AMINO-TERMINAL SEQUENCE OF PHENOBARBITAL-INDUCIBLE CYTOCHROME P-450 FROM RABBIT LIVER MICROSOMES: SIMILARITY TO HYDROPHOBIC AMINO-TERMINAL SEGMENTS OF PREPROTEINS

David A. Haugen,* Lyman G. Armes,† Kerry T. Yasunobu,† and Minor J. Coon‡

*Division of Biological and Medical Research, Argonne National Laboratory, Argonne, Illinois 60439

†Department of Biochemistry and Biophysics, University of Hawaii Medical School, Honolulu, Hawaii 96822

‡Department of Biological Chemistry Medical School, The University of Michigan, Ann Arbor, Michigan 48109

Received June 9,1977

SUMMARY: The amino-terminal sequence of two electrophoretically homogeneous forms of rabbit liver microsomal cytochrome P-450, P-450LM$_2$ and P-450LM$_4$, has been examined by automated Edman degradation. Methionine is the amino terminus of P-450LM$_2$, and 17 of the first 20 residues are hydrophobic, including two clusters of five consecutive leucines. The composition and sequence of this region are similar to those of the short-lived hydrophobic amino-terminal precursor segments of certain other proteins, especially myeloma immunoglobulin light chains and pancreatic zymogens. Multiple amino-terminal residues, including methionine, were detected for P-450LM$_4$ suggesting the presence of several highly similar forms of P-450 or that partial proteolysis had occurred.

During the past few years, several different forms of liver microsomal cytochrome P-450 (P-450LM) have been separated and purified (1-7). We have recently begun an investigation of the primary structure of P-450LM as an aid in understanding the variations in the biological and chemical properties of this family of proteins, which play an essential role in the metabolism of prostaglandins, fatty acids, and steroids as well as a host of foreign compounds, including drugs and carcinogens.

Two electrophoretically homogeneous forms of P-450LM from rabbit liver (1) were examined: P-450LM$_2$, which is induced by phenobarbital but is absent in microsomes from untreated animals, and P-450LM$_4$, which is induced by β-naptho-flavone and certain carcinogens but is also present at significant levels in microsomes from phenobarbital-treated and untreated animals. These cytochromes have subunit molecular weights of 49,000 and 55,000, respectively, and also differ in their spectral, immunochemical, and catalytic properties (1,2,8).

The data reported here confirm the purity of P-450LM$_2$, and indicate that

ISSN 0006-291X

the composition and sequence of its extremely hydrophobic amino-terminal region are remarkably similar to those of the hydrophobic amino-terminal segments present in certain preproteins, but not in the corresponding "mature" proteins (9-15). To our knowledge, this is the first observation of such a sequence appearing in a "mature" protein.

MATERIALS AND METHODS

Rabbit liver P-450LM$_2$ and P-450LM$_4$: The two electrophoretically homogeneous forms of P-450LM were isolated from liver microsomes of phenobarbital-treated rabbits as previously described (1). Two pooled preparations of P-450LM$_2$ with specific contents of 16.2 and 12.1 nmol per mg of protein were used. A single pooled preparation of P-450LM$_4$ was used which had a specific content of 11.3 nmol per mg of protein. The P-450LM$_4$ was isolated from liver microsomes of phenobarbital-induced rabbits but is not induced by this agent. Due to partial heme loss during purification of both proteins, some apoenzyme was present. The heme was removed from each preparation with acidic acetone (16), and the apoprotein was converted to the S-(β-carboxymethyl)cysteinly derivative (17).

Amino-terminal sequence and amino acid analysis: For each sequence analysis, about 110 nmol of the carboxymethylated protein was placed in a Beckman Model 890 sequencer, and automated Edman degradation was performed using a protein double cleavage program (18,19). The amount of protein used in each experiment was determined by amino acid analysis of 24-hour acid hydrolysates, assuming that the glutamic acid content was 42 and 46 mol per mol protein for P-450LM$_2$ and P-450LM$_4$, respectively (1). The yields of the phenylthiohydantoins were determined by gas chromatography (20), except for glutamic acid, arginine, and histidine, for which the yield was estimated by amino acid analysis of acid hydrolysates with 50 to 70% recovery (21). The data were corrected for background and for out of step degradation where appropriate (22). Amino acid analyses were performed as described by Spackman et al. (23) using a Beckman Model 120 analyzer.

RESULTS AND DISCUSSION

Edman degradation of P-450LM$_2$: The results of sequence analysis of P-450LM$_2$ are shown in Table I. For all residues except those enclosed by parentheses, identical unambiguous assignments were made in two experiments using different preparations of the protein. The repetitive yield was 96%.

A high degree of purity of the protein is indicated by the initial yields (Table I), the extrapolated initial yield (83%), and the observation that for Steps 1 to 4 no nonprimary residues were identified with yields greater than 3%. Similarly, treatment of the native protein with carboxypeptidase released arginine with a yield greater than 95% (1). Ingelman-Sundberg and Gustafsson (24) have reported the resolution of P-450LM$_2$ into several heme-containing forms by isoelectric focusing. In our experience, electrophoretically homogeneous preparations of P-450LM$_2$ are not obtained by the methods described in their report, and their method of SDS-polyacrylamide gel electrophoresis gives poorer resolution than the discontinuous buffer system used in our laboratory (1).

TABLE I

Automated Edman Degradation of P-450LM$_2$

Step	Amino Acid	Yield (%)	Identification Method		
			GC	HYD	TLC
1	Methionine	79	+	+	+
2	Glutamic acid	> 35		+	+
3	Phenylalanine	82	+	+	+
4	Serine	50	+		+
5	Leucine	55	+	+	
6	Leucine	56	+	+	
7	Leucine	63	+	+	
8	Leucine	60	+	+	
9	Leucine	55	+	+	
10	Alanine	53	+	+	
11	Phenylalanine	50	+	+	
12	Leucine	54	+	+	
13	Alanine	47	+	+	
14	Glycine	50	+	+	
15	Leucine	55	+	+	
16	Leucine	48	+	+	
17	Leucine	47	+	+	
18	Leucine	45	+	+	
19	Leucine	49	+	+	
20	Phenylalanine	39	+	+	
21	(Arginine)	> 10		+	
22	Glycine	38	+	+	
23	(Histidine)	> 10		+	
24	(Proline)	14	+	+	
25	X	-			
26	Alanine	28	+	+	
27	X	-			
28	Glycine	23	+	+	
29	X	-			
30	Leucine	27	+	+	

The parentheses and symbol X indicate tentative assignments and unidentified residues, respectively. The identification methods are: GC, gas chromatography (20), HYD, amino acid analysis of acid hydrolysates (21), and TLC, thin layer chromatography (18). Yields estimated by amino acid analysis were not corrected for incomplete hydrolysis or instability.

A brief report by Ozols et al. (25) has indicated that the amino terminus of the phenobarbital-inducible form of rabbit liver microsomal cytochrome P-450 is blocked, as determined by automated sequence analysis. The difference

between that finding and the results presented here could conceivably be due to cleavage of the native protein in our studies during isolation or derivatization, thus exposing a free amino terminus. However, this seems unlikely in view of the high yields of the initial residues, which would require that such a postulated cleavage be highly specific and quantitative. Electrophoresis of the carboxymethylated protein demonstrated that its subunit molecular weight was the same as for the native protein, and that no smaller polypeptides were present that were indicative of cleavage.

Comparison of P-450LM$_2$ with preproteins: The amino-terminal region of P-450LM$_2$ has 17 to 18 hydrophobic residues in the first 20 positions, including two clusters of five consecutive leucines. These unusual properties led us to examine previous reports describing the composition and sequence of hydrophobic amino-terminal precursor segments for several proteins, including mouse myeloma immunoglobulin light chains (10), dog pancreatic zymogens (11), rat proinsulin (12), bovine proparathyroid hormone (13), human placental lactogen (14), and a bacterial outer membrane lipoprotein (15). For each of these secretory proteins, a preprotein has been isolated from an *in vitro* protein synthesis system. The preproteins contain a hydrophobic amino-terminal precursor segment composed of 15 to 25 amino acids preceding the amino-terminal sequence of the native protein isolated by conventional techniques. Many of the precursor segments have methionine as the amino-terminal residue, which is likely to be the initiator methionine (26). It is believed that the precursor segments may be responsible for binding of the nascent polypeptide to the endoplasmic reticulum or other membranous sites during transport (9). During a process of maturation or transport, the precursor segment is removed in a manner that is not yet understood, although enzymatic removal of the precursor segment of placental lactogen has been observed (14).

In Fig. 1, the sequence of the first 20 residues of P-450LM$_2$ is compared with those of the precursor segments of one of five homologous dog pancreatic zymogens (protein 5, Ref. 11), and three mouse myeloma immunoglobulin light chains (10). The similarity of the sequences of these functionally unrelated proteins from diverse sources is evident. For the pancreatic zymogen, the gaps introduced at positions 1, 2 and 3, and the absence of an amino-terminal methionine may indicate that the initial tripeptide was lost before the protein was isolated from the *in vitro* system. The amino-terminal sequence of P-450LM$_2$ does not resemble those of the hydrophobic precursor segments of certain other proteins, however (12-15).

The hydrophobic region of the "mature" P-450LM$_2$ is apparently the equivalent of the relatively short-lived precursor segments described above. Almost all have an amino-terminal methionine and are similar with respect to hydro-

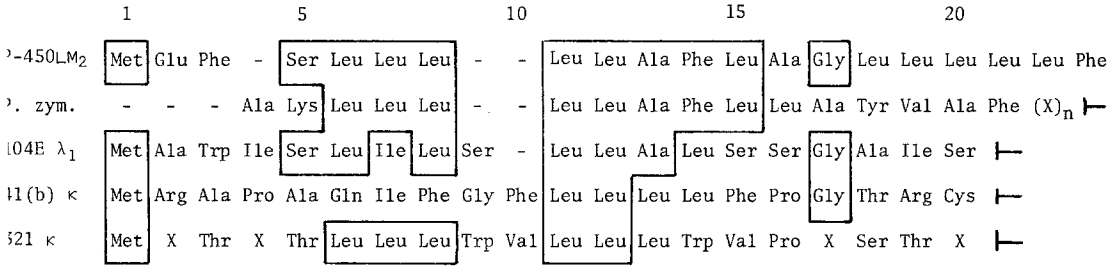

FIG. 1. Comparison of amino-terminal sequence of P-450LM$_2$ with hydrophobic precursor segments of other proteins. The proteins are: P-450LM$_2$, this report; P. zym., dog pancreas zymogen, believed to be procarboxypeptidase (protein 5, Ref. 11); 104E λ_1, 41(b) κ, and 321 κ, immunoglobulin light chains MOPC-104E λ_1, MOPC-41 (b) κ, and MOPC-321 κ, respectively (10). In the study of the pancreatic zymogens, leucine and isoleucine were not resolved (11), and leucine has been assumed to be present for the purpose of this comparison. Gaps (indicated by -) were introduced to allow maximal alignment. The symbol ⊢ indicates the beginning of the sequence of the mature protein. Residues which are homologous with P-450LM$_2$ are enclosed.

phobicity, apparent length, and high leucine content. The retention of the hydrophobic segment in P-450LM$_2$ suggests that the segment may function in the binding of the cytochrome to the endoplasmic reticulum or in its orientation in functional complexes with other components of the hydroxylation system such as NADPH-cytochrome P-450 reductase and phospholipid. Cytochrome b_5, which is also present in liver microsomes, has a hydrophobic domain that is required for binding of the protein to membranes and for the formation of functional complexes (27). However, this domain is at the carboxyl end of the protein, and its sequence does not resemble that described here for P-450LM$_2$ (27,28).

Comparison of P-450LM$_2$ and P-450$_{cam}$: The amino-terminal sequences of P-450LM$_2$ and of P-450$_{cam}$ from Pseudomonas putida (29) are compared in Fig. 2. These two proteins have similar amino acid compositions (1,31,32), immunochemical properties (32), and optical spectra (1,29), but differ greatly in their solubility, substrate specificity, and the requirement of P-450LM$_2$ for a phospholipid and P-450$_{cam}$ for an iron-sulfur protein for activity. A comparison of the amino-terminal sequences of the two proteins reveals no similarity in their composition or sequence (Fig. 2). Obviously, the data do not preclude the presence of similar regions elsewhere in the proteins.

Edman degradation of P-450LM$_4$: The first cycle of degradation of P-450LM$_4$ released glycine, methionine, serine, and alanine in the ratio 1:2:2:3 with the sum of their yields equal to approximately 85%. At each of the nine subsequent steps, two to four of the following residues were identified: methionine,

	1				5				10		
P-450LM$_2$	Met	Glu	Phe	Ser	Leu	Leu	Leu	Leu	Leu	Ala	Phe	Leu	Ala	Gly
P-450$_{cam}$	Thr	Thr	Glu	Thr	Ile	Gln	Ser	Asn	Ala	Asn	Leu	Ala	Pro	Leu

FIG. 2. Comparison of amino-terminal sequences of P-450LM$_2$ and P-450$_{cam}$. The sequence for P-450$_{cam}$ is from Tanaka et al. (30).

leucine, valine, proline, alanine, and serine. The presence of multiple amino-terminal residues may indicate that the preparation contains several proteins or several highly similar forms of the cytochrome. This seems unlikely, however, in view of the electrophoretic and immunochemical evidence for homogeneity, as well as the finding of a single carboxyl-terminal residue, lysine (1,8). An alternative conclusion is that the mixture of polypeptides represents a family of cytochromes differing only in the degree to which a hydrophobic amino-terminal segment has been degraded by proteolysis in vivo or during isolation. This view is supported by the identification of methionine as one of the amino-terminal residues, the nearly exclusive identification of hydrophobic residues for Steps 1 to 10, and by data in previous reports indicating relatively low yields of the amino-terminal methionine in some precursor segments synthesized in vitro, presumably due to hydrolysis (10,14).

ACKNOWLEDGMENTS

The authors are grateful to Dr. Marianne Schiffer (Argonne National Laboratory) for directing their attention to the previous reports describing precursor segments, and to Sylvia B. Dahl and Susan O. Krezoski (The University of Michigan) for preparation of the proteins. This research was supported by Grants PCM76-14947 from the National Science Foundation, and AM-10339 from the United States Public Health Service, to M.J.C.; Grant GM22556 from the United States Public Service, to K.T.Y.; and by the U.S. Energy Research and Development Administration, D.A.H.

REFERENCES

1. Haugen, D. A., and Coon, M. J. (1976) J. Biol. Chem. 251, 7929-7939.
2. Haugen, D. A., van der Hoeven, T. A., and Coon, M. J. (1975) J. Biol. Chem. 250, 3567-3570.
3. van der Hoeven, T. A., Haugen, D. A., and Coon, M. J. (1974) Biochem. Biophys. Res. Commun. 60, 569-575.
4. Imai, Y., and Sato, R. (1974) Biochem. Biophys. Res. Commun. 60, 8-14.
5. Hashimoto, C., and Imai, Y. (1976) Biochem. Biophys. Res. Commun. 68, 821-827.
6. Ryan, D., Lu, A.Y.H., West, S. B., and Levin, W. (1975) J. Biol. Chem. 250, 2157-2163.

7. Kawalek, J. C., Levin, W., Ryan, D., Thomas, P. E., and Lu, A. Y. H. (1975) Mol. Pharmacol. 11, 874-878.
8. Dean, W. L., and Coon, M. J. (1977) J. Biol. Chem. 252, 3255-3261.
9. Blobel, G., and Dobberstein, B. (1975) J. Cell. Biol. 67, 835-851.
10. Burstein, Y., and Schecter, I. (1977) Proc. Nat. Acad. Sci. USA 74, 716-720.
11. Devillers-Thiery, A., Kindt, T., Scheele, G., and Blobel, G. (1975) Proc. Nat. Acad. Sci. USA 72, 5016-5020.
12. Chan, S. J., Keim, P. L., and Steiner, D. F. (1976) Proc. Nat. Acad. Sci. USA 73, 1964-1968.
13. Kemper, B., Habener, J., Ernst, M. D., Potts, J. T., and Rich, A. (1976) Biochemistry 15, 15-19.
14. Birken, S., Smith, D. L., Canfield, R. E., and Boime, I. (1977) Biochem. Biophys. Res. Commun. 74, 106-112.
15. Inouye, S., Wang, S., Sekizawa, J., Halegoua, S., and Inouye, M. (1977) Proc. Nat. Acad. Sci. USA 74, 1004-1008.
16. Strittmatter, P. (1960) J. Biol. Chem. 235, 2492-2497.
17. Hirs, C. H. W. (1967) in Methods in Enzymology (C. H. W. Hirs, ed.) Vol. 11, pp. 199-203, Academic Press, New York.
18. Edman, P., and Begg, G. (1967) Eur. J. Biochem. 1, 80-91.
19. Niall, H. D. (1973) in Methods in Enzymology (C. H. W. Hirs, and S. N. Timasheff, eds.) Vol. 27D, pp. 942-1010, Academic Press, New York.
20. Pisano, J. J., and Bronzert, T. J. (1969) J. Biol. Chem. 244, 5597-5607.
21. Van Orden, H. O., and Carpenter, F. H. (1964) Biochem. Biophys. Res. Commun. 14, 399-403.
22. Smithies, O., Gibson, D., Fanning, E. M., Goodfliesh, R. M., Gilman, J. G., and Ballantyne, D. L. (1971) Biochemistry 10, 4912-4921.
23. Spackman, D. H., Moore, S., and Stein, W. H. (1958) Anal. Chem. 30, 1190-1206.
24. Ingelman-Sundberg, M., and J.-Å. Gustafsson (1977) FEBS Lett. 74, 103-106.
25. Ozols, J., Gerard, C., Imai, Y., and Sato, R. (1977) Fed. Proc. 36, 833.
26. Lodish, H. F. (1976) Ann. Rev. Biochem. 45, 39-72.
27. Ozols, J., Gerard, C., and Nobrega, F. G. (1976) J. Biol. Chem. 251, 6767-6774.
28. Corcoran, D., and Strittmatter, P. (1977) Fed. Proc. 36, 897.
29. Yu, C.-A., Gunsalus, I. C., Katagiri, M., Suhara, K., and Takemori, S. (1974) J. Biol. Chem. 249, 94-101.
30. Tanaka, M., Zeitlin, S., Yasunobu, K. T., and Gunsalus, I. C. (1976) in Iron and Copper Proteins (K. T. Yasunobu, H. F. Mower, and O. Hayaishi, eds.) pp. 263-269, Plenum Press, New York.
31. Tsai, R. L., Gunsalus, I. C., and Dus, K. (1971) Biochem. Biophys. Res. Commun. 45, 1300-1306.
32. Dus, K., Litchfield, W. J., Miguel, A. G., van der Hoeven, T. A., Haugen, D. A., Dean, W. L., and Coon, M. J. (1974) Biochem. Biophys. Res. Commun. 60, 15-21.