Molecular cloning and nucleotide sequence of cDNA for mRNA of mitochondrial cytochrome P-450(SCC) of bovine adrenal cortex
(steroid hormone/monooxygenase/molecular evolution/gene cloning/extrapeptide)

K. Morohashi†, Y. Fuji-Kuriyama*, Y. Okada†, K. Sugawa*, T. Hirose§, S. Inayama§, and T. Omura†

*Department of Biochemistry, Cancer Institute, Japanese Foundation for Cancer Research, 1-37-1 Kami-Ikebukuro, Toshima-ku, Tokyo; †Department of Biology, Faculty of Science, Kyushu University, Fukuoka; and §Pharmaceutical Institute, Keio University School of Medicine, Tokyo, Japan

Communicated by Günter Blobel, March 29, 1984

ABSTRACT We have isolated cDNA clones of the mRNA for cytochrome P-450(SCC), which catalyzes the side-chain cleavage reaction of cholesterol in bovine adrenal cortex mitochondria, by using synthetic oligonucleotides as probes. Sequence analysis of the cloned cDNAs enabled us to deduce the primary structure of the precursor form of P-450(SCC), which consisted of 520 amino acids and contained an extrapeptide of 39 amino acids at the NH₂ terminus. The amino acid sequence from the 40th to 55th amino acid residue in the predicted structure completely coincided with the sequence of the NH₂-terminal portion of purified P-450(SCC). The amino acid composition calculated from the predicted structure showed an excellent agreement with that determined with the purified protein. The extrapeptide of the precursor molecule resembles those of a few nuclear-encoded mitochondrial proteins reported so far. Although P-450(SCC) is a component of mitochondria, comparison of its primary structure with those of other forms of cytochrome P-450 shows that P-450(SCC) is structurally more related to microsomal cytochrome P-450s than to a bacterial cytochrome P-450, P-450cam. A homologous sequence observed with various forms of cytochrome P-450 is also highly conserved in the P-450(SCC) molecule. Only two cysteinyl residues are present in the mature form of P-450(SCC), one of which is located in the middle of the conserved sequence, confirming the function of this cysteinyl residue as the fifth ligand of the heme.

Cytochrome P-450(SCC) [P-450(SCC), side-chain cleavage] plays an important role in steroidogenesis of adrenal cortex, catalyzing the conversion of cholesterol to pregnenolone, the initial rate-limiting reaction in the synthesis of various steroid hormones. Unlike microsomal cytochrome P-450, this hemoprotein is synthesized on cytoplasmic free polyribosomes (1) as a larger precursor (2) and then transported into mitochondria where it functions. The extrapeptide of the precursor molecule is supposed to function as a signal for the transport of the precursor across the membrane barrier of the organelle and is removed to give the mature form during the transport process. Adrenal cortex mitochondria contain another form of cytochrome P-450, P-450(11-β), which is also synthesized as a larger precursor in the cytoplasm (3) to be imported into mitochondria. These mitochondrial cytochrome P-450s are different from their microsomal counterparts in that the former receive electrons via an iron–sulfur protein, adrenodoxin from NADPH-adrenodoxin reductase, whereas the latter receive electrons directly from a flavoprotein, NADPH-cytochrome P-450 reductase (4). Microsomal cytochrome P-450 catalyzes a wide variety of oxidation reactions of endogenous substrates as well as of xenobiotics, including drugs. On the contrary, mitochondrial cytochrome P-450s show no enzymic activities to xenobiotics. Therefore, comparison of the primary amino acid sequences of mitochondrial cytochrome P-450s with those of microsomal cytochrome P-450s will give us a deeper insight into evolutionary aspects of the cytochrome P-450 gene family as well as the structure–function relationships of cytochrome P-450 molecules.

In the present paper, we describe the isolation of cDNA clones of bovine adrenal cortex P-450(SCC) mRNA by using synthetic oligonucleotides as probes. Sequence analysis of the cloned cDNAs gave the complete amino acid sequence of P-450(SCC), including the extrapeptide consisting of 39 amino acids at the NH₂ terminus.

MATERIALS AND METHODS

Materials. Restriction enzymes were purchased from Takara Shuzo Co. (Kyoto, Japan), New England Biolabs, and Bethesda Research Laboratories. Escherichia coli DNA polymerase I (Klenow fragment) and polynucleotide kinase were purchased from Boehringer Mannheim. Bacterial alkaline phosphatase was from Worthington. Terminal deoxynucleotidyl transferase and avian myeloblastosis virus reverse transcriptase were purchased from P-L Biochemicals and Life Sciences (St. Petersburg, FL), respectively. [γ-32P]ATP (500 Ci/mmol; 1 Ci = 37 GBq) and [α-32P]dATP (500 Ci/mmol) were obtained from the Radiochemical Centre (Amersham, England).

mRNA Preparation. Total RNA and poly(A)⁺ RNA were prepared from bovine adrenal cortex as described (5). The mRNA coding for P-450(SCC) was enriched by sucrose density gradient centrifugation of the poly(A)⁺ RNA according to a published procedure (6). P-450(SCC) mRNA in RNA fractions was assayed by immunoprecipitation of in vitro synthesized P-450(SCC) from the translation products by using anti-P-450(SCC) antibody as described (5).

Cloning Procedures. Double-stranded cDNA was synthesized and inserted into the vector pBR322 at the //I site by the G C-tailing method to transform E. coli χ1776 (6). Tetra-cyclo-resistant transformants were screened by in situ colony hybridization with synthetic oligonucleotide probes as described (7). The oligonucleotides were synthesized by the triester method (8). The recombinant plasmids containing P-450(SCC) cDNA were transferred to E. coli HB101 for a large-scale preparation of the plasmid DNA. All of the cloning procedures were conducted in P-2 facilities of the Cancer Institute according to the guidelines for research involving human subjects.

Abbreviations: P-450(SCC), cytochrome P-450 that catalyzes the side-chain cleavage reaction of cholesterol in bovine adrenal cortex; EF-Tu, elongation factor Tu.

†To whom reprint requests should be addressed.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.
recombinant DNA molecules issued by the Ministry of Education, Science, and Culture of Japan.

Restriction Enzyme Mapping. Approximately 1 μg of either the plasmid or insert DNA was cleaved with various restriction endonucleases and was fractionated as described (6).

DNA Sequence Determination. Terminal DNA labeling at 5' protruding ends produced by restriction endonucleases was performed as described (9). (α-32P)ddATP and terminal deoxynucleotidyl transferase were used in labeling 3' protruding ends of appropriate DNA fragments. Sequencing of the end-labeled DNA fragments was conducted as described by Maxam and Gilbert (9).

RESULTS AND DISCUSSION

Screening of P-450(SCC) cDNA Clones by Using Synthetic Oligonucleotides as Probes. In order to synthesize oligonucleotides used as probes, two short peptides, penta- and hexapeptides, were chosen from the partial amino acid sequences of P-450(SCC) reported by Akhrem et al. (10). All possible coding sequences for the two peptides and their corresponding anticoding sequences are shown in Fig. 1. Four mixtures of oligonucleotides (I-a, I-b, II-a, and II-b) were synthesized as the probes (Fig. 1).

Of the four mixtures of synthetic oligonucleotides, I-b and II-a were selected as probes for further work because results

Table 1. Amino acid compositions of P-450(SCC) determined by amino acid analysis of purified P-450(SCC) and from the predicted sequence

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>a</th>
<th>b</th>
<th>II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asn</td>
<td>20</td>
<td>0</td>
<td>42</td>
</tr>
<tr>
<td>Asp</td>
<td>24</td>
<td>0</td>
<td>26</td>
</tr>
<tr>
<td>Thr</td>
<td>23</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>Ser</td>
<td>22</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>Gln</td>
<td>19</td>
<td>0</td>
<td>51</td>
</tr>
<tr>
<td>Glu</td>
<td>34</td>
<td>2</td>
<td>51</td>
</tr>
<tr>
<td>Pro</td>
<td>30</td>
<td>3</td>
<td>28</td>
</tr>
<tr>
<td>Gly</td>
<td>23</td>
<td>8</td>
<td>23</td>
</tr>
<tr>
<td>Ala</td>
<td>22</td>
<td>4</td>
<td>23</td>
</tr>
<tr>
<td>Val</td>
<td>29</td>
<td>3</td>
<td>28</td>
</tr>
<tr>
<td>Met</td>
<td>14</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>Ile</td>
<td>32</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td>Leu</td>
<td>51</td>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td>Tyr</td>
<td>20</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>Phe</td>
<td>30</td>
<td>0</td>
<td>28</td>
</tr>
<tr>
<td>Lys</td>
<td>33</td>
<td>1</td>
<td>34</td>
</tr>
<tr>
<td>His</td>
<td>15</td>
<td>2</td>
<td>15</td>
</tr>
<tr>
<td>Arg</td>
<td>29</td>
<td>3</td>
<td>28</td>
</tr>
<tr>
<td>Trp</td>
<td>9</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Cys</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>481</td>
<td>39</td>
<td>470</td>
</tr>
</tbody>
</table>

I and II indicate the amino acid compositions calculated from the deduced amino acid sequence and determined by amino acid analysis of the purified P-450(SCC) (II), respectively. a and b show the amino acid composition of the mature protein and that of the extra-peptide, respectively.

of RNA blot analysis using poly(A) + RNA from bovine adrenal cortex showed a clear band at the position equivalent to 2000 bases with I-b and II-a probes, but not with I-a and II-b probes (data not shown). Approximately 15,000 tetracycline-resistant transformants were screened by in situ colony hybridization with the I-b probe to obtain 11 positive clones. The II-a probe was used then for a second screening by colony hybridization. Six positive clones were selected by the second screening from the 11 clones isolated in the first screening. Of the 6 clones, 3 clones containing long cDNA inserts, pcP-450(SCC)-1, -2, and -3, were analyzed for their nucleotide sequences.

Nucleotide Sequence of Cloned cDNA for P-450(SCC) mRNA. The cDNA inserts prepared from pcP-450(SCC)-1, -2, and -3, were mapped by using several restriction enzymes as shown in Fig. 2. These cDNA inserts showed the same cleavage maps with one another in their overlapping regions.

Fig. 1. Synthetic oligonucleotides used as probes in screening cDNA clones for P-450(SCC) mRNA. Oligonucleotide probes were synthesized according to the amino acid sequences of peptides I and II reported by Akhrem et al. (10). Their positions in the predicted primary structure of P-450(SCC) are residues 80–85 and residues 186–190, respectively. All possible anti-coding sequences were synthesized as two pools for each of the two peptides.

Fig. 2. Restriction cleavage maps and sequencing strategy for P-450(SCC) cDNAs [pcP-450(SCC)-1, -2, and -3]. Restriction cleavage maps were determined by single or double digestions of the cDNA inserts with various restriction enzymes. The direction of each sequence determination is shown by horizontal arrows, starting at appropriate restriction sites. The number of bases determined in each sequence determination is indicated by the length of the arrows. Only restriction sites used for sequence analysis are shown. Acc, Acc 1; Ava, Ava 1; Bam, BamHI; Bgl, Bgl II; Fok, Fok I; Hind, HindIII; Nco, Nco I; Pst, Pst I; Pvu, Pvu II; Sau, Sau3A; bp, base pairs.
and the difference in the predicted sequence of P-450(SCC). Only one of Akhemk's sequences could not be assigned in the predicted sequence. From these considerations, we conclude that the cloned cDNAs were derived from the mRNA for P-450(SCC). This conclusion was also supported by an excellent agreement between the reported amino acid composition of purified P-450(SCC) (11) and that calculated from its predicted primary sequence (Table 1).

The predicted primary structure of P-450(SCC) indicates the presence of an extrapeptide of 39 amino acids at the NH2 terminus. The calculated molecular weight of the predicted extrapeptide is in good agreement with that estimated from the difference in the molecular weight between the precursor and the mature form of P-450(SCC) (2), suggesting that the methionine at the NH2 terminus of the predicted sequence of the precursor is the correct initiator methionine of the precursor molecule.

Consequently, it was concluded that the precursor of P-450(SCC) consists of 520 amino acids, in which 39 amino acids at the NH2-terminal region constitute the extrapeptide. The Mfs of the precursor and the mature form of P-450(SCC) were calculated to be 60,322 and 56,387, respectively. These values are consistent with those estimated from their mobilities in NaDdSO4/polyacrylamide gel electrophoresis. Following alignment of the extrapeptide with other P-450s, it was apparent that the extrapeptide was conserved in the P-450 family and that it possibly functions as a nuclear receptor for the enzyme. The existence of such a receptor is of great interest in the regulation of P-450 expression. The sequences of the extrapeptide are shown in Table 2.
The conserved nature of amino acid residues from the conserved region (HR2; ref. 19) of cytochromes P-450 \(_b\), P-450 \(_c\), P-450 \(_d\), and P-450 \(_am\) are taken from refs. 12, 13, 15, 14, and 17, respectively. The amino acids conserved in more than four forms of cytochrome P-450 are enclosed in boxes. Numbers of cysteine residues indicate the numbers of amino acid residues from the initiator methionine. Numbers under the cysteine residues in parentheses indicate the numbers of amino acid residues from the NH\(_2\) terminus of mature proteins.

lowing the termination codon TGA, a trailer sequence consisting of 251 nucleotides continues. In the trailer sequence, a typical poly(A) addition signal is present 242 bases downstream from the termination codon, followed by attachment of a poly(A) stretch 13 bases further downstream from the signal.

**Structural Comparison of Mitochondrial P-450(SCC) with Other Forms of Cytochrome P-450.** Mitochondrial P-450(SCC) shares many physicochemical and enzymatic properties in common with microsomal and bacterial cytochrome P-450s. On the other hand, mitochondrial cytochrome P-450s, including P-450(SCC), receive reducing equivalents for hydroxylation reactions from NADPH via a reducing system consisting of a flavoprotein and an iron–sulfur protein, whereas microsomal cytochrome P-450s receive electrons from NADPH via a flavoprotein (4). In this regard, mitochondrial cytochrome P-450s resemble bacterial ones, which also receive electrons via a flavoprotein and an iron–sulfur protein. Therefore, comparison of the primary structure of P-450(SCC) with those of microsomal (12-16) and bacterial cytochrome P-450s (17) would be interesting in terms of evolutionary aspects. The results of a comparison among these structures are shown in Table 2. The alignment scores indicate degrees of sequence relatedness of P-450(SCC) with other types of cytochrome P-450 in standard deviation units. All scores are well above the standard threshold of 3.0, indicating a statistically significant relatedness of P-450(SCC) to the other types of cytochrome P-450 listed. Furthermore, these values, together with percentage values of matched residues, show that P-450(SCC) is more structurally related to animal microsomal cytochrome P-450s than to bacterial cytochrome P-450 in spite of an apparent compositional similarity between the mitochondrial cytochrome P-450 system and the bacterial one as described above. This seems to suggest that the divergence leading to bacterial cytochrome P-450 \(_am\) and a eukaryotic prototype of cytochrome P-450 occurred earlier than the divergence of microsomal cytochrome P-450 and mitochondrial P-450(SCC).

**Possible Heme-Binding Portions in the P-450(SCC) Peptide.** There has been some controversy concerning the heme-binding site in the cytochrome P-450 molecules. By comparing the primary sequences of various cysteine-containing portions in microsomal and bacterial cytochrome P-450 molecules, we previously proposed that a short segment (designated as HR2), which locates close to the COOH terminus and contains a single cysteine residue in its middle, is the heme-binding site (14, 19). On the other hand, from results of homologous cytochrome residues with monooiodoacetamide, others claimed that another part of the molecule, located in the NH\(_2\)-terminal half, plays such a role (20, 21). There has been general agreement that a thiolated cysteine coordinates to the heme of cytochrome P-450 as the fifth ligand. Since the optical and magnetic properties of various forms of cytochrome P-450 are very similar, the primary structure as well as the conformation around the heme seems to be well conserved among cytochrome P-450s. The mature form of P-450(SCC) contains only two cysteine residues in the molecule and this fortuitous situation allows us to draw a definite conclusion as to the heme-binding portion in the cytochrome P-450 molecule on the assumption that the thiol group of a cysteine residue is the fifth ligand of the heme. As shown in Fig. 3, the two cysteine residues are located at the 303rd and 461st positions from the NH\(_2\) terminus of the precursor molecule, respectively, and they are proximal to the COOH terminus. We find no cysteine residue at or near the equivalent position in the NH\(_2\)-terminal half of the P-450(SCC) molecule, as other groups claimed for bacterial cytochrome P-450 \(_am\) (20) and rabbit cytochrome P-450 LM2 (21). In addition, the local amino acid sequence around the 461st cysteine of P-450(SCC) is clearly homologous to the conserved sequences at the corresponding positions of various forms of cytochrome P-450, as shown in Fig. 4. The local sequence around the 303rd cysteine of P-450(SCC) shows no homology to the conserved sequences. Therefore, we conclude that the cysteine residue in the conserved sequences that is found in all cytochrome P-450 molecules so far examined is most likely to be the fifth ligand of the heme.

**Structure of the Extrapeptide of P-450(SCC).** The presence of an extrapeptide consisting of 39 amino acids in the predicted primary structure of pre-P-450(SCC) has confirmed the previous observation that P-450(SCC) was synthesized as a larger precursor in an in vitro translation system directed by isolated mRNA (2). Nuclear gene-encoded mitochondrial....

---

**Fig. 4.** Comparison of possible heme-binding peptide sequences in several forms of cytochrome P-450. The local amino acid sequences in the conserved region (HR2; ref. 19) of cytochromes P-450\(_b\), P-450\(_c\), P-450\(_d\), and P-450\(_am\) taken from refs. 12, 13, 15, 14, and 17, respectively. The amino acids conserved in more than four forms of cytochrome P-450 are enclosed in boxes. Numbers of cysteine residues indicate the numbers of amino acid residues from the initiator methionine. Numbers under the cysteine residues in parentheses indicate the numbers of amino acid residues from the NH\(_2\) terminus of mature proteins.

**Fig. 5.** Homologous sequence in the extrapeptides of P-450(SCC), cytochrome \(_c\) peroxidase, and EF-Tu. The amino acid sequences of cytochrome \(_c\) peroxidase and EF-Tu were taken from refs. 23 and 26, respectively. The enclosed amino acids indicate the common amino acids between P-450(SCC) and at least one of the two proteins. Dashes show gaps introduced to obtain maximal homology.
drial proteins are generally synthesized as larger precursors in the cytoplasm to be transported into mitochondria (22). However, information about the primary structures of the extrapeptides of the precursor proteins is still limited. The presence of the extrapeptide at the NH₂-termini of precursor proteins and their partial or complete amino acid sequences have been shown for yeast and Neurospora enzymes, cytochrome c peroxidase (23), proteolipid subunit of ATPase (24), ATPase inhibitor (25), and elongation factor Tu (EF-Tu) (26). Although the lengths of the extrapeptides appear to be variable from 68 amino acid residues for cytochrome c peroxidase (23) to 37 amino acid residues for EF-Tu (26), a significant homology was noticed (26) between cytochrome c peroxidase and EF-Tu in the primary structures of the extrapeptides. We have compared the primary structure of the extrapeptide of pre-P-450(SCC) with those of the two yeast mitochondrial proteins. These protein molecules have different localizations in mitochondria; P-450(SCC) is an inner-membrane protein, whereas cytochrome c peroxidase and EF-Tu are soluble proteins in the inter-membrane space and matrix of the organelle, respectively. As shown in Fig. 5, the initial 20 amino acid residues of pre-P-450(SCC) show a distinct homology with those of EF-Tu and cytochrome c peroxidase—that is, basic amino acids such as arginine and lysine are distributed rather periodically in every 4 or 5 amino acid residues. It is likely that this conserved portion of the extrapeptides between yeast and animal enzymes functions as a common signal for the transport of precursor proteins across the membrane barrier to mitochondria, although the enzymes compared here are different in their sub mitochondrial localization.11

11The predicted extrapeptide sequence of pre-P-450(SCC) is in agreement with the partial radiosequencing data (Met-Leu-X-X-Leu-X-Leu-) on the NH₂-terminal portion of in vitro translated P-450(SCC) reported recently by Ogishima et al. (27).

We express sincere thanks to Dr. O. Gotoh (Saitama Cancer Center Institute, Saitama, Japan) for computer analysis of the primary sequences of cytochrome P-450. This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Science, and Culture of Japan, funds obtained under the Life Science Project from the Institute of Physical and Chemical Research, Japan, and funds from the Princess Takamatsu Cancer Research Fund.