPARα mRNA observed in Northern blots (S. Green and W. Wahli, personal communication). It is therefore probable that exon 8 extends much further in λ clone 18b, or that additional 3' non-coding exons exist further downstream. Further analysis of PPAR mRNA and cDNAs will be necessary to map the 3' end of this gene.

Mapping of the transcription start site and analysis of the 5' flanking region

The mouse genomic DNA clone λ2b contains a region homologous to the 5' end of the rat cDNA. Using complementary RNA probes generated from this genomic clone we have mapped the initiation site of the mouse PPARα mRNA by RNase protection (fig. 3a). Major products of 105 and 106 nucleotides were produced, indicating that the major initiation sites

are located on G and C residues (designated nucleotides +1 and +2 respectively, see fig. 2). In addition, there are at least two minor initiation sites apparent at nucleotides -4 and +5. In these assays, the PPAR mRNA levels were significantly higher in liver than in kidney.

The sequence in which transcription initiates is reminiscent of the start sites for many other receptors from this family. There are seven putative SP1 binding sites or GC boxes (26) in the 5' flanking sequence and one in the first exon (see fig. 2), and the entire region is extremely GC rich. No TATA or CCAAT elements are present in the 5' flanking region. The high GC content (71%) and presence of GC boxes is very common for promoters of receptor genes. These features have been noted in the promoters of the glucocorticoid, androgen, thyroid hormone, retinoic acid and progesterone receptors as well as in the NGFIB promoter region (27-35).

In the rat, PPAR expression is reported to be induced by treatment with peroxisome proliferators (5). It is therefore likely that the promoter for this receptor contains elements which resemble PPREs of other peroxisome proliferator responsive genes. However, we were not able to identify any putative PPREs in the first 1 kb of the 5' flanking sequence. Transfection studies using portions of the upstream sequence contained in λ2b will be necessary to elucidate this issue.

It is not known what factors control expression of PPAR, but it is possible that this gene is regulated by other nutritional or hormonal signals. A search for hormone response elements in the 5' flanking sequence reveals only one perfect hexanucleotide half site element (TGACCT; nucleotides -586 - -581, fig. 2), but an inverted, imperfect repeat spaced by 4 nucleotides is present (AGATCAGATTTCGCCT; nucleotides -804 - -789). There is also a complex set of 5 imperfect direct repeats starting at nucleotide -732, spaced by -1, +3, +5 and -1 nucleotides. The functional significance of these regions remains to be determined.

It is possible that species differences regarding response to peroxisome proliferators (for a review see 36) may be determined in part by the concentration of PPAR in the cell. It will be interesting to determine whether the promoters of mouse and human PPARα are regulated differently and whether the pattern and level of expression of this gene differs between rodents and man.

Analysis of the 5' untranslated region

The presence of introns in the 5' untranslated region of nuclear receptor genes is not uncommon; for example, the glucocorticoid receptor and retinoic acid receptor have separate 5' non- coding exons (32, 37-40). Since the sequence of the rat and mouse cDNAs diverge 5' to the border of exon 2, we investigated the organisation of this region of the gene in more detail. As can be seen in fig. 2, the sequence of the 5' end of the mouse cDNA is located within a predicted intron relative to the sequence homologous to the 5' end of the rat cDNA. We performed RNase protection experiments using a probe overlapping by 71 bp the one used to map the transcription start site (shown in fig. 3a) to distinguish between differential splicing of the same primary transcript or different transcription start sites (fig. 3b). A strong band of approximately 255 nucleotides was observed, corresponding to transcripts spliced at

the same position as the rat cDNA (intron 1a, fig. 2). A faint band of 332 nucleotides, corresponding to the position of the intron in the mouse cDNA (intron 1b, fig. 2), was only observed in long exposures of the film, indicating that this is a minor variant of the transcript. A rather strong band of approximately 380 nucleotides was also observed, indicating protection of the entire genomic DNA region of the probe, suggesting that further variants with splice sites downstream of the sequence given in fig. 2 also exist. There was no evidence seen for transcription start sites in addition to those mapped in fig. 3a.

In the case of the RAR genes, many cDNA isoforms are generated by differential splicing in the 5' untranslated region (32, 38-40). The first exon of PPARα is differentially spliced to give two distinct mRNAs which have been observed in cDNA clones (9, 11) and it is clear from our RNAse protection experiments that at least one further splicing variant exists. It will be interesting to see whether the expression of these different isoforms has any functional significance.

CONCLUSIONS

The position of introns within the PPARα gene adds support to the idea that this receptor is more closely related to TR and RAR than to other members of the nuclear receptor superfamily, although there are important differences which show that this gene structure is unique. The structure of the Xenopus PPARβ gene is similar to that of the mouse PPARα gene in the DNA binding and ligand binding domains (W. Wahli, personal communication). This suggests that PPAR gene structure is likely to be conserved between other subfamily members and between species.

Expression of PPAR is restricted to peroxisome proliferator responsive tissues; PPAR transcripts have been detected prenatally in liver and brown adipose tissue and post-natally also in kidney and heart (41). The clones characterised in this study provide useful material for the study of the control of PPAR expression, and will also be useful in future studies to disrupt PPAR by gene targeting aimed at obtaining information about the function of this protein in regulating lipid homeostasis.

ACKNOWLEDGMENTS

We would like to thank Zekiye Cansu, Helen Olsson and Camilla Eriksson for oligonucleotide synthesis and DNA sequencing services and Sofia Paulsson for technical help. We are also grateful to S. Green for discussion and W. Wahli for communicating results prior to publication. This work was supported by the Swedish Medical Research Council (grant no. 13X-2819) and the Swedish Cancer Society.

REFERENCES

1. Tugwood, J.D., Issemann, I., Anderson, R.G., Bundell, K.R., McPheat, W.L. and Green, S. (1992) *EMBO J.* **11**, 433-439.

2. Zhang, B., Marcus, S.L., Sajjadi, F.G., Alvares, K., Reddy, J.K., Subramani, S., Rachubinski, R.A. and Capone, J.P. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 7541-7545.
3. Issemann, I., Prince, R., Tugwood, J. and Green, S. (1992) *Biochem. Soc. Trans.* **20**, 824-827.
4. Muerhoff, A. S., Griffin, K. J. and Johnson, E. F. (1992) *J. Biol. Chem.* **267**, 19051-19053.
5. Gebel, T., Arand, M. and Oesch, F. (1992) *FEBS Letters* **309**, 37-40.
6. Marcus, S.L., Miyata, K.S., Zhang, B., Subramani, S., Rachubinski, R.A. and Capone, J.P. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 5723-5727.
7. Bardot, O., Aldridge, T.C., Latruffe, N. and Green, S. (1993) *Biochem. Biophys. Res. Commun.* **192**, 37-45.
8. Kaikaus, R.M., Chan, W.K., Lysenko, N., Ray, R., Ortiz de Montellano, P.R. and Bass, N.M. (1993) *J. Biol. Chem.* **268**, 9593-9603.
9. Issemann, I. and Green, S. (1990) *Nature* **347**, 645-650.
10. Laudet, V., Hänni, C., Coll, J., Catzeflis, F. and Stehelin, D. (1992) *EMBO J.* **11**, 1003-1013.
11. Göttlicher, M., Widmark, E., Li, Q. and Gustafsson, J.-Å. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 4653-4657.
12. Keller, H., Dreyer, C., Medin, J., Mahfoudi, A., Ozato, K. and Wahli, W. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 2160-2164.
13. Göttlicher, M., Demoz, A., Svensson, D., Tollet, P., Berge, R.K. and Gustafsson, J.-Å. (1993) *Biochem. Pharmacol.*, in press.
14. Banner, C. D., Göttlicher, M., Widmark, E., Sjövall, J., Rafter, J.J. and Gustafsson, J.-Å. (1993) *J. Lipid Research*, in press.
15. Boie, Y., Adam, M., Rushmore, T.H. and Kennedy, B.P. (1993) *J. Biol. Chem.* **268**, 5530-5534.
16. Kliewer, S.A., Umesono, K., Noonan, D.J., Heyman, R.A. and Evans, R.M. (1992) *Nature* **358**, 771-774.
17. Gearing, K.L., Göttlicher, M., Teboul, M., Widmark, E., and Gustafsson, J.-Å. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 1440-1444.
18. Dreyer, C., Krey, G., Keller, H., Givel, F., Helftenbein, G. and Wahli, W. (1992) *Cell* **68**, 879-887.
19. Sher, T., Yi, H.-F., McBride, W. and Gonzalez, F.J. (1993) *Biochemistry* **32**, 5598-5604.
20. Schmidt, A., Endo, N., Rutledge, S. J., Vogel, R., Shinar, D. and Rodan, G. A. (1992) *Mol. Endocrinol.* **6**, 1634-1641.
21. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor University Press, Cold Spring Harbor.
22. Laudet, V., Begue, A., Henry-Duthoit, C., Joubel, A., Martin, P., Stehelin, D. and Saule, S. (1991) *Nucleic Acids Res.* **19**, 1105-1112.
23. Lehmann, J.M., Hoffmann, B., and Pfahl, M. (1991) *Nucleic Acids Res.* **19**, 573-578.
24. Ritchie, H.H., Wang, L.-H., Tsai, S., O'Malley, B.W. and Tsai, M.-J. (1990)*Nucleic Acids Res.* **18**, 6857-6862.
25. Fawell, S.E., Less, J.A., White, R. and Parker, M.G. (1990) *Cell* **60**, 953-962.
26. Dynan, W.S. and Tjian, R. (1983) *Cell* **35**, 79-87.
27. Zong, J., Ashraf, J. and Thompson, E.B. (1990) *Mol. Cell Biol.* **10**, 5580-5585.
28. Encio, I.J. and Detera-Wadleigh, S.D. (1991) *J. Biol. Chem.* **266**, 7182-7188.
29. Baarends, W. M., Themmen, A.P., Blok, L.J., Mackenbach, P., Brinkmann, A.O., Meijer, D., Faber, P.W., Trapman, J. and Grootegoed, J.A. (1990) *Mol. Cell. Endoc.* **74**, 75-84.
30. Sakurai, A., Miyamoto, T. and DeGroot, L.J. (1992) *Biochem. Biophys. Res. Commun.* **185**, 78-84.
31. Laudet, V., Vanacker, J.M., Adelmant, G., Begue, A. and Stehelin, D. (1993) *Oncogene* **8**, 975-982.
32. Leroy, P., Krust, A., Zelent, A., Mendelsohn, C., Garnier, J.-M., Kastner, P., Dierich, A.and Chambon, P. (1991) *EMBO J.* **10**, 59-69.
33. Lehmann, J.M., Zhang, X.-K. and Pfahl, M. (1992) *Mol. Cell. Biol.* **12**, 2976-2985.
34. Huckaby, C.S., Conneely, O.M., Beattie, W.G., Dobson, A.D., Tsai, M.-J. and O'Malley, B. W. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 8380-8384.
35. Ryseck, R.-P., Macdonald-Bravo, H., Mattei, M.-G., Ruppert, S. and Bravo, R. (1989) *EMBO J.* **8**, 3327-3335.

36. Green, S. (1992) *Cancer Surveys* **14**, 221-232.
37. Strähle, U., Schmidt, A., Kelsey, G., Stewart, A. F., Cole., T.J., Schmid, W. and Schutz, G. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 6731-6735.
38. Krust, A., Kastner, P., Petkovich, M., Zelent, A. and Chambon, P. (1989)*Proc. Natl. Acad. Sci. USA* **86**, 5310-5314.
39. Kastner, P., Krust, A., Mendelsohn, C., Garnier, J.-M., Zelent, A. Leroy, P., Staub, A. and Chambon, P. (1990)*Proc. Natl. Acad. Sci. USA* **87**, 2700-2704.
40. Zelent, A., Mendelsohn, C., Kastner, P., Garnier, J.-M., Ruffenach, F., Leroy, P., and Chambon, P. (1991) *EMBO J.* **10**, 71-81.
41. Beck, F., Plummer, S., Senior, P.V., Byrne, S., Green, S. and Brammar, W.J. (1992) *Proc. R. Soc. Lond.* **247**, 83-87.