J Biochem 127, 805–811 (2000)

Structure of the Pig Sterol 14α-Demethylase (CYP51) Gene and Its Expression in the Testis and Other Tissues[1]

Misaki Kojima,*[2] Takeya Morozumi,† Akira Onishi,‡ and Tadayoshi Mitsuhashi*

Laboratory of Gene Function and ‡Genetic Resource Development, National Institute of Animal Industry, Norindanchi, P O Box 5, Tsukuba, Ibaraki 305-0901, and †Animal Genome Analysis Team, STAFF Institute, Ippaizuka, Kamiyokoba, Tsukuba, Ibaraki 305-0854

Received December 27, 1999, accepted February 15, 2000

A cDNA coding sterol 14α-demethylase (CYP51), which was isolated from a pig liver cDNA, contained a 1,512 bp open reading frame and a 758 bp 3′-untranslated region. The deduced amino acid sequence was 94% identical to those of human and rat CYP51s. The pig *CYP51* gene spanned about 21 kb and was divided into 10 exons. The sites of exon–intron junctions were completely identical to those in the human and rat *CYP51* genes. Five GC boxes, but not a TATA box, were found in the 5′-flanking region of the gene, and cyclic AMP and sterol responsive elements were also found in this region. The main transcription start site determined with the 5′-RACE method with poly(A)$^+$ RNA from the liver and testis was located at 143 nucleotides upstream from the initiation codon in both tissues. Northern blot analysis revealed that an approximately 2.4 kb mRNA, which is produced through the use of a polyadenylation signal (AATAAA) located at 740 nucleotides downstream of the stop codon, was expressed in all the tissues examined in pigs: The mRNA levels were much higher in the liver and testis than in the kidney, lung, and epididymis. Furthermore, after the onset of spermatogenesis, a smaller size of mRNA (about 1.8 kb) was found in the testis but not in the epididymis. The 1.8 kb mRNA was produced through the use of an unusual polyadenylation signal (AAGAAA) located at 28 nucleotides downstream of the stop codon.

Key words: CYP51, *CYP51* mRNA, gene structure, lanosterol 14α-demethylase, pig.

Sterol 14α-demethylase (CYP51) is the only cytochrome P450 (P450) enzyme, which has been widely conserved from prokaryotes to eukaryotes such as mammals, fungi, and plants, and is characterized as a housekeeping enzyme (*1–3*). The CYP51 enzyme of eukaryotes catalyzes the 14α-demethylation of 14-methylsterol, which is an essential step in sterol biosynthesis (*1–6*). Recently, the amino acid sequences of the rat and human CYP51 enzymes have been reported to exhibit 93% identity and to show 35–42% homology to those of lower eukaryotic CYP51s (*2*, *7*). Phylogenetic analysis has indicated that CYP51 might have arisen in the era of prokaryotic evolution and that it has been conserved in both prokaryotes and eukaryotes (*2*, *3*).

Human and rat *CYP51* mRNAs have been reported to be expressed in all tissues examined, especially in the testis (*7–9*). In rats, the expression pattern of *CYP51* mRNA species in the testis is different from those in other tissues; in addition to three basic transcripts produced in all tissues, there is a shorter transcript produced through the use of an unusual polyadenylation signal of the *CYP51* gene (*8*, *9*). Furthermore, a testis-specific shorter transcript, which is suggested to be more stable than basic transcripts in rats (*8*, *9*), has been reported to be present in germ cells (*8*). These reports suggest that the shorter transcript of the *CYP51* gene plays an important role in spermatogenesis. However, there is little information about the *CYP51* gene and its expression in mammals with the exception of humans and rats.

In this study, to determine whether the production of the shorter transcript of the *CYP51* gene in the testis is common to mammals or not, we used pigs, which are thought to be similar to humans from viewpoints of anatomy and physiology, and determined the cDNA sequence of CYP51 and its gene structure. Furthermore, we examined the age-dependent expression of the *CYP51* mRNA species in the testis and other tissues.

MATERIALS AND METHODS

Isolation of Pig CYP51 cDNA—A plasmid clone, 523, containing a fragment of pig CYP51 cDNA (about 1,600 bp) lacking the N-terminal region has already been obtained in our institute. To further obtain the full length cDNA, a

The sequence data of pig liver CYP51 cDNA have been entered in the DDBJ/EMBL/GenBank data bases under the accession number AB009988.

[1] This work was supported in part by the Japan Racing Association (JRA).

[2] To whom correspondence should be addressed. Tel/Fax: +81-298-38-8662, E-mail: misaki@niai affrc.go.jp

[3] The accession numbers of the amino acid sequences used for the homology search are X13768 (pig CYP11A), M14565 (human CYP-11A), J05156 (rat CYP11A), U37312 (pig CYP19), J04127 (human CYP19), M33986 (rat CYP19), M63507 (pig CYP17), M14564 (human CYP17), and M31681 (rat CYP17).

Abbreviations bp, base pair(s), CRE, cyclic AMP responsive element, G3PDH, glyceraldehyde 3-phosphate dehydrogenase, kbp, kilobase pair(s); RACE, rapid amplification of cDNA ends, SRE-1, sterol responsive element-1

TABLE I **PCR primer sets used for amplification of introns of the pig *CYP51* gene.**

Intron	Sense primer	Antisense primer
A	5′-ACCTGTTCCGCCAAGCCATC-3′	5′-CCCAAATGCTATAGCATGTCC-3′
B	5′-AGCATTTGGGAAAAGTCCAATTG-3′	5′-CCCCAGAAGGTAGGTAAACG-3′
C	5′-CTACAGTCGCCTGACAACACC-3′	5′-GTGGGCTATGTTAAGGCCAC-3′
D	5′-GGAATACTTTCAAAGTTGGGG-3′	5′-CCATGTAAACAATGGCTTGCTG-3′
E	5′-GTGGTTTTAGCCACGCAGCC-3′	5′-CTTCTGATTGTCTGCGTTTC-3′
F	5′-GAAACGCAGACAATCAGAAG-3′	5′-AGTCCAATAAGCATGCCGGC-3′
G	5′-GGCCAGAGACAAAACACTTC-3′	5′-CGTTTCTTTTATGCAGCGATC-3′
H	5′-TTCGACCTCCGATAATGACC-3′	5′-GGAGAAACACACACCTGATG-3′
I	5′-AGCATCAGGAGAGAAGTTTG-3′	5′-ATTGTCTTGATTTGAACATAGGC-3′

cDNA library constructed from liver poly(A)$^+$ RNA of an adult female pig (Landrace × Large White × Duroc; LWD) with a ZAP-cDNA Synthesis Kit (Stratgene) was screened by colony hybridization. A fragment (about 1,300 bp) obtained on digestion of the clone 523 plasmid DNA with *Eco*RI and *Sty*I was digoxygenin-labeled and then used as a probe. A positive clone containing the longest insert was selected and subjected to sequence analysis.

Isolation of Pig CYP51 Genomic DNA—CYP51 genomic DNA was isolated from a pig genomic bacterial artificial chromosome (BAC) library, which was constructed using high molecular weight DNA from a male LWD pig kidney by the method of Asakawa *et al.* (*10*). Screening was carried out by means of two-step polymerase chain reaction (PCR) as described by Asakawa *et al.* (*10*). A PCR primer set (sense, 5′-ATGCAACACCTCCCAGTG-3′; antisense, 5′-CTTTTATACTTGCAGATCTAG-3′) amplifying 315 bp of the 3′-untranslated region of pig CYP51 cDNA was used for the screening.

To determine the structure of the *CYP51* gene, we performed the PCR using primer sets (Table I) encompassing speculated exon-intron junctions based on the human and rat *CYP51* gene structures (*9*, *11*). Amplification of introns A, C, D, E, F, G, and H was performed with AmpliTaq Gold polymerase as follows: 1 cycle at 94°C for 10 min; 35 cycles at 94°C for 1 min, 57°C for 1 min and 72°C for 3 min; and 1 cycle at 72°C for 3 min. For introns B and I, Ex-Taq DNA polymerase (Takara Shuzo, Kyoto) was used, and the amplification conditions were 1 cycle at 94°C for 1 min; 30 cycles at 98°C for 20s and 60°C for 5 min; and 1 cycle at 72°C for 10 min. To confirm the products and the junction sequences, direct sequencing of the PCR products was performed using PCR primer sets as sequence primers. The lengths of introns were determined by electrophoresis on a 1.0% agarose gel.

Sequence Analysis—Sequence analysis was performed with an ABI377 sequencer and a Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems).

5′- and 3′- Rapid Amplification of cDNA End (5′- and 3′-RACE)—Double stranded cDNAs were synthesized by use of a Marathon cDNA synthesis kit (Clontech) with 2 μg of poly(A)$^+$ RNA from the liver and testis of a 6-month-old male LWD pig. Both the 5′- and 3′-ends of the cDNA were ligated to the 44-mer adapter containing two overlapping primer sets for AP1 (5′-CCATCCTAATACGACTCACTATAGCGGC-3′) and AP2 (5′-ACTCACTATAGGGCTCGAGCGGC-3′). For 5′-RACE, PCR was performed with the cDNA using KlenTaq Polymerase (Clontech), an antisense primer, AS1 (5′-ACAGGTGTTGTCAGGCGACTGTAG-3′), and the AP1 primer as a 5′-end primer. The PCR products were used for the 2nd PCR using a nested antisense primer, AS2 (5′-CTTGCAGTAAGCCCAGCAACAC-3′), and the nested 5′-end primer AP2. For 3′-RACE, the 1st PCR was carried out with a sense primer, S1 (5′-CTACAGTCGCCTGACAACACCTG-3′), and AP1 as a 3′-end primer, and the 2nd PCR was primed with a nested sense primer, S2 (5′-AGCATCAGGAGAGAAGTTTG-3′), and AP2 as a nested 3′-end primer. PCR was performed according to the instruction manual. The second PCR products were cloned into the pCR 2.1 vector with a TA-cloning kit (Invitrogen) and then subjected to sequence analysis.

Northern Blot Analysis—Total RNAs were prepared from several pig tissues with Trizol reagent (Life Technologies). A portion of the RNA preparation (40 μg) from each tissue was electorophoresed on a 1.0% agarose gel containing 2.2 M formaldehyde, and then the separated RNA on the gel was transferred to a nylon membrane. The membrane was hybridized at 42°C overnight with the ^{32}P-labeled 530 bp cDNA fragment of nucleotide positions 833 to 1362 (Fig. 1) in a solution containing 5× SSPE (0.75M NaCl, 50 mM NaH_2PO_4, 5 mM EDTA, pH 7.4), 5× Denhardt's, 0.1% SDS, sonicated salmon sperm DNA (50 μg/ml), and 50% formamide. The membrane was washed with 2 × SSC containing 0.5% SDS and then with 0.2× SSC containing 0.5% SDS, and finally exposed to X-ray film at −80°C for 48 h.

RESULTS

Isolation of CYP51 cDNA—Pig CYP51 cDNA was isolated by screening of the liver cDNA library. The nucleotide sequence of the cDNA clone containing the longest insert was determined. The resultant cDNA was 2,370 bp-long and contained a 1,512 bp-long open reading frame (ORF) encoding 503 amino acids and a 758 bp-long 3′-untranslated region including three typical polyadenylation signals, AATAAA (Fig. 1). The heme binding region was also found between deduced amino acids 442 and 452 (Fig. 1). The pig CYP51 ORF sequence showed 91.7 and 86.3% homology to those of humans and rats, respectively, and the deduced amino acid sequence showed 94% homology to those of humans and rats.

Organization of the CYP51 Gene—To determine the structure of the pig *CYP51* gene, its genomic DNA was obtained by screening of a pig genomic BAC library. Two clones, 267A10 and 282F10, were isolated as positive clones. Southern blot analysis of BAC DNAs from both the clones indicated that these clones contained the same part of the pig *CYP51* gene (data not shown). Therefore, clone 267A10 was used for further analysis.

As shown in Table II and Fig. 1, the pig *CYP51* gene consisted of 10 exons spanning about 21 kb. The nucleotide sequences of the exons were completely identical to the

CYP51 cDNA sequence (Fig. 1), and the donor and acceptor sites of the 9 introns followed the GT-AG rule (*12*). Southern blotting of pig genomic DNA digested with *Bam*HI, *Eco*RI, *Pst*I, or *Hin*dIII, for which pig CYP51 cDNA has no sites, was performed using a probe, *i.e.* a 315 bp-fragment amplified with the primer set used for the screening of the pig genomic BAC library or a fragment recognizing exon 1. In all cases, only one band was detected (data not shown). This indicates that the pig *CYP51* gene is a single copy one. To obtain further information about the pig *CYP51* gene expression, the 5′-flanking sequence of the gene was analyzed. As shown in Fig. 2, neither a TATA nor a CAAT box was found in the region of 880 bp upstream from the initiation codon. Five GC boxes were present at nucleotide positions −206, −222, −391, −462, and −550. Furthermore, five CRE elements were found at nucleotide positions −22, −188, −348, −520, and −688, and a SRE-1 element was located at nucleotide position −502.

The transcription start site of the pig *CYP51* gene was estimated by means of the 5′-RACE method with poly(A)$^+$ RNAs from the liver and testis. The nucleotide sequences of the inserts in all clones obtained were identical to the 5′-

```
CGCCCATCTGCTGGCTAGTCTCCCTCAGTGGTTTCGGAGGAGCCGCGGCGACCTCGGCCTCCAGAATCCCCAGCAAACTGACGTGAGGGTGGCGGCCGGG   100
ATGGTGTTGCTGGGCTTACTGCAAGCTGGGGGGTCGGTGCTTGGGCAGGCGATGGAGCAGGTAACTGGCGTCAATCTCCTATCCAGTCTGCTCCTCGCCT   200
 M  V  L  L  G  L  L  Q  A  G  G  S  V  L  G  Q  A  M  E  Q  V  T  G  V  N  L  L  S  S  L  L  L  A    33
                                                                          ▼A
GCGCCTTCACGCTCATCTTAGTTTACCTGTTCCGCCAAGCCATCGGCCACCTGGCCCCACTGCCGGCTGGAGCGAAAAGTCCACCATACATTTTCTCTCC   300
C  A  F  T  L  I  L  V  Y  L  F  R  Q  A  I  G  H  L  A  P  L  P  A  G  A  K  S  P  P  Y  I  F  S  P    67
                                                                ▼B
AATTCCATTTCTTGGACATGCTATAGCATTTGGGAAAAGTCCAATTGAATTCCTAGAAAATGCTTATGAGAAGTACGGACCTGTATTTAGTTTTACCATG   400
  I  P  F  L  G  H  A  I  A  F  G  K  S  P  I  E  F  L  E  N  A  Y  E  K  Y  G  P  V  F  S  F  T  M   100
GTGGGCAAAACGTTTACCTACCTTCTGGGGAGCGATGCCGCTGCACTGCTTTTTAATAGTAAAAATGAAGATTTGAATGCAGAAGATGTCTACAGTCGCC   500
 V  G  K  T  F  T  Y  L  L  G  S  D  A  A  A  L  L  F  N  S  K  N  E  D  L  N  A  E  D  V  Y  S  R   133
                                                ▼C
TGACAACACCTGTGTTTGGGAAGGGAGTCGCATATGATGTGCCTAATCCAGTTTTCTTGGAGCAGAAGAAGATGTTAAAAAGTGGCCTTAACATAGCCCA   600
L  T  T  P  V  F  G  K  G  V  A  Y  D  V  P  N  P  V  F  L  E  Q  K  K  M  L  K  S  G  L  N  I  A  H   167
                                                                ▼D
CTTTAGACAGCATGTTTCTATAATTGAAAAAGAAACAAAGGAATACTTTCAAAGTTGGGGAGAAAGTGGAGAAAGAAATTTGTTTGAAGCTCTTTCTGAG   700
  F  R  Q  H  V  S  I  I  E  K  E  T  K  E  Y  F  Q  S  W  G  E  S  G  E  R  N  L  F  E  A  L  S  E   200
CTCATAATTTTAACAGCAAGCCATTGTTTACATGGAAAGGAAATCAGAAGTCAGCTTAACGAGAAGGTGGCACAACTGTATGCGGATTTGGATGGTGGTT   800
 L  I  I  L  T  A  S  H  C  L  H  G  K  E  I  R  S  Q  L  N  E  K  V  A  Q  L  Y  A  D  L  D  G  G   233
                                                ▼E
TTAGCCACGCAGCCTGGCTCTTACCAGGCTGGCTTCCCTTGCCCAGTTTCAGACGCAGGGACAGAGCACATCGAGAGATCAAGAATATTTTCTATAAGGC   900
F  S  H  A  A  W  L  L  P  G  W  L  P  L  P  S  F  R  R  R  D  R  A  H  R  E  I  K  N  I  F  Y  K  A   267
                                                                ▼F
AATCCAGAAACGCAGACAATCAGAAGAAAAAATTGATGACATTCTCCAAACCTTACTAGATTCTACTTACAAGGATGGGCGTCCTTTGACAGATGATGAA  1000
  I  Q  K  R  R  Q  S  E  E  K  I  D  D  I  L  Q  T  L  L  D  S  T  Y  K  D  G  R  P  L  T  D  D  E   300
GTAGCCGGCATGCTTATTGGACTGCTCTTGGCGGGGCAGCACACATCCTCAACTACCAGTGCCTGGATGGGCTTCTTTTTGGCCAGAGACAAAACACTTC  1100
 V  A  G  M  L  I  G  L  L  L  A  G  Q  H  T  S  S  T  T  S  A  W  M  G  F  F  L  A  R  D  K  T  L   333
                                                                ▼G
AAGAAAAATGTTATTTGGAACAGAAAACAGTTTGTGGAGAGGATCTGCCACCATTAACTTATGACCAGCTCAAGGATCTAAACTTACTTGATCGCTGCAT  1200
Q  E  K  C  Y  L  E  Q  K  T  V  C  G  E  D  L  P  P  L  T  Y  D  Q  L  K  D  L  N  L  L  D  R  C  I   367
                                                ▼H
AAAAGAAACGTTAAGACTTCGACCTCCGATAATGACCATGATGAGAATGGCTAAAACTCCTCAGACTGTGGCAGGGTACACTATTCCTCCAGGACATCAG  1300
  K  E  T  L  R  L  R  P  P  I  M  T  M  M  R  M  A  K  T  P  Q  T  V  A  G  Y  T  I  P  P  G  H  Q   400
GTGTGTGTTTCTCCCACTGTCAATCAAAGACTTAAAGACTCATGGGTAGAACGACTGGACTTTAATCCTGATCGCTACTTACAGGACAATCCAGCATCAG  1400
 V  C  V  S  P  T  V  N  Q  R  L  K  D  S  W  V  E  R  L  D  F  N  P  D  R  Y  L  Q  D  N  P  A  S   433
                                   ▼I
GAGAGAAGTTTGCCTATGTGCCATTTGGAGCTGGGCGTCATCGTTGTATTGGGGAGAATTTTGCCTATGTTCAAATCAAGACAATTTGGTCCACTATGCT  1500
G  E  K  F  A  Y  V  P  F  G  A  G  R  H  R  C  I  G  E  N  F  A  Y  V  Q  I  K  T  I  W  S  T  M  L   467
TCGTTTATATGAATTTGATCTCATTGATGGATATTTTCCCACTGTGAATTATACAACCATGATTCACACCCCTGAAAACCCAGTTATCCGATACAAACGA  1600
  R  L  Y  E  F  D  L  I  D  G  Y  F  P  T  V  N  Y  T  T  M  I  H  T  P  E  N  P  V  I  R  Y  K  R   500
AGATCAAAGTGAAAAAGGCTGCAAGGAACTAAAGTGTGAAAGAAACATCATTGTAAACTGCAAAAGCATTGAAGAGAATGTTTATGCAACACCTCCCAGT  1700
 R  S  K  *                                                                                             503
GTATTCTCTTTGTGAGTGTGTAATTTTAAAAGACAGCCAGCTCAACTTAGGGTTTTTGTATTTTAACTGAATGGTTCCATAAAATCTAATGGCCTTCTAG  1800
AAATTTCCTAATAGTTTTATATTGGTTATTACATATGTTTTATGTAAGGAAATCAAGCTCATTAGTTTATGGAGGGGATCTTTGACAATTGGCAGAGTGT  1900
AAATAACACAATTTCCAGGTATTAAAGACCATATACATAGCTCTTAGCAAGTAAATTCAGAGTGTTCAGATCTTGGCCTAGATCTGCAAGTATAAAAGCC  2000
CAGTAAACTCTTTGGAAGTCTTAGGTTATTTATTCATTTACTGGGTAATGTTGTGTTTAACGATGAGTGGGAAGGGAATAAACCCACAGTTAGAGCCTTG  2100
GATGAGAGAATCCTATTGTTTTGACATGATGATCCTTAATGATGGGGCTATGTTGTAGGAAAATATGTCCTTTGAGATTTTAACTAAAAGCACTGTCTGT  2200
AGGAATCTGAAGAAAGATCTATGACAATGACTTAATTAACTTTCCTACAAAATAATAAAGACTGCTCTTGTAACTTGATCTAATTACTAAATCTTACACA  2300
TTTATCTGCCTGGTAATCGCTCCACATGGTTGCATCCATGTGCAAGTCCAAAATAAAGTCCTTGAGGGAT                                  2370
```

Fig. 1. **The pig CYP51 cDNA and predicted amino acid sequences.** CYP51 cDNA was isolated from a pig liver cDNA library. The nucleotide sequence has been registered in the DDBJ/EMBL/GenBank (accession number, AB009988). The predicted amino acid sequence is indicated below the corresponding nucleotide sequence. The numbers of nucleotides and amino acids are indicated at the right. Three typical polyadenylation signals (AATAAA) are underlined. A heme binding region is denoted by a wavy line. Intron positions are indicated by arrowheads (see Table II for details).

flanking sequence of the *CYP51* gene determined with BAC clone 267A10 (Fig 2). The transcription start sites of the liver-derived clones were concentrated in the region comprising nucleotide positions −138 to −146 (Fig. 3), nucleotide position −143 being most frequently used as a transcription start site. Similar results were obtained for the testis-derived clones (data not shown). These results strongly suggest that position −143 is a main transcription start site of the *CYP51* gene in both the liver and testis of pigs.

Expression Pattern of CYP51 mRNA Species in Various Tissues—The expression patterns of *CYP51* mRNA species in various tissues (liver, kidney, lung, epididymis, and testis) of pigs were investigated by Northern blotting. In all tissues examined, 2.4 kb *CYP51* mRNA, which corresponds to the product surmised from the main transcription start site (nucleotide position −143), was observed (Fig. 4). A high level of *CYP51* mRNA was observed in the liver and testis, and a lower level of *CYP51* mRNA was expressed in the kidney, lung, and epididymis. In the mature pig (6-month-old), 1.8 kb *CYP51* mRNA was observed together with the 2.4 kb mRNA in the testis, but not in the epididymis (Fig. 4A). Since no 1.8 kb *CYP51* mRNA was observed in the testis of immature pigs (10-week-old) (Fig. 4B), we further examined the age-dependent expression of *CYP51* mRNA species in the testis. Ubiquitously expressed 2.4 kb mRNA was observed even in newborn pigs, and its level did not significantly change at least up to 6-month-old (Fig. 4C). On the other hand, testis-specific 1.8 kb mRNA was undetectable in immature pigs and appeared after the age of 4 months.

The production of the 1.8 and 2.4 kb mRNAs was thought to be due to the difference in the polyadenylation site, because their main transcription start sites were the same, as described above. Therefore, we further examined the polyadenylation site in the testis by means of the 3′-RACE method with poly(A)$^+$ RNA from the testis of a mature male pig. The polyadenylation sites in 30 clones derived from the testis poly(A)$^+$ RNA were examined, and the results are shown in Fig. 5. The polyadenylation sites in 24 of the 30 clones examined were at (or around) nucleotide position 1693, and the resultant transcript corresponded to

TABLE II **Sequences of exon/intron junctions of the pig *CYP51* gene.** The sequences of exon/intron junctions were determined by sequencing the products amplified by PCR using BAC clone 267A10 DNA as a template, the PCR primer sets used being shown in Table I The sizes of the introns, with the exception of introns F and H, were determined by 1 0% agarose gel electrophoresis, those of introns F and H being determined by sequence analysis. The nucleotides of exons and introns are shown in capitals and small letters, respectively

Exon no	5′ donor	Intron, size (bp)	3′ acceptor	Exon no
		A (~1,900)		
1	GCTGGAGCG gtacgtcacc - - - - - - - - - - - - - - -		tttcctgcag AAAAGTCCA	2
		B (~4,500)		
2	TATGAGAAG gtaagtcttt - - - - - - - - - - - - - - -		tctctcttag TACGGACCT	3
		C (~1,600)		
3	CCTAATCCA gtaagtgacc - - - - - - - - - - - - - - -		tattttgaag GTTTTCTTG	4
		D (~1,900)		
4	GAGAAAGAA gtaagcgaaa - - - - - - - - - - - - - - -		-ttccttccag ATTTGTTTG	5
		E (~2,300)		
5	CAGTTTCAG gtatggataa - - - - - - - - - - - - - - -		tgaattacag ACGCAGGGA	6
		F (384)		
6	TACTTACAA gtaagagctc - - - - - - - - - - - - - - -		cattctgtag GGATGGGCG	7
		G (~1,400)		
7	TATGACCAG gtttgttgta - - - - - - - - - - - - - - -		ctcttttag CTCAAGGAT	8
		H (828)		
8	ACTCCTCAG gtaagtctcc - - - - - - - - - - - - - - -		tcatttgtag ACTGTGGCA	9
		I (~4,500)		
9	TTGGAGCTG gtaagatggt - - - - - - - - - - - - - - -		ttctttatag GGCGTCATC	10

```
-883 TGGTCCTAAA AAGCTAGAAT AATTGTCGTT CTTGGGAAAT GTTGCCTGAA ATAGGTGACG
-823 ACGGCCTTGA ACATTAAAGT TGGTTCGGAA AACCTGTTTT TGTATAACGC AGAATACTCA
-763 ACAAGCGAGA CTTGTTACCT TGAAGCGCGC TTTTGAACAT CTATCAGACT GGTTTTACCC
-703 GGGGTGAGGT GGGACTGACA TCAAATGTCC CACCTTTCAG TGTCTGGACG GCGTCTTTCA
-643 CAAACTCCAT TTTCATAAAA CGTAGAGACT CCCCCAAATC CCAGGAAGCA GCCCAGGAGC
-583 AGGTTCGGCT GTGGCGGGTC CGCGCGCCCG GGCAAGGGGC GGTGCTTCAG ACACCCGCCC
-523 AGCTGACGCG ACATAGGCCG AGATCACCTC ACGCGCGCGC GGTGCAATCA CGAAGCGCGC
-463 TCCCCCTGCC CCACCTTCAC AACCCTCGCC TTGAGATCCT TCCGCCCATC TCATCCCGGT
-403 CGCGTGACCT GGCGCCCCGC ACCCAGCACA CCCAAGTTGG GGTGGACCGC GATTCTGAGG
-343 TTATCTCCTC TTAGGCCACA TTCTGGTCTG TCTCTGTACC ACCCGTGCGA CAGGTGCAGC
-283 ATTATTCTTC AGGTCCTCTC TCCTGCGTAA GGAGATGACC GCGGCACGAA GTACAATTTA
-223 TGGGGCGTGG GCATAGGAAG CCCCTCCCAC TCCTGTGACG TACAGGGTGG TGCGTGAGGG
-163 AGCCGGCAGT AGGGGGTGCG GCTGGGTTCA GCAGGGGGCT TGCGGGCCGG CGCCGCTTTC
-103 CGACGCCCAT CTGCTGGCTA GTCTCCCTCA GTGGTTTCGG AGGAGCCGCG GCGACCTCGG
 - 43 CCTCCAGAAT CCCCAGCAAA CTGACGTGAG GGTGGCGGCC GGGatggtgt tgctgggctt
```

Fig. 2. **Primary structure of the 5′-flanking region of the pig *CYP51* gene.** The nucleotides of the *CYP51* gene are negatively numbered from the initiation codon indicated by underlining, and the numbers are shown at the left. GC boxes, and CRE, and SRE-1 elements are indicated by boxes, double underlining and wavy underlining, respectively. The main transcription start site determined by the 5′-RACE method is indicated by an asterisk.

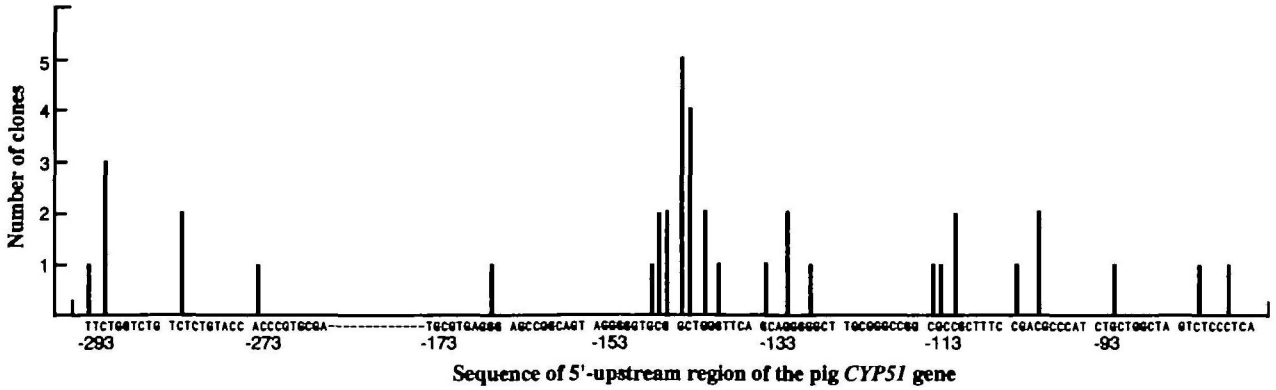

Fig 3 **Determination of the transcription start sites of the *CYP51* gene in the pig liver by the 5′-RACE method.** Transcription start sites were determined by the 5′-RACE method as described under "MATERIALS AND METHODS." Black bars show the transcription start sites, and their heights show the numbers of clones using each transcription start site

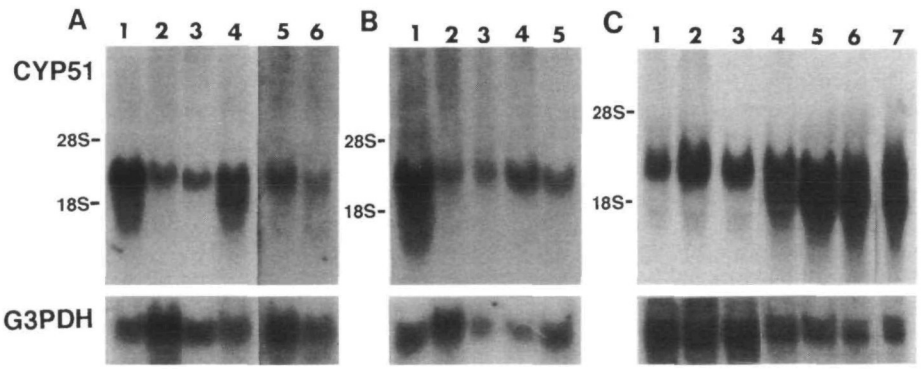

Fig. 4. **Northern blot analysis of *CYP51* mRNA in the testis and other tissues of pigs.** Total RNA preparations (40 μg per lane) were electrophoresed on a 1.0% agarose gel containing 2.2 M formaldehyde. The separated RNA fractions were transferred to a nylon membrane and then hybridized with a ^{32}P-labeled fragment of pig CYP51 cDNA as described under "MATERIALS AND METHODS". G3PDH mRNA was simultaneously examined. The expression pattern of *CYP51* mRNA species was examined in several tissues of 6-month-old (A) and 10-week-old (B) male pigs. Lanes 1 to 6 in both A and B represent liver, kidney, lung, testis, epididymis (head), and epididymis (tail), respectively. C shows the age-dependent expression of *CYP51* mRNA species in the testis. Lane 1, at birth; lane 2, 3-week-old; lane 3, 3-month-old, lanes 4 and 5, 4-month-old; lanes 6 and 7, 6-month-old

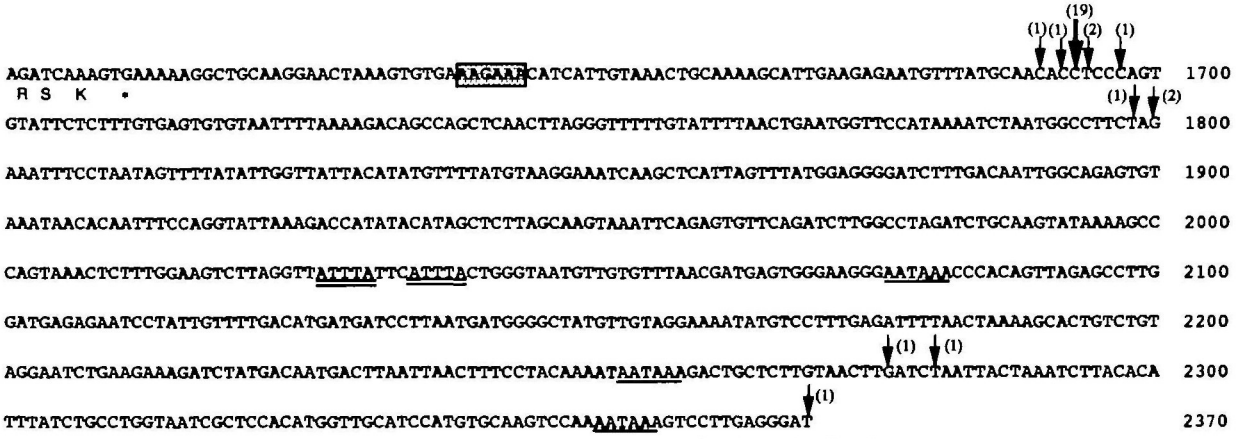

Fig. 5. **Determination of the polyadenylation sites of *CYP51* mRNA in the pig testis by the 3′-RACE method.** Polyadenylation sites were determined by the 3′-RACE method as described under "MATERIALS AND METHODS." Arrows indicate the polyadenylated sites, and the numbers in parentheses represent the numbers of clones showing polyadenylation at each site. Three typical polyadenylation signals (AATAAA) are underlined, and another polyadenylation signal (AAGAAA) is denoted by a hatched box. The ATTTA sequence, which is a destabilizer element of mRNA, is double-underlined.

1.8 kb *CYP51* mRNA, as shown in Fig. 4. Accordingly, the 1.8 kb mRNA might be produced through the use of a variant polyadenylation signal (AAGAAA) located at nucleotide positions 1640–1645 (Fig. 5). On the other hand, most of the polyadenylation sites in the case of the liver poly(A)$^+$ RNA were at nucleotide position 2370, *i.e.* after the third typical polyadenylation signal located at nucleotide positions 2352–2357, and the product sizes were about 2.4 kb (data not shown). These results indicated that the difference in the production of *CYP51* mRNA species between the testis and other tissues in pigs was dependent on the difference in the polyadenylation site used.

DISCUSSION

We demonstrated herein that the pig *CYP51* gene consists of 10 exons spanning approximately 21 kb, and that its 5′-upstream region contains GC boxes, and CRE and SRE-1 elements, and that 1.8 kb *CYP51* mRNA, together with ubiquitously expressed 2.4 kb mRNA, is only found in the testis of mature male pigs. Furthermore, the production of different sizes of *CYP51* mRNAs, 1.8 and 2.4 kb, was demonstrated to be due to the use of different polyadenylation signals.

The homology (94%) of the CYP51 amino acid sequence among pigs, humans and rats was much higher than those (62–80%) of other CYPs (CYP11A, CYP19, and CYP17) related to the synthesis of steroid hormones.[3] The structure of the pig *CYP51* gene also showed high similarity to those of humans and rats, as follows: (i) the insertion sites of the introns in the pig *CYP51* gene (Fig. 1 and Table II) were completely identical to those in humans and rats (*9*, *11*), and (ii) the first and fifth GC boxes, the second and fourth CRE elements, and the SRE-1 element in the 5′-flanking sequence of the *CYP51* gene in pigs (Fig. 2) were also observed in humans and rats. These findings suggest that the conserved regulatory elements, GC box, CRE, and SRE-1, might control the expression of the *CYP51* gene. The similarities in the structure of the *CYP51* gene and its deduced amino acid sequence among pigs, humans and rats confirmed that CYP51 is a unique P450 highly conserved during evolution (*2*, *3*).

Intronless *CYP51* pseudogenes have been observed in humans (*13*) and rats (*9*), while in the present study, no *CYP51* pseudogene was observed in pigs on screening of the BAC library using PCR. Furthermore, Southern blotting of *Bam*HI-, *Eco*RI-, *Pst*I-, or *Hin*dIII-treated pig genomic DNA revealed only one band. If there are any pseudogenes in pigs, more than two bands should be obtained on Southern blotting. Accordingly, these results suggest that there is no *CYP51* pseudogene in pigs.

Northern blotting demonstrated that, in pig tissues, only one transcript of the *CYP51* gene (2.4 kb) was ubiquitously expressed (Fig. 4), as well as in humans (3.8 kb) (*7*), while in rats, three different *CYP51* mRNAs (3.1, 2.7, and 2.3 kb) were ubiquitously expressed (*8*, *9*). These facts suggest that there is species difference in the regulation of the *CYP51* gene between rats and pigs/humans. In pigs, a high level of the *CYP51* mRNA was expressed in the liver and testis but not in the kidney or lung (Figs. 4, A and B), although in humans, a high level of *CYP51* mRNA was expressed in the kidney and lung as well as the liver and testis (*7*). Furthermore, as shown in Fig. 4C, a high level of 2.4 kb *CYP51* mRNA was found in the testis of newborn pigs and the level did not significantly change, at least up to 6-month-old (Fig. 4), whereas in immature and mature rat testes, three ubiquitously expressed *CYP51* mRNAs were found only at low levels (*8*, *9*). Namely, we demonstrated here that there are species and tissue differences in the expression of the *CYP51* gene.

In the testis of mature male pigs, 1.8 kb *CYP51* mRNA was expressed together with 2.4 kb *CYP51* mRNA (Fig. 4C). The appearance of the 1.8 kb mRNA might be correlated with the onset of spermatogenesis, because spermatogenesis in pigs has been reported to be completed after 4-month-old (*14*). On the other hand, in spite of the presence of sperm in the epididymis, no 1.8 kb *CYP51* mRNA was observed (Fig. 4A). Likewise, no 1.8 kb *CYP51* mRNA was observed in the testis of 3-month-old pigs, which contains spermatogonia and pachyten spermatocytes (*14*). These findings suggest that a testis-specific 1.8 kb *CYP51* mRNA in pigs is only expressed at specific stage(s) of spermatogenesis, and might be present in postmeiotic spermatids. This idea might be supported by the previous finding (*8*) that the testis-specific transcript of the *CYP51* gene was most abundant in elongating spermatids in rats. Recently, it was demonstrated that the expression of testis-specific *CYP51* mRNA is regulated by CREMτ, a cAMP/cAMP-responsive element modulator, which shows binding activity toward CRE elements in the promoter region of the human *CYP51* gene (*15*). Because there are CRE elements in the promoter region of the *CYP51* genes of humans, rats and pigs, the CREMτ-dependent expression of testis-specific *CYP51* mRNA might be common in mammals.

The biological significance of a shorter transcript of the *CYP51* gene in the testis has not been clarified yet. Since the testis-specific shorter transcript appears after maturation in pigs as well as rats (*8*, *9*), it seems to contribute to the enhancement of the synthesis of cholesterol, a precursor for testosterone biosynthesis. As another possibility, it was thought that the appearance of the shorter transcript might be related to the initiation of meiosis in the testis, because the CYP51 enzyme catalyzes the conversion of lanosterol to 4,4-dimethylcholesta-8,14,24-trienol (meiosis-activating sterol; MAS), which reinitiates meiosis in mouse oocytes (*16*). However, a testis-specific shorter transcript was abundant in spermatids at the postmeiotic stage rather than the meiotic stage (*8*), suggesting that CYP51-mediated MAS production in the testis does not necessarily contribute to the initiation of meiosis in male germ cells. Thus, the function of MAS as to meiosis in male germ cells remains unclear.

The testis-specific CYP51 transcript (1.8 kb *CYP51* mRNA) in pigs was demonstrated to be produced through the utilization of a variant polyadenylation signal (AAGAAA) at nucleotide positions 1640–1645. Likewise, a variant polyadenylation signal (AATGAA) in rats has been reported to be used for the production of a testis-specific *CYP51* mRNA (*8*, *9*). Since the testis-specific 1.8 kb *CYP51* mRNA did not contain an AUUUA element, which is present in 2.4 kb *CYP51* mRNA at nucleotide positions 2028 and 2036 (Fig. 5), and destabilizes the mRNA (*17*), the testis-specific mRNA should be more stable than the 2.4 kb mRNA. Furthermore, it has been demonstrated that a rat testis-specific shorter transcript was translated to the intact CYP51 protein, and that the increase in the tran-

script level was closely correlated with that in CYP51 enzyme activity (*8*) . These findings suggest that the testis-specific shorter transcript of the *CYP51* gene in pigs might also be translated to the corresponding enzyme having activity.

In the present study, we determined the structure of the pig *CYP51* gene and its transcripts, and further demonstrated the tissue difference in the expression of *CYP51* mRNA species, especially a testis-specific shorter mRNA. However, the biological significance of the tissue-dependent expression of *CYP51* mRNA species has not been clarified yet. Therefore, further study will be necessary to determine the biological significance of each *CYP51* mRNA species in the testis and other tissues.

REFERENCES

1. Yoshida, Y (1993) Sterol biosynthesis in *Cytochrome P-450* 2nd ed. (Omura, T., Ishimura, Y., and Fujii-Kuriyama, Y, eds.) pp 93–101, Kodansha, Tokyo, and VCH, Weinheim
2 Aoyama, Y., Noshiro, M., Gotoh, O , Imaoka, S , Funae, Y , Kurosawa, N , Horiuchi, T., and Yoshida, Y. (1996) Sterol 14-demethylase P450 (P45014DM) is one of the most ancient and conserved P450 species. *J. Biochem.* **119**, 926–933
3 Yoshida, Y., Noshiro, M , Aoyama, Y., Kawamoto, T., Horiuchi, T., and Gotoh, O. (1997) Structural and evolutionary studies on sterol 14-demethylase P450 (CYP51), the most conserved P450 monooxygenase: II Evolutionary analysis of protein and gene structures. *J Biochem.* **122**, 1122–1128
4 Trzaskos, J M., Bowen, W.D , Shafiee, A., Fischer, R.T., and Gaylor, J.L (1984) Cytochrome P-450-dependent oxidation of lanosterol in cholesterol biosynthesis. Microsomal electron transport and C-32 demethylation *J. Biol. Chem.* **259**, 13402–13412
5. Aoyama, Y., Yoshida, Y., and Sato, R. (1984) Yeast cytochrome P-450 catalyzing lanosterol 14α-demethylation: II. Lanosterol metabolism by purified $P\text{-}450_{14DM}$ and by intact microsomes. *J Biol. Chem.* **259**, 1661–1666
6 Cabello-Hurtado, F , Zimmerlin, A., Rahier, A., Taton, M , DeRose, R , Nedelkina, S , Batard, Y , Durst, F , Pallett, K.E., and Werck-Reichhart, D. (1997) Cloning and functional expression in yeast of a cDNA coding for obtusifoliol 14α-demethylase (CYP51) in wheat. *Biochem Biophys. Res. Commun.* **230**, 381–385
7. Strömstedt, M , Rozman, D., and Waterman, M R. (1996) The ubiquitously expressed human CYP51 encodes lanosterol 14α-demethylase, a cytochrome P450 whose expression is regulated by oxysterol. *Arch Biochem. Biophys.* **329**, 73–81
8. Strömstedt, M., Waterman, M R , Haugen, T.B , Tasken, K., Parvinen, M., and Rozman, D (1998) Elevated expression of lanosterol 14α-demethylase (CYP51) and the synthesis of oocyte meiosis-activating sterols in postmeiotic germ cells of male rats. *Endocrinology* **139**, 2314–2321
9 Noshiro, M , Aoyama, Y., Kawamoto, T , Gotoh, O , Horiuchi, T., and Yoshida, Y. (1997) Structural and evolutionary studies on sterol 14-demethylase P450 (CYP51), the most conserved P450 monooxygenase. I Structural analyses of the gene and multiple sizes of mRNA. *J Biochem.* **122**, 1114–1121
10 Asakawa, S , Abe, I., Kudoh, Y , Kishi, N., Wang, Y., Kubota, R., Kudoh, J , Kawasaki, K., Minoshima, S , and Shimizu, N (1997) Human BAC library construction and rapid screening. *Gene* **191**, 69–79
11 Rozman, D., Strömstedt, M , Tsui, L-C., Scherer, S W. and Waterman, M R. (1996) Structure and mapping of the human lanosterol 14α-demethylase gene (CYP51) encoding the cytochrome P450 involved in cholesterol biosynthesis; Comparison of exon/intron organization with other mammalian and fungal CYP genes. *Genomics* **38**, 371–381
12. Rogers, J and Wall, R. (1980) A mechanism for RNA splicing. *Proc. Natl. Acad Sci. USA* **77**, 1877–1879
13. Rozman, D., Strömsted, M , and Waterman, M.R (1996) The three human cytochrome P450 lanosterol 14α-demethylase (CYP51) genes reside on chromosome 3, 7, and 13: Structure of the two retrotransposed pseudogenes, association with a line-1 element, and evolution of the human CYP51 family *Arch. Biochem Biophys.* **333**, 466–474
14 Colenbrander, B , Fankenhuis, M.T , and Wensing C J.G. (1982) Male sexual development in *Control of Pig Reproduction* (Cole, D J.A. and Foxchroft, G R., eds.) pp 3–24, Butterworth Scientific, London
15 Rozman, D , Fink, M , Fimia, G.M , Sassone-Corsi, P., and Waterman, M R. (1999) Cyclic adenosine 3′,5′-monophosphate (cAMP)/cAMP-responsive element modulator (CREM)-dependent regulation of cholesterogenic lanosterol 14α-demethylase (CYP51) in spermatids. *Mol Endocinol* **13**, 1951–1962
16. Byskov, A.G , Andersen, C Y., Nordholm, L , Thøgersen, H., Guoliang, X., Wassmann, O , Andersen, J V., Guddal, E., and Rded, T (1995) Chemical structure of sterols that activate oocyte meiosis. *Nature* **374**, 559–562
17 Shaw, G and Kamen, R (1986) A conserved AU sequence from the 3′ untranslated region of GM-CSF mRNA mediates selective mRNA degradation *Cell* **46**, 659–667