(London) New Biol. 229, 10–13.
Stefanowska, B., & Ledochowski, A. (1978) Rocz. Chem. 46, 1637.
Wakelin, L. P. G. (1986) Med. Res. Rev. 6, 275–340.
Wirth, M. Buchardt, O., Koch, T., Nielsen, P. E., & Nordén, B. (1987) J. Am. Chem. Soc. (submitted for publication).
Wu, H.-M., & Crothers, D. M. (1984) Nature (London) 308, 509–513.

# Microheterogeneity in the Major Phenobarbital-Inducible Forms of Rabbit Liver Microsomal Cytochrome P-450 As Revealed by Nucleotide Sequencing of Cloned cDNAs†

Masayuki Komori,‡ Yoshio Imai,* Susumu Tsunasawa, and Ryo Sato
*Institute for Protein Research, Osaka University, Suita, Osaka 565, Japan*
*Received April 13, 1987; Revised Manuscript Received July 30, 1987*

ABSTRACT: We have isolated one full-length cDNA clone, termed pHP1, and a number of clones of shorter insert lengths, tentatively called b14, b46, etc., all encoding phenobarbital- (PB-) inducible forms of rabbit liver microsomal cytochrome P-450, and determined their nucleotide sequences. The polypeptides encoded by these cDNAs can be classified into five types, represented by HP1, b14, b46, b52, and b54, the deduced amino acid sequences of which are more than 95% similar to one another. Amino acid differences among them total 24 positions, which are distributed over the entire sequence, in contrast to the microheterogeneity observed in two PB-inducible rat liver microsomal cytochromes P-450 (P-450b and P-450e). The primary structure deduced for the HP1 protein is 97% similar to that determined for rabbit P-450 LM2 (form 2), which has been purified by Coon and co-workers [van der Hoeven, T. A., Haugen, D. A., & Coon, M. J. (1974) *Biochem. Biophys. Res. Commun. 60*, 569–675; Haugen, D. A., & Coon, M. J. (1976) *J. Biol. Chem. 251*, 7929–7939] as the major PB-inducible form of rabbit liver microsomal cytochrome P-450. The amino acid sequence of P-450$_1$, which we have purified as the major PB-inducible rabbit liver cytochrome P-450, was partially determined with the sequence reported for P-450 LM2 as a reference. The two sequences are closely similar to each other, but at least two amino acid differences can be detected between them. The primary structure predicted from the nucleotide sequence of b14 cDNA, which covers only the carboxyl-terminal 40% of the entire sequence, is identical with that determined for P-450$_1$. It is concluded that the liver of PB-treated rabbits expresses at least five distinct but closely similar forms of cytochrome P-450, all of which are encoded by genes belonging to the rabbit P450IIB (major PB-inducible) subfamily.

Cytochromes P-450 play a crucial role in the oxidative metabolism of a large variety of xenobiotics as well as a number of endogenous substrates. It is now established that multiple forms of cytochrome P-450 occur in liver microsomes and administration of drugs to animals induces the synthesis of a specific form(s) of the cytochrome depending on the drug administered. The multiplicity and inducibility have been recognized and confirmed by purification of distinct forms of cytochrome P-450 from liver microsomes of untreated and variously drug-treated animals and in recent years also by isolation and structural analysis of cDNA and genomic clones encoding different forms of the cytochrome.

The first successful isolation of homogeneous preparations of hepatic microsomal cytochrome P-450 was reported by two groups. Thus, we purified a form, called P-450$_1$, from liver microsomes of phenobarbital- (PB-)[1] treated rabbits (Imai & Sato, 1974; Imai et al., 1980), whereas Coon and co-workers isolated a preparation, termed P-450 LM2 or form 2 (hereafter will be referred to as LM2), from the same source (van der Hoeven et al., 1974; Haugen & Coon, 1976). Since the molecular, spectral, and catalytic properties of P-450$_1$ and LM2 were hardly distinguishable from each other, we concluded that they are identical with each other and represent the major PB-inducible form in rabbit liver microsomes (Imai et al., 1980). On the other hand, two forms of cytochrome P-450, called P-450b and P-450e, have been purified from liver microsomes of PB-treated rats (Ryan et al., 1979, 1982). These two forms have been shown to be immunochemically indistinguishable and exhibit only minor functional differences (Ryan et al., 1982; Vlasuk & Walz, 1982). Furthermore, their amino acid sequences, deduced from cDNA nucleotide sequences, have been found to be 97% similar to each other (Fujii-Kuriyama et al., 1982; Mizukami et al., 1983a). Despite such close similarities, evidence has been presented that they are actually encoded by different genes (Mizukami et al., 1983a; Suwa et al., 1985).

In this study we examined whether microheterogeneity can also be found in major PB-inducible forms of cytochrome P-450[2] in rabbit liver. For this purpose, we isolated one

† This work was supported in part by a grand-in-aid (5806002) from the Ministry of Education, Science and Culture of Japan.
* Author to whom correspondence should be addressed.
‡ Present address: Faculty of Pharmaceutical Sciences, Hokkaido University, Sapporo 060, Japan.

[1] Abbreviations: PB, phenobarbital; LM2, P-450 LM2 or rabbit form 2; SDS, sodium dodecyl sulfate; Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride; EDTA, ethylenediaminetetraacetic acid.

0006-2960/88/0427-0073$01.50/0 

full-length cDNA clone and a number of clones of shorter insert lengths, all coding for P-450$_1$- or LM2-like species of rabbit liver cytochrome P-450. Nucleotide sequencing of these cDNAs and partial determination of the primary structure of purified P-450$_1$ have shown that microheterogeneity in fact exists in the major PB-inducible forms of rabbit liver cytochrome P-450, all of which are encoded by genes belonging to the cytochrome P-450 gene subfamily called P450IIB according to the nomenclature recently recommended (Nebert et al., 1987).

## Materials and Methods

*Materials.* A bacterial clone harboring the plasmid pcP-450pb4 that contains a cDNA sequence for rat P-450b (Fujii-Kuriyama et al., 1982) was kindly supplied by Dr. Y. Fujii-Kuriyama of the Cancer Institute, Japanese Foundation for Cancer Research, Tokyo. The recombinant plasmid was purified as described by Taniguchi et al. (1980). The following materials were obtained from the sources indicated: avian myeloblastosis virus reverse transcriptase (Life Science); Klenow fragment of *Escherichia coli* DNA polymerase I and *Pst*I-cut, oligo(dG)-tailed pBR322 (Bethesda Research Laboratories); $S_1$ nuclease (Sankyo); oligo(dT)$_{12-18}$ and oligo-(dT)–cellulose (P-L Biochemicals); M13 cloning, sequencing, and nick translation kit (Amersham International); restriction enzymes (New England Biolabs, Nippon Gene, Takara Shuzo, and Bethesda Research Laboratories); nitrocellulose filters (S & S). The following labeled compounds were from Radiochemical Centre, Amersham: [$\alpha$-$^{32}$P]dCTP (3000 Ci/mmol); [$\gamma$-$^{32}$P]ATP (3000 Ci/mmol), and [$\alpha$-$^{35}$S]dATP$\alpha$S (650 Ci/mmol).

*PB Treatment of Rabbits.* Male white rabbits (~2 kg) were maintained on a laboratory chow and injected i.p. with PB dissolved in saline (120 mg per capita) 20 h prior to sacrifice (for isolation of RNA) or daily for 7 days (for purification of P-450$_1$). The animals were fasted for 24 h and then killed.

*Isolation of RNA.* Total RNA was extracted from liver microsomes with SDS–phenol–chloroform (Palmiter, 1974). Poly(A$^+$) RNA, prepared by oligo(dT)–cellulose column chromatography (Aviv & Leder, 1972), was size-fractionated on a 5–25% sucrose gradient in 10 mM Tris-HCl (pH 7.5) containing 1 mM EDTA and 0.2% SDS. RNA in each fraction was translated in a wheat germ cell free system in the presence of [$^{35}$S]methionine (Erickson & Blobel, 1983), and the proteins synthesized were immunoprecipitated with guinea pig antibodies raised against purified P-450$_1$ as described (Mihara & Blobel, 1980). The immunoprecipitate was analyzed by SDS–polyacrylamide gel electrophoresis and subsequent fluorography. Fractions rich in mRNA encoding P-450$_1$-like proteins were pooled and used for cDNA synthesis. Upon RNA blot hybridization, these fractions were shown to hybridize with the insert of pcP-450pb4, a cDNA clone for rat P-450b.

*Construction and Screening of cDNA Libraries.* The first strand of cDNA was synthesized from the enriched mRNA by reverse transcriptase with oligo(dT)$_{12-18}$ as a primer (Kawajiri et al., 1983). Double-stranded cDNA (ds-cDNA) was generated by two methods. In method I, ds-cDNA was synthesized by reverse transcriptase with the 3′-end hairpin loop of the first strand as a primer, followed by Klenow fragment (Kawajiri et al., 1983). The hairpin loop was then removed by $S_1$ nuclease. In method II, the RNA strand of the RNA–DNA hybrid was replaced by DNA with the aid of the repair system of Okayama and Berg (1982). The ds-cDNAs synthesized by both methods were size-fractionated by Sephadex CL-4B gel filtration, and fractions containing cDNAs longer than 500 bp were pooled. The pooled cDNAs were oligo-(dC)-tailed (Gordon et al., 1978) and annealed with *Pst*I-cut, oligo(dG)-tailed pBR322 (Clarke & Carbon, 1975). The recombinant plasmids were used to transform *E. coli* DH1 or HB101 to obtain a cDNA library. The library prepared by method I was screened by in situ colony hybridization (Grunstein & Hogness, 1975) with a 390-bp *Bgl*II fragment or a 183-bp *Hin*dIII/*Bgl*II fragment of pcP-450pb4 as a probe. The prewashed filter was hybridized with the probe (labeled by nick translation) and then washed twice with 2× SSC and 0.1% sodium dodecyl sarcosinate at 55 °C for 30 min. The plasmids in positive transformants were confirmed by Southern blot analysis with the same probe, and the cDNA inserts of several strongly positive clones were preliminarily sequenced. Since the insert of b11[3] was found to contain a sequence identical with that of the HR2 region of rabbit LM2, the cDNA library was rescreened by in situ colony hybridization with the b11 insert as a probe under more stringent conditions (0.1× SSC and 0.1% sodium dodecyl sarcosinate at 65 °C for 30 min). The cDNA library constructed by method II was also screened with the b11 insert as a probe under the same stringent conditions as above. Positive transformants thus selected were again confirmed by Southern blot analysis.

*Filter Blot Hybridization.* For Southern blot analysis, plasmid cDNA was digested with an appropriate restriction enzyme, subjected to agarose gel electrophoresis, and transferred to a nitrocellulose filter (Southern, 1975). The filter-bound DNA was hybridized with a nick-translated cDNA probe and washed as described (Maniatis et al., 1975). For RNA blot analysis, partially purified mRNA was subjected to agarose gel electrophoresis under denaturing conditions (Kawajiri et al., 1983) and transferred to a nitrocellulose filter. The filter-bound RNA was hybridized with a nick-translated cDNA probe and washed as described (Thomas, 1980; Kawajiri et al., 1983).

*Restriction Mapping and Nucleotide Sequencing.* For restriction mapping, the DNA sample was digested with various restriction enzymes, and the sizes of the fragments thus produced were analyzed by agarose gel (1–1.5%) electrophoresis. The nucleotide sequences of cDNA inserts were determined by the dideoxy termination method (Sanger et al., 1980) after their restriction fragments were subcloned in vectors M13mp10 and M13mp11. Sequencing was conducted with [$\alpha$-$^{35}$S]-dATP$\alpha$S as a radioactive label, essentially according to the instructions provided by the kit supplier. The sequences of several regions, where the dideoxy method gave ambiguous results, were confirmed by the chemical cleavage method (Maxam & Gilbert, 1980).

*Purification of P-450$_1$.* P-450$_1$ was purified from liver microsomes of a PB-treated rabbit as described (Imai et al., 1980). The purified protein gave a single protein band on SDS–polyacrylamide gel electrophoresis and was eluted as a single symmetrical peak on HPLC using a Baker bond wide-pore butyl column (0.46 × 25 cm, $C_4$, 330A; U. T. Baker

[2] From purification studies, there is little doubt that P-450$_1$ or LM2, which are closely related and may be regarded as allelic variants, is the most abundant form of cytochrome P-450 in liver microsomes of PB-treated rabbits. In this paper, all cytochromes P-450, the primary structures of which are more than 95% similar to that of P-450$_1$ or LM2, are called "major PB-inducible forms of cytochrome P-450" disregarding their actual levels in the liver of PB-treated rabbits.

[3] In this paper, cDNA clones that do not contain a full-length coding sequence for a cytochrome P-450 are tentatively termed b11, b14, etc.

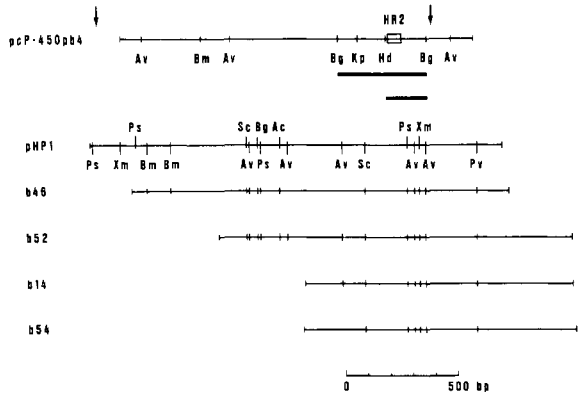


FIGURE 1: Restriction maps of the inserts of five representative cDNA clones for PB-inducible forms of rabbit liver microsomal cytochrome P-450. The map of rat P-450b cDNA in pcP-450pb4 is also shown at the top. The *Bgl*II fragment and the *Hin*dIII/*Bgl*II fragment of pcP-450pb4 cDNA used as probes are shown by closed boxes. HR2, the conserved heme-binding region, is indicated by an open box. The positions of both initiation and termination codons are indicated by vertical arrows. Ac, *Acc*I; Av, *Ava*I; Bg, *Bgl*II; Bm, *Bam*HI; Kp, *Kpn*I; Hd, *Hin*d III; Ps, *Pst*I; Pv, *Pvu*II; Sc, *Sac*I; Xm, *Xma*I.

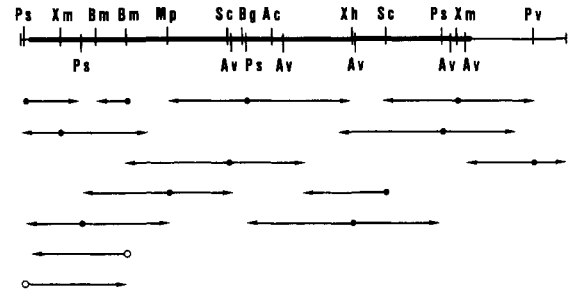


FIGURE 2: Sequencing strategy for the cDNA insert of pHP1. A restriction map of the pHP1 insert is shown at the top. The *arrows with closed circle* indicate the fragments that were subcloned in either M13mp10 or M13mp11 and sequenced by the dideoxy termination method in the direction and to the extent shown. The fragments sequenced by the chemical cleavage method are indicated by the *arrows with open circle*. Mp, *Msp*I; Xh, *Xho*II. The other abbreviations of restriction enzymes are the same as in Figure 1.

Chemical) when elution was conducted with a linear gradient of 1–60% 2-propanol/acetonitrile (7:3 v/v) in 0.1% trifluoroacetic acid for 1 h at a flow rate of 1 mL/min.

*Partial Determination of Primary Structure of P-450*$_1$. After Emulgen 913 was removed from the purified P-450$_1$ preparation as described previously (Imai et al., 1980), the protein was digested with *Achromobacter* lysylendopeptidase (Wako Pure Chemicals) in 10 mM Tris-HCl (pH 9.0) at 37 °C for 6 h at a substrate to peptidase molar ratio of 100:1. The peptide fragments produced were separated by HPLC using a Baker bond wide-pore butyl column. Elution was conducted with a linear gradient of 0–60% 2-propanol/acetonitrile (7:3 v/v) in 0.1% trifluoroacetic acid at a flow rate of 1 mL/min. The peptides thus separated were hydrolyzed with 5.6 N HCl at 110 °C in vacuo for 24 h, and the hydrolysates were analyzed for amino acid compositions in a Hitachi 835S amino acid analyzer. The amino-terminal residue for each peptide was determined by the method of Gray (1972). Amino acid sequences were determined in an Applied Biosystems 470S protein sequencer with a standard program (Hunkapiller et al., 1983). PTH-amino acids were determined by reverse-phase HPLC with isocratic elution as described (Tsunasawa et al., 1985).

## RESULTS

*cDNA Cloning and Restriction Mapping.* A cDNA library was constructed by method I (see Materials and Methods) with enriched mRNA from liver microsomes of a PB-treated rabbit as template and screened for clones carrying cytochrome P-450 cDNAs by in situ colony hybridization. For this purpose, a 390-bp *Bgl*II fragment or a 183-bp *Hin*dIII/*Bgl*II fragment of pcP-450pb4 (a cDNA clone for rat P-450b) was used as a probe (see Figure 1), because these fragments contained the sequence encoding the HR2 region that has been shown to be highly conserved among the various species of cytochrome P-450 so far sequenced (Gotoh et al., 1983; Morohashi et al., 1984). Upon Southern hybridization analysis of the 16 positive clones thus selected, b11 hybridized with the probe most strongly. Sequence analysis of a 300-bp *Pst*I fragment of b11 cDNA indicated that it contained an HR2 sequence that is identical with that of LM2, the primary structure of which has been determined by Tarr et al. (1983). Preliminary sequencing of the other clones giving weaker hybridization signals, such as b32 and b43, indicated that their sequences were considerably different from that of LM2. The sequences of some of them will be reported in the following paper (Imai et al., 1988). Rescreening of the same library was then carried out under more stringent conditions with the b11 insert as a probe, resulting in the isolation of many clones, such as b14 and b52. These clones as well as b11 were, however, found to contain inserts that were shorter than those expected for cytochrome P-450 mRNAs. To obtain clones carrying full-length cDNAs for PB-inducible forms of cytochrome P-450, a second library was constructed by method II. Screening of this library with the b11 insert as a probe under the same stringent conditions as above led finally to the isolation of several cDNA clones, such as pHP1 and b46, which contained inserts of nearly 2 kbp long.

Figure 1 shows the restriction maps of cDNA inserts of five representative clones isolated in this study. pHP1, b52, and b14 cDNAs exhibited the same restriction map in their overlapping regions, but preliminary sequencing indicated the occurrence of nucleotide differences among the three cDNAs at 20–40 nucleotides upstream from the sequence corresponding to the HR2 region. The maps of b46 and b54 cDNAs were identical with each other and similar to those of the above three cDNAs, though one *Ava*I site was missing in the maps of the former group. As will be reported below, the nucleotide sequences of b46 and b54 differ from each other. Figure 1 shows further that the pHP1 insert extends furthest toward the 5′-direction and was expected to contain the initiation codon for a cytochrome P-450.

*Sequence Analysis of Cloned cDNAs.* The strategy adopted for sequence analysis of pHP1 cDNA is shown in Figure 2, and the complete nucleotide sequence thus determined is presented in Figure 3. This cDNA (1802 bp) contains a 37-bp 5′-flanking region and a 289-bp 3′-noncoding sequence [which lacks a poly(A) sequence] in addition to an open reading frame encoding a protein consisting of 491 amino acid residues. The primary structure deduced for this protein is also shown in Figure 3. This structure contains two highly conserved regions called HR1 (residues 142–160) and HR2 (residues 429–449) (Gotoh et al., 1983; Morohashi et al., 1984). In the HR2 region, the conserved cysteinyl residue that binds to the heme (Gotoh et al., 1983; Poulos et al., 1985) is found at residue 436. The molecular weight calculated for the deduced primary structure is 55 682, a value that is considerably higher than that estimated by SDS–polyacrylamide gel electrophoresis for P-450$_1$ (49 000) (Imai et al., 1980). Similar discrepancies have been reported for other microsomal cytochromes P-450 (Fujii-Kuriyama et al., 1982; Kawajiri et al., 1983). The amino acid sequence predicted from the nucleotide sequence of pHP1

Table I: Amino Acid Differences Observed in the Forms of Cytochrome P-450 Belonging to the Rabbit P450IIB Gene Subfamily[a]

| residue | HP1 | b14 | b46 | b52 | b54 | LM2 | $P\text{-}450_1$ |
|---|---|---|---|---|---|---|---|
| 20 | Leu (CTC) | | | | | Phe | Phe |
| 35 | Pro (CCC) | | | | | Ser | Ser |
| 53 | His (CAC) | | Arg (=G=) | | | Arg | nd |
| 57 | Gln (CAG) | | Arg (=GC) | | | Arg | nd |
| 104 | Leu (TTG) | | Met (A==) | | | Val | Val |
| 114 | Thr (ACC) | | Ile (=T=) | | | Ile | Ile |
| 174 | Val (GTC) | | Val (===) | | | Ile | nd |
| 209 | Thr (ACC) | | Ile (=T=) | Ile (=T=) | | Ile | nd |
| 221 | Ser (TCG) | | Ser (===) | Pro (C==) | | Pro | nd |
| 248 | Ser (AGC) | | Ser (===) | Thr (=C=) | | Ser | Ser |
| 251 | Arg (AGG) | | Lys (=A=) | Lys (=A=) | | Lys | Lys |
| 254 | Glu (GAA) | | Ala (=C=) | Ala (=C=) | | Ala | Ala |
| 286 | Gln (CAG) | | Gln (===) | Arg (=G=) | | Gln | Gln |
| 294 | Thr (ACG) | | Ser (T==) | Thr (===) | | Ser | Ser |
| 314 | Leu (CTG) | Leu (===) | Leu (===) | Met (A==) | Met (A==) | Met | Leu |
| 363 | Ile (ATC) | Ile (===) | Ile (===) | Val (G==) | Ile (===) | Ile | Ile |
| 365 | Leu(TTA) | Phe (==C) | Phe (==C) | Phe (==C) | Phe (==C) | Phe | Phe |
| 367 | Val (GTG) | Val (===) | Val (===) | Ala (=C=) | Val (===) | Val | Val |
| 370 | Met (ATG) | Thr (=C=) | Thr (=C=) | Met (===) | Thr (=C=) | Thr | Thr |
| 417 | Asp (GAC) | Asn (A==) | Asn (A==) | Asp (===) | Asn (A==) | Asn | Asn |
| 420 | Leu (CTG) | Met (A==) | Leu (===) | Leu (===) | Leu (===) | Leu | Met |
| 423 | Ile (ATT) | Asn (=A=) | Asn (=A=) | Asn (=A=) | Asn (=A=) | Asn | Asn |
| 425 | Gly (GGC) | Gly (===) | Cys (T==) | Gly (===) | Gly (===) | Gly | Gly |
| 480 | Val (GTG) | Val (===) | Val (===) | Val (===) | Leu (C==) | Val | Val |

[a] The sequences for HP1, b14, b46, b52, and b54 were deduced from their nucleotide sequences, whereas those for LM2 (Tarr et al., 1983) and $P\text{-}450_1$ were chemically determined. The same nucleotide as found in HP1 is shown by a sign of equality (=). nd, not determined. Residue numbers covered by the respective polypeptides are as follows: HP1, 1–491; b14, 301–491; b46, 52–491; b52, 182–491; b54, 297–491; LM2, 1–491; $P\text{-}450_1$, 1–491.

cDNA is 97% similar to that determined for LM2 (Tarr et al., 1983), and 17 amino acid differences are detectable between them (see Table I).

The shorter inserts of the other cDNA clones that seemed to encode $P\text{-}450_1$- or LM2-like proteins were also sequenced. It was thus found that all the primary structures deduced from their nucleotide sequences can be classified into five groups, represented by pHP1, b14, b46, b52, and b54, though these clones, except for pHP1, do not cover the entire cytochrome P-450 sequence. Amino acid differences noticed in their deduced primary structures are summarized in Table I, whereas silent nucleotide substitutions are shown in Table II. The amino acid sequences deduced for the five groups of protein are more than 95% similar to one another. None of these structures are identical with that determined for LM2 (Tarr et al., 1983), but the structure of LM2 is very similar to that of the b14 polypeptide. Only two amino acid differences can be seen between them (Met-314 in LM2 to Leu and Leu-420 in LM2 to Met) in their overlapping region, i.e., the carboxyl-terminal 191-residue segment. Amino acid differences among the five groups and LM2 total 24 positions that are distributed over the entire primary structure, in contrast to the case of rat P-450b and P-450e where the differences are seen only in the carboxyl-terminal region (Figure 4).

*Amino Acid Sequencing of* $P\text{-}450_1$. To see if any of the cDNAs cloned in this study corresponds to mRNA encoding $P\text{-}450_1$, the amino acid sequence of this protein was partially analyzed with the primary structure determined for LM2 (Tarr et al., 1983) as a reference. Purified $P\text{-}450_1$ was digested with *Achromobacter* lysylendopeptidase, and 16 peptide fragments (out of 21 expected from the specificity of the endopeptidase) were isolated in almost pure form by HPLC. These peptides were identified by determining their amino acid compositions and amino-terminal residues. By a comparison with the primary structure of LM2, it was possible to assign the isolated 16 peptides to peptides 1, 2, 5–7, and 11–21 in Figure 5. The carboxyl-terminal 191-residue portion, which is present in common in all the five representative primary structures predicted, was completely recovered as peptides 14–21. Since the amino acid compositions of peptides 14, 17, 19, 20, and 21 seemed to be somewhat inconsistent with those expected from the sequence of LM2, the amino acid sequences of these five peptides were determined. It was thus found that the sequence of $P\text{-}450_1$ is identical with that deduced for b14

Table II: Nucleotide Substitutions That Result in No Amino Acid Changes in the Nucleotide Sequences of cDNAs Coding for Cytochromes P-450 Belonging to the Rabbit P450IIB Gene Subfamily[a]

| residue | HP1 | b14 | b46 | b52 | b54 |
|---|---|---|---|---|---|
| 92 | GCT (Ala) | | ==A | | |
| 102 | GCT (Ala) | | ==C | | |
| 103 | GTA (Val) | | ==G | | |
| 106 | CCA (Pro) | | ==G | | |
| 117 | AAT (Asn) | | ==C | | |
| 188 | TTC (Phe) | | === | ==T | |
| 197 | CGC (Arg) | | A=G | A=G | |
| 217 | TTT (Phe) | | === | ==C | |
| 227 | TTT (Phe) | | ==C | === | |
| 230 | ACG (Thr) | | === | ==A | |
| 258 | CCT (Pro) | | ==C | ==C | |
| 263 | GAC (Asp) | | === | ==T | |
| 265 | ATT (Ile) | | ==C | ==C | |
| 275 | GAT (Asp) | | ==C | === | |
| 279 | CCG (Pro) | | ==A | === | |
| 289 | ATC (Ile) | | === | ==T | |
| 307 | CTC (Leu) | ==T | === | === | === |
| 324 | GTC (Val) | ==T | === | === | ==T |
| 350 | GAT (Asp) | ==C | ==C | ==C | ==C |
| 359 | CTC (Leu) | === | ==G | === | ==G |
| 362 | CTC (Leu) | === | === | ==T | === |
| 366 | GGG (Gly) | === | === | ==T | === |
| 372 | ACC (Thr) | ==A | ==A | === | ==A |
| 375 | ACA (Thr) | ==G | ==G | ==G | ==G |
| 380 | TAC (Tyr) | ==T | ==T | ==T | ==T |
| 386 | ACG (Thr) | === | ==A | === | ==A |
| 396 | CTC (Leu) | === | ==G | === | ==G |
| 397 | CAC (His) | ==T | === | ==T | === |
| 407 | ACC (Thr) | ==T | === | === | === |
| 419 | GCC (Ala) | === | === | ==A | === |
| 429 | TTC (Phe) | === | === | ==T | === |

[a] The same nucleotide as found in HP1 is shown by a sign of equality (=).

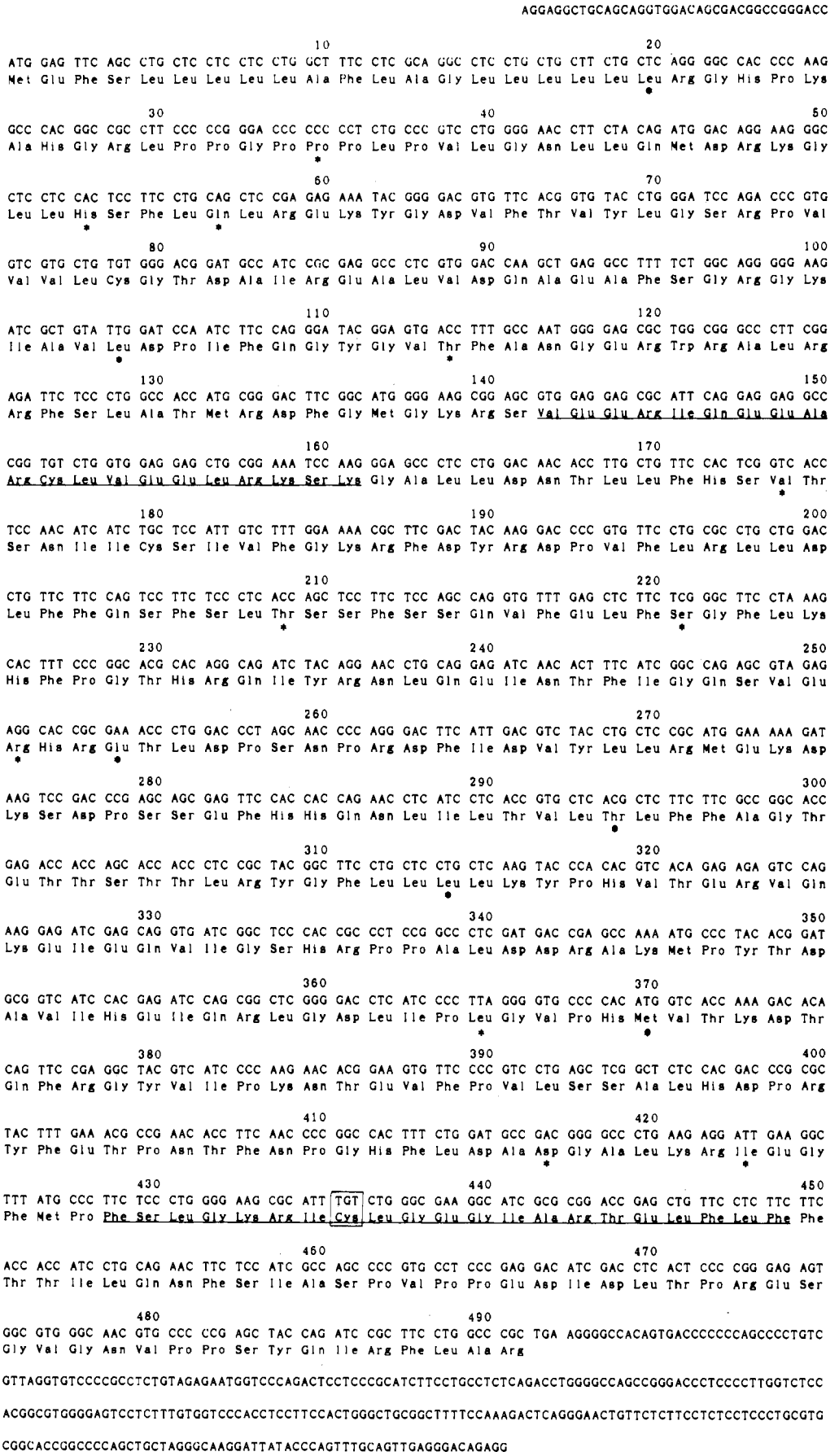


FIGURE 3: Complete nucleotide sequence of the pHP1 cDNA and the amino acid sequence deduced therefrom. Amino acid residues are numbered from the initiator methionine. The cysteinyl residue that acts as the fifth ligand to the heme iron is boxed. Two highly conserved regions, HR1 (upper) and HR2 (lower), are underlined. Residues at which amino acid differences are noticed between the HP1 protein and LM2 are asterisked.

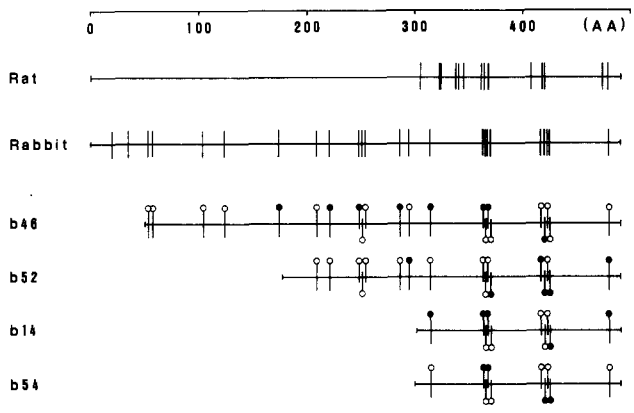


FIGURE 4: Positions of amino acid differences noticed among the primary structures deduced for the five representative cDNAs isolated in this study and that protein chemically determined for LM2. The positions where amino acid differences are noticed are indicated by vertical lines. Lines with open and closed circles, respectively, indicate that the residues are identical with and different from those in the HP1 sequence. The positions of amino acid differences between two rat PB-inducible forms of cytochrome P-450 (P-450b and P-450e) are also shown (horizontal line marked with "Rat").

cDNA at least in the carboxyl-terminal 191-residue region. It is, however, to be noted that upon sequence analysis of the peptides derived from P-450$_1$ one or two minor amino acids were detected at nine positions in addition to the major one. For instance, at position 287 a small but clearly detectable amount of glutamine was found together with a large amount of leucine. The other positions and minor amino acids detected there are as follows: 290 (+Ile), 334 (+Met), 407 (+Ala, Val), 409 (+Leu), 411 (+Ile), 412 (+Gln), 464 (+Gln), and 467 (+Ile). These findings indicate that the P-450$_1$ preparation used, though homogeneous on SDS–polyacrylamide gel electrophoresis, was in fact not completely pure but contaminated by one or several minor polypeptides possessing very similar but distinct sequences. In the amino-terminal 300-residue portion, which is missing in the b14 polypeptide, 14 amino acid differences are seen between the sequence of LM2 and that of the HP1, b46, or b52 protein (Table I). Since 9 of these 14 differences could be located in the peptides isolated from P-450$_1$, the amino acid sequences of these peptides were also determined. It was thus found that all the residues examined in the P-450$_1$ peptides are identical with those of LM2. As far as the sequenced regions of P-450$_1$ are concerned, only two amino acid differences are detectable between the sequences of LM2 and P-450$_1$.

## Discussion

The cDNA insert of pHP1, isolated from the liver of a PB-treated rabbit, contains an open reading frame encoding a form of cytochrome P-450 consisting of 491 amino acid residues. The primary structure of the encoded protein, HP1, predicted from the nucleotide sequence is not identical with but 96.5% similar to that reported for LM2 (Tarr et al., 1983), which has been purified as the major PB-inducible form of rabbit liver microsomal cytochrome P-450 (van der Hoeven et al., 1974; Haugen & Coon, 1976). Only 17 amino acid differences are seen between the 491-residue sequences of these two proteins. On the other hand, the amino acid sequence reported by Heinemann and Ozols (1983) for the major PB-inducible rabbit liver microsomal cytochrome P-450 is composed of 489 residues, in contrast to the fact that both LM2 and the HP1 protein consist of 491 residues. Moreover, the Heinemann–Ozols sequence is only 94% similar to that of the HP1 protein (28 amino acid differences). The reason for such discrepancies is unclear.

At any rate, sequence analysis of the other cDNA clones carrying shorter inserts has shown that the proteins encoded by these clones and pHP1 can be classified into five types, represented by the HP1, b14, b46, b52, and b54 proteins, on the basis of a sequence comparison of their overlapping regions. The primary structures of these five types of cytochrome P-450 are 96–98% similar to one another. Partial determination of the amino acid sequence of P-450$_1$, which has been purified by us as the major PB-inducible form of rabbit liver cytochrome P-450 (Imai & Sato, 1974; Imai et al., 1980), indicates that this form is identical with the b14 polypeptide at least in the carboxyl-terminal 191-residue portion. The sequence of P-450$_1$ is also very similar to that of LM2. As far as the sequenced regions of P-450$_1$ are concerned, there are only two amino acid differences between the two sequences. Such a minor difference suggests that they are alleric variants, although other possibilities cannot be ruled out. In any case, these findings indicate that at least five genes encoding cy-

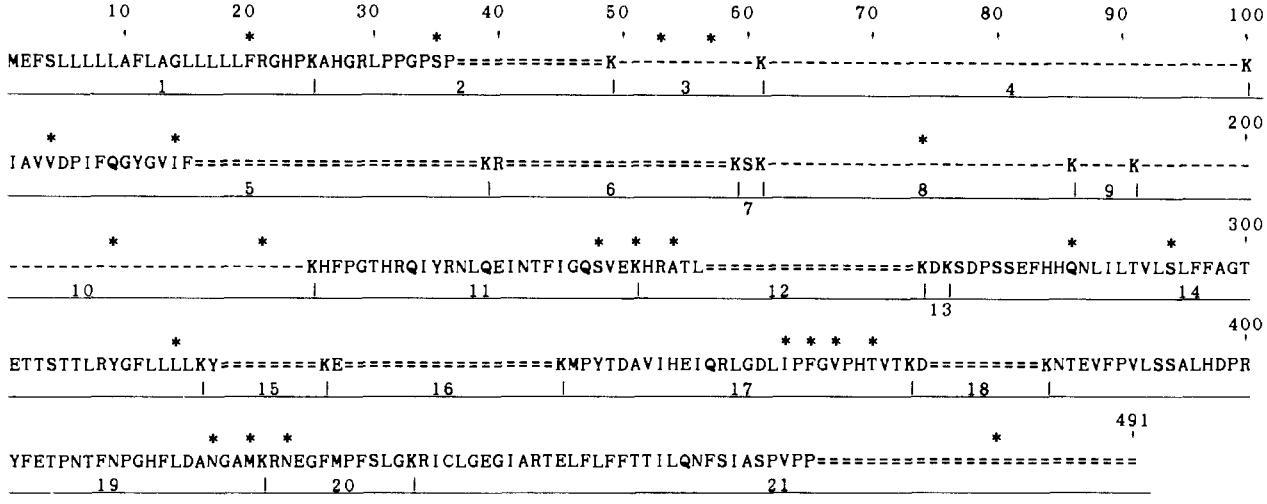


FIGURE 5: Alignment of peptides isolated from the *Achromobacter* lysylendopeptidase digest of P-450$_1$. Amino acid residues actually sequenced, the amino-terminal residue determined, and lysyl residues that should be located at the carboxyl termini of peptide fragments (from the specificity of the endopeptidase) are shown by one-letter abbrevations. Double dotted lines indicate portions that were thought to be identical with the corresponding portions of LM2 from their amino acid compositions. Peptides indicated by single dotted lines were not isolated in pure form and therefore not examined. The alignment was made based on a comparison with the primary structure determined for LM2 (Tarr et al., 1983). Asterisks show the positions where amino acid differences are noticed among the five representative deduced sequences.

tochromes P-450, which are structurally closely related to P-450$_1$ (LM2), are expressed in the liver of PB-treated rabbits. It is likely that microheterogeneity in this group of cytochrome P-450 is more extensive than this in view of the finding that the P-450$_1$ preparation used in this study was contaminated by one or more proteins possessing closely similar amino acid sequences. According to the nomenclature recently recommended by Nebert et al. (1987) for families and subfamilies within the P450 gene superfamily, rabbit LM2 as well as rat P-450b and P-450e is assigned to the P450IIB subfamily. It can be said that at least five genes belonging to the rabbit P450IIB subfamily are expressed in the liver of PB-treated rabbits. It should be stressed in this connection that in the liver of PB-treated rabbits many genes that are not members of this subfamily are also expressed, as will be reported in the accompanying paper (Imai et al., 1988).

Microheterogeneity has also been detected in two PB-inducible forms of rat liver microsomal cytochrome P-450, P-450b, and P-450e (Fujii-Kuriyama et al., 1982; Mizukami et al., 1983a,b). These two forms have been shown to differ from each other by only 14 amino acid substitutions. As shown in Figure 4, these amino acid differences are restricted to the carboxyl-terminal portion consisting of some 200 residues (Suwa et al., 1985). Structural analysis of the corresponding genomic clones has further shown that the substitutions occur most frequently in exon 7 (Suwa et al., 1985), a finding that has been explained by assuming that extensive gene conversion has occurred in the region including exon 7 (Atchison & Adesnik, 1986). However, this gene conversion mechanism does not seem to account for the substitutions observed in the rabbit counterparts, because they are distributed over the entire sequence (Figure 4). Isolation and structural analysis of rabbit genomic clones corresponding to the five cDNAs examined in this study are needed to elucidate the evolutionary origin of the microheterogeneity in the members of the rabbit P450IIB gene subfamily. In view of the multiplicity of genes of the rabbit P450IIB subfamily, it is likely that the rat P450IIB subfamily also contains members other than P-450b and P-450e. In fact, it has been reported that the rat genome contains at least six genes that are closely related to the P-450b gene (Mizukami et al., 1983b).

From purification studies, it is clear that P-450$_1$ (or LM2) is the most abundant cytochrome P-450 species in liver microsomes of PB-treated rabbits. In cDNA cloning from the same source, however, b14 cDNA (identical with P-450$_1$ cDNA) has not been detected most frequently. Instead, the five types of cDNA have been cloned with similar frequency. It is likely that this discrepancy is due to the fact that the mRNA used for cDNA synthesis was isolated from the liver 20 h after a single injection of PB, whereas P-450$_1$ was purified from the liver after 7 daily treatments of a rabbit with the drug. It is possible that mRNAs encoding different cytochromes P-450 exhibit different time-dependent responses to the PB treatment. Another possibility is to assume that P-450$_1$ is much more stable in vivo than the other forms of this group. If this is so, then it is expected that P-450$_1$ accumulates in liver microsomes even though the level of its mRNA is not high. As a third possibility, it can be assumed that P-450$_1$ mRNA is stable in vivo. This will also lead to the selective accumulation of P-450$_1$.

The finding that the purified P-450$_1$ preparation, though gel electrophoretically homogeneous, was not pure but contaminated by one or more minor proteins having closely similar primary structures may partly account for the remarkably broad substrate specificity of this preparation (Sato et al., 1982). In a reconstituted system containing purified P-450$_1$ and NADPH–cytochrome P-450 reductase, it has been reported that the reduction of the cytochrome by NADPH proceeds in a biphasic fashion (Taniguchi et al., 1979). The possibility exists that this biphasic reaction kinetics is also due to the heterogeneity of the P-450$_1$ preparation used. These and other properties of the purified P-450$_1$ preparation should be reexamined with a completely homogeneous preparation, which may be obtained by expression of a full-length P-450$_1$ cDNA in yeast or other suitable cells.

ACKNOWLEDGMENTS

We thank Dr. Y. Fujii-Kuriyama of the Cancer Institute, Japanese Foundation for Cancer Research, Tokyo, for kindly supplying us with pcP-450pb4.

**Registry No.** HP1 (DNA), 111769-75-6; HP1 (peptide), 111769-73-4; cytochrome P-450, 9035-51-2.

REFERENCES

Atchison, M., & Adesnik, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 2300–2304.

Aviv, H., & Leder, P. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 1408–1412.

Clarke, L., & Carbon, J. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 4361–4365.

Erickson, A., & Blobel, G. (1983) *Methods Enzymol. 96*, 38–50.

Fujii-Kuriyama, Y., Mizukami, Y., Kawajiri, K., Sogawa, K., & Muramatsu, M. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 2793–2797.

Gordon, J. I., Burns, A. T. H., Christmann, J. L., & Deeley, R. G. (1978) *J. Biol. Chem. 253*, 8629–8639.

Gotoh, O., Tagashira, Y., Iizuka, T., & Fujii-Kuriyama, Y. (1983) *J. Biochem. (Tokyo) 93*, 807–817.

Gray, W. R. (1972) *Methods Enzymol. 25*, 121–138.

Grunstein, M., & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 3961–3965.

Haugen, D. A., & Coon, M. J. (1976) *J. Biol. Chem. 251*, 7929–7939.

Heinemann, F. S., & Ozols, J. (1983) *J. Biol. Chem. 258*, 4195–4201.

Hunkapiller, M. W., Hewick, R. M., Dryer, W. J., & Hood, L. E. (1983) *Methods Enzymol. 91*, 399–413.

Imai, Y., & Sato, R. (1974) *Biochem. Biophys. Res. Commun. 60*, 8–14.

Imai, Y., Hashimoto-Yutsudo, C., Satake, H., Girardin, A., & Sato, R. (1980) *J. Biochem. (Tokyo) 88*, 489–503.

Imai, Y., Komori, M., & Sato, R. (1988) *Biochemistry* (following paper in this issue).

Kawajiri, K., Sogawa, K., Gotoh, O., Tagashira, Y., Muramatsu, M., & Fujii-Kuriyama, Y. (1983) *J. Biochem. (Tokyo) 94*, 1465–1473.

Maniatis, T., Jeffrey, A., & Kleid, D. G. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 1184–1189.

Maxam, A. V., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499–560.

Mihara, K., & Blobel, G. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 4160–4164.

Mizukami, Y., Sogawa, K., Suwa, Y., Muramatsu, M., & Fujii-Kuriyama, Y. (1983a) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3958–3962.

Mizukami, Y., Fujii-Kuriyama, Y., & Muramatsu, M. (1983b) *Biochemistry 22*, 1223–1229.

Morohashi, K., Fujii-Kuriyama, Y., Okada, Y., Sogawa, K., Hirose, T., Inayama, S., & Omura, T. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 4647–4651.

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., Levin, W., Phillips, I. R., Sato, R., & Waterman, M. R. (1987) *DNA 6*, 1-11.
Okayama, H., & Berg, P. (1982) *Mol. Cell. Biol. 2*, 161-170.
Palmiter, R. D. (1974) *Biochemistry 13*, 3606-3615.
Poulos, T. L., Finzel, B. C., Gunsalus, I. C., Wagner, G. C., & Kraut, J. (1985) *J. Biol. Chem. 260*, 16122-16130.
Ryan, D. E., Thomas, P. E., Koreniwski, D., & Levin, W. (1979) *J. Biol. Chem. 254*, 1365-1374.
Ryan, D. E., Thomas, P. E., & Levin, W. (1982) *Arch. Biochem. Biophys. 216*, 272-288.
Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J., & Roe, B. A. (1980) *J. Mol. Biol. 143*, 161-178.
Sato, R., Aoyama, T., & Imai, Y. (1982) in *Oxygenases and Oxygen Metabolism* (Ishimura, Y., Yamamoto, S., & Nozaki, M., Eds.) pp 321-332, Academic, New York.
Southern, E. (1975) *J. Mol. Biol. 98*, 503-518.
Suwa, Y., Mizukami, Y., Sogawa, K., & Fujii-Kuriyama, Y. (1985) *J. Biol. Chem. 260*, 7980-7984.
Taniguchi, H., Imai, Y., Iyanagi, T., & Sato, R. (1979) *Biochim. Biophys. Acta 550*, 341-355.
Taniguchi, T., Ohno, S., Fujii-Kuriyama, Y., & Muramatsu, M. (1980) *Gene 10*, 11-15.
Tarr, G. E., Black, S. D., Fujita, V. S., & Coon, M. J. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 6552-6556.
Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 5201-5205.
Tsunasawa, S., Kondo, J., & Sakiyama, F. (1985) *J. Biochem. (Tokyo) 97*, 701-704.
Van der Hoeven, T. A., Haugen, D. A., & Coon, M. J. (1974) *Biochem. Biophys. Res. Commun. 60*, 569-675.
Vlasuk, W., & Walz, F. G., Jr. (1982) *Biochemistry 21*, 789-793.

# Comparison of Primary Structures Deduced from cDNA Nucleotide Sequences for Various Forms of Liver Microsomal Cytochrome P-450 from Phenobarbital-Treated Rabbits†

Yoshio Imai,* Masayuki Komori,‡ and Ryo Sato
*Institute for Protein Research, Osaka University, Suita, Osaka 565, Japan*
*Received April 13, 1987; Revised Manuscript Received July 30, 1987*

ABSTRACT: cDNA clones, termed pHP2, b32-3, b43, and b43-1, encoding cytochromes P-450 that are expressed in the liver of phenobarbital- (PB-) treated rabbits were isolated, and their nucleotide sequences were determined. pHP2 cDNA contains an open reading frame for a 490-residue protein and is a full-length counterpart of pP-450PBc2 [Leighton, J. K., Debrunner-Vossbrinck, B. A., & Kemper, B. (1984) *Biochemistry 23*, 204-210]. The b32-3 insert has a sequence for a protein whose primary structure is 91% similar to that of progesterone 21-hydroxylase P-450 1, though this cDNA lacks the sequence encoding the amino-terminal 110 residues. The overlapping clones b43 and b43-1 together encode an ethanol-inducible form of cytochrome P-450, though the amino-terminal five or more residues are missing in the composite b43/b43-1 sequence. Northern blot analysis showed that the b43/b43-1 protein is more strongly inducible by polycyclic aromatic hydrocarbons and isosafrole than by PB, in contrast to the case of the HP2 and b32-3 proteins. A comparison of the primary structures of eight forms of cytochrome P-450, including the HP2, b32-3, and b43/b43-1 proteins, that are expressed in the liver of PB-treated rabbits showed that 149 out of 487-492 amino acid residues are conserved in these cytochromes P-450. The eight forms can be assigned to three rabbit cytochrome P-450 gene subfamilies, P450IIB, P450IIC, and P450IIE. It was also shown that the members of the rabbit P450IIC subfamily can be further classified into three subgroups on the basis of their sequence similarity.

In attempts at elucidating the biological significance of the multiplicity of liver microsomal cytochrome P-450 at the DNA level, a number of cDNA and genomic clones for different forms of cytochrome P-450 have been isolated and their structures analyzed [e.g., Fujii-Kuriyama et al. (1982), Mizukami et al. (1983), Kawajiri et al. (1984), Sogawa et al. (1984, 1985), Gonzalez et al. (1984, 1985, 1986a,b), Leighton et al. (1984), Suwa et al. (1985), Tukey et al. (1985), Govind et al. (1986), Song et al. (1986), Hardwick et al. (1987), Imai (1987), and Kagawa et al. (1987)]. These studies have provided evidence that there are several distinct cytochrome P-450 gene families, such as polycyclic aromatic hydrocarbon inducible, phenobarbital- (PB-)[1] inducible, steroid-inducible, and peroxisome proliferator inducible ones. Nebert et al. (1987) have recently proposed a classification of the cytochrome P-450 gene superfamily, comprising not only liver cytochromes P-450 but also those from all other sources, into families and subfamilies and recommended a nomenclature for them.

In the preceding paper (Komori et al., 1988), we reported that microheterogeneity exists in major PB-inducible forms of rabbit liver cytochrome P-450, which belong to the rabbit P450IIB gene subfamily. Leighton et al. (1984) and Govind et al. (1986) have isolated cDNA and genomic clones, respectively, encoding rabbit liver cytochromes P-450 that are distinct from major PB-inducible ones but expressed in the

† This work was supported in part by a grant-in-aid (5806002) from the Ministry of Education, Science and Culture of Japan.
* Author to whom correspondence should be addressed.
‡ Present address: Faculty of Pharmaceutical Sciences, Hokkaido University, Sapporo 060, Japan.

[1] Abbreviations: PB, phenobarbital; SDS, sodium dodecyl sulfate.

0006-2960/88/0427-0080$01.50/0 