Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 313–317, January 1986
Biochemistry

Complete nucleotide sequence of cDNA and deduced amino acid sequence of rat liver catalase

(peroxisomes/recombinant DNA/heme site/organelle biogenesis)

SHUICHI FURUTA*, HIROAKI HAYASHI†, MAKOTO HIJIKATA*, SHOKO MIYAZAWA*, TAKASHI OSUMI*, AND TAKASHI HASHIMOTO*

*Department of Biochemistry, Shinshu University School of Medicine, Matsumoto, Nagano 390, Japan; and †Department of Protein Chemistry, Institute of Endocrinology, Gunma University, Maebashi, Gunma 371, Japan

Communicated by N. E. Tolbert, September 16, 1985

ABSTRACT **We have isolated five cDNA clones for rat liver catalase (hydrogen peroxide:hydrogen peroxide oxidoreductase, EC 1.11.1.6). These clones overlapped with each other and covered the entire length of the mRNA, which had been estimated to be 2.4 kilobases long by blot hybridization analysis of electrophoretically fractionated RNA. Nucleotide sequencing was carried out on these five clones and the composite nucleotide sequence of catalase cDNA was determined. The 5′ noncoding region contained 83 bases and was followed by 1581 bases of an open reading frame that encoded 527 amino acids. The 3′ noncoding region was 831 bases long and contained long repeats of the unit AC. The amino acid sequence deduced from the nucleotide sequence of the cDNAs showed about 90% homology with the reported primary structure of bovine liver catalase. The molecular weight of rat liver catalase was calculated to be 59,758 from the predicted amino acid sequence. The amino acid residues in contact with the heme group are completely identical for bovine liver and rat liver catalases. The amino acid sequence at the COOH terminus was confirmed by the results of carboxypeptidase P treatment of the protein purified from rat liver in the presence of leupeptin. Rat liver catalase has no cleavable signal peptide for translocation of the enzyme into peroxisomes.**

Catalase (hydrogen peroxide:hydrogen peroxide oxidoreductase, EC 1.11.1.6) is a characteristic enzyme of rat liver peroxisomes. The biosynthesis of the following peroxisomal enzymes of animal tissues have been studied *in vitro*: catalase (1, 2), three enzymes of peroxisomal β-oxidation (3–6), uricase (urate:oxygen oxidoreductase, EC 1.7.3.3) (2), and carnitine octanoyltransferase (7). In all cases, no cleavable signal peptides for transport of the enzymes into peroxisomes were found. The amounts of the enzymes for peroxisomal β-oxidation and of their translatable mRNA in rat liver are increased >10-fold by the administration of di(2-ethylhexyl)-phthalate, a peroxisome proliferator (8). Under similar conditions, the level of catalase is increased only slightly. We started cloning of cDNAs for peroxisomal enzymes to predict the amino acid sequence of each protein and to find out the mechanisms that regulate gene expression of peroxisomal enzymes. We have already reported cloning of cDNA for rat liver acyl-CoA oxidase (9), the complete nucleotide sequence encoding the bifunctional peroxisomal enzyme enoyl-CoA hydratase [(*S*)-3-hydroxyacyl-CoA hydro-lyase, EC 4.2.1.17]/3-hydroxyacyl-CoA dehydrogenase [(*S*)-3-hydroxyacyl-CoA: NAD^+ oxidoreductase, EC 1.1.1.35] (10), and a partial nucleotide sequence of cDNA for rat liver catalase (11).

The cloning and nucleotide sequences of cDNAs for catalase were reported for rat liver (11) and human fibroblast (12), but the two cDNAs lack the nucleotide sequences encoding the NH_2-terminal region of the enzyme. The complete primary structure of bovine liver catalase (13) and partial amino acid sequence of human erythrocyte catalase (14) were determined by amino acid sequencing of the enzyme. However, COOH-terminal amino acids predicted from the nucleotide sequence of cDNAs (11, 12) were not found in the purified preparations of bovine liver (20 amino acids) and human erythrocyte (6 amino acids) catalases (13, 14). When rat liver catalase was purified without any protection from possible proteolytic modification during the procedure, the preparation had somewhat lower molecular weight than that of the *in vitro* translation product (1).

In this paper, we report the entire nucleotide sequence of cDNA for rat liver catalase and its amino acid sequence and provide evidence that proteolytic modification occurs at the COOH terminus of rat liver catalase during purification.

MATERIALS AND METHODS

Materials. Restriction endonucleases and other nucleic acid-modifying enzymes were from Bethesda Research Laboratories and Takara Shuzo (Kyoto, Japan); radioactive nucleotides and nick-translation kit, from Amersham; M13 sequencing kits, from Amersham and Takara Shuzo; carboxypeptidase P and leupeptin, from Peptide Institute (Osaka, Japan).

Construction of Catalase cDNA Clones. cDNA clones of rat liver catalase were constructed and screened as described (11). The cDNA clones that correspond to the 5′-terminal region of catalase mRNA were constructed by the method of Land *et al.* (15).

Purification of Catalase. Catalase was purified from rat liver according to the method of Price *et al.* (16). This preparation was designated as modified catalase, because in the absence of a protease inhibitor during the purification, catalase was subjected to cleavage by proteases (1, 17). In order to protect against proteolytic modification of catalase, purification was carried out in the presence of leupeptin throughout the procedure. Livers of rats were perfused and homogenized with five volumes of a solution containing 0.25 M sucrose, 0.1% ethanol, 20 mM glycylglycine (pH 7.5), and 0.04 mM leupeptin. The homogenate was centrifuged at 700 × *g* for 5 min, and the supernatant was centrifuged at 22,000 × *g* for 15 min. The precipitate was suspended in 50 mM potassium phosphate, pH 7.5/0.4 mM leupeptin. Catalase was purified from this fraction according to the method described (17), except that all buffers for dissolving the precipitates contained 0.4 mM leupeptin. This preparation was designated as unmodified catalase.

Digestion of Catalases by Carboxypeptidase P. Catalase (1.2 mg) was dialyzed against water, lyophilized, and dissolved in

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

400 μl of 0.1 M acetic acid. The reaction was started by addition of 40 μg of carboxypeptidase P in 40 μl of 0.1 M acetic acid. The reaction mixture was incubated at 37°C. Samples were removed at appropriate intervals and boiled for 5 min to stop the digestion. The boiled samples were lyophilized, dissolved in 0.02 M HCl, and subjected to amino acid analysis. As controls, catalase and carboxypeptidase P were incubated separately under the same conditions and their amino acid compositions were analyzed.

Other Methods. DNA sequencing was carried out by the procedure of Maxam and Gilbert (18) or by the dideoxy chain-termination method of Sanger *et al.* (19), after subcloning of restriction fragments in pUC9 vector instead of M13 phage vector (20). $NaDodSO_4$/PAGE (21) and polyacrylamide gel isoelectric focusing (22) were carried out as described. Amino acid composition of catalase was determined after hydrolysis at 110°C for 24 hr in 6 M HCl.

RESULTS AND DISCUSSION

Isolation of cDNA Clones for Rat Liver Catalase. Besides the partial cDNA clones for rat liver catalase, pMJ501 and pMJ504 (11), new clones with inserts containing the 5′-terminal sequence of the catalase mRNA were selected from our original cDNA library (9). A clone, pMJ512, was selected (Fig. 1). This clone extended about 100 nucleotides further in the 5′ direction than pMJ504 and contained the entire 3′ noncoding sequence, but it did not cover the complete 5′-terminal region. We attempted to synthesize new cDNAs using a *Pst* I–*Pst* I restriction fragment (nucleotides 1116–1281) of pMJ512 as a primer. This fragment was hybridized with rat liver mRNA according to the method of Fujii-Kuriyama *et al.* (23). cDNA was synthesized according to the method described by Land *et al.* (15) and inserted into the *Pst* I site of pBR322 by the dG·dC-tailing method. *Escherichia coli* strain χ1776 was transformed with the recombinant plasmids and the transformants were screened with a ^{32}P-labeled probe derived from a *Rsa* I–*Pst* I fragment (nucleotides 692–1115) of pMJ512. We screened about 2400 colonies and selected two clones, pMJ1005 and pMJ1010 (Fig. 1), which had the longest inserts among the hybridization-positive colonies. Fig. 1 shows the restriction maps of the cDNAs and the strategy used to determine their nucleotide sequences.

Nucleotide Sequences of Rat Catalase cDNAs. Fig. 2 shows the composite nucleotide sequence of catalase cDNA. The 5′ noncoding region consisted of 83 nucleotides. From the first ATG codon to the TAA termination codon (nucleotides 1582–1584), 1581 base pairs of an open reading frame coding for 527 amino acids were determined. The 3′ noncoding region contained 831 base pairs. A consensus polyadenylylation signal, AATAAA (nucleotides 2394–2399), is located 14 residues upstream of the poly(A) tract. Long repeats of the AC unit are found between nucleotides 1974 and 2162. Korneluk *et al.* (12) determined the nucleotide sequence of a cDNA clone of human fibroblast catalase. This clone, pCAT 1, encoded 451 amino acids and lacked the sequence corresponding to the NH_2-terminal 75 amino acids when compared with the primary structure of human erythrocyte catalase reported by Schroeder *et al.* (14). The 1353 base pairs of coding sequence of pCAT 1 show 85.1% homology to the corresponding parts of the nucleotide sequence of cDNA for rat liver catalase. No deletion or insertion was observed between the two cDNAs.

Deduced Amino Acid Sequence. Based on the nucleotide sequence of the cDNA, 527 residues of the amino acid sequence of rat liver catalase were deduced (Fig. 2). The molecular weight of the enzyme was calculated to be 59,758. The homology of amino acid sequences among three catalases was about 90%. Murthy *et al.* (24) have examined the structure of bovine liver catalase by x-ray analysis. The polar amino acid residues that participate in ionic interactions within bovine liver catalase were highly conserved in the rat liver and human enzymes. Only 4 of 49 residues of bovine liver catalase were changed in rat (Glu-437, Asp-488, and Gln-499) and in human (Glx-13). Complete conservation was found for 39 residues that are in contact with heme (Fig. 2). No difference was detected in the heme-environment resi-

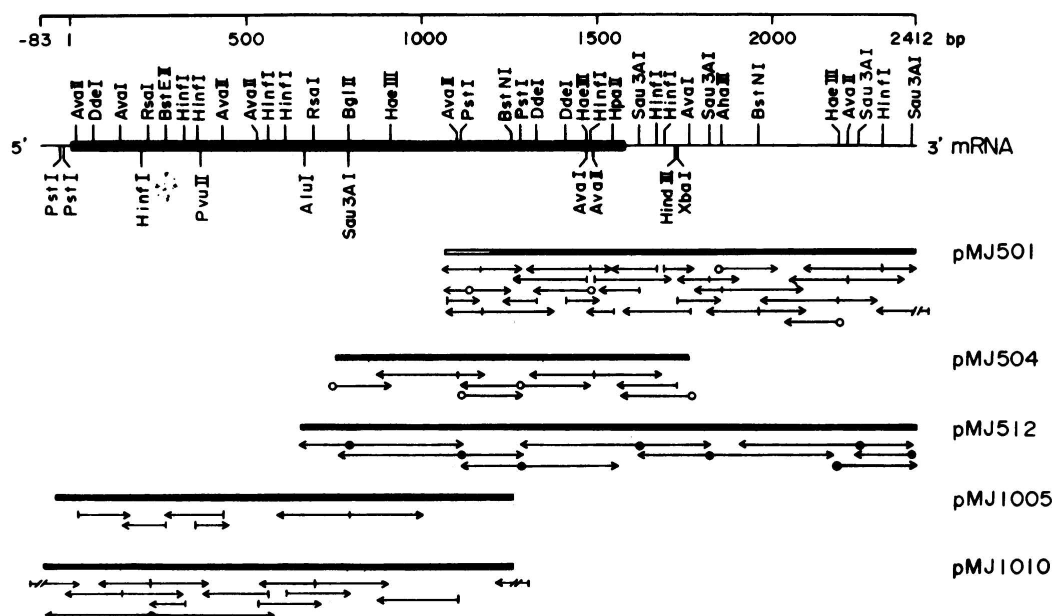

FIG. 1. Restriction map and sequencing strategy of the cDNAs for rat liver catalase. Only the restriction sites used for sequence analysis are shown. Direction and extent of sequence determinations are indicated by arrows; short vertical bars and open circles represent the sites of 5′ and 3′ end labeling for sequencing by the chemical-degradation method (18); closed circles represent the starting points of dideoxy chain-termination sequencing (19); double slash marks mean that the labeled end was located on the vector DNA. The protein-coding region of the mRNA is indicated by the thick line. The open bar in pMJ501 indicates sequence that does not code for any amino acids related to catalase (11). bp, Base pairs.

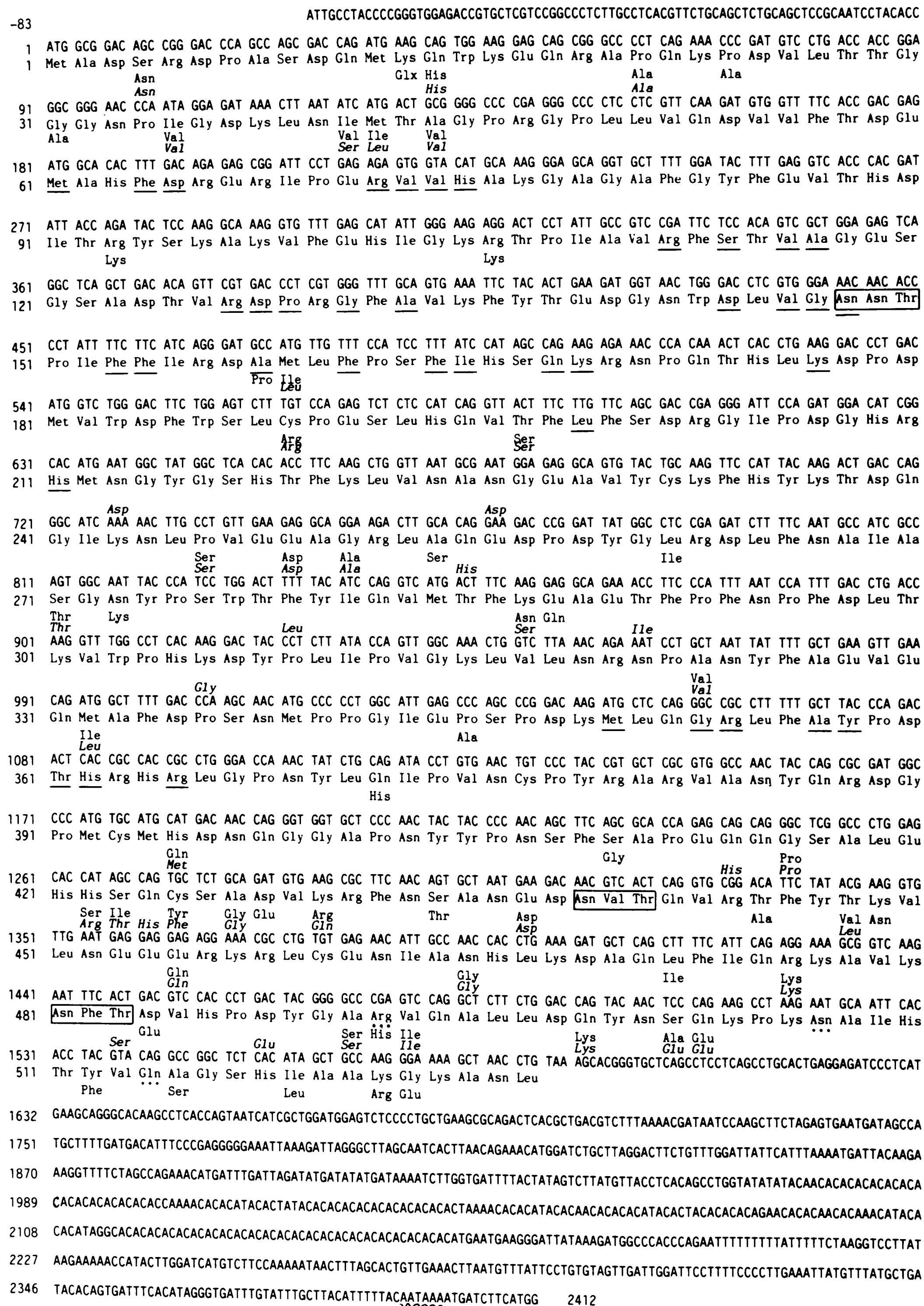

FIG. 2. Composite nucleotide sequence of the cDNA for rat liver catalase and its deduced amino acid sequence. Nucleotides are numbered in the 5′ to 3′ direction, beginning with the first residue of the initiator codon ATG. Nucleotides on the 5′ side of residue 1 are designated by negative numbers. Amino acids are numbered from the initiator methionine. The amino acids of human catalases [residues 2–74 determined by amino acid sequencing of erythrocyte catalase (14), and residues 75–527 predicted from the nucleotide sequence of cDNA for fibroblast catalase (12)] are shown below the rat sequence only where different from the corresponding residue of rat enzyme. Amino acids of bovine catalase (13) that are different from the corresponding rat residues are given in italics below the human sequence. Initiator methionine was not present in bovine and human catalases, where the NH_2-terminal amino acid is a blocked alanine at position 2 (13, 14). Underlined residues are in contact with heme in bovine liver catalase (24). Boxes indicate the sequences Asn-Xaa-Thr (Asn-148, -439, and -481) which are the potential asparagine-linked glycosylation sites. Dots (Arg-492, Asn-507, Gln-514) indicate the COOH-terminal residues predicted from the analyses of carboxypeptidase P digestion of modified catalase. Polyadenylylation signal in the 3′ noncoding region is indicated by a wavy line. In pMJ501, one nucleotide was different from the corresponding residue of other clones (at position 1301, A instead of G).

dues between rat and bovine liver catalases. Only one difference was found, at residue 158, which is alanine in rat and bovine catalases and is proline in the human enzyme. Thus, the amino acid sequences essential for structure and

catalytic activity of catalase are well conserved in these three mammalian species.

Catalases from many sources give multiple protein bands following native gel electrophoresis and exhibit different pIs following native isoelectric focusing (25, 26). Both catalases shown in Fig. 3 (subunits of unmodified and modified rat liver catalases) had multiple protein bands when examined by native polyacrylamide gel isoelectric focusing (data not shown). Jones and Masters (27) proposed that catalase includes sialic acid and that the multiple forms of catalase are due to the degree of sialylation. Catalase from *Aspergillus niger* has been reported to be a glycoprotein (28, 29), and Furuta *et al.* (30) reported the presence of a sugar moiety in a purified preparation of mouse liver catalase. In Fig. 2, three potential asparagine-linked glycosylation sites were indicated (asparagines at positions 148, 439, and 481). Phosphorylation or acetylation of the protein might also be considered as the cause for the multiple molecular forms of catalase.

COOH-Terminal Amino Acids of Purified Catalase. Fig. 3 shows the $NaDodSO_4$/polyacrylamide gel electropherogram of unmodified catalase and modified catalase. The unmodified catalase clearly showed lower mobility than that of modified catalase. Robbi and Lazarow (1) reported that the *in vitro* translation product of rat liver catalase was about 4000 daltons larger than the enzyme purified in a conventional manner (16). Crane *et al.* (31) also observed that mouse liver catalase purified in the absence of protease inhibitors was somewhat smaller than that purified in the presence of the inhibitors. Moreover, the reported primary structure of bovine liver catalase (13) lacks 20 amino acids of its COOH terminus that were predicted from the nucleotide sequence of cDNA for rat liver catalase (11) and human fibroblast catalase (12). Purified preparations of bovine and human erythrocyte catalase lack 6 and 9 amino acids of their COOH termini, respectively (13, 14).

In order to identify the amino acid sequence of the COOH termini of the unmodified catalase and modified catalase, both forms of the enzyme from rat liver were digested with carboxypeptidase P. Carboxypeptidase P is an exopeptidase that liberates amino acids sequentially from the COOH end of a protein (32). Fig. 4 shows the time course of release of free amino acids from the two catalases incubated with carboxypeptidase P. The time course for the unmodified catalase (Fig. 4*A*) matched reasonably well the predicted COOH-terminal sequence: -Ala-Ala-Lys-Gly-Lys-Ala-Asn-Leu. On the other hand, the result of carboxypeptidase P digestion of the modified catalase was more complicated (Fig. 4*B*). The main amino acids liberated after 15 min were

a b

FIG. 3. $NaDodSO_4$/10% PAGE of the purified catalases (0.2 μg per lane). Lane a: catalase purified from rat liver in the presence of leupeptin (unmodified catalase). Lane b: catalase purified from rat liver in the absence of leupeptin (modified catalase). Gels were stained with 0.25% Coomassie brilliant blue R.

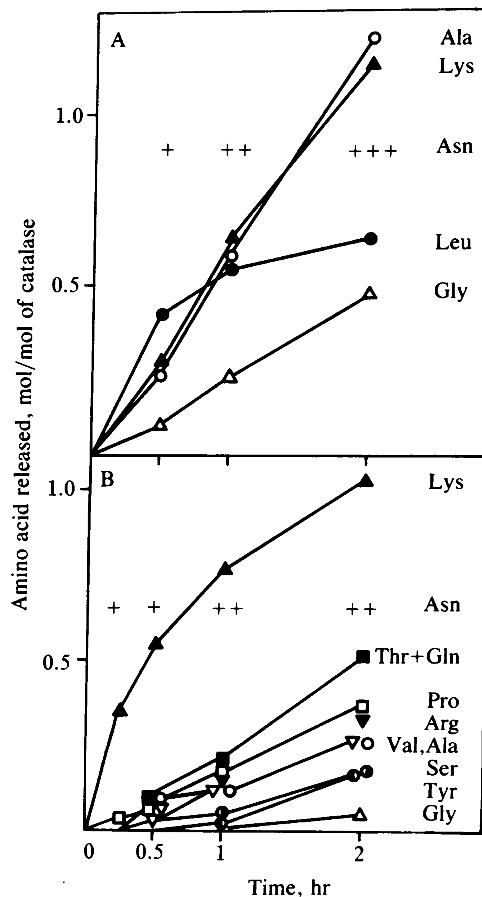

FIG. 4. Rate of release of amino acids from unmodified catalase (*A*) and from modified catalase (*B*) by carboxypeptidase P. +, Asn; ●, Leu; ○, Ala; ▲, Lys; △, Gly; ■, Thr + Gln; □, Pro; ▼, Arg; ▽, Val; ◑, Ser; ◐, Tyr. Asparagine was not determined quantitatively, because of the limitation of our system of amino acid analysis. Therefore, the relative height of peaks are represented as + < ++ < +++.

asparagine, lysine, and proline. Based on this result, it is possible that the COOH-terminal residue of modified catalase is asparagine at position 507 and that the sequence is -Gln-Lys-Pro-Lys-Asn. This asparagine is coincident with the reported COOH-terminal residue, asparagine at position 507, of bovine liver catalase (13). The amino acid-release pattern suggests that this modified enzyme preparation is heterogeneous, containing catalase molecules with different COOH ends. These COOH termini could be attributed to Arg-492 and Gln-514. In Fig. 3, a faint protein band is visible below the main band in the electropherogram of the modified catalase (lane b). This is probably due to a minor component of catalase which has a different COOH terminus.

We also tried to identify the NH_2-terminal amino acid residue by the Edman degradation method for both the unmodified and modified enzymes. However, we could not detect a significant amount of amino acid after several cycles of analysis for either form of the enzyme, suggesting that their NH_2 termini are masked. The NH_2-terminal amino acid of bovine liver and erythrocyte (13) and human erythrocyte (14) catalase was determined to be a blocked alanine, corresponding to Ala-2 of the deduced amino acid sequence of rat liver catalase (Fig. 2). From these results, we propose that the NH_2-terminal residue of rat liver catalase is a blocked alanine at position 2 (Fig. 2), for both the unmodified and modified forms.

The modified rat liver catalase described in this paper had no difference in its catalytic properties and absorption spec-

trum compared with those of the unmodified catalase. According to the three-dimensional structure of bovine liver catalase determined by Murthy *et al.* (24), the COOH-terminal amino acids are not buried in the protein molecule and thus do not participate in subunit association or interaction with heme. Therefore, it is probable that the COOH-terminal amino acids of catalase are attacked easily and removed from the enzyme by proteases during its purification but that this change has no significant effect on the structure and function of the enzyme.

From these results, we conclude that after its synthesis, catalase is transported into and accumulated within peroxisomes without any proteolytic removal of its NH_2- and COOH-terminal amino acids.

We thank Dr. S. Hattori for the protocol of nucleotide sequencing by the chain-termination method using pUC plasmid and Dr. Y. Yabusaki for advice in NH_2-terminal amino acid sequencing. This work was supported in part by research grants from the Ministry of Education, Science and Culture of Japan.

1. Robbi, M. & Lazarow, P. B. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4344–4348.
2. Goldman, B. M. & Blobel, G. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 5066–5070.
3. Furuta, S., Hashimoto, T., Miura, S., Mori, M. & Tatibana, M. (1982) *Biochem. Biophys. Res. Commun.* **105**, 639–646.
4. Miura, S., Mori, M., Takiguchi, M., Tatibana, M., Furuta, S., Miyazawa, S. & Hashimoto, T. (1984) *J. Biol. Chem.* **259**, 6397–6402.
5. Rachubinski, R. A., Fujiki, Y., Mortensen, R. M. & Lazarow, P. B. (1984) *J. Cell Biol.* **99**, 2241–2246.
6. Fujiki, Y., Rachubinski, R. A., Mortensen, R. M. & Lazarow, P. B. (1985) *Biochem. J.* **226**, 697–704.
7. Ozasa, H., Miyazawa, S. & Osumi, T. (1983) *J. Biochem. (Tokyo)* **94**, 543–549.
8. Furuta, S., Miyazawa, S. & Hashimoto, T. (1982) *J. Biochem. (Tokyo)* **92**, 319–326.
9. Osumi, T., Ozasa, H. & Hashimoto, T. (1984) *J. Biol. Chem.* **259**, 2031–2034.
10. Osumi, T., Ishii, N., Hijikata, M., Kamijo, K., Ozasa, H., Furuta, S., Miyazawa, S., Kondo, K., Inoue, K., Kagamiyama, H. & Hashimoto, T. (1985) *J. Biol. Chem.* **260**, 8905–8910.
11. Osumi, T., Ozasa, H., Miyazawa, S. & Hashimoto, T. (1984) *Biochem. Biophys. Res. Commun.* **122**, 831–837.
12. Korneluk, R. G., Quan, F., Lewis, W. H., Guise, K. S., Willard, H. F., Holmes, M. T. & Gravel, R. A. (1984) *J. Biol. Chem.* **259**, 13819–13823.
13. Schroeder, W. A., Shelton, J. R., Shelton, J. B., Robberson, B., Apell, G., Fang, R. S. & Bonaventura, J. (1982) *Arch. Biochem. Biophys.* **214**, 397–421.
14. Schroeder, W. A., Shelton, J. R., Shelton, J. B., Apell, G., Evans, L., Bonaventura, J. & Fang, R. S. (1982) *Arch. Biochem. Biophys.* **214**, 422–424.
15. Land, H., Grez, M., Hauser, H., Lindenmaier, W. & Schutz, G. (1981) *Nucleic Acids Res.* **9**, 2251–2266.
16. Price, V. E., Sterling, W. R., Tarantola, V. A., Hartley, R. W., Jr., & Rechcigl, M., Jr. (1962) *J. Biol. Chem.* **237**, 3468–3475.
17. Mainferme, F. & Wattiaux, R. (1982) *Eur. J. Biochem.* **127**, 343–346.
18. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
19. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
20. Messing, J., Crea, R. & Seeberg, P. H. (1981) *Nucleic Acids Res.* **9**, 309–321.
21. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
22. Doi, Y. & Nishida, T. (1983) *J. Biol. Chem.* **258**, 5840–5846.
23. Fujii-Kuriyama, Y., Taniguchi, T., Mizukami, Y., Sakai, M., Tashiro, Y. & Muramatsu, M. (1981) *J. Biochem. (Tokyo)* **89**, 1869–1879.
24. Murthy, M. R. N., Reid, T. J., III, Sicignano, A., Tanaka, N. & Rossman, M. G. (1981) *J. Mol. Biol.* **152**, 465–499.
25. Holmes, R. S. & Masters, C. J. (1972) *Arch. Biochem. Biophys.* **148**, 217–223.
26. Jones, G. L. & Masters, C. J. (1974) *Arch. Biochem. Biophys.* **161**, 601–609.
27. Jones, G. L. & Masters, C. J. (1975) *Arch. Biochem. Biophys.* **169**, 7–21.
28. Wasserman, B. P. & Hultin, H. O. (1981) *Arch. Biochem. Biophys.* **212**, 385–392.
29. Kikuchi-Torii, K., Hayashi, S., Nakamoto, H. & Nakamura, S. (1982) *J. Biochem. (Tokyo)* **92**, 1449–1456.
30. Furuta, S., LeDonne, N. & Tolbert, N. E. (1984) in *International Cell Biology*, eds. Seno, S. & Okada, Y. (Jpn. Soc. Cell Biol., Tokyo, Japan), p. 284.
31. Crane, D., Holmes, R. & Masters, C. (1982) *Biochem. Biophys. Res. Commun.* **104**, 1567–1572.
32. Yokoyama, S., Oobayashi, A., Tanabe, O., Sugawara, S., Arai, E. & Ichishima, E. (1974) *Appl. Microbiol.* **27**, 953–960.