Biochemical and Biophysical Research Communications **278,** 63–71 (2000)
doi:10.1006/bbrc.2000.3774, available online at http://www.idealibrary.com on IDEAL®

Identification of Novel First Exons in *Ad4BP/SF-1* (*NR5A1*) Gene and Their Tissue- and Species-Specific Usage

Rina Kimura,*,† Hironori Yoshii,* Masatoshi Nomura,‡ Naoe Kotomura,§,¶ Tokuo Mukai,‡,§,¶ Satoru Ishihara,§,¶,‖ Koichi Ohba,‡ Toshihiko Yanase,‡ Osamu Gotoh,** Hajime Nawata,‡,¶ and Ken-ichirou Morohashi§,¶,‖

**Department of Molecular Biology and ‡Department of Medicine and Bioregulatory Science, Graduate School of Medical Science, Kyushu University, Maidashi, Fukuoka 812-8582, Japan; †Liver Cancer Research Division, Kurume University Research Center for Innovative Cancer Therapy, Asahi-machi, Kurume 830-0011, Japan; §Division of Cell Differentiation, Department of Developmental Biology, National Institute for Basic Biology, Myodaiji-cho, Okazaki 444-8585, Japan; ¶CREST, JST; ‖Department of Molecular Biomechanics, School of Life Science, Graduate University for Advanced Studies, Myodaiji-cho, Okazaki 444-8585, Japan; and **Department of Biochemistry, Saitama Cancer Research Center Institute, Ina-machi, Saitama 362-0806, Japan*

Received September 30, 2000

It has been demonstrated that the mammalian *Ad4BP/SF-1* (*NR5A1*) gene is regulated precisely in sex, tissue, and developmental stage specific manners. To clarify the complex transcriptional regulation, we investigated in the present study whether the gene transcription is regulated by multiple promoters accompanied by noncoding first exons. Novel first exons (Io and Ig) were identified downstream of the already identified exon Ia. Nucleotide sequences revealed that Ia and Ig exons were well conserved, whereas Io exon was less conserved among the mouse, rat, and human genes. Interestingly, the splice donor of the mouse and human Io and human Ig exons do not satisfy the consensus sequence. Transcripts containing Ia, Io, and Ig were detected in all rat tissues examined, while the transcript containing Io was undetectable in the corresponding tissues of mice. The lack of exon Io usage in the mouse was confirmed by transient transfection assays with cultured cells. Quantitative RT-PCR analysis revealed that the transcript containing Ig exon was the main product in the pituitary but significantly less in the spleen, suggesting that the regulation of *Ad4BP/SF-1* gene transcription in the pituitary and spleen is distinct from that of other tissues. The above findings, together with the structural abnormality at the splice donor site, suggest that acquisition of the multiple first exons enables the *Ad4BP/SF-1* gene to be regulated differentially in different animal species and in different tissues in the same animal. © 2000 Academic Press

Key Words: **Ad4BP/SF-1; nuclear receptor; transcription factor; steroidogenic tissue; tissue specificity; multiple exon.**

Investigation of transcription factors that regulate the steroidogenic *CYP* (*P450*) genes identified Ad4BP/SF-1 encoded by the *NR5A1* (1) gene as a steroidogenic tissue (adrenal cortex and gonads) specific transcription factor. Extensive analyses of the functions of Ad4BP/SF-1 have clarified that the transcription factor governs the expression of several genes necessary for the functions of the steroidogenic tissues such as the *steroidogenic CYP*, *3β-HSD*, *StAR*, *MIS*, *Dax-1*, etc. (2, 3). In addition to these tissues, the expression of Ad4BP/SF-1 has been confirmed in the pituitary gonadotroph, ventromedial hypothalamic nucleus (4–6) and recently in the spleen (7, 8). Although the functional significance of Ad4BP/SF-1 in the latter two tissues remain to be clarified, that in the gonadotroph is known to regulate gonadotropin, *LH* and *FSH*, gene transcription (9–11). Another prominent function of Ad4BP/SF-1 was clearly demonstrated by the phenotype of the gene disrupted mice that suffer from simultaneous agenesis of the adrenal gland and gonads from the early stages of tissue differentiation. These mice also display structural and functional abnormalities in the pituitary, ventromedial hypothalamic nucleus, and spleen (6, 7, 12, 13). Thus, there is a general agreement that *Ad4BP/SF-1* is one of the genes essential for the proper development of those tissues. Evidence for such crucial functions of Ad4BP/SF-1 was provided in part by the finding that the expression of Ad4BP/SF-1 was detectable in the primordial cells of the steroidogenic tissues (14–16). Although genes downstream of Ad4BP/SF-1 in the primordial cells have not yet been identified, it is likely that the transcription factor is implicated in regulation of a set of genes essential for tissue differentiation. In this regard, it is noteworthy

0006-291X/00 $35.00

that Ad4BP/SF-1 regulates critically, both *in vitro* and *in vivo*, the expression of the *Dax-1* gene (*NR0B1*) (17–19), which was identified as the responsible gene for adrenal hypoplasia congenita and hypogonadotropic hypogonadism (20, 21). The expression profile of Ad4BP/SF-1 is also characterized by its sexual dimorphism. Namely, higher amounts of Ad4BP/SF-1 were present in the fetal testes than in the corresponding stages of the ovaries after gonad sex differentiation, while such sexually dimorphic expression was undetectable in the adrenal cortex and the sexually indifferent gonads (14, 15).

These observations concerning the function and expression of Ad4BP/SF-1 strongly suggested that understanding the mechanisms underlying the spacio-temporal and sexually dimorphic expression is the next issue to be addressed. Studies from this aspect demonstrated that an E-box and CCAAT box (22–25) at the 5′ upstream of the *Ad4BP/SF-1* gene is functional as a crucial, probably basic, element for the transcription. To our knowledge, however, there are no studies that have previously elucidated the mechanism of the spacio-temporal and sexually dimorphic expression. In the present study, we demonstrated the presence of novel first exons in the *Ad4BP/SF-1* gene and provided evidences that the first exons are used in species-specific and tissue-specific manners.

MATERIALS AND METHODS

5′-RACE. Various adult and fetal tissues including the adrenal glands, testes, ovaries, hypothalamuses, pituitaries, spleens, and livers were isolated from rats and mice. Total RNAs were prepared from the tissues by the acid-guanidium thiocyanate-phenol-chloroform method (26). A proportion of the RNAs prepared from the rat fetal tissues were subjected to 5′-RACE using 5′-AmpliFINDER RACE Kit (Clontech, Palo Alto, CA). Primer, AdEL-1R2, corresponding to a nucleotide sequence in the third exon of the rat *Ad4BP/SF-1* gene was used to prime a reverse transcriptase reaction. After amplification of the cDNA using the method recommended by the supplier, the DNA was cloned into pUC19 for nucleotide sequencing.

RT-PCR. Synthesis of the first stranded cDNAs was performed as described previously (27). For amplification of the transcript containing exon Ia of the rat *Ad4BP/SF-1* gene, 25 cycles of PCR (94°C, 60°C, and 72°C each for 30 s) were performed with a set of primers, RAP1 and RR3. For amplification of the transcripts containing exon Io and Ig, 35 cycles of PCR (94°C, 60°C, and 72°C each for 30 s) were performed with sets of primers, ROP1 and RR3, and RGP1 and RR3, respectively. For the mouse gene transcript containing exon Ia, 25 cycles of PCR (94°C, 62°C, and 72°C each for 30 s) were performed with primers, MAP1 and MR22. For the transcripts containing exon Ig, 35 cycles of PCR (94°C, 70°C, and 72°C each for 30 s) were performed with primers, MGP1 and MR22. For transcripts containing exon Io, 35 cycles of PCR (94°C, 66°C, and 72°C each for 30 s) were performed with primers, MOP1 and MR22. The PCR products were electrophoresed, followed by Southern blot analyses probed with digoxygenin (DIG) (Boehringer Mannheim, Mannheim, Germany) labeled oligonucleotides; RMAP1 for the rat and mouse exons Ia, RMGP1 for the rat and mouse exons Ig, RGP2 for the rat exon Io, and MOP2 for the mouse exon Io. As a control, β-actin transcript was amplified with a set of primers, β A1 and β A1R, by 25 cycles of PCR (94°C, 60°C, and 72°C each for 30 s). The amplified DNA was analyzed by probing with DIG-labeled β A2. The sequences of the primers and probes are described below.

Quantitative analyses by RT-PCR. For quantitative RT-PCR, standard DNA fragments containing exon Ia and Ig of the mouse *Ad4BP/SF-1* were prepared as follows. DNA fragments containing exon Ia and Ig were amplified by PCR with MAP2 corresponding to a sequence in exon Ia and MR32 corresponding to a sequence in the third exon, and with MGP4 corresponding to a sequence in the exon Ig and MR32, respectively. After the PCR products were cloned into pGEM-T (Promega, Madison, WI), the fragments were recovered from agarose gel. The amounts of DNA fragments were photometrically determined. Serial dilutions ranging from 10^3 to 10^8 molecules in reaction mixtures were prepared as the standards for quantitative PCR.

One microgram of the RNA was reverse-transcribed into cDNA with a primer, MR33, corresponding to a sequence in the third exon, and Superscript II RNaseH(−) reverse transcriptase (Life Technologies, Rockville, MD) in a final volume of 20 μl according to the protocol provided by the supplier. The first-stranded cDNA was diluted with 80 μl of 10 mM Tris–HCl (pH 7.5) containing 1 mM EDTA, and then used as templates for the following quantitative PCR. PCR amplification was performed using LightCycler (Roche Diagnostics, Mannheim, Germany). In this assay, the amounts of PCR products are monitored at every cycle of the reaction by detecting increasing fluorescence. The reaction mixture consisted of certain amount of the cDNA, 0.5 μmol of a set of primers (MR22 and MAP1, and MR22 and MGP1 for the transcript carrying exon Ia and exon Ig, respectively), 1× SYBR Green I master mix (Roche Molecular Biochemicals) containing dNTP, Taq DNA polymerase, 3 mM $MgCl_2$, and TaqStart antibody (Clontech). All amplification reactions were performed under identical conditions (95°C for 0 s and 68°C for 20 s) after denaturation at 95°C for 2 min. After each elongation at 68°C, fluorescence of SYBR Green I was measured at 88°C for 3 s. Purity of the products was examined by generating a melting curve according to the protocol recommended by the manufacture and by agarose gel electrophoresis. As standards, a dilution series of cDNA fragments containing Ia, Io, or Ig was used in each PCR run. Quantitative analyses were performed with the LightCycler quantification software v3.0 according to instructions provided by the manufacturer. RNA extraction, cDNA preparation and quantitative PCR were all performed in triplicate.

Transfection of Ad4BP/SF-1 gene and analysis of product. The 3.9-kb-long segment upstream from the second exon of the mouse *Ad4BP/SF-1* gene was ligated into p00CAT (28) to produce mAd4ECAT. rAd4ECAT2.0K contains an upstream region from the second exon of the rat gene as described (22). These CAT reporter gene constructs were transfected into mouse Y-1 adrenocortical tumor cells (29) to investigate which of the first exons is transcribed. RT-PCR using an RNA prepared from cells transfected with rAd4ECAT2.0K was performed with sets of primers, RAP1 and CAT1, ROP1 and CAT1, and RGP1 and CAT1, to amplify the rat reporter gene transcript containing exon Ia, Io, and Ig, respectively. For analyses of the transcripts from mAd4ECAT, RT-PCR was performed with sets of primers, MAP1 and CAT1, MOP1 and CAT1, and MGP1 and CAT1, to amplify the Ia, Io, and Ig exons of the mouse reporter gene products, respectively. The mRNAs containing exon Ia, Io, and Ig transcribed from the endogenous *Ad4BP/SF-1* gene of mouse Y-1 cells were amplified using three sets of primers, MAP1 and MR22, MOP1 and MR22, and MGP1 and MR22, respectively. Thirty-five cycles of PCR (94°C, 60°C, and 72°C each for 30 s) were performed with all the PCRs. The resultant DNAs were subjected to Southern blot analyses probed with a DIG-labeled CAT3 oligonucleotide for detection of transcripts from the CAT reporter gene constructs while probed with a DIG-labeled MR2S oligonucleotide for detection of the transcript from the endogenous *Ad4BP/SF-1* gene.

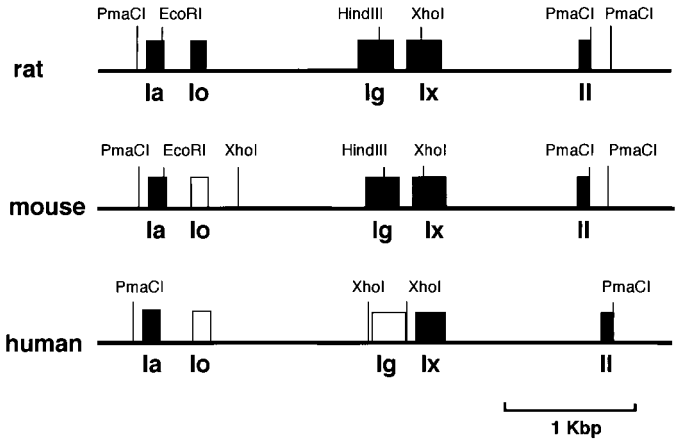

FIG. 1. Structures of the rat, mouse, and human *Ad4BP/SF-1* genes. Genomic organization covering the first (Ia, Io, Ig, and Ix) to the second exons (II) of the *Ad4BP/SF-1* gene of rat, mouse, and human is shown. Boxes indicate the locations and sizes of exons. As described in the text, the mouse Io and human Io and Ig exons shown as open boxes are unlikely active exons because of the disrupted splice donor sequences at their 3′ termini.

Probes and primers. All oligonucleotides used for primers and probes are as follows: RAP1, AGTGTCCACCCTTATCCG; ROP1, CTGAGAGTATGGTGTTCCGTG; RGP1, TTGCCCACCACCCCACTCATG; RGP2, CAGAGGCAAGCCAGACACTC; RR3, TCTTGTCGATTTTGCAGCTC; RMAP1, CACCCTTATCCGGCTGAGAAT; RMGP1, GCTTCGGAG GATCGAGCTG; MAP1, CCGCTGCTGGGTGAAGAAGTT; MAP2, CCCGCCCGCTGCTGGGTGAAG; MOP1, GGGTGGTTGGTGGTGTCTTAG; MOP2, AGGGTGTGGTGTTCTGTGTC; MGP1, AGGCCTCGGCACCCCTAAACTG; MGP4, TCCAGTTTTTCCTTGCTCACC; MR22, CCCGTAGTGGTAGCCCGACAC; MR32, GCCACCGTCAGCACTTCTGA; MR33, CGGTGCACGTGTAATGCTTGT; MR2S, TCCAGGTCCTCGTCGTAC; βA1, GCTGTATTCCCCTCCATCGTG; βA1R, CGGTTGGCCTTAGGGTTCAGG; βA2, TCTACAATGAGCTGCGTGTG; CAT1, CATATCACCAGCTCACCGTC; CAT III, TAGCTCCTGAAAATCTCGCC.

RESULTS

Identification of Novel First Exons in Ad4BP/SF-1 Gene

We have previously demonstrated that the rat *Ad4BP/SF-1* gene has a noncoding first exon (Ia) at approximately 3 kb upstream from the second exon where translation starts (22). In addition, embryonal long-terminal repeat binding protein (ELP), which is a product from the same *Ad4BP/SF-1* gene, was transcribed from a distinct start site directly upstream from the second exon (Fig. 1) (30). Further studies by Ninomiya *et al.* (31) clarified the presence of another first exon (referred to as Ix in Fig. 1) upstream from the second exon. Considering complex transcription start sites are present in the gene, we first investigated whether the gene has first exons other than those identified so far. DNA fragments obtained by 5′-RACE with the RNAs prepared from adrenal glands, testes, and ovaries of rat fetuses were subjected to nucleotide sequence analyses. Although the majority of DNA fragments contained exon Ia, two novel nucleotide sequences were found with certain frequency. The newly identified first exons were tentatively designated Io and Ig. Thus, together with the first exons identified by the previous studies, the *Ad4BP/SF-1* gene has at least five distinct first exons, Ia, Io, Ig, Ix, and directly upstream from the second exon.

To locate these exons and compare their nucleotide sequences among animal species, we determined the sequences upstream from the second exon of the rat, mouse, and human *Ad4BP/SF-1* genes (accession numbers are AB009575 for the rat, AB009576 for the mouse, and AB009577 for the human gene). Based on the homology of the nucleotide sequences, the locations of these first exons of the mouse and human were predicted from the rat gene (Fig. 2). As a structural feature common to the three animal species, G/C rich regions, where CpG dinucleotide appeared at a high frequency, were observed at a region covering from exon Ig to the second exons (data not shown). With respect to the sequence homology among the animal species, the second exons encoding amino acids show high homology, as expected, while noncoding exons Ia, Ig, and Ix are moderately conserved. Exon Io, however, shows less homology compared with the other first exons (Table I). A marked difference is noticed at the splice donor sequences. Although the splice donor consensus "GTPuPuPu" (32) is completely satisfied in the case of exon Ia (GTAAG for the three animals) and probably satisfied in Ix (GTGCG for the three animals) of all these animal genes, the corresponding regions of the other first exons do not always satisfy the consensus. Namely those of the human and mouse exons Io (GGCTC for the human and GGAGG for the mouse) and the human exon Ig (GGATG) have sequences distinct from the consensus (Fig. 2).

First Exon Usage in Cultured Cells

Considering that abnormal splice donor sequences occurred at exons Io and Ig, the next question was whether these exons of the mouse and human genes are functional similar to the corresponding exons of the rat gene. To address this issue, we transfected the CAT reporter gene constructs carrying all the first exons of the rat (Ad4ECAT2.0K) and mouse (mAd4ECAT) *Ad4BP/SF-1* genes (Fig. 3A) into mouse adrenocortical Y-1 cells, in which the endogenous *Ad4BP/SF-1* gene is kept active for transcription (33). After transfection of the CAT reporter gene constructs, RNAs prepared from Y-1 cells were amplified by RT-PCR with primers for each first exon and the CAT gene. Subsequently, the products were subjected to blotting analyses probed with oligonucleotides, CAT3 for the transgenes and MR2S for the endogenous gene (Fig. 3A). The endogenous *Ad4BP/SF-1* gene transcript in Y-1 cells was ex-

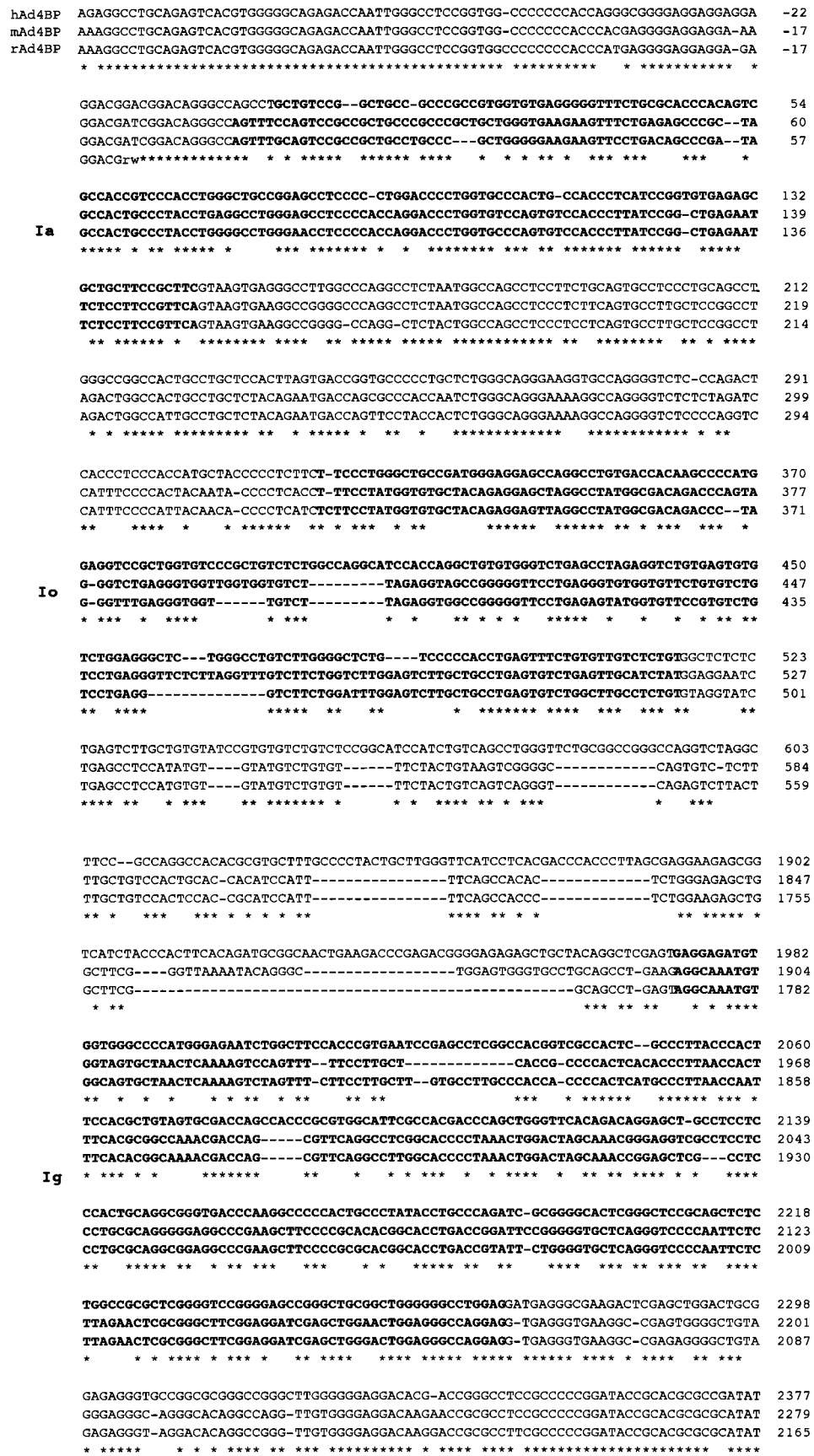

FIG. 2. Alignment of the nucleotide sequences of exons Ia, Io (upper), and Ig (lower) of the human (top), mouse (middle), and rat (bottom) Ad4BP/SF-1 genes with their flanking regions. These sequences are aligned to show maximum homology. Asterisks indicate nucleotides identical among the three animals. Exons are shown in bold letters and the numbers at the left are relative to the transcriptional start site of exon Ia.

TABLE I

Sequence Homology of the Multiple First Exons of the Ad4BP/SF-1 Gene between Human and Mouse (Hu-Mo), Human and Rat (Hu-Rt), and Mouse and Rat (Mo-Rt)

	Ia	Io	Ig	Ix	II
Hu-Mo	69.5	46.5	55.4	66.0	92.9
Hu-Rt	65.5	41.9	58.2	62.5	94.5
Mo-Rt	92.0	79.9	89.9	98.7	98.4

amined with RNA prepared from untransfected cells. As shown in Fig. 3B, the endogenous transcripts contained exons Ia and Ig but not Io. When the rat gene construct, Ad4ECAT2.0K, was transfected, all RNA transcripts containing Ia, Io, and Ig were detected, indicating that exon Io of the rat gene can be transcribed even in mouse Y-1 cells. In contrast, when the mouse gene construct, mAd4ECAT, was transfected, the RNA transcript containing exon Io was hardly detectable, whereas transcripts containing exons Ia and Ig were clearly observed.

Expression of mRNA Isoforms in Vivo

As described above, it was unlikely that the exon Io of the mouse *Ad4BP/SF-1* gene is detectable in tissues. We then investigated whether the mRNA containing Io exon is expressed *in vivo* using RNAs prepared from rat and mouse tissues. The RNAs from the adult or fetal tissues were amplified using primers specific for exon Ia, Io, or Ig and for the second exon of each animal (Fig. 4A). As shown in Fig. 4B, the RNAs from the adult and fetal steroidogenic tissues of rats (the adrenal glands, testes, and ovaries) contained exons Ia, Io, and Ig while all three transcripts were undetectable in the liver. When RNAs prepared from the mouse tissues were used, the transcripts containing exon Ia and Ig were detected in the steroidogenic tissues of both adults and fetuses, whereas the RNA containing exon Io was hardly detectable (Fig. 4C). Since it has been established that the *Ad4BP/SF-1* is expressed in nonsteroidogenic tissues such as spleen, hypothalamus, and pituitary, the RNAs from these nonsteroidogenic tissues of mice were examined. As shown in Fig. 4D, RNA containing exon Ia was present in the adult tissues and fetal spleen, but was undetectable in the hypothalamus and pituitary of the fetuses. Ig containing RNA was expressed both in adult and fetal tissues although the intensity of the signal varied widely among RNA samples; strong in the adult hypothalamus and pituitary but weak in the adult spleen and fetal tissues. As in the case of the steroidogenic tissues, exon Io was never utilized in the mouse nonsteroidogenic tissues. The integrity of RNAs subjected to the analyses was confirmed by detection of β-actin.

Differential Usage of the First Exons

As indicated above, it was likely that the first exons were used with preferences in tissues and developmental stage. Such possibility seemed to be quite interesting when considering the promoter functions upstream from the multiple first exons. However, since the above RT-PCR analyses were qualitative but not quantitative, it was impossible to definitely conclude whether the exon usage varied among tissues. Thus, a quantitative RT-PCR procedure was developed as described under Materials and Methods. Before analyzing the

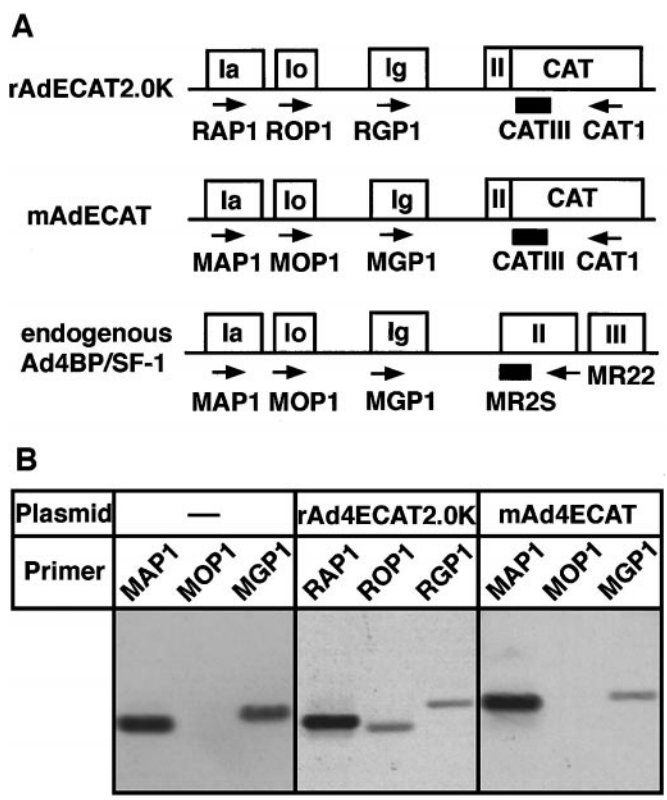

FIG. 3. Investigation of RNA species transcribed from endogenous and transfected *Ad4BP/SF-1* genes in mouse Y-1 adrenocortical tumor cells. Two constructs, rAd4ECAT2.0K and mAd4ECAT, carrying all the first exons of the rat and mouse genes, respectively, were transiently transfected into Y-1 cells. RNAs were prepared from the transfected and untransfected cells, and subsequently used for RT-PCR analyses. (A) Schematic presentation of the experimental design for PCR analysis. Open boxes indicate exons I, II, and III, and CAT reporter gene. The arrows accompanied by RAP1, ROP1, RGP1, MAP1, MOP1, MGP1, CAT1, and MR22 indicate approximate locations of oligonucleotide primers for PCR. Small closed boxes indicate approximate locations of oligonucleotide probes used for hybridization. These probes were designed to distinguish the transcripts derived from *CAT* reporter gene plasmids from those of endogenous *Ad4BP/SF-1* gene. (B) Analyses of the transcripts in Y-1 cells untransfected and transfected with constructs, rAd4ECAT2.0K and mAd4ECAT, by RT-PCR. RNAs prepared from Y-1 cells were subjected to RT-PCR using primers indicated at each lane. Thereafter, the amplified products were characterized by blotting analyses with specific probes indicated in A.

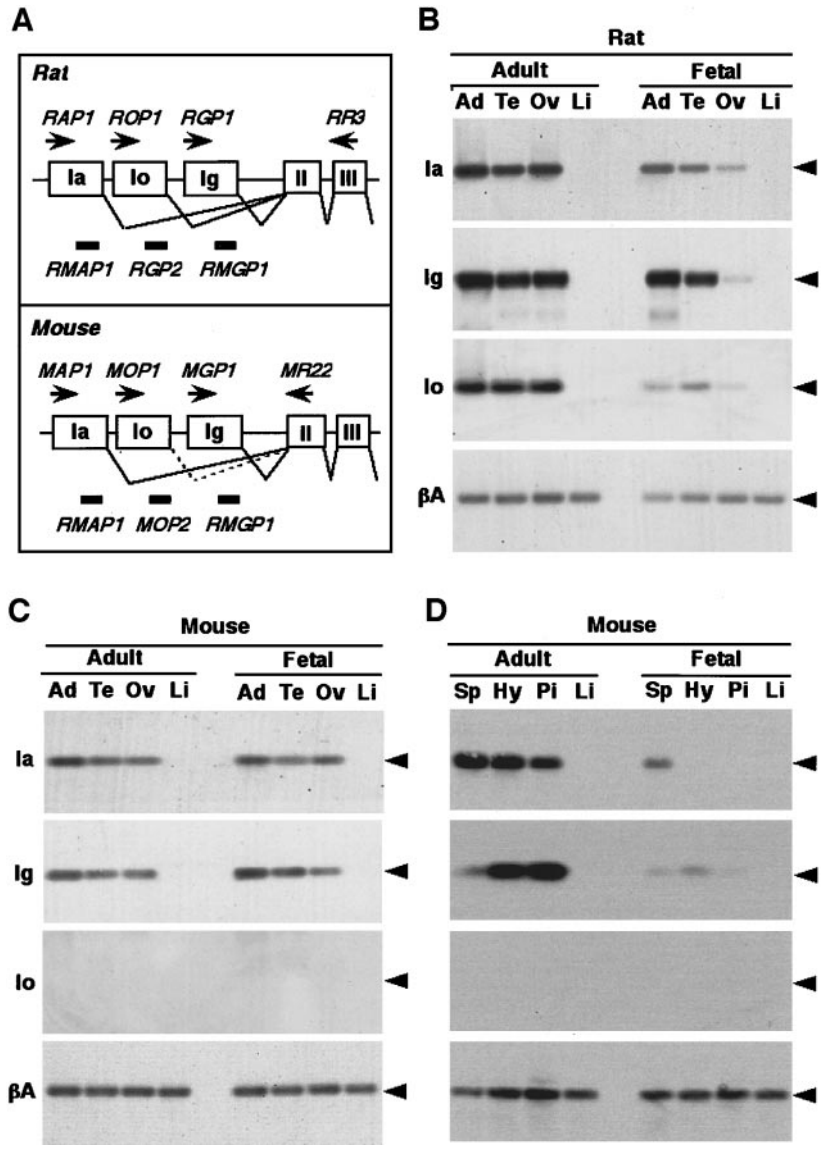

FIG. 4. Expression profiles of each first exon in rat and mouse tissues. RNAs prepared from rat and mouse tissues of adults and fetuses were used for RT-PCR analyses. (A) Schematic presentation of the experimental design for RT-PCR analysis. Open boxes indicate the first (Ia, Io, and Ig), second (II), and third (III) exons of the rat (upper panel) and mouse (lower panel) Ad4BP/SF-1 genes. Approximate locations of oligonucleotides used for PCR (RAP1, ROP1, RGP1, and RR3 for the rat, and MAP1, MOP1, MGP1, and MR22 for the mouse) and hybridization probes (RMAP1, RGP2, RMGP1 for the rat, and RMAP1, MOP2, and RMGP1 for the mouse) are indicated by arrows and closed boxes, respectively. (B, C, and D) Detection of mRNA species containing Ia, Io, or Ig. RNAs were prepared from the adrenal gland (Ad), testis (Te), ovary (Ov), spleen (Sp), hypothalamus (Hy), pituitary (Pi), and liver (Li) of adults (10 weeks after birth) and fetuses (E18.5) of rats (B). Tissues isolated from adult (10 weeks after birth) and fetal (E16.5 to E18.5) mice (C and D) were used for RNA preparation. Then the RNAs were subjected to RT-PCR. The amplified DNAs were identified by blotting analyses with primers specific for exon Ia, Io, Ig, and β-actin (βA).

RNA samples, linear ranges for PCR were determined with an authentic template cDNA. As shown in Figs. 5A and 5B, in the assay condition developed, 10^3 to 10^8 copies of the cDNAs containing Ia and Ig were confirmed to be in the range of linear amplification. Then the RNAs prepared from the steroidogenic and non-steroidogenic tissues of adult mice were examined to determine the amount of mRNA containing Ia or Ig exon in these tissues. As indicated in Fig. 5C, Ia exon was predominantly used in the adrenal glands, testes, ovaries, spleen, and hypothalamus. The amount of Ia-containing mRNA is approximately five- to ten-fold higher than Ig-containing mRNA in tissues, except for the spleen. It was interesting to note that Ig-containing transcript was hardly detected in the spleen in the quantitative range. In contrast, the pituitary was a unique tissue in which Ig-containing transcript was predominantly expressed.

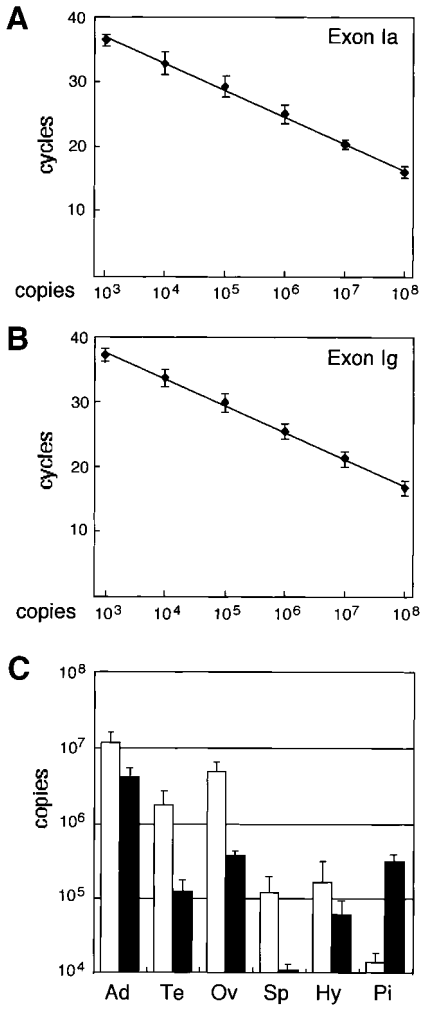

FIG. 5. Quantitative PCR analysis of exons Ia and Ig. (A and B) Standard curves of exon Ia-containing (A) and exon Ig-containing (B) mRNA. 10^3 to 10^8 copies of the corresponding cDNA molecules (indicated on the abscissa) were amplified as described under Materials and Methods. Larger amounts of amplified DNAs were observed with increasing strength of fluorescence, which was monitored at each PCR cycle. Ordinates indicate PCR cycle numbers. The amounts of amplified DNAs are plotted against the PCR cycle numbers. Standard reaction was performed at the same time with the assay. Data are mean ± SD of three experiments. (C) Determination of the amount of mRNAs containing Ia and Ig expressed in mouse tissues. RNAs prepared from adult mouse tissues, the adrenal gland (Ad), testis (Te), ovary (Ov), spleen (Sp), hypothalamus (Hy), and pituitary (Pi) were subjected to quantitative PCR. Open and closed bars indicate Ia- and Ig-containing mRNAs, respectively. Data are mean ± SD of three experiments. The copy numbers on the ordinate are plotted using a logarithmic scale. $P < 0.02$ in Ad, Te, Ov, Sp, and Pit. $P = 0.141$ in Hy.

DISCUSSION

In the present study we identified novel first exons, Io and Ig, in the rat *Ad4BP/SF-1* (*NR5A1*) gene in addition to Ia (22, 34) and Ix (the first exon used in ELP3) (31) reported previously. Sequence analyses revealed that these first exons are localized between exon Ia and the second exon. Based on the homology of the nucleotide sequence, it was expected that these first exons identified in the rat gene are also functional in the mouse and human. In fact, mRNA containing exon Ig was present in mouse as well as rat tissues. Interestingly, however, the transcript containing exon Io was undetectable in the mouse. The latter finding was further confirmed by examining the endogenous transcripts of mouse Y-1 adrenocortical tumor cells. A similar species-specific usage of exon Io was observed when the promoter constructs with the rat and mouse genes were transfected into the cultured cells. Similar to endogenous genes, exon Io of the mouse gene was never observed in mRNA transcribed from the transfected construct. These findings seem quite reasonable when considering the nucleotide sequences around splice donor sites of exons Io. Although the rat sequence satisfies the splice donor consensus, the mouse gene has an unusual sequence, which is unlikely to act as the splice donor. Since such alteration occurs in the corresponding region of the human gene, it is unlikely as in the case of the mouse gene that exon Io in the human gene is transcribed. In addition, an altered splice donor site at the human Ig strongly suggested that the human Ig exon is also inactive, in addition to Io. Interestingly, however, our recent search of DNA database identified a human EST clone (AA720791) containing a sequence corresponding to the human Ig exon. This finding strongly indicates that the human Ig is functional in at least synovial sarcoma, which was used as the mRNA source for the EST database. Since the clone contained a sequence extending from the presumptive splice junction of exon Ig to the 3′ downstream but did not contain the second exon, the splice junction between Ig and second exon is unknown. However, according to the sequence at the 3′ side of the clone, it is likely that a potential splice donor is present at more downstream region. Taken together, it seems that the usage of multiple first exons of the *Ad4BP/SF-1* gene is quite different among animal species probably due to their structures.

Such species-specific exon usage has been reported in several genes including the *CYP19* gene encoding aromatase P450 (35, 36). The genes, *Ad4BP/SF-1* and *CYP19*, have similar features as described below. Structural analyses of the *CYP19* genes of several animal species revealed that they have multiple first exons, and interestingly the first exons are different in their numbers among animal species. Another similarity is that all the first exons of the *CYP19* genes are noncoding, indicating that the gene products are identical in the amino acid sequence even though they are differentially utilized from the multiple first exons in combination with splicing. Considering that alternative usage and splicing of the first exons fails to provide distinct protein in terms of the structure and function,

it was reasonable to assume that the gene is endowed with a certain mechanism underlying the gene expression by acquisition of multiple first exons. In fact, it has been elucidated that the multiple first exons of *CYP19* are utilized by showing tissue-preference probably mediated by a distinct promoter upstream from each first exon (37, 38).

We have recently demonstrated that the *Ad4BP/SF-1* gene is expressed in several tissues such as steroidogenic tissues and nonsteroidogenic pituitary, hypothalamus, and spleen (8). In addition, studies from our laboratories and those of others have also shown the gene transcription is regulated in a developmental stage-specific manner (15, 39). These regulations are obviously realized by the precise function of the promoter. Therefore, as in the case of the *CYP* gene regulation, we anticipate that *Ad4BP/SF-1* gene transcription is controlled by the function of the regions upstream from the multiple first exons. In fact, we confirmed that the usage of the first exons, Ia and Ig, shows tissue preference in mouse tissues. In particular, it is interesting to note that Ig exon is predominantly utilized in the pituitary but hardly detectable in the spleen. It is known that Ad4BP/SF-1 is expressed in specialized cell types in these tissues, gonadotrophs in the pituitary (4, 6) and vascular endothelial cells in the spleen (7). Both of these cell types are quite different from steroidogenic cells with respect to their ontogeny and functions. Therefore, we consider that different mechanisms regulate the differential Ad4BP/SF-1 gene transcription among tissues. Establishment of multiple promoters accompanied by the first exons is obviously one of the structural bases for acquisition of such mechanisms.

Assuming the significance of multiple exons in the *Ad4BP/SF-1* gene as described above, it seems unusual that all first exons identified in the rat gene are not always active in other animal species. The human gene is the most unusual since the splice donor sequences of exons Io and Ig are both unlikely to be active, indicating that exon Ig used predominantly in the mouse pituitary gland is probably nonfunctional in human. However, in addition to exon Ig, it was previously reported that exon Ix (ELP3) is also utilized in the mouse (31). Thus, in addition to the possible existence of yet unidentified exons, we cannot exclude the possible usage of exon Ix of the human gene in the pituitary. Although the exact mechanism of gene regulation through these first exons should be identified, these differences can potentially provide the genomic basis for the development of sophisticated gene regulatory mechanisms in each animal.

ACKNOWLEDGMENTS

This study was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports, and Culture of Japan and grants from Asahi Glass Foundation and Senri Life Science Foundation.

REFERENCES

1. Nuclear Receptors Nomenclature Committee (1999) A unified nomenclature system for the nuclear receptor superfamily. *Cell* **97,** 161–163.
2. Morohashi, K., and Omura, T. (1996) Ad4BP/SF-1, a transcription factor essential for the transcription of steroidogenic cytochrome P450 genes and for the establishment of reproductive function. *FASEB J.* **10,** 1569–1577.
3. Parker, K. L., and Schimmer, B. P. (1997) Steroidogenic factor 1: A key determinant of endocrine development and function. *Endocrine Rev.* **18,** 361–377.
4. Ingraham, H. I., Lala, D. S., Ikeda, Y., Luo, X., Shen, W-H., Nachtingal, M. W., Abbu,d R., Nilson, J. H., and Parker, K. L. (1994) The nuclear receptor steroidogenic factor 1 acts at multiple levels of the reproductive axis. *Genes Dev.* **8,** 2302–2312.
5. Ikeda, Y., Luo, X., Abbud, R., Nilson, J. H., and Parker, K. L. (1995) The nuclear receptor steroidogenic factor 1 is essential for the formation of the ventromedial hypothalamic nucleus. *Mol. Endocrinol.* **9,** 478–486.
6. Shinoda, K., Lei, H., Yoshii, H., Nomura, M., Nagano, M., Shiba, H., Sasaki, H., Osawa, Y., Ninomiya, Y., Niwa, O., Morohashi, K., and Li, E. (1995) Developmental defects of the ventromedial hypothalamic nucleus and pituitary gonadotroph in the *Ftz-F1* disrupted mice. *Dev. Dynam.* **204,** 22–29.
7. Morohashi, K., Tsuboi-Asai, H., Matsushita, S., Suda, M., Nakashima, M., Sasano, H., Hataba, Y., Li, C-L., Fukuta, J., Irie, J., Watanabe, T., Nagura, H., and Li, E. (1999) Structural and functional abnormality in the spleen of an *mFtz-F1* gene-disrupted mouse. *Blood* **93,** 1586–1594.
8. Morohashi, K. (1999) Gonadal and extragonadal functions of Ad4BP/SF-1: Developmental aspects. *Trends Endocrinol. Metab.* **10,** 169–173.
9. Barnhart, K. M., and Mellon, P. (1994) The orphan nuclear receptor, steroidogenic factor-1, regulates the glycoprotein hormone α-subunit gene in pituitary gonadotropes. *Mol. Endocrinol.* **8,** 878–885.
10. Drean, Y. L., Liu, D., Wong, A. O. L., Xiong, F., and Hew, C. L. (1995) Steroidogenic factor 1 and estradiol receptor act in synergism to regulate the expression of the salmon gonadotropin IIβ subunit gene. *Mol. Endocrinol.* **10,** 217–229.
11. Lee, S. L., Sadovsky, Y., Swirnoff, A. H., Polish, J. A., Goda, P., Gavrilina, G., and Milbrandt, J. (1996) Luteinizing hormone deficiency and female infertility in mice lacking the transcription factor NGFI-A (Egr-1). *Science* **273,** 1219–1221.
12. Luo, X., Ikeda, Y., and Parker, K. L. (1994) A cell-specific nuclear receptor is essential for adrenal and gonadal development and sexual differentiation. *Cell* **77,** 481–490.
13. Sadovsky, Y., Crawford, P. A., Woodson, K. G., Polish, J. A., Clements, M. A., Tourtellotte, L. A., Simburger, K., and Milbrandt, J. (1995) Mice deficient in the orphan receptor steroidogenic factor 1 lack adrenal glands and gonads but express P450 side-chain-cleavage enzyme in the placenta and have normal embryonic serum levels of corticosteroids. *Proc. Natl. Acad. Sci. USA* **92,** 10939–10943.
14. Ikeda, Y., Shen, W., Ingraham, H. A., and Parker, K. L. (1994) Developmental expression of mouse steroidogenic factor-1, an essential regulator of the steroid hydroxylases. *Mol. Endocrinol.* **8,** 654–662.
15. Hatano, O., Takayama, K., Imai, T., Waterman, M. R., Takakusu, A., Omura, T., and Morohashi, K. (1994) Sex-dependent expression of a transcription factor, Ad4BP, regulating steroido-

genic P-450 genes in the gonads during prenatal and postnatal rat development. *Development* **120,** 2787–2797.

16. Hatano, O., Takakusu, A., Nomura, M., and Morohashi, K. (1996) Identical origin of adrenal cortex and gonad revealed by expression profiles of Ad4BP/SF-1. *Genes Cells* **1,** 663–671.
17. Vilain, E., Guo, W., Zhang, Y-H., and McCabe, E. R. B. (1997) DAX1 gene expression upregulated by steroidogenic factor 1 in an adrenocortical carcinoma cell line. *Biochem. Mol. Med.* **61,** 1–8.
18. Yu, R. N., Ito, M., and Jameson, J. L. (1998) The murine Dax-1 promoter is stimulated by SF-1 (steroidogenic factor-1) and inhibited by COUP-TF (chicken ovalbumin upstream promoter-transcription factor) via a composite nuclear receptor-regulatory element. *Mol. Endocrinol.* **12,** 1010–1022.
19. Kawabe, K., Shikayama, T., Tsuboi, H., Oka, S., Oba, K., Yanase, T., Nawata, H., and Morohashi, K. (1999) Dax-1 as one of the target genes of Ad4BP/SF-1. *Mol. Endocrinol.* **13,** 1267–1284.
20. Zanaria, E., Muscatelli, F., Bardoni, B., Strom, 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) An unusual member of the nuclear hormone receptor superfamily responsible for X-linked adrenal hypoplasia congenita. *Nature* **372,** 635–641.
21. Muscatelli, F., Strom, 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., Camerino, G., Meitinger, T., and Monaco, A. P. (1994) Mutations in the *DAX-1* gene give rise to both X-linked adrenal hypoplasia congenita and hypogonadotropic hypogonadism. *Nature* **372,** 672–676.
22. Nomura, M., Bartsch, S., Nawata, H., Omura, T., and Morohashi, K. (1995) An E box element is required for the expression of the *ad4bp* gene, a mammalian homologue of *ftz-fl* gene, which is essential for adrenal and gonadal development. *J. Biol. Chem.* **270,** 7453–7461.
23. Woodson, K. G., Crawford, P. A., Sadovsky, Y., and Milbrandt, J. (1997) Characterization of the promoter of SF-1, an orphan nuclear receptor required for adrenal and gonadal development. *Mol. Endocrinol.* **11,** 117–126.
24. Harris, A. N., and Mellon, P. L. (1998) The basic helix-loop-helix, lucine zipper transcription factor, USF (upstream stimulatory factor), is a key regulator of SF-1 (steroidogenic factor-1) gene expression in pituitary gonadotrope and steroidogenic cells. *Mol. Endocrinol.* **12,** 714–726.
25. Daggett, M. A. F., Rice, D. A., and Heckert, L. L. (2000) Expression of steroidogenic factor 1 in the testis requires an E box and CCAAT box in its promoter proximal region. *Biol. Reprod.* **62,** 67–679.
26. Chomczynski, P., and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162,** 156–159.
27. Honda, S., Morohashi, K., Nomura, M., Takeya, H., Kitajima, M., and Omura, T. (1993) Ad4BP regulating steroidogenic *P-450* gene is a member of the steroid hormone receptor superfamily. *J. Biol. Chem.* **268,** 7494–7502.
28. Araki, E., Shimada, F., Shichiri, M., Mori, M., and Ebina, Y. (1988) pSV00CAT: Low background CAT plasmid. *Nucleic Acids Res.* **16,** 1627.
29. Kirita, S., Hashimoto, T., Kitajima, M., Honda, S., Morohashi, K., and Omura, T. (1990) Structural analysis of multiple bovine P-450(11β) genes and their promoter activities. *J. Biochem.* **108,** 1030–1041.
30. Tsukiyama, T., Ueda, H., Hirose, S., and Niwa, O. (1992) Embryonal long terminal repeat-binding protein is a murine homolog of FTZ-F1, a member of the steroid receptor superfamily. *Mol. Cell Biol.* **12,** 1286–1291.
31. Ninomiya, Y., Okada, M., Kotomura, N., Suzuki, K., Tsukiyama, T., and Niwa, O. (1995) Genomic organization and isoforms of the mouse ELP gene. *J. Biochem.* **118,** 380–389.
32. Senapathy, P., Shapiro, M. B., and Harris, N. L. (1990) Splice junctions, branch point sites, and exons: Sequence statistics, identification, and applications to genome project. *Methods Enzymol.* **183,** 252–278.
33. Morohashi, M., Zanger, U. M., Honda, S., Hara, M., Waterman, M. R., and Omura, T. (1993) Activation of *CYP11A* and *CYP11B* gene promoters by the steroidogenic cell-specific transcription factor, Ad4BP. *Mol. Endocrinol.* **7,** 1196–1204.
34. Ikeda, Y., Lala, D. S., Luo, X., Kim, E., Moison, M-P., and Parker, K. L. (1993) Characterization of the mouse *FTZ-F1* gene, which encodes a key regulator of steroid hydroxylase gene expression. *Mol. Endocrinol.* **7,** 852–860.
35. Means, G. D., Kilgore, M. W., Mahendroo, M. S., Medelson, C. R., Simpson, E. R. (1991) Tissue-specific promoters regulate aromatase cytochrome P450 gene expression in human ovary and fetal tissues. *Mol. Endocrinol.* **5,** 2005–2013.
36. Harada, N. (1992) A unique aromatase (P-450AROM) mRNA formed by alternative use of tissue-specific exons 1 in human skin fibroblasts. *Biochem. Biophys. Res. Commun.* **189,** 1001–1007.
37. Harada, N., Utsumi, T., and Takagi, Y. (1993) Tissue-specific expression of the human aromatase cytochrome P-450 gene by alternative use of multiple exons 1 and promoters, and switching of tissue-specific exons 1 in carcinogenesis. *Proc. Natl. Acad. Sci. USA* **90,** 11321–11316.
38. Toda, K., Simpson, E. R., Mendelson, C. R., Shizuta, Y., and Kilgore, M. W. (1994) Expression of the gene encoding aromatase cytochrome P450 (CYP19) in fetal tissues. *Mol. Endocrinol.* **8,** 210–217.
39. Shen, W-H., Moore, C. C. D., Ikeda, Y., Parker, K. L., and Ingraham, H. A. (1994) Nuclear receptor steroidogenic factor 1 regulates the Mullerian inhibiting substance gene: A link to the sex determination cascade. *Cell* **77,** 651–661.