The Flavin-Containing Monooxygenase Expressed in Pig Liver: Primary Sequence, Distribution, and Evidence for a Single Gene

Rodolfo Gasser,[‡] Ronald E. Tynes,[‡] Michael P. Lawton,[‡,§] Katy K. Korsmeyer,[∥] D. M. Ziegler,[∥] and Richard M. Philpot*[,‡]

Laboratory of Cellular and Molecular Pharmacology, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina 27709, Toxicology Program, North Carolina State University, Raleigh, North Carolina 27695, and Clayton Foundation Biochemical Institute, Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712

Received June 21, 1989; Revised Manuscript Received August 1, 1989

ABSTRACT: The primary sequence of the flavin-containing monooxygenase expressed in pig liver has been derived from the nucleotide sequence of cloned cDNA. The derived sequence is composed of 532 amino acids and represents a protein having a molecular weight of 58 952. The complete sequence was obtained from a single clone containing 2070 bases. A second clone, obtained from an independent library, yielded an identical sequence for the 1374 bases present. The amino acid composition compiled from the derived sequence is very similar to that obtained previously from the purified protein. In addition, a 10 amino acid sequence in a peptide formed from the purified protein by digestion with V8 protease exactly matches the derived sequence for residues 309–318. The flavin-containing monooxygenase expressed in pig liver is also expressed in pig lung and kidney as determined by analysis of both microsomal proteins and mRNA. The ratio of mRNA to protein for the enzyme in kidney is about 5 times greater than the same ratio for liver and about twice the ratio for lung. The reasons for these differences are not understood. Southern analysis of genomic DNA indicates that there is a single gene encoding the flavin-containing monooxygenase expressed in pig liver. Therefore, the broad activity of this enzyme in liver appears to be the result of the catalytic diversity of a single protein.

Microsomal monooxygenases are among the most catalytically diverse mammalian enzymes associated with the metabolism of xenobiotics. With the cytochrome P-450 monooxygenase system, this diversity is a consequence of multiple enzymes with overlapping, but distinct, substrate specificities. The various forms of cytochrome P-450, shown to be products of different genes, constitute a very complex family of proteins (Nebert et al., 1989). The substrate specificity of the microsomal flavin-containing monooxygenase (FMO;[1] EC 1.14.13.8), although somewhat narrower than that of the cytochrome P-450 system, is also remarkably broad. The enzyme from pig liver has been shown to catalyze the oxidation of numerous drugs, pesticides, industrial chemicals, and naturally occurring plant toxins containing nitrogen, sulfur, selenium, or phosphorus (Ziegler, 1988).

Flavoproteins with properties similar to those of the pig liver FMO have also been purified from rat (Kimura et al., 1983), mouse (Sabourin et al., 1984), and rabbit (Tynes & Hodgson, 1985) liver, and preparations of these enzymes appear to contain single proteins when examined by polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate. Hepatic microsomal samples from these species, as well as from humans, dogs, guinea pigs, and hamsters, all react with antibodies to the FMO from pig liver (Dannan & Guengerich, 1982; Tynes & Philpot, 1987) to form single bands on immunoblots. Recently, a second form of the enzyme, which does not cross-react with antibodies to the enzyme expressed in liver, has been purified from lungs of rabbits (Williams et al., 1984; Tynes et al., 1985) and mice (Tynes et al., 1985). While expression of the "hepatic" enzyme has been observed in lung and kidney of several species (Dannan & Guengerich, 1982; Tynes & Philpot, 1987), expression of the "pulmonary" enzyme in liver has not been detected (Tynes & Philpot, 1987). These results suggest that the broad substrate specificity associated with the FMO in hepatic microsomal preparations may reflect the activity of a single enzyme. However, this evidence does not rule out the existence of multiple, physically similar but catalytically distinct, FMO enzymes in liver.

On the other hand, the oxidation of structurally diverse compounds by a single enzyme is consistent with the catalytic mechanism of the pig liver FMO. Results of kinetic studies (Poulsen & Ziegler, 1979; Beaty & Ballou, 1980) indicate that formation of a 4α-hydroperoxyflavin occurs prior to substrate addition and that any soft nucleophile gaining access to the hydroperoxyflavin is oxidized. Unlike all other monooxygenases requiring an external reductant, FMO apparently discriminates between essential and physiologically useless nucleophiles by excluding the former rather than selectively binding the latter (Ziegler, 1988). Further insights into the molecular basis of this catalytic process have been hampered by lack of information on the structure of the enzyme. Although first purified in 1972 (Ziegler & Mitchell, 1972), the enzyme, which apparently has a blocked N-terminus, has proven difficult to sequence. Therefore, we have used techniques of molecular biology to approach this problem and can now report a primary sequence for the FMO expressed in pig liver.

Materials and Methods

Animals and Tissues. Livers, lungs, and kidneys of adult pigs were obtained from local slaughterhouses in Waelder, TX,

* Author to whom correspondence should be addressed.
[‡] National Institutes of Health.
[§] North Carolina State University.
[∥] The University of Texas at Austin.

[1] Abbreviations: FMO, flavin-containing monooxygenase; PVDF, poly(vinylidene difluoride).

0006-2960/90/0429-119$02.50/0

and Garner, NC. Enzyme for protein sequence determinations was purified from livers obtained in Texas, and enzyme for antibody production was purified from livers obtained in North Carolina. Microsomal fractions, mRNA, and DNA were isolated from tissues obtained in North Carolina. Tissues for mRNA and DNA isolation were placed in liquid nitrogen immediately following removal from the animal.

Preparation of Microsomes, Purified Enzyme, and Antibodies. The microsomal fractions from pig liver, lung, and kidney were prepared by standard procedures (Tynes & Philpot, 1987). Pig liver FMO was purified by the method of Poulsen and Ziegler (1979) for sequence determinations and by the method of Sabourin et al. (1984) for preparation of antibodies. Antibodies used in this study were characterized previously (Tynes & Philpot, 1987).

Isolation of mRNA and DNA. Total RNA was obtained from pig liver, lung, and kidney by a modification (Gasser & Philpot, 1988) of the methods of Chirgwin et al. (1979) and Glisin et al. (1974). Polyadenylated mRNA was isolated by chromatography on oligo(dT)–cellulose (Aviv & Leder, 1972). Standard procedures (Enrietto et al., 1983; Gross-Bellard et al., 1973) were used for the isolation of DNA from pig liver.

Cloning of cDNA. Double-stranded cDNA was synthesized from pig liver mRNA by sequential incubation with cloned Moloney murine leukemia virus reverse transcriptase (M-MLV RT, BRL, Gaithersburg, MD) followed by DNA polymerase I plus RNase H (Gubler & Hoffman, 1983; Okayama & Berg, 1982). Following attachment of synthetic *Eco*RI linkers, cDNA species of greater than 2000 base pairs were isolated by agarose gel electrophoresis and ligated into λ ZAP (Stratagene, La Jolla, CA). About 6×10^6 recombinant clones were obtained per microgram of cDNA.

Isolation of Clones Encoding the FMO. Initially, a commercial pig liver λgt11 cDNA library (Clontech, San Diego, CA) was screened for clones encoding the FMO. Positive clones were identified by detection of the fused β-galactosidase–antigen product with antibody to pig liver FMO, and the largest of these clones (A) was subcloned into plasmid Bluescript (Stratagene, La Jolla, CA) and subsequently used as a probe for screening the library we constructed.

Nucleotide Sequence Analysis. Fragments of purified cDNA clones were formed by sonication (Deininger, 1983) or enzymic restriction (Messing, 1983), subcloned into M13, and sequenced by the dideoxy chain-termination method (Sanger et al., 1977; Biggin et al., 1983). Sequence data were analyzed with the Sequence Analysis Software Package from the University of Wisconsin (Devereux et al., 1984).

Analysis of mRNA and DNA. Samples of liver, lung, and kidney mRNA were electrophoresed in agarose (FMC Bio-Products, Rockland, ME) containing methylmercury (Bailey & Davidson, 1976), transferred to a nylon membrane (Nytran; Schleicher & Schuell, Keene, NH), and hybridized with cDNA encoding pig liver FMO. Relative amounts of mRNA were determined by densiometric scanning (Joyce Loebl, Gateshead, England). Restriction fragments of hepatic genomic DNA were separated by electrophoresis, transferred to nylon membrane, and analyzed by the method of Southern (1975). Analysis of DNA was done with both partial and full-length cDNAs into which [^{32}P]dCTP was incorporated ($>10^9$ cpm/μg) by the nick-translation (Maniatis et al., 1975) or oligo primer directed (Feinberg & Vogelstein, 1983) methods.

Protein Sequence Determination. Apparently homogeneous pig liver FMO was further purified by electrophoresis on single-well polyacrylamide (10%) minigels by the method of Laemmli (1970). The gels were stained with a solution (100 mL) containing Coomassie blue (0.25 g), methanol (45.5 mL), and glacial acetic acid (9 mL) for 2 min, rinsed with water, and destained in two changes of methanol (45%)–glacial acetic acid (10%) for 3 min each. The FMO band was cut from the gel and immersed for 30 min in Tris buffer (0.125 M, pH 6.8) containing EDTA (1 mM) and sodium dodecyl sulfate (1 mg/mL). Attempts to obtain sequence data from the N-terminus of the intact protein were unsuccessful. Partial proteolysis with V8 protease (Sigma, St. Louis, MO) of FMO present in the gel slices, and separation of peptides by electrophoresis, was carried out according to Cleveland et al. (1977). Following electrophoresis, the gels were soaked for 5 min in 3-(cyclohexylamino)-1-propanesulfonic acid (10 mM, pH 11.0) containing 10% methanol (transfer buffer). The peptides were then electroeluted onto poly(vinylidene difluoride) (PVDF) sheets. The PVDF sheets were rinsed in water for 5 min, stained with Coomassie blue (0.1%) in methanol (50%) for 5 min, destained with methanol (50%)–glacial acetic acid (10%), and rinsed with several changes of water. Peptide bands resolved clearly by this procedure were cut from the PVDF, dried by spin lyophilization, and stored at −20 °C. Isolated peptides were sequenced (Applied Biosystems 447A gas-phase sequencer) by the Edman degradation procedure as modified by Schroeder (1967).

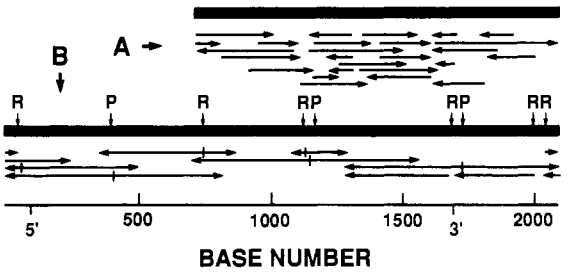

FIGURE 1: Strategy used for sequencing the cDNAs encoding the flavin-containing monooxygenase expressed in pig liver. Graphical representations of clones A (1384 bases) and B (2070 bases) are shown along with the fragments from each that were subcloned and sequenced. Fragments from clone A were obtained by sonication and those from clone B by restriction with *Pst*I (P) and *Rsa*I (R).

Electrophoresis, Immunoblotting, and Immunostaining. Polyacrylamide (7.5%) gel electrophoresis of microsomal proteins in the presence of sodium dodecyl sulfate was done by the method of Laemmli and Favre (1973). The separated proteins were then transferred to nitrocellulose and stained by a modification (Domin et al., 1984) of the method of Towbin et al. (1979). Conditions for these procedures were those reported by Tynes and Philpot (1987).

Materials. All reagents not specifically noted above were obtained at the appropriate purity from commercial sources.

RESULTS

Isolation and Sequencing of cDNA Clones Encoding Pig Liver FMO. Several clones identified by detection with antibodies to the pig liver FMO were isolated from the commercial cDNA library. The largest of these (A), about 1400 base pairs, was used as a probe for detection of positive clones in the library constructed with cDNA sized to exclude species of less than approximately 2000 base pairs. One of the cDNA clones (B) isolated in this manner was restricted with *Pst*I and *Rsa*I, and the resulting fragments were subcloned and sequenced. The cDNA insert present in clone A was also sequenced following random fragmentation by sonication. The strategies used for sequencing A and B are shown in Figure 1. Clone B contained a cDNA insert with an open reading

Table I: Nucleotide Sequence of cDNA Encoding the Flavin-Containing Monooxygenase Expressed in Pig Liver and Derived Primary Sequence of the Protein

```
0001      GTGCTTGGGGACAGTAAAGTCAGATATTATCAAGTCTTCTGTGTACACTGAAGAACATGGCCAAGCGAGTTGCAATTGTGGGAGCT  0086
0001                                                    MetAlaLysArgValAlaIleValGlyAla              0010

0087  GGGGTCAGTGGCCTGGCCTCCATCAAGTGCTGCCTGGAGGAAGGGCTGGAGCCCACCTGCTTTGAGAGGAGCGATGACCTTGGGGGGCTG  0176
0011  GlyValSerGlyLeuAlaSerIleLysCysCysLeuGluGluGlyLeuGluProThrCysPheGluArgSerAspAspLeuGlyGlyLeu  0040

0177  TGGAGATTCACTGAACATGTTGAAGAAGGCAGAGCCAGCCTCTACAAGTCTGTGGTTTCCAACAGCTGCAAGGAGATGTCTTGTTACCCA  0266
0041  TrpArgPheThrGluHisValGluGluGlyArgAlaSerLeuTyrLysSerValValSerAsnSerCysLysGluMetSerCysTyrPro  0070

0267  GACTTTCCATTCCCAGAAGATTACCCAAACTATGTGCCAAATTCTCACTTCCTGGAATATCTCAGAATGTATGCAAACCAGTTCAACCTT  0356
0071  AspPheProPheProGluAspTyrProAsnTyrValProAsnSerHisPheLeuGluTyrLeuArgMetTyrAlaAsnGlnPheAsnLeu  0100

0357  CTGAAGTGCATTCAATTCAAGACTAAAGTCTGCAGTGTAACAAAACATGAAGATTTTAATACCACCGGTCAATGGGATGTAGTCACTCTG  0446
0101  LeuLysCysIleGlnPheLysThrLysValCysSerValThrLysHisGluAspPheAsnThrThrGlyGlnTrpAspValValThrLeu  0130

0447  TGTGAAGGAAAGCAAGAGTCAGCCGTCTTTGATGCTGTCATGGTCTGCACTGGTTTTCTCACTAACCCATATTTGCCATTGGACTCCTTC  0536
0131  CysGluGlyLysGlnGluSerAlaValPheAspAlaValMetValCysThrGlyPheLeuThrAsnProTyrLeuProLeuAspSerPhe  0160

0537  CCAGGCATAAATACTTTTAAAGGCCAGTATTTCCATAGCCGACAGTATAAACATCCAGATATATTTAAGGACAAGAGTGTTCTTGTGGTT  0626
0161  ProGlyIleAsnThrPheLysGlyGlnTyrPheHisSerArgGlnTyrLysHisProAspIlePheLysAspLysSerValLeuValVal  0190

0627  GGAATGGGAAATTCGGGCACAGACATTGCTGTGGAGGCCAGCCACCTGGCAAAAAAGGTGTTCCTCAGCACCACTGGAGGAGCATGGGTG  0716
0191  GlyMetGlyAsnSerGlyThrAspIleAlaValGluAlaSerHisLeuAlaLysLysValPheLeuSerThrThrGlyGlyAlaTrpVal  0220

0717  ATCAGCCGTGTCTTTGACTCAGGGTACCCATGGGACATGGTGTTCATGACACGATTTCAGAACATGTTCAGAAATTCTCTCCCAACTCCA  0806
0221  IleSerArgValPheAspSerGlyTyrProTrpAspMetValPheMetThrArgPheGlnAsnMetPheArgAsnSerLeuProThrPro  0250

0807  ATTGTGAATTGGTTGATAGCAAAAAAGATGAACAGCTGGTTCAATCATGCAAATTATGGCTTAATACCCGAAGACAGGATACAACTAAGA  0896
0251  IleValAsnTrpLeuIleAlaLysLysMetAsnSerTrpPheAsnHisAlaAsnTyrGlyLeuIleProGluAspArgIleGlnLeuArg  0280

0897  GAGCCTGTGCTGAATGATGAGCTCCCAGGCCGTATCATCACTGGGAAAGTGCTCATCAAGCCAAGTATAAAGGAGGTGAAGGAAAACTCC  0986
0281  GluProValLeuAsnAspGluLeuProGlyArgIleIleThrGlyLysValLeuIleLysProSerIleLysGluValLysGluAsnSer  0310

0987  GTGGTATTTAACAGCTCCCCAGAGGAAGAGCCCATTGATATCATTGTCTTTGCCACTGGATACACCTTTGCTTTCCCCTTCCTTGATGAG  1076
0311  ValValPheAsnSerSerProGluGluGluProIleAspIleIleValPheAlaThrGlyTyrThrPheAlaPheProPheLeuAspGlu  0340

1077  AGTGTAGTGAAGGTTGAAGATGGCCAGGCATCACTATACAAGTACATCTTCCCTGCACATCTGCAGAAGCCAACTCTGGCTGTTATTGGC  1166
0341  SerValValLysValGluAspGlyGlnAlaSerLeuTyrLysTyrIlePheProAlaHisLeuGlnLysProThrLeuAlaValIleGly  0370

1167  CTCATCAAACCCTTGGGCTCCTTACTACCCACAGGGGATACACAAGCTCGGTGGGCTGTTCGGGTCCTGAAAGGTGTGAATAAGTTACCA  1256
0371  LeuIleLysProLeuGlySerLeuLeuProThrGlyAspThrGlnAlaArgTrpAlaValArgValLeuLysGlyValAsnLysLeuPro  0400

1257  CCATCAAGTGTCATGATACAGGAAGTTAATACAAGAAAAGAAAACAAGCCCAGTGGGTTTGGCTTGTGTTACTGCAAAGCTTTACAATCC  1346
0401  ProSerSerValMetIleGlnGluValAsnThrArgLysGluAsnLysProSerGlyPheGlyLeuCysTyrCysLysAlaLeuGlnSer  0430

1347  GATTATATTGCATACATAGATGAACTGCTGACCTATATCGATGCAAAGCCCAACATGTTTTCTCTGCTCCTGACGGATCCACACCTGGCT  1436
0431  AspTyrIleAlaTyrIleAspGluLeuLeuThrTyrIleAspAlaLysProAsnMetPheSerLeuLeuLeuThrAspProHisLeuAla  0460

1437  TTGACCATCTTCTTTGGCCCATGCACACCATACCAGTTCCGCCTGACTGGTCCAGGGAAATGGGAAGGAGCCAGGAATGCTATCATGACC  1526
0461  LeuThrIlePhePheGlyProCysThrProTyrGlnPheArgLeuThrGlyProGlyLysTrpGluGlyAlaArgAsnAlaIleMetThr  0490

1527  CAGTGGGACCGAACATTCAAGGTCACCAAAACTCGAATTGTAAAAGAATCCCCATCTCCCTTTGCAAGCTTACTTAAACTCTTTAGCTTT  1616
0491  GlnTrpAspArgThrPheLysValThrLysThrArgIleValLysGluSerProSerProPheAlaSerLeuLeuLysLeuPheSerPhe  0520

1617  CTAGCCTTGCTTGTGGCCATTTTCCAGATTTTCCTCTAAGTATTACTGAATGGATTATTAAATGCACTGAGTAGACGGTTAGTACTCTGG  1706
0521  LeuAlaLeuLeuValAlaIlePheGlnIlePheLeuEnd                                                      0532

1707  CTCCTCCACTGCAGCAGCATTGCCTCCCCAGCTATGAATCATATTTGAGTAGTGAAAGCCACCCATGCCCTTCCCTGGCTAGCTTAAGCA  1796

1797  TTACCACTAATACTCCCAAGTGTCCTCAGCTCTACCTGCTTCCAATGCTAGAGGAAGATAACCAAAATTTCTGTGCATTTGAAAGCTGCT  1886

1887  AGAAGGTTCAAGGTTCATTTTAGAAGGGAGGGCTGTCCCAGACAGCACTCTGAGCAATTAGGCCTCTTTCCTATTTTCATAGATCTCGAC  1976

1977  TAAAATTCCCTACTTCACAGAAGATTTTGTTGTACTAAAATCAGTTGATTTTGTTGTACTAAAATGGTATCAGTTGATTAAAAATAAAAA  2066

2067  GAAAGGAAAGTCTA                                                                               2080
```

frame of 1596 bases, a 5′-flanking region of 56 bases, and a 3′-flanking region of 418 bases. Clone A contained 1384 bases of which 956 were in the coding region and 428 were in the 3′-flanking region. Identical sequences were obtained for the overlapping areas (1374 bases) of A and B. The complete base sequence of B is shown in Table I along with the derived amino acid sequence of the pig liver FMO. The combined molecular weight of the 532 residues making up this sequence is 59 952. The amino acid composition of this protein compiled from the derived sequence and that obtained directly from the protein by Poulsen and Ziegler (1979) are compared in Table II.

Sequencing of Protein and Comparison with Derived Sequences. The pattern of peptides formed from the purified FMO by digestion with V8 protease for 30 min at 24–25 °C was reproducible, and three of the more intense bands were clearly separated by electrophoresis. However, only one of these fragments (apparent M_r = 17 000) yielded a reproducible, definitive sequence for at least 10 residues from its amino-terminal end. Identification of the amino acids in this fragment was unambiguous. The sequence of this peptide exactly coincides with the sequence derived from the cDNA for positions 309–318 (Table III). In addition, an appropriate cleavage residue (Glu) is present at position 308 of the derived sequence. The sequence obtained from a second fragment (apparent M_r = 28 000) was not completely reproducible and was based on weak signals. This sequence, including positions for which variable results were obtained, is also shown in Table III. The amino acids in the nine positions for which a single

Table II: Amino Acid Composition of the Hepatic Flavin-Containing Monooxygenase from Pig

amino acid	residues/molecule[a]	% total
alanine	30 (31)	5.6
arginine	19 (21)	3.6
asparagine	23 } (52)	4.3
aspartate	24 }	4.5
cysteine	12 (09)	2.3
glutamine	16 } (52)	3.0
glutamate	29 }	5.5
glycine	33 (37)	6.2
histine	09 (10)	1.7
isoleucine	30 (29)	5.6
leucine	51 (58)	9.6
lysine	36 (39)	6.8
methionine	12 (10)	2.3
phenylalanine	37 (37)	7.0
proline	34 (42)	6.4
serine	37 (38)	7.0
threonine	31 (33)	5.8
tryptophan	09 (09)	1.7
tyrosine	19 (19)	3.6
valine	41 (44)	7.7

[a] The number of residues from the derived sequence are shown. The numbers in parentheses were determined from the protein and are taken from Poulsen and Ziegler (1979).

Table III: Sequences of Peptides Obtained from Purified Protein and Derived from cDNA Nucleotide Sequences

source	sequence
cDNA Derived (Residue #)	GLU ASN SER VAL VAL PHE ASN SER SER PRO GLU 308 309 310 311 312 313 314 315 316 317 318
PROTEIN (pmol yield)	ASN SER VAL VAL PHE ASN SER SER PRO GLU 22 20 9 9 11 9 11 11 2 5
cDNA Derived (Residue #)	GLU SER ANA VAL PHE ASP ALA VAL MET VAL CYS THR GLY PHE LEU THR 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151
PROTEIN (pmol yield)	(SER)ALA VAL PHE ASP ALA(VAL MET)VAL(GLY)THR(GLY PHE LEU)THR (VAL) 9 5 6 4 8 (SER ARG) 4 (THR) 4 (THR SER ARG) 5 (ASP)

residue was resolved show a perfect match with those in positions 138–142, 145, 147, and 151 of the derived sequence. Of the seven remaining positions, for which variable results were obtained, only the one corresponding to position 146 shows no match (Table III). The glutamine at position 136 of the derived sequence is consistent with cleavage by V8 protease between residues 136 and 137. No reproducible sequence was obtained for the third protein fragment (apparent M_r = 15000) that was isolated. The purified FMO was also digested with trypsin, and several peptides formed in this manner were purified to apparent homogeneity by HPLC or FPLC. None of these peptides yielded unambiguous sequence information.

Expression of Hepatic FMO in Liver, Lung, and Kidney. Antibodies to the pig liver FMO were used for the detection of protein in microsomal preparations from pig liver, lung, and kidney (Figure 2). Immunoreactive protein of the same mobility was detected in all three tissues, with the largest amount present in liver. Similar amounts of protein (about 3 times less than found in liver) were present in lung and kidney. The same tissues were also used for the preparation of mRNA, which was subsequently probed with cDNA for the pig liver FMO (Figure 3). All three tissues contained mRNA that hybridized with the FMO cDNA probe. The amount of mRNA detected in the liver sample was similar to the amount detected in the lung sample but only about half the amount detected in the kidney sample.

Analysis of Pig Genomic DNA with cDNA for Hepatic FMO. Analysis of genomic DNA was carried out with a probe

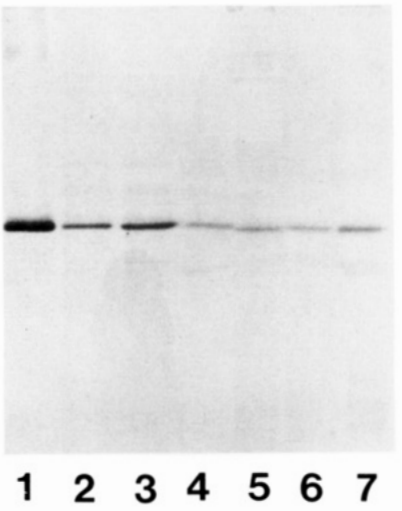

FIGURE 2: Detection of microsomal protein with antibodies to the flavin-containing monooxygenase purified from pig liver. Samples of microsomal protein from pig liver (lane 2, 5 μg; lane 3, 10 μg), kidney (lane 4, 10 μg; lane 5, 30 μg), and lung (lane 6, 10 μg; lane 7, 30 μg) and purified FMO (lane 1, 0.2 μg) were analyzed by immunoblotting with antibodies to the FMO as described under Materials and Methods.

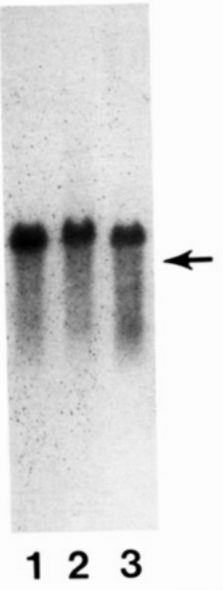

FIGURE 3: Northern analysis of mRNA from pig tissues with cDNA encoding the flavin-containing monooxygenase expressed in pig liver. Samples of mRNA from pig kidney (lane 1, 5 μg), lung (lane 2, 5 μg), and liver (lane 3, 5 μg) were subjected to electrophoresis in agarose, transferred to nylon membrane, and probed with cDNA to the FMO as described under Materials and Methods. The arrow shows the location of mRNA for cytochrome P-450 IIB (2.4 kilobases).

constructed from clone B by restriction with *Pst*I. The 5′ *Pst*I fragment (bases 1–388 of clone B; see Table I) was isolated and labeled with ^{32}P by the oligo primer directed method (Feinberg & Vogelstein, 1983). Samples of genomic DNA (10 μg) from liver were restricted with *Xho*I, *Cla*I, *Bam*HI, *Kpn*I, *Sca*I, *Pst*I, or *Eco*RI and hybridized with the 5′ probe following electrophoresis in agarose and transfer to a nylon membrane (Figure 4). Autoradiography resulted in the visualization of a single band of DNA with each sample examined. The apparent sizes of these bands ranged from about 4400 to 23000 bases. It should be noted that none of the enzymes used restrict the cDNA between bases 1 and 388 but that four of the enzymes do restrict the cDNA at or to the 3′ of base 388: *Pst*I at bases 388, 1139, and 1718, *Sca*I at

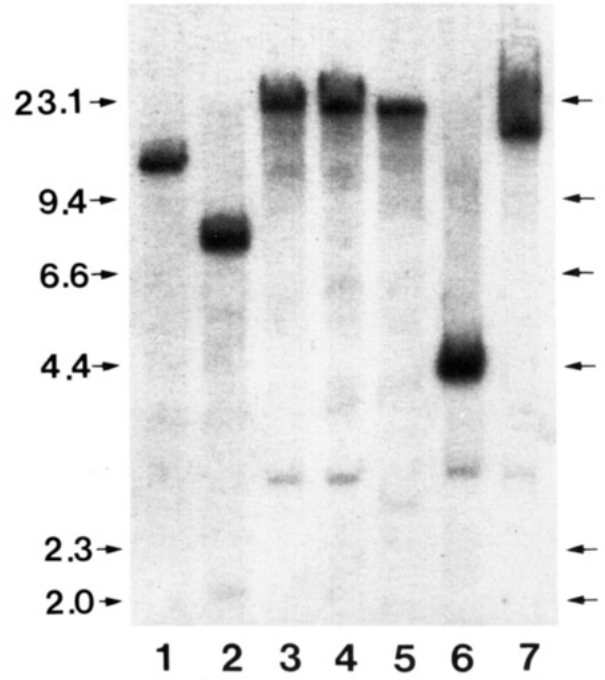

FIGURE 4: Southern analysis of genomic DNA from pig liver with cDNA encoding the flavin-containing monooxygenase expressed in pig liver. Samples of genomic DNA (10 μg) from liver were restricted with *Xho*I (lane 1), *Cla*I (lane 2), *Bam*HI (lane 3), *Kpn*I (lane 4), *Sca*I (lane 5), *Pst*I (lane 6), or *Eco*RI (lane 7), subjected to electrophoresis in agarose, transferred to nylon membrane, and probed with a 5′ fragment (bases 1–388) of the cDNA encoding the FMO. The sizes of standards are shown in kilobases.

base 1700, *Kpn*I at base 740, and *Bam*HI at base 1422. The cDNA contains no restriction sites for *Eco*RI, *Cla*I, or *Xho*I. Hybridization of genomic DNA with full-length cDNA indicated the presence of a second, partially related gene, which could not be detected under conditions of high stringency (not shown). Similar cross-reactivity was not observed with the 5′ probe.

Discussion

The complete primary structure of the flavin-containing monooxygenase (EC 1.14.13.8) expressed in liver and other tissues of pig has been derived from a cDNA nucleotide sequence. The majority of this sequence was duplicated by analysis of cDNA clones isolated from independent libraries. The nucleotide sequence extends 56 bases to the 5′, and 428 bases to the 3′, of the coding region which contains 1596 bases encoding 532 amino acids. The mRNA detected by hybridization with cDNA encoding the FMO contains about 2600 bases, which indicates that clone B is missing about 500 bases, most likely from the 3′ end.

Evidence that the cDNAs isolated encode the FMO is provided by the manner in which the original clone was detected (antibody recognition), the marked similarity between the amino acid compositions derived from the cDNA sequence and that determined directly from the protein (Poulsen & Ziegler, 1979), and the match between sequences of peptides isolated from the protein and present in the derived primary structure. The last evidence has been particularly difficult to obtain, as a result of both a blocked N-terminal residue and the problems encountered with automated gas-phase sequencing of isolated peptides. Ambiguous sequence data were obtained for peptides produced by digestion with trypsin and purified to apparent homogeneity by HPLC or FPLC, as well as for most of those separated by electrophoresis on polyacrylamide gels. Why this protein is so refractory to sequencing is not clear, but the problem likely has to do with the formation of hydrophobic peptides that are not completely separated by the methods employed.

Because the amino terminus of the FMO is blocked, the presence of methionine at this position in the active form of the enzyme remains to be proven. Some processing of the protein at the amino terminus and subsequent modification of a residue other than methionine are possible. However, the extent to which this can occur is likely limited by the putative FAD-binding site that begins with the valine at position 5 and extends to at least the leucine at position 15. The sequence of these 11 residues with the exception of a valine-isoleucine substitution at residue 8 is identical with the sequence of the FAD-binding domain of tryptophan 2-monooxygenase from *Pseudomonas savastanoi* (Yamada et al., 1985) and contains the glycine-x-glycine-x-x-glycine (positions 9–14) characteristic of the FAD-binding sites present in a number of proteins (Untucht-Grau et al., 1981).

Analysis of genomic DNA indicates that there is a single gene encoding the hepatic FMO. Therefore, the broad substrate specificity associated with the FMO in hepatic microsomes is likely due to the catalytic activity of a single enzyme since the other form of FMO that has been described is not expressed in liver (Tynes & Philpot, 1987; Williams et al., 1984). It should be noted that these results do not suggest that the pulmonary FMO is absent in pig. However, they do indicate that the genes for the pulmonary and hepatic enzymes are not highly related, which is consistent with the lack of immunochemical identity between the two proteins. Immunochemical analysis of microsomal samples shows that antibodies to the pig liver FMO recognize proteins expressed in lung and kidney, a finding that is true for all species examined except rabbit (Tynes & Philpot, 1987). The presence of "hepatic" FMO in lung and kidney is further substantiated by detection of mRNA by hybridization with cDNA encoding the liver enzyme. The lack of correlation between the relative amounts of protein and mRNA associated with the FMO in the three tissues remains to be explained.

Acknowledgments

The assistance of personnel at the J & B Sausage Co., Waelder, TX, in securing fresh hog liver for enzyme preparation is gratefully acknowledged.

Registry No. Flavin-containing monooxygenase, 9038-14-6; DNA (pig liver flavin-containing monooxygenase messenger RNA complementary), 123962-86-7; flavin-containing monooxygenase (pig liver protein moiety reduced), 123962-90-3.

References

Aviv, H., & Leder, P. (1972) *Proc. Natl. Acad. Sci. U.S.A. 80*, 1408–1412.

Bailey, J. M., & Davidson, N. (1976) *Anal. Biochem. 70*, 75–85.

Beaty, N. B., & Ballou, D. P. (1980) *J. Biol. Chem. 255*, 3817–3819.

Biggin, M. D., Gibson, T. J., & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3963–3965.

Chirgwin, J. M., Przybyla, A. E., Macdonald, R. J., & Rutter, W. J. (1979) *Biochemistry 18*, 5294–5299.

Dannan, G. A., & Guengerich, F. P. (1982) *Mol. Pharmacol. 22*, 787–794.

Deininger, P. L. (1983) *Anal. Biochem. 129*, 216–223.

Devereux, J., Haeberli, P., & Smithies, O. (1984) *Nucleic Acids Res. 12*, 387–395.

Domin, B. A., Serabjit-Singh, C. J., & Philpot, R. M. (1984) *Anal. Biochem. 136*, 390–396.

Enrietto, P. J., Payne, L. N., & Hayman, M. J. (1983) *Cell 35*, 369-379.
Feinberg, A., & Vogelstein, B. (1983) *Anal. Biochem. 132*, 6-10.
Gasser, R., Negishi, M., & Philpot, R. M. (1988) *Mol. Pharmacol. 32*, 22-30.
Glisin, V., Crkvenjakov, R., & Byus, C. (1974) *Biochemistry 13*, 2633-2637.
Gross-Bellard, M., Oudet, P., & Chambon, P. (1973) *Eur. J. Biochem. 26*, 32-38.
Gubler, U., & Hoffman, B. J. (1983) *Gene 25*, 263-269.
Kimura, T., Kodama, M., & Nagata, C. (1983) *Biochem. Biophys. Res. Commun. 110*, 640-645.
Laemmli, U. K. (1970) *Nature 227*, 680-685.
Laemmli, U. K., & Favre, M. (1973) *J. Mol. Biol. 80*, 575-590.
Maniatis, T., Jeffery, A., & Kleid, D. G. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 1184-1189.
Messing, J. (1983) *Methods Enzymol. 57*, 20-78.
Nebert, D. W., Nelson, D. R., 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. (1989) *DNA 8*, 1-13.
Okayama, H., & Berg, P. (1982) *Mol. Cell. Biol. 2*, 161-170.
Poulsen, L. L., & Ziegler, D. M. (1979) *J. Biol. Chem. 254*, 6449-6455.
Sabourin, P. J., Smyser, B. P., & Hodgson, E. (1984) *Int. J. Biochem. 16*, 713-720.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463-5467.
Schroeder, W. A. (1967) *Methods Enzymol. 11*, 445-461.
Towbin, H., Staehelin, T., & Gordon, J. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4350-4354.
Tynes, R. E., & Hodgson, E. (1985) *Arch. Biochem. Biophys. 240*, 77-93.
Tynes, R. E., & Philpot, R. M. (1987) *Mol. Pharmacol. 31*, 569-574.
Tynes, R. E., Sabourin, P. J., & Hodgson, E. (1985) *Biochem. Biophys. Res. Commun. 126*, 1069-1075.
Untucht-Grau, R., Schirmer, R. H., Schirmer, I., & Krauth-Siegel, R. L. (1981) *Eur. J. Biochem. 120*, 407-419.
Williams, D. E., Ziegler, D. M., Nordin, D. J., Hale, S. E., & Masters, B. S. S. (1984) *Biochem. Biophys. Res. Commun. 125*, 116-122.
Yamada, T., Palm, C. J., Brooks, B., & Kosuge, T. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 6522-6526.
Ziegler, D. M. (1988) *Drug Metab. Rev. 6*, 1-32.
Ziegler, D. M., & Mitchell, C. H. (1972) *Arch. Biochem. Biophys. 150*, 116-125.

Structure of the Gene for Human Butyrylcholinesterase. Evidence for a Single Copy[†,‡]

Martine Arpagaus,[§] Matthew Kott, Kostas P. Vatsis, Cynthia F. Bartels, Bert N. La Du, and Oksana Lockridge*
Department of Pharmacology, University of Michigan Medical School, Medical Science Building I, Ann Arbor, Michigan, 48109-0626

Received May 19, 1989; Revised Manuscript Received August 16, 1989

ABSTRACT: We have isolated five genomic clones for human butyrylcholinesterase (BChE), using cDNA probes encoding the catalytic subunit of the hydrophilic tetramer [McTiernan et al. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6682-6686]. The BChE gene is at least 73 kb long and contains four exons. Exon 1 contains untranslated sequences and two potential translation initiation sites at codons -69 and -47. Exon 2 (1525 bp) contains 83% of the coding sequence for the mature protein, including the N-terminal and the active-site serine, and a third possible translation initiation site (likely functional), at codon -28. Exon 3 is 167 nucleotides long. Exon 4 (604 bp) codes for the C-terminus of the protein and the 3′ untranslated region where two polyadenylation signals were identified. Intron 1 is 6.5 kb long, and the minimal sizes of introns 2 and 3 are estimated to be 32 kb each. Southern blot analysis of total human genomic DNA is in complete agreement with the gene structure established by restriction endonuclease mapping of the genomic clones: this strongly suggests that the BChE gene is present in a single copy.

Two types of cholinesterases exist in vertebrates: acetylcholinesterase (AChE,[1] EC 3.1.1.7) and butyrylcholinesterase (BChE, EC 3.1.1.8). Although the two enzymes present parallel series of molecular forms including monomers, dimers, tetramers, and complex oligomers with collagen-like elements, they differ in their substrate specificity and inhibitor sensitivity [reviews in Massoulié and Bon (1982), Massoulié and Toutant (1988), and Chatonnet and Lockridge (1989)]. They are distinct enzymes encoded by two different but related genes.

The physiological function of AChE is to hydrolyze acetylcholine at cholinergic synapses, but the role of BChE remains unclear. The serum BChE, however, plays a key role in the hydrolysis of the muscle relaxant drug succinylcholine which is administered during anaesthesia (Kalow & Gunn, 1957). Patients with abnormal genetic variants of BChE can

[†] This work was supported by U.S. Army Medical Research and Development Command Contract DAMD17-86-C-6037 (to O.L.) and by National Institutes of Health Grant GM27028 (to B.N.L.).
[‡] The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J02879.
* To whom correspondence should be addressed.
[§] Permanent address: INRA, Phytopharmacie, Route de Saint Cyr, 78000 Versailles, France.

[1] Abbreviations: AChE, acetylcholinesterase; BChE, butyrylcholinesterase; bp, base pair(s); nt, nucleotide(s); kb, kilobase(s).

0006-2960/90/0429-124$02.50/0