Structural Analysis of the Cysteine-Containing Peptides from the Major 3-Methylcholanthrene-Induced Isozyme of Cytochrome P-450 (P-450c) in Rat Liver Microsomes†

Mitsuru Haniu, Pau-Miau Yuan,‡ Dene E. Ryan, Wayne Levin, and John E. Shively*

ABSTRACT: Cytochrome P-450c, the major 3-methylcholanthrene-inducible isozyme of cytochrome P-450 in rat liver microsomes, was subjected to proteolytic digestion after S-carboxymethylation of the protein, and the peptides were resolved by high-pressure liquid chromatography. Since it is now recognized that cytochromes P-450 contain a thiolate as the axial fifth ligand of the heme, seven peptides containing eight cysteines were subjected to microsequence analysis. One cysteine-containing peptide (Tsa-56) was shown to possess 46–69% homology with a common peptide found in five other cytochromes P-450 but is *not* anticipated to be the heme-binding segment on the basis of X-ray crystallographic results obtained with *Pseudomonas putida* cytochrome P-450$_{cam}$. Analysis of the other cysteine-containing peptides in cytochrome P-450c revealed two peptides (Tsa-54 and T-46) of only limited homology with the highly conserved region of cytochromes P-450$_{cam}$, P-450$_{LM2}$, P-450b, and P-450e that are all presumed to contain the heme-binding cysteine. Another peptide that contained two cysteines and a stretch of hydrophobic residues (Tsa-47) had limited sequence homology with a similar peptide found in several other cytochromes P-450. This domain is located a short distance from the proposed heme-binding cysteine in other cytochromes P-450. Sequence analysis of a cysteine-containing peptide (T-30) from another 3-methylcholanthrene-inducible rat liver cytochrome P-450 (cytochrome P-450d) revealed 91% homology with peptide T-46 from cytochrome P-450c, but this peptide shows no significant homology with any of the cysteine-containing peptides from other cytochromes P-450. This very high degree of homology was not totally unanticipated since the two isozymes (cytochromes P-450c and P-450d) are coinducible by a variety of xenobiotics and they possess some shared immunochemical determinants.

Hepatic microsomal cytochrome P-450 is the terminal oxidase of an electron transport pathway that catalyzes a host of oxidation reactions involving a large variety of xenobiotic and endogenous substrates (Conney, 1967). The existence of multiple isozymes of cytochrome P-450, differing in substrate specificity, primary structure, electrophoretic mobility, and immunological properties, is well established (Coon, 1981; Guengerich et al., 1982; Ryan et al., 1982a). The total cytochrome P-450 in hepatic endoplasmic reticulum is increased after treatment of animals with a variety of xenobiotic substances (Conney, 1967), and certain of these inducers increase specific isozymes by as much as 50-fold (Thomas et al., 1979, 1983). For example, treatment of rats with phenobarbital leads to a dramatic increase in cytochromes P-450b and P-450e whereas 3-methylcholanthrene treatment markedly induces cytochrome P-450c and, to a lesser extent, cytochrome P-450d. Isosafrole induces cytochrome P-450d to a greater degree than cytochrome P-450c and also induces cytochromes P-450b and P-450e (Thomas et al., 1983). Changes in the amounts and relative proportions of these isozymes in animals markedly influences the pharmacological activity and toxicity of foreign chemicals (Conney & Levin, 1982).

Recent molecular studies have elucidated the total amino acid sequences of several cytochromes P-450 including rat hepatic cytochromes P-450b and P-450e (Fujii-kuryama et al., 1982; Yuan et al., 1983b) rabbit hepatic cytochrome P-450$_{LM2}$ (Heinemann & Ozols, 1983; Tarr et al., 1983), and *Pseudomonas putida* cytochrome P-450$_{cam}$ (Haniu et al., 1982a,b). Comparisons of sequence homology among these cytochromes P-450 have revealed the presence of two highly conserved cysteine-containing peptides, one near the NH_2-terminal region and one near the COOH-terminus. Interest in these portions of the primary sequences is based on evidence that a thiolate is the fifth ligand of the heme (White & Coon, 1980; Poulous et al., 1983), the site at which molecular oxygen is activated and substrate is hydroxylated. It is assumed that the heme-binding site, which imparts the unusual Soret maximum for the CO-reduced hemoprotein, would be evolutionarily conserved among various cytochromes P-450. A recent sequence analysis of the cysteine-containing peptides from porcine adrenocortical microsomal cytochrome P-450 (steroid 21-hydroxylase) revealed that three of these peptides exhibited sequence homology with cysteine-containing peptides of rat hepatic microsomal cytochromes P-450b and P-450e and *Pseudomonas putida* cytochrome P-450$_{cam}$ (Yuan et al., 1983a). However, the cysteine-containing peptide with the greatest sequence homology is unlikely to be the heme–thiolate ligand on the basis of the analogy to cytochrome P-450$_{cam}$ (Haniu et al., 1982c). These observations have made the sequence analysis of cysteine-containing peptides an important goal of our ongoing structural studies of microsomal cytochromes P-450.

We report here the amino acid sequences of the cysteine-containing peptides from rat hepatic cytochrome P-450c. This isozyme is of particular interest since it is the most efficient catalyst in the conversion of many polycyclic aromatic hydrocarbons to ultimate carcinogenic metabolites (Thakker et al., 1982; Levin et al., 1982) and is distinct from rat hepatic cytochromes P-450b and P-450e in its immunological prop-

† From the Division of Immunology, Beckman Research Institute of the City of Hope, Duarte, California 91010 (M.H., P.-M.Y., and J.E.S.), and the Department of Biochemistry and Drug Metabolism, Hoffmann-La Roche Inc., Nutley, New Jersey 07110 (D.E.R. and W.L.).
Received August 3, 1983; revised manuscript received December 30, 1983. J.E.S. and M.H. are members of the City of Hope Research Cancer Center supported by NCI Grant CA 33572.
‡ Present address: Beckman Instruments, Spinco Division, Palo Alto, CA 94304.

0006-2960/84/0423-2478$01.50/0

erties, substrate specificity, peptide maps, and NH_2- and COOH-terminal sequences and is under different regulatory control (Ryan et al., 1982a).

Experimental Procedures

Materials. Acetonitrile (HPLC grade)[1] was obtained from J. T. Baker Chemical Co. Trifluoroacetic acid was distilled as described previously (Yuan et al., 1983a). Iodoacetic acid (Sigma) was recrystallized from petroleum ether under subdued light and iodo[1-^{14}C]acetic acid (50 μCi/0.7 mg) was obtained from New England Nuclear. TPCK-treated trypsin, and carboxypeptidases A and B were obtained from Worthington Biochemical Co. *o*-Phthaldialdehyde, *Staphylococcus aureus* protease (strain V8, specific activity 500–700 units/mg), and carboxypeptidase Y were from Pierce Chemical Co.

Purification of Cytochromes P-450. Liver microsomal cytochromes P-450c and P-450d were purified to homogeneity from rats treated with 3-methylcholanthrene and isosafrole, respectively, as described previously (Ryan et al., 1980, 1982b).

Isolation of Cysteine-Containing Tryptic Peptides. The proteins (500 μg of cytochromes P-450c and P-450d) were S-carboxymethylated first with iodo[1-^{14}C]acetic acid and then with an excess of unlabeled iodoacetic acid in 6 M guanidinium chloride containing 0.1% 2-mercaptoethanol. The sample was dialyzed against 0.2 M ammonium bicarbonate (pH 8.0), treated with TPCK–trypsin (0.5 mg/mL, 0.2 M ammonium bicarbonate, 24 h at 37 °C, enzyme to protein ratio 1/50 w/w), and centrifuged to remove the insoluble tryptic core. The soluble fraction was chromatographed either on an Ultrasphere C-8 (250 × 4.6 mm) column or on a SynChropak RP C-8 (250 × 4.6 mm) column as described previously (Yuan et al., 1983a) with a linear gradient from 100% solvent I (0.1% TFA) to 70% solvent II (0.1:9.9:90 = TFA–H_2O–CH_3CN). The insoluble tryptic core material of cytochrome P-450c was further digested with *Staphylococcus aureus* V8 protease (1 mL of 0.2 M ammonium bicarbonate, 500–700 units of enzyme, 24 h at 37 °C). The resulting peptides (Tsa) were separated by HPLC on a SynChropak RP C-8 column (250 × 4.6 mm) as described above. Peptides were detected by absorbance at 206 nm and manually collected, and 20 μL from each peak fraction was counted for radioactivity. (Carboxymethyl)cysteine-containing peptides were detected as radioactive peaks and were quantitated by amino acid analysis.

Amino Acid Analysis. Amino acid analysis was performed essentially by the method of Del Valle & Shively (1979).

NH_2-Terminal Sequence. Approximately 1 nmol of each peptide was subjected to automated Edman degradations on an updated, modified Beckman 890C sequencer, as previously described by Shively (1981). PTH derivatives were identified by HPLC on an Ultrasphere (Altex) ODS column with a Waters Associates chromatograph (Hawke et al., 1982).

COOH-Terminal Sequence. COOH-Terminal sequence analysis of peptides was performed by time course analysis of amino acid released by carboxypeptidase A, B, or Y. The amino acids were monitored as their *o*-phthaldialdehyde derivatives by HPLC according to the method of Jones et al. (1981). Although this method gives a poor response for lysine and no response for proline, the reproducibility of the lysine response for standards was good (with the inclusion of Brij-35 in the OPA reaction mixture), and the presence or absence of proline was determined by conventional ninhydrin based amino acid analysis (see Amino Acid Analysis).

[1] Abbreviations: NADPH, nicotinamide adenine dinucleotide phosphate (reduced form); TPCK, tosylphenyl chloromethyl ketone; HPLC, high-performance liquid chromatography; ODS, octadecylsilane; PTH, phenylthiohydantoin; TFA, trifluoroacetic acid.

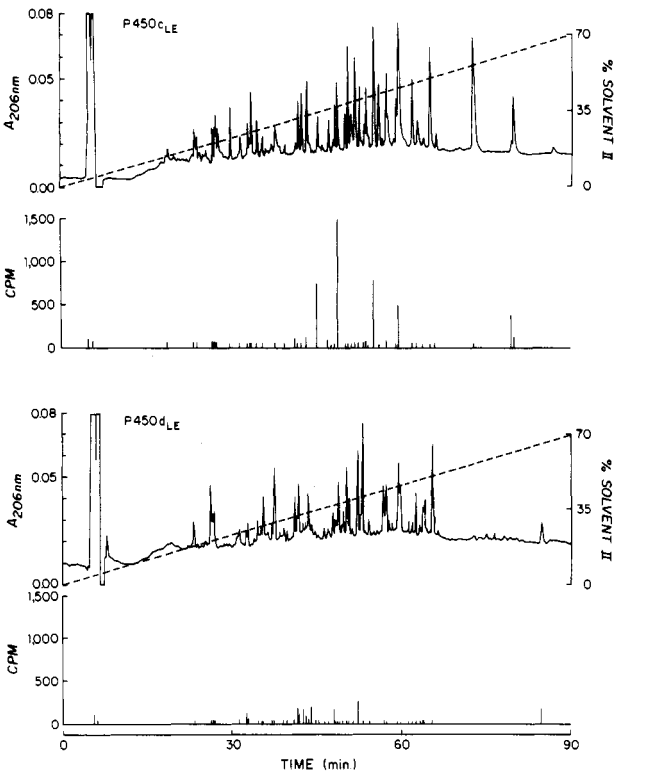

FIGURE 1: Comparative tryptic maps of cytochromes P-450c (top) and P-450d (bottom). Approximately 2 nmol of the soluble tryptic peptides of the S-carboxymethylated isozymes was chromatographed on an Ultrasphere C-8 column (250 × 4.6 mm, 5 μm particle size). A linear gradient program was used from 100% solvent I (0.1% TFA) to 70% solvent II (TFA–H_2O–CH_3CN, 0.1:9.9:90 v/v/v) at a flow rate of 0.9 mL/min. Fractions were collected manually, and a 20-μL aliquot from each fraction was counted for ^{14}C radioactivity (shown on the bottom portion of each map).

Results and Discussion

Cysteine-Containing Peptides. Five to seven nanomoles of S-carboxymethylated cytochromes P-450c and P-450d were digested with TPCK-treated trypsin, and the soluble peptides were separated by reverse-phase HPLC as described under Experimental Procedures. The comparative results for the two isozymes (performed under identical analytical conditions) are shown in Figure 1. On the basis of the relative retention times and peak height ratios, there are apparently no common peptides or fragmentation patterns between these two isozymes. This result is in contrast to the comparative HPLC maps of cytochromes P-450b and P-450e which have been shown to possess over 90% sequence homology (Yuan et al., 1983b). Thus, in spite of their immunological relatedness, these two isozymes must possess enough amino acid sequence changes so that they possess no identical tryptic peptides.

In order to prepare peptides for sequence analysis, preparative maps for each of the isozymes was performed on different reverse-phase columns as shown in Figure 2. Four cysteine-containing peptides were identified from cytochrome P-450c, whereas seven to eight cysteine-containing peptides were obtained from cytochrome P-450d. In order to completely resolve the additional peptides obtained from cytochrome P-450d, it was necessary to use slightly different chromatographic conditions. In contrast to cytochrome P-450c, P-450d gave less insoluble tryptic core. Thus, it was concluded that all of the cysteine-containing peptides from cytochrome P-450d were obtained by a single proteolytic digestion. Three additional cysteine-containing peptides were isolated by treatment of the insoluble tryptic core fraction of cytochrome P-450c with *S. aureus* protease (Figure 3). The

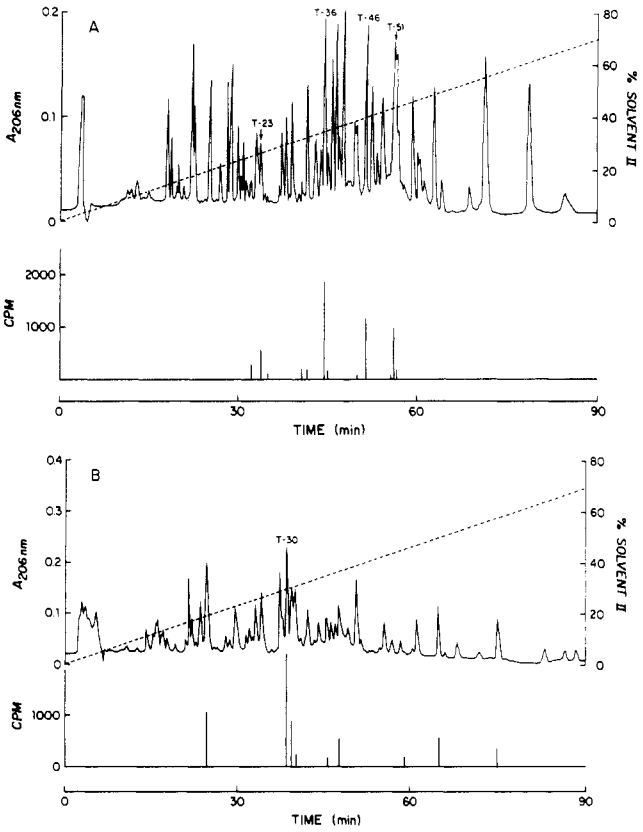

FIGURE 2: HPLC mapping of soluble tryptic peptides from cytochrome P-450c (A) and P-450d (B). Approximately 5 nmol of a tryptic digest of cytochrome P-450c was centrifuged to remove the insoluble core material, and the soluble material was loaded onto an Ultrasphere C-8 column (250 × 4.6 mn, 5 μm particle size). The tryptic peptide map of cytochrome P-450d was obtained in a similar manner except for the use of a SynChropak RP C-8 column (250 × 4.1 mm, 10 μm particle size). The peptides sequenced are indicated for both maps.

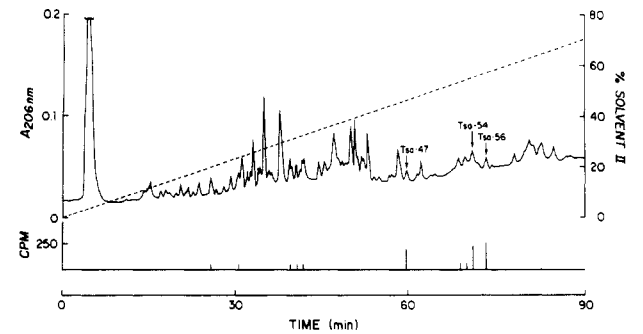

FIGURE 3: Peptide mapping of *S. aureus* proteolytic digest of tryptic core material from cytochrome P-450c. The insoluble fraction from the tryptic digest was further digested with *S. aureus* protease and applied to a Synchropak RP C-8 column (250 × 4.1 mn, 10 μm particle size). Chromatographic conditions and the detection method of cysteine-containing peptides were the same as in Figure 1. The peptides sequenced are indicated.

overall yield of cysteine-containing peptides from the double proteolytic digest was low presumably due to their relatively low solubility, but sufficient quantities were obtained for further structural analysis. The sequences of the cysteine-containing peptides were established by amino acid compositional analysis, NH_2-terminal microsequence analysis, and a time course of carboxypeptidase digestion [the data are available in Tables I and III in the supplementary material (see paragraph at end of paper regarding supplementary material)]. The sequences are summarized in Figure 4. Sufficient material was not available to perform COOH-terminal analysis on the three peptides obtained from the in-

Peptide	Sequence
	1 5 9
T-23	CEHFQVQVR
	1 5 10 13
T-36	DITDSLIEHCQDR
	1 5 10 15 20 22
T-46	SFSIASDPTLASSCYLEEHVSK
	1 5 10 15 20 25
T-51	GHCVFVNQWQVNHDQELWGDPNEFR
	1 5 10 15
Tsa-47	YLVVSVANVICAICF ...
	1 5 10 13
Tsa-56	CIGETIGRLEVFL...
	1 5 10 12
Tsa-54	CLTSSGTLDK(H)L ...

FIGURE 4: Sequence of cysteine-containing peptides from cytochrome P-450c. Wavy lines show the hydrophobic or nonpolar amino acid regions. Cysteine residues are shown in boldface.

P-450c (Tsa-56)	CIGETIGRLEVFL
P-450b/e (436-448)	CLGEGIARNELFL
P-450LM2 (436-448)	CLGEGIARTELFL
P-450c21 (Tsa-23)	CLGEPLARLELFV
P-450cam (355-367)	CLGQSLARREIIV

FIGURE 5: Cysteine-containing peptides with high sequence homology in cytochromes P-450. Sequences are from Fujii-Kuriyama et al. (1982), Yuan et al. (1983a,b), Tarr et al. (1983), and Haniu et al. (1982b). Isoleucine and leucine are considered equivalent in this comparison.

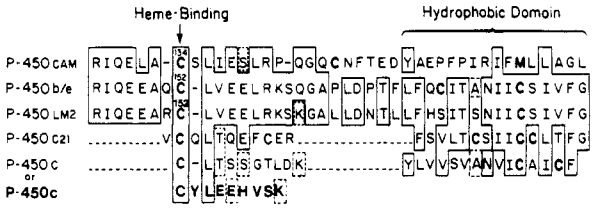

FIGURE 6: Comparison of the proposed heme-binding site of cytochromes P-450. Cys-152 (cytochromes P-450b and P-450e), Cys-152 (cytochrome $P\text{-}450_{LM2}$), Cys-134 (cytochrome $P\text{-}450_{cam}$), and the corresponding cysteine in cytochromes $P\text{-}450_{C21}$ and P-450c are proposed to be heme-binding thiol groups. Due to the low degree of homology, either one of two peptides is a candidate for heme-binding in cytochrome P-450c. Nonpolar and hydrophobic regions around Cys-180 (cytochrome P-450b), Cys-180 (cytochrome $P\text{-}450_{LM2}$), and corresponding cysteine residues in cytochromes $P\text{-}450_{C21}$ and P-450c are proposed to be adjacent to the heme-binding regions.

soluble tryptic core. In all cases, however, the location of the cysteine residue(s) in each peptide was determined.

The peptide with the highest extent of homology to other cytochromes P-450 is Tsa-56 (Figure 5). This peptide has 9/13 identical residues (Leu/Ile substitution is considered neutral) to cytochromes P-450b and P-450e and 6/13 to cytochrome $P\text{-}450_{cam}$. In spite of this high degree of homology, it is unlikely that this domain is involved in either substrate or heme binding, since X-ray crystallographic data (Poulos et al., 1983) and S-alkylation under mild conditions (Haniu et al., 1982c) place this domain on the exposed surface of cytochrome $P\text{-}450_{cam}$. The function of this domain is unknown but may be common to all cytochromes P-450 (based on homology) and may include a protein–protein contact with NADPH–cytochrome P-450 reductase in the case of the microsomal enzymes.

A second cysteine peptide (Tsa-47) with more limited homology is shown in Figure 6. This peptide contains a preponderance of hydrophobic amino acids including up to two aromatic residues and the homologous location of one cysteine residue.

P-450c (T-46)	SFSIASDPT	LA	SSCYLEEHVSK	
P-450d (T-30)	SFSIASDPT	SV	SSCYLEEHVSK	

FIGURE 7: Sequence homology between cytochromes P-450c and P-450d. The percentage of homology is approximately 90%. The substitutes Leu–Ser and Ala–Val are possible single base codon changes.

The average hydrophobicity of this region for each cytochrome P-450 roughly corresponds to the change in hydrophobicity for their respective substrates, but since the microsomal enzymes are able to metabolize a wide range of substrates, this correlation should not be over emphasized. The comparison of this domain to residues 152–167 for cytochrome P-450_{cam} is also speculative since it possesses the least degree of homology with the other enzymes. It would be premature on the basis of these data alone to suggest a role for this domain in substrate binding, but it is interesting to observe that this domain is adjacent to the proposed heme binding sulfhydryl group for several of the isozymes (Figure 6).

The heme-binding region of cytochrome P-450_{cam} (Cys-134, Figure 6) from X-ray crystallographic data (Poulos et al., 1983) has a high degree of sequence homology to the major phenobarbital-induced liver isozymes (rat cytochromes P-450b and P-450e and rabbit cytochrome P-450_{LM2}). Black et al. (1982) proposed from sequence homology that Cys-152 in cytochrome P-450b, the corresponding Cys in cytochrome P-450_{LM2} and Cys-134 in cytochrome P-450_{cam} might serve an essential function such as providing the axial ligand to the heme iron atom. However, this proposal was based primarily on the suggestion that Cys-355 and Cys-436 in cytochromes P-450_{cam} and P-450b (Figure 5), respectively, were replaced with a Glu in cytochrome P-450_{LM2}, and thus, this Cys was not highly conserved. More recently, Tarr et al. (1983) showed that P-450_{LM2} does indeed have a Cys at position 436 (Figure 5). Despite the high degree of sequence homology in this region, Tarr et al. (1983) favor Cys-152 as the source of the axial ligand to the heme iron from their predicted membrane model of the protein. Less sequence information is available for the adrenocortical microsomal steroid 21-hydroxylase and rat hepatic cytochrome P-450c, but on the basis of their limited homology only one peptide from cytochrome P-450 C21 and either of two candidates (Tsa-54 or T-46) from cytochrome P-450c can be placed in the heme-binding domain (Figure 6). The major assumption in this model is that all of the enzymes have a somewhat conserved sequence surrounding the heme-binding regions and that cytochrome P-450c and the steroid 21-hydroxylase heme binding domains are adjacent to the hydrophobic domain (Figure 6).

Indirect evidence that the cysteine residue in peptide Tsa-54 maybe involved in heme binding is provided by its low reactivity with S-alkylating agents. We previously found that little or no iodo[^{14}C]acetic acid was incorporated into Cys-152 of cytochromes P-450b and P-450e (unpublished results), and a similar result was obtained for the Cys in Tsa-54 from cytochrome P-450c. This residue is eventually S-alkylated, but only after the cold chase with excess iodoacetic acid. If this deduction is correct, then the proposal by Fujii-Kuriyama et al. (1982) and Heinemann & Ozols (1982, 1983) that the highly conserved domain shown in Figure 5 is the heme-binding domain is incorrect.

The immunological relatedness of cytochromes P-450c and P-450d (Ryan et al., 1980; Reik et al., 1982) suggests the possibility of more extensive sequence homology between these two isozymes compared to several other isozymes. In spite of this likelihood, the NH_2- and COOH-terminal sequences are different (Botelho et al., 1979, 1982), and their respective tryptic maps show no evidence of sequence homology (Figure 1). In order to further investigate their structural relatedness, we performed sequence analyses on one of the cysteine-containing peptides from cytochrome P-450d. The peptide T-30 from the map shown in Figure 2B had 20/22 identical residues compared to peptide T-46 from cytochrome P-450c (Figure 7). This result is especially noteworthy since neither peptide has sequence homology with cysteine-containing peptides from the other isozymes studied to date. Although further studies are required, it is reasonable to predict that cytochromes P-450c and P-450d will contain sufficient sequence homology to explain their immunological relatedness. These particular peptides (T-46 from P-450c and T-30 from P-450d) may contain at least one candidate for a common antigenic determinant. It is also noteworthy that if peptide T-46 from cytochrome P-450c contains the heme-binding cysteine residue, then the corresponding peptide in cytochrome P-450d is highly conserved between these two isozymes but is less conserved compared to the other isozymes sequenced so far.[2]

Acknowledgments

We gratefully acknowledge the technical assistance of Dave Hawke, Chad Miller, Kassu Legesse, and Bill Hempel.

Supplementary Material Available

Amino acid compositions (Table I), microsequence analysis (Table II), and COOH-terminal sequence analysis (Table III) of cysteine-containing peptides (5 pages). Ordering information is given on any current masthead page.

Registry No. Cytochrome P-450, 9035-51-2; cysteine, 52-90-4; heme, 14875-96-8.

References

Black, S. D., Tarr, G. E., & Coon, M. J. (1982) *J. Biol. Chem. 257*, 14616–14619.

Botelho, L. H., Ryan, D. E., & Levin, W. (1979) *J. Biol. Chem. 254*, 5635–5640.

Botelho, L. H., Ryan, D. E., Yuan, P.-M., Kutny, R., Shively, J. E., & Levin, W. (1982) *Biochemistry 21*, 1152–1155.

Conney, A. H. (1967) *Pharmacol. Rev. 19*, 317–366.

Conney, A. H., & Levin, W. (1982) in *Advances in Pharmacology and Therapeutics II* (Yoshida, H., Hagihara, Y., & Ebashi, S., Eds.) Vol. 4, pp 99–118, Pergamon Press, Oxford and New York.

Coon, M. J. (1981) *Drug Metab. Dispos. 9*, 1–4.

Del Valle, U., & Shively, J. E. (1979) *Anal. Biochem. 96*, 77–83.

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

Guengerich, F. P., Dannan, G. A., Wright, S. T., Martin, M. V., & Kaminsky, L. S. (1982) *Biochemistry 21*, 6019–6030.

Haniu, M., Tanaka, M., Yasunobu, K. T., & Gunsalus, I. C. (1982a) *J. Biol. Chem. 257*, 12657–12663.

Haniu, M., Armes, L. G., Yasunobu, K. T., Shastry, B. S., & Gunsalus, I. C. (1982b) *J. Biol. Chem. 257*, 12664–12671.

Haniu, M., Yasunobu, K. T., & Gunsalus, I. C. (1982c) *Biochem. Biophys. Res. Commun. 107*, 1075–1081.

[2] We have subsequently determined the sequences of six cysteine-containing peptides from cytochrome P-450d. Peptide T-30 has the highest percent homology with these peptides to any cysteine-containing peptides in cytochrome P-450c.

Hawke, D., Yuan, P.-M., & Shively, J. E. (1982) *Anal. Biochem. 120*, 302–311.
Heinemann, F. S., & Ozols, J. (1982) *J. Biol. Chem. 257*, 14988–14999.
Heinemann, F. S., & Ozols, J. (1983) *J. Biol. Chem. 257*, 4195–4201.
Jones, B. N., Paabo, S., & Stein, S. (1981) *J. Liq. Chromatogr. 4*, 565–586.
Levin, W., Wood, A., Chang, R., Ryan, D., Thomas, P., Yagi, H., Thakker, D., Vyas, K., Boyd, C., Chu, S.-Y., Conney, A. H., & Jerina, D. M. (1982) *Drug Metab. Rev. 13*, 555–580.
Poulos, T. L., Alden, R. A., Wagner, G. C., & Gunsalus, I. C. (1983) *J. Mol. Biol.* (in press).
Reik, L. M., Levin, W., Ryan, D. E., & Thomas, P. E. (1982) *J. Biol. Chem. 257*, 3950–3957.
Ryan, D. E., Thomas, P. E., & Levin, W. (1980) *J. Biol. Chem. 255*, 7941–7955.
Ryan, D. E., Thomas, P. E., Reik, L. M., & Levin, W. (1982a) *Xenobiotica 11*, 727–744.
Ryan, D. E., Thomas, P. E., & Levin, W. (1982b) *Arch. Biochem. Biophys. 216*, 272–288.
Shively, J. E. (1981) *Methods Enzymol. 79*, 31–48.
Tarr, G. E., Black, S. D., Fujita, V. S., & Coon, M. J. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 6552–6556.
Thakker, D. R., Levin, W., Yagi, H., Conney, A. H., & Jerina, D. M. (1982) in *Biological Reactive Intermediates II* (Snyder, R., Parke, D., Kocsis, J. J., Jollow, D. J., Gibson, G., & Witmer, C., Eds.) pp 501–523, Plenum Press, New York.
Thomas, P. E., Korzeniowski, D., Ryan, D. E., & Levin, W. (1979) *Arch. Biochem. Biophys. 192*, 524–532.
Thomas, P. E., Reik, L. M., Ryan, D. E., & Levin, W. (1983) *J. Biol. Chem. 258*, 4590–4598.
White, R. E., & Coon, M. J. (1980) *Annu. Rev. Biochem. 49*, 315–356.
Yuan, P.-M., Nakajin, S., Haniu, M., Shinoda, M., Hall, P. F., & Shively, J. E. (1983a) *Biochemistry 22*, 143–149.
Yuan, P.-M., Ryan, D. E., Levin, W., & Shively, J. E. (1983b) *Proc. Natl. Acad. Sci. U.S.A. 80*, 1169–1173.

Amino Acid Sequence of Human D of the Alternative Complement Pathway†

Marilyn A. Niemann,* Ajit S. Bhown, J. Claude Bennett, and John E. Volanakis

ABSTRACT: The primary structure of human D, the serine protease activating the C3 convertase of the alternative complement pathway, has been deduced by sequencing peptides derived from various chemical (CNBr and *o*-iodosobenzoic acid) and enzymatic (trypsin, lysine protease, *Staphylococcus aureus* V8 protease, and chymotrypsin) cleavages. Carboxypeptidase A was also used to confirm the COOH-terminal sequence. The peptides were purified by high-pressure liquid chromatography. The proposed sequence of human D contains 222 amino acids and has a calculated molecular weight of 23 748. It exhibits a high degree of homology with other serine proteases, especially around the NH_2-terminus as well as the three residues corresponding to the active-site His-57, Asp-102, and Ser-195 (chymotrypsinogen numbering). This sequence homology is highest (40%) with plasmin, intermediate (35%) with pancreatic serine proteases, such as elastase, trypsin, chymotrypsin, and kallikrein, and least (30%) with the serum enzymes thrombin and factor X. D, however, exhibits only minimal amino acid homology with the other sequenced complement serine proteases, C1r (25%) and Bb (20%). The substitution of a basic lysine for a neutral amino acid three residues NH_2-terminal to the active-site serine as well as a small serine residue for a bulky aromatic amino acid at position 215 (chymotrypsinogen numbering) in the binding pocket may be important in determining the exquisite substrate specificity of D. The presence of His-40 which interacts with Asp-194 (chymotrypsinogen numbering) to stabilize other serine protease zymogens [Freer, S. T., Kraut, J., Robertus, J. D., Wright, H. T., & Xuong, N. H. (1970) *Biochemistry 9*, 1997] argues in favor of such a D precursor molecule.

D, the protease responsible for activating the C3 convertase of the alternative pathway, is one of the least abundant of the complement components in human serum. It consists of a single polypeptide chain with a molecular weight of approximately 23 000 (Volanakis et al., 1977) and is irreversibly inhibited by diisopropyl fluorophosphate (Fearon et al., 1974) which led to its characterization as a serine protease. Its proteolytic attack is directed exclusively against a single arginyl–lysyl peptide bond of B (Lesavre et al., 1979) which becomes susceptible to D only when B is complexed with C3b. The resulting C3bBb complex is the C3 convertase of the alternative complement pathway.

The study of D has been hampered by its low serum concentration and by the difficulty in removing contaminants of similar apparent molecular weight and behavior on ion-exchange chromatography (Johnson et al., 1980). Recently, however, several reports have appeared in the literature detailing the isolation of purified D in sufficient quantities to carry out partial amino acid sequence analysis (Davis et al., 1979; Volanakis et al., 1980; Davis, 1980; Johnson et al., 1980; Reid et al., 1981). We now report the determination of the essentially complete amino acid sequence of this key complement component.

† From the Department of Medicine, Division of Clinical Immunology and Rheumatology, and the Department of Pathology, University of Alabama in Birmingham Medical Center, Birmingham, Alabama 35294. *Received August 29, 1983.* Supported by U.S. Public Health Service Grants AM 03555 and AI 17825 and the Arthritis Foundation.
* Correspondence should be addressed to this author at the Department of Medicine, Division of Clinical Immunology and Rheumatology.

0006-2960/84/0423-2482$01.50/0