

Cytochrome P450c17 (steroid 17 α -hydroxylase/17,20 lyase): Cloning of human adrenal and testis cDNAs indicates the same gene is expressed in both tissues

(steroid hormone/monooxygenase/DNA sequence/protein sequence/androgens)

BON-CHU CHUNG*[†], JAMES PICADO-LEONARD*, MITSURU HANIU[‡], MICHAEL BIENKOWSKI[‡], PETER F. HALL[§], JOHN E. SHIVELY[‡], AND WALTER L. MILLER*[¶]

*Department of Pediatrics and the Metabolic Research Unit, University of California, San Francisco, CA 94143; [†]Division of Immunology, Beckman Research Institute of the City of Hope, Duarte, CA 91010; and [§]Worcester Foundation for Experimental Biology, Shrewsbury, MA 01545

Communicated by Yuet Wai Kan, October 8, 1986 (received for review August 20, 1986)

ABSTRACT P450c17 is the single enzyme mediating both 17 α -hydroxylase (steroid 17 α -monooxygenase, EC 1.14.99.9) and 17,20 lyase activities in the synthesis of steroid hormones. It has been suggested that different P450c17 isozymes mediate these activities in the adrenal gland and testis. We sequenced 423 of the 509 amino acids (83%) of the porcine adrenal enzyme; based on this partial sequence, a 128-fold degenerate 17-mer was synthesized and used to screen a porcine adrenal cDNA library. This yielded a 380-base cloned cDNA, which in turn was used to isolate several human adrenal cDNAs. The longest of these, λ hac17-2, is 1754 base pairs long and includes the full-length coding region, the complete 3'-untranslated region, and 41 bases of the 5'-untranslated region. This cDNA encodes a protein of 508 amino acids having a predicted molecular weight of 57,379.82. High-stringency screening of a human testicular cDNA library yielded a partial clone containing 1303 identical bases. RNA gel blots and nuclease S1-protection experiments confirm that the adrenal and testicular P450c17 mRNAs are indistinguishable. These data indicate that the testis possesses a P450c17 identical to that in the adrenal. The human amino acid sequence is 66.7% homologous to the corresponding regions of the porcine sequence, and the human cDNA and amino acid sequences are 80.1 and 70.3% homologous, respectively, to bovine adrenal P450c17 cDNA. Both comparisons indicate that a central region comprising amino acid residues 160-268 is hypervariable among these species of P450c17. Comparison of the amino acid sequence of P450c17 with two other human steroidogenic cytochromes P450 show much greater homology with P450c21 (28.9%), another microsomal enzyme, than with P450scc (12.3%), a mitochondrial enzyme.

Steroid 17 α -hydroxylase (steroid 17 α -monooxygenase, EC 1.14.99.9) converts pregnenolone to 17-hydroxypregnenolone and converts progesterone to 17-hydroxyprogesterone. These 17-hydroxylated steroids may then be converted by 17,20-lyase to dehydroepiandrosterone and androstenedione, respectively. These latter two steroids are precursors of testosterone and estrogen synthesis while 17-hydroxyprogesterone is a key precursor of cortisol synthesis. Although steroid 17 α -hydroxylase and 17,20 lyase activities can be readily distinguished by examination of circulating venous steroidal products (1, 2), studies in both the guinea pig (3) and pig (4) show that both activities reside in a single protein, P450c17. Thus, the P450c17 enzyme is a key branch point in human steroid hormone synthesis, as 17 α -hydroxylase activity distinguishes between synthesis of mineralocorticoids (aldosterone) and glucocorticoids (cortisol) and as 17,20 lyase

activity distinguishes between synthesis of glucocorticoids and sex steroids. P450c17 is encoded by a gene or genes now termed P450XVII (5). P450c17 mRNA accumulation is regulated hormonally (6, 7) and developmentally (8). Like P450c21 (steroid 21-hydroxylase), P450c17 is bound to the endoplasmic reticulum and accepts electrons from NADPH via a flavoprotein (9-11). By contrast, the two other adrenal steroidogenic cytochromes P450, P450scc (cholesterol side-chain cleavage enzyme) and P450c11 (11 β /18 hydroxylase), reside in mitochondria and employ a different electron transfer chain (12). cDNAs have been cloned for two human steroidogenic enzymes, P450c21 (ref. 13; also K. J. Matteson, B.-c.C., W.L.M., unpublished results) and P450scc (14, 15). The sequence of bovine adrenal P450c17 cDNA was reported (16) but the human cDNA and amino acid sequence are not known.

Slight differences have been reported in porcine P450c17 isolated from the adrenal and testis (17). These proteins were distinguished by molecular weight, amino acid composition, and N-terminal amino acid sequence but were enzymologically and immunologically indistinguishable, raising the possibility of tissue-specific isozymes of P450c17; the existence of such P450c17 isozymes remains unconfirmed. We now report the partial amino acid sequence of porcine adrenal P450c17. This amino acid sequence was used to produce a degenerate-sequence oligonucleotide, which in turn was used to isolate a small porcine adrenal cDNA. This porcine cDNA was then used to isolate the full-length human adrenal cDNA. A human testis cDNA library contains an identical sequence, and nuclease S1-protection experiments confirm that the same P450c17 mRNA is found in human adrenal gland and testis.

MATERIALS AND METHODS

Porcine adrenal P450c17 was isolated as described (18). Approximately 2 mg of protein was *S*-carboxymethylated, digested with trypsin or *Staphylococcus aureus* V8 protease, and sequenced as described (19, 20). The sequences of 50 tryptic fragments and 28 V8 protease fragments were ordered by overlapping regions and by alignment with the human cDNA (below) and bovine cDNA (16). Double-stranded porcine adrenal cDNA was synthesized (21), ligated to pBR327 by dC·dG tailing, and cloned in *Escherichia coli* MC 1061. The 128-fold degenerate 17-base probe 3' AC_GCAXC-

[†]Present address: Institute of Molecular Biology, Academia Sinica, Nankang, Taipei 11529, Taiwan.

[¶]To whom reprint requests should be addressed at: Room 677-S, Department of Pediatrics, University of California, San Francisco, CA 94143.

CXCT^ATAC^TTA 5' was synthesized by phosphoramidite chemistry.

The human adrenal cDNA library has been described (14); the human testis cDNA library was from Clontech (Palo Alto, CA) (15). The porcine cDNA was nick-translated (specific activity, >10⁸ cpm/μg) and used to screen the human adrenal cDNA library under conditions of low stringency: 5× SSC (1× SSC = 0.15 M NaCl/0.015 M sodium citrate, pH)/20% (vol/vol) formamide/denatured salmon sperm DNA (100

μg/ml)/0.2% polyvinylpyrrolidone/0.2% Ficoll/0.2% bovine serum albumin at 37°C. High stringency screening was done at 42°C in the same solution containing 50% (vol/vol) formamide. Cloned cDNA was cleaved from λgt10, cloned directly into pUC13 and pUC18, and appropriate fragments were subcloned into M13mp11 and mp18 for dideoxy sequencing (22). Sequence comparisons were done as described (23).

S1-nuclease mapping was done as described (24). In the 5'

pAA hAA hCDNA	ACTCCACTGCTGTCTATCTTGCCTGCCGGCACCCAGCCACC	Met Trp Val Leu Leu Val Phe Phe Leu Leu	10	Thr Leu Thr Tyr Leu Phe Trp Pro Lys	20	Thr Leu Thr Tyr Leu Phe Trp Pro Lys	30	Leu Pro Val Leu Pro Val Val Gly Ser Leu	40	Pro Phe Leu Pro Arg Arg His Gly His Met	50	His Thr Phe Leu Phe Lys Leu Gln Asp Lys Phe	60	Gly Pro Ile Phe Ser Phe Arg Leu Gly Thr	70	Lys Thr Thr Val Val Val Ile Gly Asp His	80	Gln Leu Ala Lys Glu Val Leu Leu Lys Lys Gly	90	Lys Glu Phe Ser Gly Arg Pro Lys Val Met	100	Thr Leu Thr Asp Ile Leu Ser Asp Asn Gln	110	Lys Lys Gly Ile Ala Phe Ala Asp Glu Gly Thr	120	Trp Glu Leu His Arg Asp Leu Ala Met Gly	130	Thr Phe Thr Ser Leu Phe Lys Asp Gly Thr	140	Gln Lys Glu Thr Ser Thr Ile Leu Ser Asp Asn	150	Gln His Thr Thr Thr Thr Thr Thr Thr Thr Thr	160	His Thr Thr Thr Thr Thr Thr Thr Thr Thr	170	Leu Thr Thr Thr Thr Thr Thr Thr Thr Thr	180	Thr Thr Thr Thr Thr Thr Thr Thr Thr	190	Thr Thr Thr Thr Thr Thr Thr Thr Thr	200	Thr Thr Thr Thr Thr Thr Thr Thr Thr	210	Thr Thr Thr Thr Thr Thr Thr Thr Thr	220	Thr Thr Thr Thr Thr Thr Thr Thr Thr	230	Thr Thr Thr Thr Thr Thr Thr Thr Thr	240	Thr Thr Thr Thr Thr Thr Thr Thr Thr	250	Thr Thr Thr Thr Thr Thr Thr Thr Thr	260	Thr Thr Thr Thr Thr Thr Thr Thr Thr	270	Thr Thr Thr Thr Thr Thr Thr Thr Thr	280	Thr Thr Thr Thr Thr Thr Thr Thr Thr	290	Thr Thr Thr Thr Thr Thr Thr Thr Thr	300	Thr Thr Thr Thr Thr Thr Thr Thr Thr	310	Thr Thr Thr Thr Thr Thr Thr Thr Thr	320	Thr Thr Thr Thr Thr Thr Thr Thr Thr	330	Thr Thr Thr Thr Thr Thr Thr Thr Thr	340	Thr Thr Thr Thr Thr Thr Thr Thr Thr	350	Thr Thr Thr Thr Thr Thr Thr Thr Thr	360	Thr Thr Thr Thr Thr Thr Thr Thr Thr	370	Thr Thr Thr Thr Thr Thr Thr Thr Thr	380	Thr Thr Thr Thr Thr Thr Thr Thr Thr	390	Thr Thr Thr Thr Thr Thr Thr Thr Thr	400	Thr Thr Thr Thr Thr Thr Thr Thr Thr	410	Thr Thr Thr Thr Thr Thr Thr Thr Thr	420	Thr Thr Thr Thr Thr Thr Thr Thr Thr	430	Thr Thr Thr Thr Thr Thr Thr Thr Thr	440	Thr Thr Thr Thr Thr Thr Thr Thr Thr	450	Thr Thr Thr Thr Thr Thr Thr Thr Thr	460	Thr Thr Thr Thr Thr Thr Thr Thr Thr	470	Thr Thr Thr Thr Thr Thr Thr Thr Thr	480	Thr Thr Thr Thr Thr Thr Thr Thr Thr	490	Thr Thr Thr Thr Thr Thr Thr Thr Thr	500	Thr Thr Thr Thr Thr Thr Thr Thr Thr	508	Thr Thr Thr Thr Thr Thr Thr Thr Thr	516	Thr Thr Thr Thr Thr Thr Thr Thr Thr
---------------------	---	---	----	-------------------------------------	----	-------------------------------------	----	---	----	---	----	---	----	---	----	---	----	---	----	---	-----	---	-----	---	-----	---	-----	---	-----	---	-----	---	-----	---	-----	---	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------	-----	-------------------------------------

hCDNA CATTCTTCCTACTCCCAACCCACTCGCTCTTTTCAGCTGTGGCAATGCCAGTGATGCGAATAAAGAGTTTTTTTTTTTCAAAAAAGAGAGATACAAACCCCCACCTTCTCCGC

FIG. 1. Sequences of P450c17. The upper line, designated pAA, shows the 423 residues of porcine amino acid sequence determined by peptide sequencing, except residues 491 and 492 are inferred from the porcine cDNA. The second line, designated hAA, shows the human amino acid sequence derived from the human adrenal cDNA sequence (hCDNA) shown on the third line. The fourth line, designated pcDNA, shows the sequence of the small porcine cDNA clone. The fifth line, designated bAA, shows only those bovine amino acid residues (16) that differ from the human sequence. The overlined region of 17 bases (amino acids 442-447) indicates the region corresponding to the 128-fold degenerate 17-base porcine oligonucleotide probe. The arrows between amino acids 99 and 100 and between the 76th and 77th bases in the 3'-untranslated region indicate the extent of the human testicular P450c17 cDNA. The "ATAAA" polyadenylation signal is underlined.

probe, P450c17 mRNA protects the 461 bases from the 5' terminus of the cDNA to the *Sau3A* site at codons 139–140 (Fig. 1). In the 3' probe, P450c17 mRNA protects the 297 bases from the *Pvu* II site at codons 472–473 to the 3' end of the cDNA.

RESULTS

Amino Acid Sequence of Porcine Adrenal P450c17. Microsequencing of proteolytic fragments of the available porcine adrenal P450c17 enzyme yielded the sequence of 423 amino acids; by analogy with the human (below) and bovine (16) sequences, this represents 83% of the expected 508 or 509 amino acids. A 128-fold degenerate 17-mer was synthesized corresponding to amino acids 442–447, where residue 442 is the thiolate cysteine of the heme-binding site. With this probe six porcine adrenal cDNA clones were identified. One of these, *ppa17*, was chosen at random and partially sequenced. The sequence of these 325 bases, shown in Fig. 1, corresponds precisely with the amino acid sequence of several overlapping proteolytic peptides of porcine P450c17. By alignment with the human (below) and bovine (16) P450c17 sequences, this cDNA encodes amino acids 409–502.

Identification and Sequence of Human Adrenal P450c17 cDNA. Hybridization of the porcine P450c17 fragment to our human adrenal cDNA library yielded about 300 positive clones. Two clones were chosen at random, subcloned, and partially sequenced. One clone contained about 650 base pairs (bp), the other contained 950 bp. Both had identical restriction maps and identical nucleotide sequences in the regions examined, including the entire 3'-untranslated region of both clones. The 5' end of the longer clone was subcloned and used to reprobe the library, identifying about 150 clones. Sixty of these were picked and grown in pools of six without plaque purification, cleaved with *EcoRI*, electrophoresed through agarose, and examined by Southern blotting. The pool containing the largest observed hybridizing band was then broken down to its individual members, and plaques of the appropriate clone were purified. The clone, designated λ hac17-3, was sequenced by the strategy shown in Fig. 2, showing that the cDNA contained 1754 bases plus the *EcoRI* linkers used in constructing the library. By homology with the porcine amino acid sequence, this cDNA encodes 41 bases of the 5'-untranslated sequence, the whole protein of 508 amino acids, the complete 171-base 3'-untranslated region, and part of the poly(A) tail (Fig. 1). The predicted molecular weight of human P450c17 is 57,379.82.

Comparison of Human Adrenal and Testicular P450c17 mRNAs. The report of tissue-specific P450c17 isozymes in the pig describes very minor differences in the amino acid sequence and composition of these two forms (17), indicating the sequences of their mRNAs would be very similar. Therefore, to determine if similar P450c17 mRNAs are found in the human adrenal and testis, we used stringent hybridization conditions to probe an RNA gel blot of human adrenal

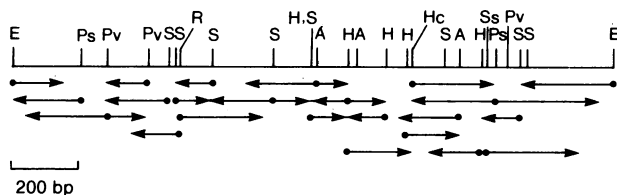


FIG. 2. Sequencing strategy for λ hac17-3. Only restriction sites used in sequencing are shown. Each arrow indicates the direction and extent of deoxy sequencing. Each reaction was performed at least twice. Restriction sites indicated are as follows: A, *Alu* I; E, *EcoRI*; H, *Hae* III; Hc, *HincII*; Ps, *Pst* I; Pv, *Pvu* II; R, *Rsa* I; S, *Sau3A*; Ss, *Sst* I.

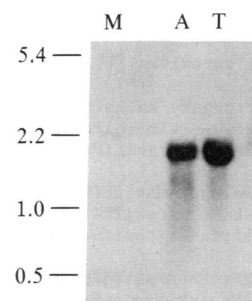


FIG. 3. RNA gel blot. Lane M, molecular size markers [*HindIII*-cleaved bacteriophage PM2: 5.4, 2.2, and 1.0 kilobases (kb)]. Lane A, 35 μ g of total RNA from a human adrenal carcinoma; lane T, 35 μ g of total RNA from a human Leydig cell testicular adenoma. The blot was probed with 32 P-labeled human adrenal P450c17 cDNA.

and testicular RNA with the adrenal P450c17 cDNA clone. As shown in Fig. 3, bands of equivalent migration and hybridization intensity are seen in the RNA from each tissue, suggesting the P450c17 mRNA in each tissue is very similar, if not identical.

To determine if the P450c17 in human testis is identical to that in the human adrenal, the 950-base adrenal P450c17 cDNA was used to screen a human testicular cDNA library under high-stringency conditions. Three clones were identified out of 150,000 screened, and one was sequenced, showing it contained the 1303 bases encoding amino acid 100 through the 76th base of the 3'-untranslated region (Fig. 1). All 1303 bases in the testicular clone are identical to the

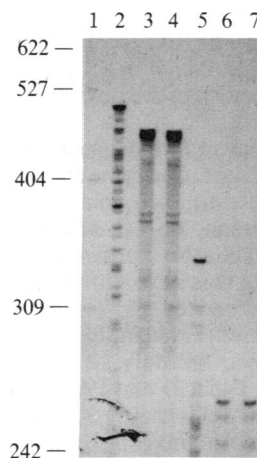


FIG. 4. S1-nuclease protection of 5' and 3' ends of adrenal and testicular P450c17 mRNA. A 5% acrylamide/7 M urea gel was prepared. Lane 1, molecular size markers (*Hpa* II-cleaved pBR322) as indicated in bp. Lane 2, the 498-base 5' probe corresponding to the first 461 bp of the cDNA plus 27 bp of M13 and the 10-bp *EcoRI* linker used in cloning, incubated with 50 μ g of tRNA and no S1 nuclease. The smaller bands in this and other lanes represent fragmented probe due to 32 P decay, as the probes were synthesized 1 week before use. Lanes 3 and 4, S1-nuclease protection of the 5' probe by adrenal and testicular RNA, respectively (these RNAs are the same as those used in Fig. 3). Lane 5, the 377-base 3' probe corresponding to the terminal 297 bp of the cDNA plus the 10-bp *EcoRI* linker and 70 bp of M13 to the *Pvu* II site at base 6356 of M13. This probe migrates at 345 bases on the gel due to "snap-back" self-hybridization of the 18-base poly(A) tail to the 12 thymidine residues at the end of the 3'-untranslated region (Fig. 1). Lanes 6 and 7, S1-nuclease protection of the 3' probe by adrenal and testicular RNA, respectively; the protected fragment is 265 bp long—i.e., equal to the 297 bases of cDNA minus the self-hybridizing "snap-back" region. The existence of this "snap-back" hybridization was proven by S1 digestion of probe unprotected by RNA and electrophoresis on a 15% acrylamide gel (data not shown).

```

P450c17                                     MweLvaLLLLLTLayl
P450c21                                     M-LLlgLLLLLpLlag
P450scc  mlakglpprsvlvkgytflsapreglgrlrvptgegagistrsprfneipspgdngwnlyhfwretgthkvhlhvhqnfqkygpiyreklgnvesvyvidpe

P450c17  fwpkrccpgakyPkSLLs1PLVgslpflprhghmhn-nfKlQkKyGPIYsvrmtKttVivghhqlakEvlIKKgkDFsGRPqmaT---ldiaSnNrkGiafaD
P450c21  arllwnwv---klrSLh1PPLapgfhlLl-qpd1PI-yLLgLtqKfGPIYr1h1G1qDvVvLNskrtieEAmvKKwaDFaGRPeptT---yklvSKNypd1s1GD
P450scc  dvaLl1fksegnpPerfLlPpWVayhqy----yqrPIgvLLKksaaw-----kKDRValNqevmapEAtknflplldavsrdfsvvlhrrikKagsGnysGD

P450c17  sgahWqlHrrLamatfaLfkdgdqkLEkiicQeistlCdm1athnGqsid1Sfpvfvavtnv1s1cFntsyKngdpeLnv1qynvEgfidnlskds1-VdlvPw
P450c21  yS1lWkaHkkLtrsallLGiRdsM-EpVv-eQltQeFCermraqpGtpvaIeeefsl1tcsiIcyltFgdkiKddnLmpayykciqEvlKTWshwsiqiVDv1Pf
P450scc  iSddlfrfafesitnvifGeRqgM1EEVnpea-QrF-----IdalyqmfhtsvpmlLppdl1r1frtKTKWkdhvaawD-----

P450c17  Lk1FPNktLekLKshvkiRnd1lnkiLenyKEkfrsdsitNM1DtLmqakmnsdngnagpdqdSeLLsDnHilttigDIfgaGvETTsvVkwTlaFLLHnPg--
p450c21  LrffFPNpLrrLKqaiekRdhivemqLrqhKESlvagqwrMmDym1Q----gVaqsmeegSgqLlegHvhMaavD11igGTETTant1sWavvFLLHHPe--
P450scc  -----vifSkadiytqNfyweLrQ---kgsVhdyrgmlyrLLgDsk--MsfeDikanvTEmlAggVdtTsmTlQwhlyem

P450c17  -vkkKLyEEiDqnVG---fSRtPtisDRnRL1LLeAtIRVLR1RPVaPmLiPHkaNvdSS1gefavdkGTeVIiNLwALHhnEkeWhqPdqFmPFRFLNpagt
P450c21  -1qqrLQEELDhE1GpgassSRvP-ykDRaRLPLLnAT1aEVLRLRPVvP1a1PHrttrpSS1sgYdIPeGTviIpNLqgaHldETvWerPheFwPDRFLPgKN
P450scc  arnlKvQdmLraEV1aarhqaqgdmatmlqlvPLLKAsIkEtLrLhP1svtLqrylvNd1vlrd-YmIPakT1Vqva1yALgrefTfffdPenFdpTRwLskDKN

P450c17  qlispsvsvyLpFGaGpRsc1GEiLARqELFlimawLLQrFdEv---pDDGqLPSLegiPkvvflidsfkvikVRQawreaqaegst
P450c21  -----sRaLaFGcCapvCLGepLARELEFvvLtrLLQaFtLlp---SgDa-LPSLqplPhcsvilkmqPFq--VR1Qprgmahspgqntq
P450scc  ---ityfRnLgFGwGvRqCLGrr1AeLEmtifLinmLenFrveIqhlSDvGtftnLilmpkpsiftfwPFnqeatQQ

```

FIG. 5. Relationship of human P450c17 to other human steroidogenic P450 sequences. The P450c21 sequence is as reported by Higashi *et al.* (25) and the P450scc sequence is from our laboratory (15). The alignments and insertions of gaps were done manually as described (23). The single letter code (26) is used; capital letters designate residues identical in two or more of the sequences; nonhomologous residues are shown in lower case.

corresponding bases in the adrenal clone. Rescreening these "plaque lifts" under low-stringency conditions did not identify additional hybridizing clones.

To examine the homology between human adrenal and testicular P450c17 mRNAs in the regions not found in the testicular cDNA, we performed S1-nuclease protection experiments. A 5' fragment of the adrenal cDNA encompassing the first 461 bases and a 3' fragment encompassing the last 297 bases of the cDNA were labeled and used as probes. Both human adrenal and testis RNA protect identical fragments of both probes from digestion with S1 (Fig. 4). Thus, it is likely that the uncloned regions of the testicular cDNA are identical to the adrenal cDNA.

Comparison of Human Steroidogenic P450s. P450c17 is the third human steroidogenic enzyme for which the amino acid sequence has been determined, joining P450c21 (refs. 13 and 25, and K. J. Matteson, B.-c.C. and W.L.M., unpublished data) and P450scc (14, 15). To investigate the evolutionary relationships among these cytochromes P450, we compared their amino acid sequences (23, 24). Of the 508 amino acids in P450c17, 147 (28.9%) are identical in P450c21. By contrast, P450c17 and P450scc are only 12.3% homologous (64/521 amino acids) while P450c21 and P450scc are 15.7% homologous (82/521 amino acids); only 36 amino acids (6.9%) are identical in all three proteins (Fig. 5). Somewhat surprisingly, the region of greatest similarity between P450c17 and P450c21 is not the heme binding site at residues 433-453 (13/21 identical amino acids), but is concentrated in residues 346-366 (18/21 identical amino acids). The possible role of this region in binding steroid hormones is uncertain, although the corresponding regions of rat liver P450c and P450d determine the substrate specificity of those enzymes (27).

DISCUSSION

Studies of P450c17 enzymology, amino acid composition, and N-terminal amino acid sequences suggested that pigs might have two similar but distinct organ-specific isozymes of P450c17 (17). We have shown (28) that humans have two or more P450XVII gene sequences and that the gene encoding the cDNA we have sequenced from both human adrenal and testis lies on chromosome 10. As this is the functional gene, it is termed P450XVIIA1 (5). Testicular steroidogenic Leydig cells and adrenocortical cells share a common embryologic origin (29) and adrenal "rest" cells are often found in the testis (29, 30). Thus, the expression of the "adrenal" form of

P450c17 in the testis is not surprising and does not rule out the possibility of another P450XVII gene encoding a testis-specific P450c17 isozyme. Although it is attractive to hypothesize that one of the other gene sequences detected by low-stringency probing (28) might encode another isozyme of P450c17, no additional testicular P450c17 cDNA clones were detected at these low-stringency conditions. The equivalent intensities of the bands of "adrenal" probe protected from S1-nuclease digestion by adrenal and testicular RNA indicates this is a major, if not the only, form of P450c17 mRNA in the human testis.

Of the corresponding 423 amino acids of the human and porcine sequences, only 282 (66.7%) are identical. By contrast, the human and bovine (16) amino acid sequences are 70.5% homologous (358/508 amino acids) (80.1% nucleotide homology); if only those bovine and human amino acids corresponding to the sequenced porcine amino acids are compared, the homology is 72.6% (307/423). It is surprising to find greater differences between the human and porcine sequences than between the human and bovine sequences, as the evolutionary precursors to primates and the precursors to Artiodactyla (including cattle and pigs) diverged about 85 million years ago, while cattle and pigs diverged from each other more recently, about 60 million years ago (26, 31, 32). The closer evolutionary relationship between cattle and pigs would suggest these sequences would be significantly more homologous, yet the corresponding amino acids of these two sequences are only 74.2% homologous (314/423)—i.e., not significantly greater than the 72.6% bovine/human homology. These findings reemphasize that the accumulation of point mutations and amino acid changes is only very roughly linear with evolutionary time, indicating that the "evolutionary clock" may run at variable speeds (25), even for a single gene family.

The nucleotide and amino acid differences among the three species are scattered throughout the sequence and are not clustered in discrete regions, unlike the clustered differences in human and bovine P450scc (15). However, the region between amino acids 160 and 268 contains 50 differences in the human and bovine sequences (53.7% homology). Although only 80 of the 108 residues in this region are known in the porcine sequence (Fig. 1), these 80 residues contain 47 amino acid changes (41.25% homology). If one looks at these same 80 residues in the bovine and human sequences, there are only 36 differences, or 55.0% homology, indistinguishable from the 53.7% homology for the whole region. Thus,

residues 160–268 represent a hypervariable region in the P450c17 sequence. Although the function of this region is not known, it is unlikely to be related to the 17 α -hydroxylase and 17,20 lyase activities common to the P450c17 of all three species.

The amino acid sequence of P450c17 is much more similar to that of P450c21 (28.9%) than to that of P450scc (12.3%). This suggests the P450XVII genes and P450XXI genes (encoding P450c21) diverged in evolution more recently than either did from P450XII gene encoding P450scc. Both P450c17 and P450c21 are bound to endoplasmic reticulum and interact directly with a flavoprotein carrying electrons from NADPH (9–11). By contrast, P450scc (and another steroidogenic enzyme, P450c11) reside in mitochondria and employ different electron-transport intermediates (12). Thus, we suggest that a very early event in *P450* gene evolution was divergence of nuclear genes encoding mitochondrial P450s from genes encoding microsomal P450s. The P450XVII and P450XXI genes then subsequently diverged from the ancestral microsomal *P450* gene while the P450XII gene and P450XI gene (encoding P450c11) diverged from the ancestral mitochondrial gene. The greater homology between microsomal P450s from different species than between microsomal and mitochondrial P450s from the same species strongly supports this model (refs. 24, 25, and this paper).

When residues 160–268 in P450c17 are compared to other human steroidogenic P450 enzymes, far less homology is found in this hypervariable region than in other regions. The human P450c17 and P450c21 sequences have only 23 amino acids (13.6%) in common in this region. This region corresponds closely to the boundaries of exons 4, 5, and 6 of the P450XXI gene (13, 24, 25). By contrast, the amino acid sequence encoded by exon 8 of the P450XXI gene is 50.8% homologous to the corresponding region of P450c17 cDNA. The available sequence data and analyses (13, 24, 25) suggest that evolutionarily ancient *P450* genes had far more exons than their present day counterparts; thus, we suggest that the P450XVII gene will be structurally unique compared to other *P450* genes, but that some exons will bear fairly close resemblance to some P450XXI exons. Such findings would suggest that one of the mechanisms of evolution of modern steroidogenic *P450* genes might be crossover events between homologous sequences.

We thank Raimo Voutilainen for the RNA blot in Fig. 3, Nhung Huyuh for technical assistance, and Carol Dahlstrom for typing. This work was supported by Grant 6-396 from the March of Dimes and by a grant from the University of California, San Francisco, Academic Senate to W.L.M., a predoctoral fellowship from the ARCS Foundation to J.P.-L., and Grant HD 14900 from the National Institutes of Health to J.E.S.

- Zachmann, M., Vollmin, J. A., Hamilton, W. & Prader, A. (1972) *Clin. Endocrinol.* **1**, 369–385.
- Goebelsmann, U., Zachmann, M., Davajan, V., Israel, R. Mestman, J. H. & Mishell, D. R. (1976) *Gynecol. Invest.* **7**, 138–156.
- Kominami, S., Shinzawa, K. & Takemori, S. (1982) *Biochem. Biophys. Res. Commun.* **109**, 916–921.
- Nakajin, S., Shinoda, M. & Hall, P. F. (1983) *Biochem. Biophys. Res. Commun.* **111**, 512–517.
- Nebert, D. W., Adesnik, M., Coon, M. J., Estabrook, R. W., Gonzalez, F. J., Guengerich, F. P., Gunsalus, I. C., Johnson, E. F., Kemper, B., Phillips, I. R., Sato, R. & Waterman, M. R. (1987) *DNA* **6**, 1–12.
- Voutilainen, R., Tapanainen, J., Chung, B., Matteson, K. J. & Miller, W. L. (1986) *J. Clin. Endocrinol. Metab.* **63**, 202–207.
- Di Blasio, A. M., Voutilainen, R., Jaffe, R. B. & Miller, W. L. (1986) *Endocrinology*, Suppl. 118, 182 (abstr. 605).
- Voutilainen, R. & Miller, W. L. (1986) *J. Clin. Endocrinol. Metab.* **63**, 1145–1150.
- Baron, J., Taylor, W. E. & Masters, B. S. S. (1972) *Arch. Biochem. Biophys.* **150**, 105–115.
- Kominami, S., Ochi, H., Kobayashi, Y. & Takemori, S. (1980) *J. Biol. Chem.* **255**, 3386–3394.
- Miller, W. L. (1986) *N. Engl. J. Med.* **314**, 1321–1322.
- Rosenfield, R. L. & Miller, W. L. (1983) in *Hirsutism and Virilism*, eds. Mahesh, V. B. & Greenblatt, R. B. (Wright-PSG, Littleton, MA), pp. 87–119.
- White, P. C., New, M. I. & Dupont, B. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 5111–5115.
- Matteson, K. J., Chung, B., Urdea, M. S. & Miller, W. L. (1986) *Endocrinology* **118**, 1296–1305.
- Chung, B., Matteson, K. J., Voutilainen, R., Mohandas, T. K. & Miller, W. L. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 8962–8966.
- Zuber, M. X., John, M. E., Okamura, T., Simpson, E. R. & Waterman, M. R. (1986) *J. Biol. Chem.* **261**, 2475–2482.
- Nakajin, S., Shonoda, M., Haniu, M., Shively, J. E. & Hall, P. F. (1984) *J. Biol. Chem.* **259**, 3971–3976.
- Nakajin, S., Shinoda, M. & Hall, P. F. (1983) *Biochem. Biophys. Res. Commun.* **111**, 512–517.
- Hawke, D. H., Harris, D. C. & Shively, J. E. (1985) *Anal. Biochem.* **147**, 315–330.
- Haniu, M., Ryan, D. E., Levin, W. & Shively, J. E. (1986) *Arch. Biochem. Biophys.* **244**, 323–337.
- Gubler, U. & Hoffman, B. J. (1983) *Gene* **25**, 263–269.
- Biggin, M. D., Gibson, P. H. & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3963–3967.
- Miller, W. L., Coit, D., Baxter, J. D. & Martial, J. A. (1981) *DNA* **1**, 37–50.
- Chung, B., Matteson, K. J. & Miller, W. L. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 4243–4247.
- Higashi, Y., Yoshioka, H., Yamane, M., Gotoh, O. & Fujii-Kuriyama, Y. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 2841–2845.
- Dayhoff, M. O., ed. (1972) *Atlas of Protein Sequence and Structure* (Natl. Biomed. Res. Found., Washington, DC), Vol. 5.
- Sakaki, T., Shibata, M., Yabusaki, Y. & Ohkawa, H. (1987) *DNA* **6**, 33–40.
- Matteson, K. J., Picado-Leonard, J., Chung, B., Mohandas, T. K. & Miller, W. L. (1986) *J. Clin. Endocrinol. Metab.* **63**, 789–791.
- Dahl, E. V. & Bahn, R. C. (1962) *Am. J. Pathol.* **40**, 587–598.
- Wilkins, L., Fleishmann, W. & Howard, J. E. (1940) *Endocrinology* **26**, 385–395.
- Dayhoff, M. O., ed. (1976) *Atlas of Protein Sequence and Structure* (Natl. Biomed. Res. Found., Washington, DC), Suppl. 2.
- Dayhoff, M. O., ed. (1978) *Atlas of Protein Sequence and Structure* (Natl. Biomed. Res. Found., Washington, DC), Suppl. 3.