BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **226,** 261–267 (1996)
ARTICLE NO. 1343

Structural Characterization of Human *Ad4bp (SF-1)* Gene

Koichi Oba,* Toshihiko Yanase,*,[1] Masatoshi Nomura,* Ken-ichirou Morohashi,† Ryoichi Takayanagi,* and Hajime Nawata*

**Third Department of Internal Medicine and †Department of Molecular Biology, Graduate School of Medical Science, Faculty of Medicine, Kyushu University, Maidashi 3-1-1, Higashi-ku, Fukuoka 812, Japan*

Received July 19, 1996

Ad4BP (or steroidogenic factor 1, SF-1) has been implicated to be an essential transcriptional factor for steroidogenesis as well as for the development of the reproductive axis. We elucidated the structure of the human *Ad4BP* gene. The gene is about 30 kb long and is split into 7 exons including a non-coding exon 1. The deduced amino acid sequence of the human Ad4BP consists of 461 amino acid residues and was highly homologous to those of other mammalian species. © 1996 Academic Press, Inc.

Ad4BP (or steroidogenic factor 1:SF-1) was originally identified as a steroidogenic tissue-specific transcription factor which regulates the expression of steroidogenic P450 hydroxylases in adrenocortical cells (1-4). This factor is a mammalian homologue of *Drosophila* fushi tarazu factor (FTZ-F1) (5) and structually belongs to a member of the nuclear receptor superfamily (6). The Ad4BP binds to the sequences PyCAAGGPyC or PuPuAGGTCA (1,7) of various steroidogenic P450 genes (1, 7, 8) and the steroidogenic acute regulatory protein (*StAR*) gene (9), thus leading to the hypothesis that Ad4BP might be a shared transcriptional factor that determines the expression of steroidogenic genes. Indeed, the actual involvement of Ad4BP in both cAMP-responsive and the tissue-specific transcription of various steroidogenic genes has been reported (8). In addition, embryonal long-terminal repeat-binding protein (ELP), originally identified as a repressor of the retroviral gene expression in embryonal carcinoma cells (10) is produced from the same gene encoding Ad4BP by alternative promoter usage and splicing (4, 11, 12). However, the functional significance of ELP in mammarian species is unknown. Most importantly, the disruption of the mouse *Ftz-f1* has been shown to cause a complete lack of adrenal glands and gonads (13,14) as well as a selective loss of gonadotropin-specific markers such as LH-β, FSH-β, and GnRH receptor in the pituitary (15). Ad4BP is, thus, thought to be essential for the development of a functioning hypothalamus-pituitary-adrenal and gonadal axis. The structure of FTZ-F1 cDNA or *FTZ-F1* gene has been so far reported in xenopus (5), mouse (4,12), rat (11) and bovine (3). Although there has been only two reports showing the chromosomal location of human *Ad4BP* gene on 9q33 (16) and the partial amino acid sequence of human Ad4BP corresponding to the regions of exon 2 and 3 (14), the complete structure of human Ad4BP cDNA or *Ad4BP* gene has not yet been elucidated. The clarification of the human *Ad4BP* gene structure is thus expected to greatly increase our general understanding for the mechanism of sexual differentiation and related diseases. In the present study, we cloned a human *Ad4BP* gene and characterized its structure.

The nucleotide sequence data reported in this paper will appear in the DDJB, EMBL, and GenBank DNA databases with accession numbers, D84206, D84207, D84208, D84209, D84210, and D84211.

[1] To whom correspondence should be addressed at Third Department of Internal Medicine, Faculty of Medicine, Kyushu University, Fukuoka 812, Japan. Fax: (092) 633-3367.

0006-291X/96 $18.00

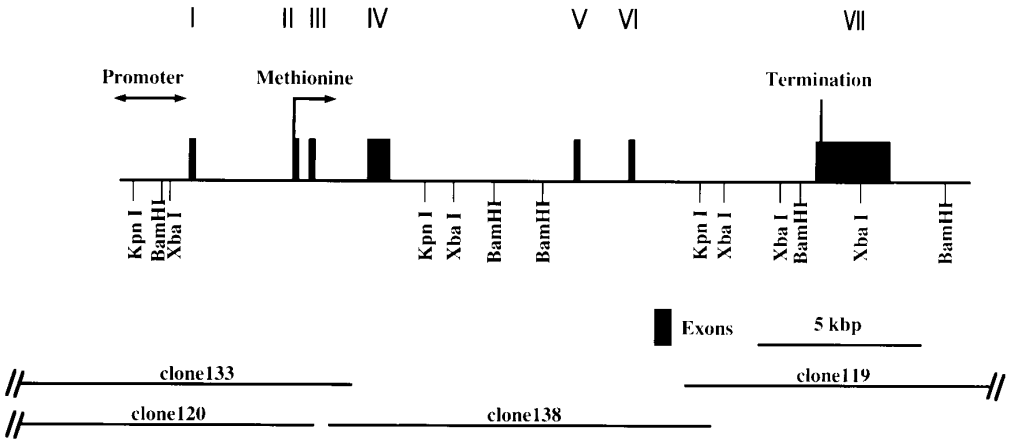

FIG. 1. A restriction enzyme map of the human *Ad4BP* gene. Four different clones, 133, 120, 138, and 119, overlapped each other and all contained whole *Ad4BP* gene. The closed boxes indicate the locations and sizes of the exons. Methionine in exon II indicate the translation initiation site, whereas Termination in exon VII indicates the translation termination site.

MATERIALS AND METHODS

Cloning of the human Ad4BP gene. Genomic DNAs from human peripheral lymphocytes were partially digested with Sau3A I and ligated into EMBL3 (Stratagene, CA, USA). This genomic library was screened by plaque hybridization with randomly labeled full length bovine Ad4BP cDNA fragment (3) and the 5′-RACE product containing non-coding exon 1 of the rat *Ad4 BP* gene (11). A restriction enzyme map of the positive clones was analyzed using several types of restriction endonuclease. The sequencing of genomic fragments subcloned into pUC19 was performed by the dideoxynucleotide method (17) using universal, riverse or appropriate internal primers.

Determination of transcriptional start site. To determine the transcriptional start site of the human *Ad4BP* gene, a RNase protection assay was carried out as previously described (18). Total and poly (A) RNA was prepared from the human adrenal gland obtained from a patient with renal carcinoma who underwent a hemilateral nephrectomy and from an adrenocortical adenoma obtained from a patient with Cushing's syndrome, using Isogen (Waco Junyaku Ltd, Osaka, Japan) and mRNA purification kit (Pharmacia Biotech, Tokyo, Japan), respectively. Figure 4 shows the schematic outline of the experiment. A 284 bp of Pst I fragment of clone 120 (see Fig.2) was at first subcloned into pUC19, after which 300 bp of the SmaI-HindIII fragment was subcloned into Bluescript SK(−). The plasmid was linealized with HindIII and then used as a template for sense [^{32}P]-labelled riboprobe synthesis using an *in vitro* transcription system (Clonetech, CA USA). A [^{32}P]-labelled RNA probe (359 nt) complementary to the 284 bp DNA template plus 59 bp of Bluescript and 16 bp of multicloning site of pUC19 was prepared using T3 RNA polymerase in the presence of [α^{32}P]-UTP (Amersham Corp). After digestion of the DNA template by RNase-free DNase I, the RNA probes were purified by CHROMA SPIN column-100 (Clonetech). A mixture of 10 μg of poly(A) RNA and the cRNA probe (4×10^5 cpm) was heated at 90 ℃ for 10 min and then annealed in 30 μl of hybridization buffer (18) at 45 ℃ for 12 h, after which 350 μl of a solution containing RNase A (50 U/ml) and RNase T1 (3 U/ml) (18) was added and incubated at 37 ℃ for 30 min. The reaction was stopped by the addition of 10 μl of 20% SDS and 10 μl of proteinase K (10 μg/ml) followed by incubation at 37 ℃ for 15 min. After phenol extraction and ethanol precipitation, the samples were subjected to electrophoresis on a denaturing polyacrylamide gel (5 %) and the results were visualized by autoradiography.

Southern blot analysis. Genomic DNA was extracted from the blood leukocytes of a normal individual as previously described (19). Ten μg of the genomic DNA was digested with several restriction enzymes, fractionated by electrophoresis in 0.8% agarose gel and then transferred to a nylon filter. Nick-translated EcoRI insert of bovine Ad4BP cDNA clone (3) as well as 0.28 kb of Pst I fragment of human *Ad4BP* gene clone, 120 which contained a non-coding exon 1 were used as a probe for hybridization. Both the hybridization and washing of the filters were carried out by conventional procedures (20).

RESULTS AND DISCUSSION

Four different positive clones measuring more than 10 kb length, No 120, 133, 138 and 119 were identified by plaque hybridization. A subsequent analysis to determine the structure

FIG. 2. The sequence of the human *Ad4BP* gene. The nucleotide sequence and the deduced amino acid sequence of the human *Ad4BP* gene are shown. The nucleotide sequence of the exons in the human *Ad4BP* gene are written in bold letters. The numbers at the left are the numbers of the exons. The positions indicated by wide arrows represent the initiation of methionine and termination codon. Regions I, II, and III are indicated with wide lines above the nucleotide letters. DNA elements are indicated by the wide underlines. The primer sequences used for sequencing as a size marker in the experiment of RNase protection assay are underlined with thin arrow.

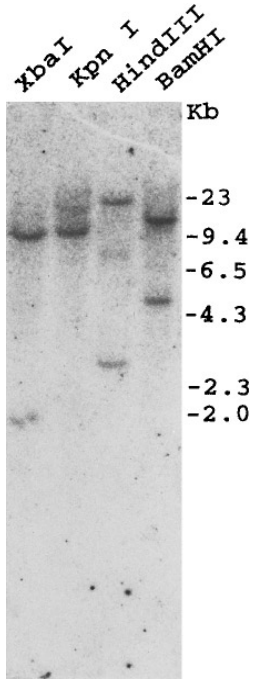

FIG. 3. A Southern blot analysis of genomic DNA digested with Xba I, KpnI, Hind III, or BamHI using the [^{32}P]-labelled bovine Ad4BP cDNA (3) and human *Ad4BP* exon 1 fragments as probes. Size markers are HindIII λ DNA fragments.

of each clone revealed these clones to overlap each other and together contain all 7 exons and 6 introns of the human *Ad4BP* gene, spanning about 30 kb (Fig.1). From the known structures of the *Ad4BP* genes and cDNAs of other mammarian species (3,4,11,12), the exon-intron organization of human *Ad4BP* gene were easily predicted. The presence of non-coding exon 1 and the structural validity of exon-intron junction was directly confirmed by the cDNA sequences generated by riverse transcriptase (RT)-PCR using total RNA from human adrenal gland (data not shown). The exonic sequences as well as the sequences at the exon/intron boundaries and at the 5′ flanking regions of the human *Ad4BP* gene are shown in Fig.2. The all exon-intron boundaries follows the ''GT-AG rule'' for the splice donor and acceptor. Southern blot analysis after the digestion of human genomic DNA with several restriction enzymes almost supported the gene organization of Fig.1 while the simple pattern of digestion also suggested the human *Ad4BP* gene to be a single copy gene (Fig.3).

The coding sequence of the human *Ad4BP* gene was highly conserved, since it was 88, 88, and 93% identical at nucleotide level and 93, 93, and 94% identical at amino acid level to those of the rat, mouse and bovine, respectively. The deduced amino acid sequence of the human Ad4BP consists of 461 amino acid residues and the predicted molecular weight was about 52,000 which was almost identical to the predicted protein size, 53 kDa by a Western blot analysis using human adrenocortical tissue specimens (21). A clarification of the primary structure of mouse, rat, bovine and human Ad4BP revealed two mammarian variants with 461 or 462 amino acids, which differ based on the absence or presence of a single glutamine residue at the amino acid position 206. The human and bovine Ad4BP proteins belong to a group without the glutamine residue, while the mouse and rat Ad4BP proteins belong to a group with the glutamine residue.

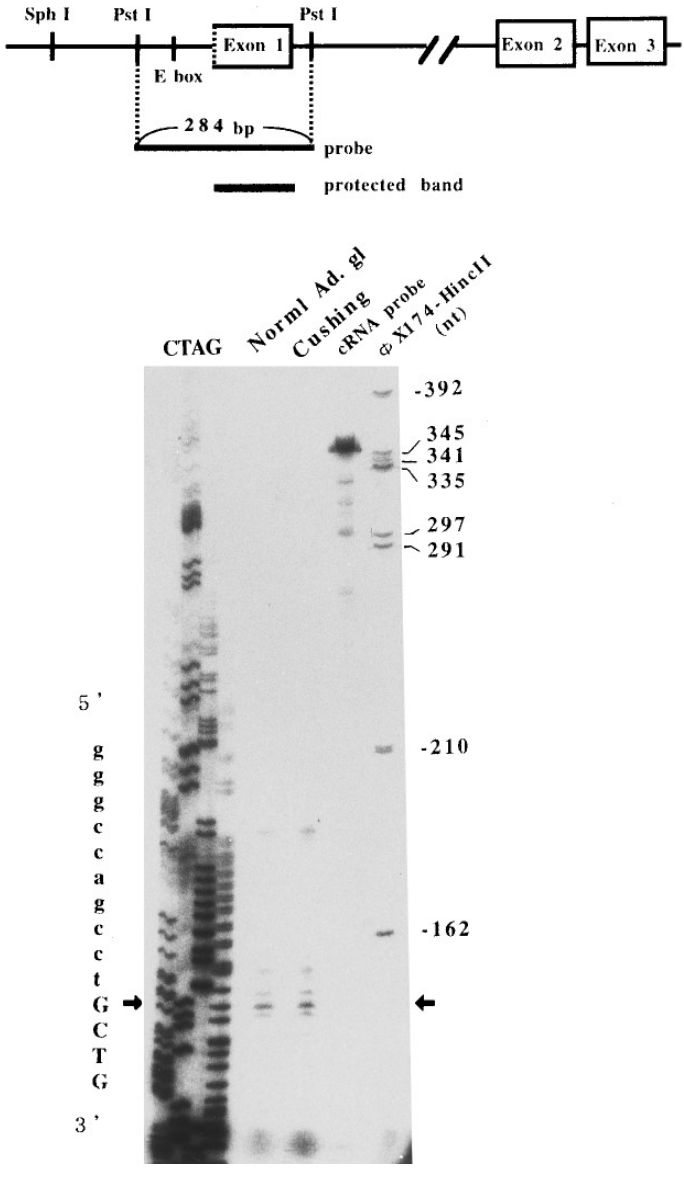

FIG. 4. Determination of the transcriptional initiation sites of the human *Ad4BP* gene by a RNase protection assay. After the [^{32}P]-labelled RNA probe was hybridized with each RNA, the mixture was digested with RNase as described in Materials and Methods. Normal ad.gl and Cushing indicate the poly(A) RNA from normal adrenal gland and adrenocortical adenoma from a patient with Cushing's syndrome, respectively. The sequencing ladder of 284 bp of the Pst I fragment generated by a phosphorylated primer (5′-CGAAGCGGAAGCAGCGCTCT-3′:P1R)(see Fig.2) was used as a size marker. The [^{32}P]-endlabelled Φ X174-Hinc II fragment was also used as a size marker.

By using a RNase protection assay, the major transcription start site of the human *Ad4BP* gene in both the normal human adrenal gland and the adrenocortical adenoma from a patient with Cushing's syndrome was mapped around the guanine nucleotide (Fig.4). The start site was 5 bases downstream to that of the rat (11) and a few bases upstream to that of the mouse (12). The promoter sequence of human *Ad4BP*, upstream of non-coding exon 1 was also highly homologous to those of the rat and the mouse, since about 85 % interspesies homology was

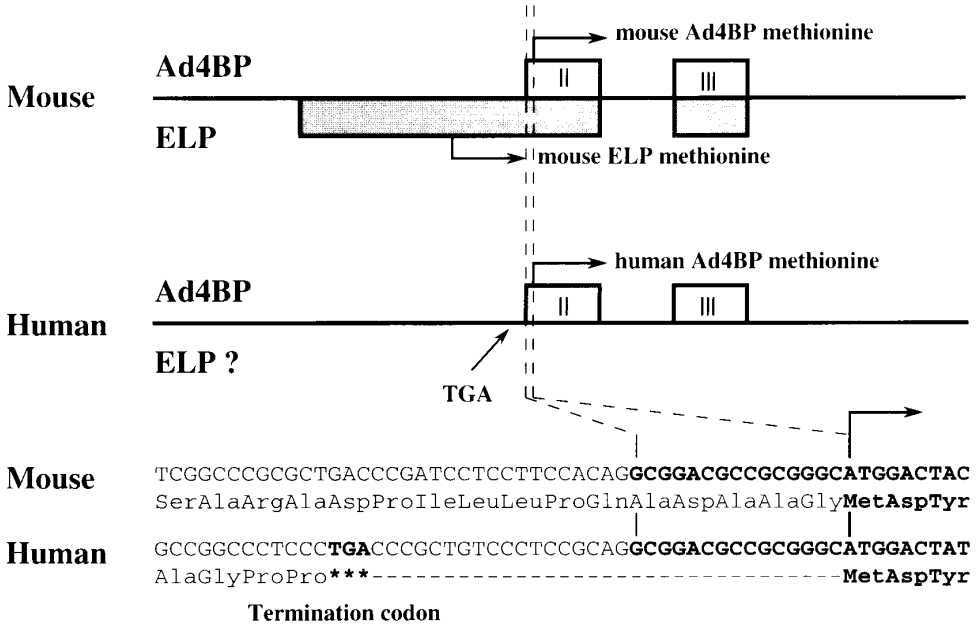

FIG. 5. Genetic difference of the nucleotide sequences around the translation initiation site between the human and mouse *ELP* genes.

observed in the region of −1 to −250 bp. Most of the consensus sequences for the binding sites of such transcriptional factor as GATA-1, E box, AP-2 and the CCAAT box were preserved (Fig.2), suggesting a shared, and essential regulatory mechanism in the expression of this gene beyond species. It is especially noteworthy that the E box sequence (CACGTG), and its binding protein has been reported to be essential for the expression of rat *Ad4BP* gene (11). The deletion analysis is currently under investigation to define the functional significance of the 5′-flanking region of the human *Ad4BP* gene.

In contrast, a predicted structure of embryonal long-terminal repeat-binding protein (ELP) in humans was rather different from those in other species. It has been reported that original ELP in mouse was transcribed within the *Ad4BP* first intron and 77 more amino acid residues were translated preceeding the Ad4BP initiation methionine (11). However, the analysis of the human *Ad4BP* gene in the present study revealed an in frame stop codon, TGA 36-bp before the first ATG of Ad4BP (Fig.5). Therefore, ELP1 (the original ELP isolate) and another isoform, ELP2 in mice (12) are unlikely to be present in humans. In addition, even the nucleotide sequence of *Ad4BP* intron 5 in human was quite different from those in the mouse which was reported to be transcribed to the carboxy-terminal sequence of ELP. A similar kind of structural alteration of ELP transcripts by premature termination due to a 22 nucleotide deletion has also been demonstrated in the rat *Ad4BP* gene (11). Although we can not rule out the possibility that other isoforms of ELP transcripts are produced by alternative promoter usage and splicing of intron 1 in human, the big structural divergence of *ELP* gene among species may thus suggest that ELP demonstrates less physiological significance.

The*Ftz-f1*-disrupted mouse was devoid of its normal adrenal glands and gonads (13,14) as well as a selective loss of gonadotropin-specific markers in the pituitary (15). Recently, a new nuclear hormone receptor gene, *DAX-1* was isolated from the Xp21 region (22) and was shown to be fatally mutated in patients with X-linked congenital adrenal hypoplasia and hypogonadotropic hypogonadism (23, 24). However, the molecular basis of exceptional cases

with congenital adrenal hypoplasia and hypogonadotropic hypogonadism, such as female patients and patients with autosomal recessive inheritance remains unclear (24). The*Ftz-f1*-disrupted mouse phenotype is therefore considered to be closely similar to that observed in the patients with *DAX-1* abnormalities. The human *Ad4BP* gene could thus be a candidate gene underlying such exceptional cases. The elucidation of the structure of the human *Ad4BP* gene has now enabled us to examine the molecular basis of such patients and such investigations are currently underway in our laboratory.

ACKNOWLEDGMENTS

We are grateful to Dr. Y. Nishi for the technical assistance. This work was supported in part by a Grant-in-Aid for General Scientific Research from the Ministry of Education, Science and Culture, Japan (No. 06671037).

REFERENCES

1. Morohashi, K., Honda, S., Inomata, Y., Hanada, H., and Omura, T. (1992) *J. Biol. Chem.* **267,** 17913–17919.
2. Lala, D. S., Rice, D. A., and Parker, K. L. (1992) *Mol. Endocrinol.* **6,** 1249–1258.
3. Honda, S., Morohashi, K., Nomura, M., Takeya, H., Kitajima, M., and Omura, T. (1993) *J. Biol. Chem.* **268,** 7494–7502, 7494–7502.
4. Ikeda, Y., Lala, D. S., Luo, X., Kim, E., Moisan, M-P., and Parker, K. L. (1993) *Mol. Endocrinol.* **7,** 852–860.
5. Lavorgna, G., Ueda, H., Clos, J., and Wu, C. (1991) *Science* **252,** 848–851.
6. Evans, R. M. (1988) *Science* **240,** 889–895.
7. Rice, D. A., Mouw, A. R., Bogerd, A. M., and Parker, K. L. (1991) *Mol. Endocrinol.* 1552–1561.
8. Parker, K. L., and Schimmer, B. P. (1993) *Trends. Endocrinol. Metab.* **4,** 46–50.
9. Clark, B. J., Soo, S-C., Caron, K. M., Ikeda, Y., Parker, K. L., and Stocco, D. (1995) *Mol. Endocrinol.* **9,** 1346–1355.
10. Tsukiyama, T., Ueda, H., Hirose, S., and Niwa, O. (1992) *Mol. Cell. Biol.* **12,** 1286–1291.
11. Nomura, M., Bartsch, S., Nawata, H., Omura, T., and Morohashi, K. (1995) *J. Biol. Chem.* **270,** 7453–7461.
12. Ninomiya, Y., Okada, M., Kotomura, N., Suzuki, K., Tsukiyama, T., and Niwa, O. (1995) *J. Biochem.* **118,** 380–389.
13. Luo, X., Ikeda, Y., and Parker, K. L. (1994) *Cell* **77,** 481–490.
14. Sadovsky, Y., Crawford, P. A., Woodson, K. G., Poish, J. F., Clements, M. A., Tourtellotte, L. M., Simburger, K., and Milbrandt, J. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 10939–10943.
15. Ingraham, H. A., Lala, D. S., Ikeda, Y., Luo, X., Shen, W-H., Nachtigal, M. W., Abbud, R., Nilson, J. H., and Parker, K. L. (1994) *Genes & Development* **8,** 2302–2312.
16. Taketo, M., Parker, K. L., Howard, T. A., Tsukiyama, T., Wong, M., Niwa, O., Morton, C. C., Miron, P. M., and Seldin, M. F. (1995) *Genomics* **25,** 565–567.
17. Sanger, F., Nicklen, S., and Coulson, A. R. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 7652–7656.
18. Yanase, T., Kagimoto, M., Suzuki, S., Hashiba, K., Simpson, E. R., and Waterman, M. R. (1989) *J. Biol. Chem.* **264,** 18076–18082.
19. Bell, G. I., Karam, J. H., and Rutter, W. J. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 5759–5763.
20. Sambrook, J. F., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
21. Sasano, H., Shizawa, S., Suzuki, T., Takayama, K., Fukaya, T., Morohashi, K., and Nagura, H. (1995) *J. Clin. Endocrinol. Metab.* **80,** 2378–2380.
22. Zanaria, E., Muscatelli, F., Bardoni, B., Storm, T. M., Guioli, S., Guo, W., Lalli, E., Moser, C., Walker, A. P., McCabe, E. R. B., Meitinger, T., Monaco, A. P., Sassone-Corsi, P., and Camerino, G. (1994) *Nature* **372,** 635–641.
23. Muscatelli, F., Storm, T. M., Walker, A. P., Zanaria, E., Recan, D., Meindl, A., Bardoni, B., Guioli, S., Zehetner, G., Rabl, W., Schwarz, H. P., Kaplan, J-C., and Camerino, G. (1994) *Nature* **372,** 672–676.
24. Yanase, T., Takayanagi, R., Oba, K., Nishi, Y., Ohe, K., and Nawata, H. (1996) *J. Clin. Endocrinol. Metab.* **81,** 530–535.