# **Cloning, Sequence Analysis, and Expression in** *Escherichia coli* **of the Gene Encoding the** *Candida utilis* **Urate Oxidase (Uricase)**

# **Yasuji Koyama,' Toshio Ichikawa, and Eiichi Nakano**

*Research and Development Division, Kikkoman Corporation, Noda 399, Noda, Chiba 278*

Received for publication, June 17, 1996

**A urate oxidase (uricase) gene was cloned from** *Candida utilis* **with an oligonucleotide probe based on the amino acid sequence of cyanogen bromide-cleaved uricase. The uricase gene contains 909 base pairs and encodes a protein with a predicted mass of 34,193 Da.** *Candida* **uricase was similar (49% match in amino acid sequence) to the uricase from** *A8pergillus flavus.* **The uricase from** *Candida utilis* **has four cysteines and one of them, Cysl68, participates in the enzyme activity. This enzyme was expressed to a level of about 20% of total cellular protein in an** *Escherichia coli* **cell as a soluble and functional form.**

**Key words:** *Candida utilis,* **nucleotide sequence analysis, overexpression in** *E. coli,* **oxidation of cysteine residue, uricase.**

Urate oxidase [EC 1.7.3.3] (uricase), is an enzyme in the purine degradation pathway and catalyzes the oxidation of uric acid to allantoin. Humans lack uricase and the accumulation of uric acid in blood causes gout symptoms *(1, 2).* Uricase, therefore, is used for the enzymatic determination of uric acid to diagnose gout. Uricase has been found in various sources and the corresponding genes cloned from some mammals (3-5), *Drosophila (6),* soybean (7), *Aspergillus (8, 9),* and *Bacillus (10),* but the gene encoding uricase of *Candida utilis,* which is widely available commercially, has not been cloned. The *Candida* uricase does not require any cofactor for the enzymatic oxidation. So it is important to improve the producibility and to investigate the structure-function relationship of this enzyme. In this report we describe the cloning of the uricase gene from C. *utilis* and a role of cysteine residues.

# MATERIALS AND METHODS

*Strains—C. utilis* IFO0988 *(11)* was used as enzyme source and DNA donor. *Escherichia coli* LE392 and P2392 *(12),* obtained from Funakosi (Tokyo), were used as host strains for a genetic library constructed in the EMBL4 bacteriophage vector *(13). E. coli* JM109 *(14)* obtained from Takara Shuzo (Kyoto) was used as a host for DNA sequencing and gene expression.

*Materials*—Restriction enzymes and DNA-modifying enzymes were obtained from Takara Shuzo and Boehringer Mannheim (Tokyo). Gigapack II Gold for *in vitro* packaging was purchased from Funakosi.  $[\gamma^{32}P]ATP$  was from Amersham (Tokyo). All other reagents were of analytical grade.

*Purification and Amino Acid Sequencing of Uricase from C. utilis IFO0988-C. utilis* was grown at 30°C for 48 h in a medium containing 5% glucose, 1% polypeptone, 0.5% yeast extract,  $0.85\%$  urea,  $0.3\%$  (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>,  $0.2\%$  K<sub>2</sub>HPO<sub>4</sub>, and  $0.05\%$  MgSO<sub>4</sub> and was harvested by centrifugation at  $5,000 \times g$  for 10 min. For induction of uricase, the cells were suspended in a medium containing 5% glucose, 0.06% uric acid,  $0.2\%$  K<sub>2</sub>HPO<sub>4</sub>, and  $0.05\%$  MgSO<sub>4</sub> and incubated at 30\*C for 4 h. After induction, the cells were harvested by centrifugation, resuspended in 0.2 M sodium carbonate buffer (pH 10) and extracted with chloroform at 30°C for 20 h. The enzyme was absorbed to DEAE-cellulose in a batch system and eluted with 0.15 M NaCl. The eluate was purified to homogeneity by HPLC on a hydroxyapatite column and a reverse-phase column CAPCELL PAK C18 (Shiseido, Tokyo). The enzyme was digested with cyanogen bromide *(15)* and the resulting peptides were separated by reverse-phase HPLC and sequenced with a model 470A protein sequencer (Perkin Elmer Japan, Urayasu).

*Oligonucleotides*—Oligonucleotides were synthesized on a DNA model 382 synthesizer (Perkin Elmer Japan). Oligonucleotide Ul (CCNCARAAYCCNAARAA, where N is A, C, G, or T, R is A or G, and Y is C or T), which was used to screen the DNA library, was based on the peptide sequence PQNPKK. Oligonucleotides UN (ATGTCAACA-ACGGTCTCATCATCCACCTAC) and UC (TTACAACTT-GGTCTTCTCCTTACGGACAAC), based on the sequences of amino and carboxyl terminals, respectively, were used as primers for amplification of the uricase-coding region by PCR.

*DNA Manipulation—*Isolation of genomic DNA and mRNA from *C. utilis* and hybridization were carried out as previously described *(16).* Genomic DNA was partially digested with Sau3AI. Fragments of 9 to 20 kb were purified by agarose gel electrophoresis and ligated to BamHI-digested EMBL4 vector. About 20,000 plaques were screened by plaque hybridization using the U1 probe labeled with  $\lceil \gamma \cdot {}^{32}P \rceil$ ATP.

A  $4.5$ -kb  $EcoRI$  fragment of a positive plaque was subcloned into pUC118 *(17)* to construct pUO501 and sequenced by the dideoxy chain termination method using Kilo-Sequence Deletion Kit (Takara Shuzo) and Dye Primer Cycle Sequencing Kit (Perkin Elmer Japan).

<sup>1</sup> To whom correspondence should be addressed. Phone: + 81-471-23- 5571, Fax: +81-471-23-5959

Computer analysis was done using DNASIS-Mac (Hitachi Software Engineering, Yokohama).

*Expression of the Uricase Gene in E. coli—*For expression of PCR-amplified fragments in *E. coli,* the expression vector pUTE200 was used (*18).* The 912 - bp DNA fragment encoding uricase was amplified by PCR using pUO501 as a template. This fragment was inserted into the *Hpal* site of pUTE200 to obtain pUOX201.

*Site-Directed Mutagenesis*—Substitution of Cys to Ser was carried out by the method of Kunkel *et al. (19)* with a Muta-Gene™ *in vitro* mutagenesis kit (Bio-Rad Japan, Tokyo) using a single-strand DNA prepared from pUO501 harboring JM109. The synthetic oligonucleotides used are listed in Table I. After the mutant structures had been confirmed by DNA sequence determination, DNA fragments containing mutated sites were cut off with suitable restriction enzymes and replaced into pU0X201.

*Uricase Assay and Protein Analysis*—*E. coli* carrying pUOX201 was cultured in L broth containing 1 mM EPTG

TABLE I. **Oligonucleotides for site-directed mutagenesis.** Asterisks indicate Ser anticodon. Each primer was designed to generate a new restriction site to facilitate screening of mutant.

Position	Mutagenic oligonucleotide		
39	5 '-AACCCACCTTCAAGCAGACTAGTGACGGTGGCCTCCATAA *** Spel		
168	5'-GTGAAGTCACTCTTGTTGTAGCCGTAGAACATCGATCCGGTGGACT CIAT $+ + +$		
250	5'-GAAACCGAGTAGACAGAGCTCGCCTTTTCCAAGATCTGAG $SacI$ ***		
293	5 '-TCTCCTTACGGACAACAGTCGACTTGATCAACCCATTTGG $S$ al $I***$		

and 100  $\mu$ g/ml ampicillin at 37°C for 16 h and disrupted by ultrasonication to obtain a crude extract. A reaction mixture containing 2 ml of 0.001% uric acid in 0.1 M borate buffer (pH 8.5), 0.7 ml of water, and 0.1 ml of crude extract was incubated at 25'C for 5 min and the reaction was stopped by addition of 0.2 ml of 20% KOH. The activity of uricase was calculated from the absorbance decrease of uric acid at 293 nm. One unit was defined as the amount of enzyme necessary to transform  $1 \mu$ mol of uric acid into allantoin in 1 min at 25"C and pH 8.5. To examine thermal stability, crude extract was incubated at 60°C for 15 min. After centrifugation, residual activity of supernatant was measured.

*Nucleotide Sequence Accession Number*—The nucleotide sequence data reported in this paper will appear in the GSDB, DDBJ, EMBL, and NCBI nucleotide sequence databases with the accession number D32043.

#### RESULTS

*Cloning of the Uricase Gene*—We concluded that the amino terminus of this protein was blocked since the amino-terminal sequence of the uricase purified from C. *utilis* IFO 0988 could not be determined. After treatment with cyanogen bromide, four peptide fragments could be collected and sequenced. Probe Ul designed from one of the peptide sequences, PQNPKK, was synthesized and used for plaque hybridization studies. About 20,000 plaques were screened with probe Ul and three positive clones were obtained. An *EcoBI* fragment of about 4.5 kb which hybridized to the probe was subcloned into pUC118 and the resultant plasmid was designated pUO501.

*Sequence of the Uricase Gene—We* sequenced 1,641 bp



**region.** The deduced amino acid sequence is shown under the DNA PQNPKK, which is double-underlined. A putative TATA box is sequence. The sequences of four peptides obtained by cyanogen boxed. The putative copper-binding site (H-X-H) is indicated by<br>bromide cleavage and confirmed by peptide sequencing are under- asterisks. bromide cleavage and confirmed by peptide sequencing are under-



Fig. 2. **(A) Structure of the uricase expression plasmid pUOX-201.** A coding region amplified by PCR was inserted into a *Hpal* site of pUTE200 *(18).* **(B) SDS-PAGE analysis of the soluble protein produced by transformants.** Crude extracts were prepared by sonication and centrifugation from the culture of *E. coli* JM109 transformed with pUTE200 (lane 1), pUOX201 (lane 2), pUOX201-

TABLE II. Thermostability of recombinant uricase from wild **type and mutants.** The activities are indicated as U/mg protein of cell extracts. Crude extract was incubated at 60'C for 15 min and residual activity was measured.

Enzyme	No treatment	$60^{\circ}$ C, 15 min	
		$-ME$	$+ME$ <sup>*</sup>
Wild type	4.55	0.66	4.40
C39S	4.17	0.37	3.66
C <sub>168</sub> S	0.09	0.06	0.07
C <sub>25</sub> 0 <sub>S</sub>	5.60	0.49	4.69
C293S	5.40	0.41	5.48

"2-Mercaptoethanol was added to final concentration of 10 mM before heat treatment.

beginning at a *Stul* site and identified an open reading frame of 909 bp (Fig. 1). The amino acid sequence deduced from the nucleotide sequence was consistent with the amino acid sequences of the four cyanogen bromide-cleaved peptides of the uricase purified from *C. utilis* (Fig. 1, underlined sequences). The protein encoded by the uricase gene was composed of 303 amino acid residues and had a molecular weight of 34,146.

*Expression of the Uricase Gene in E. coli*—To express the uricase gene in *E. coli,* we constructed an expression vector, pUTE200. A DNA fragment consisting only of the uricase-coding region was synthesized by PCR using the UN and UC oligonucleotides as primers and inserted into  $HpaI$ -cleaved pUTE200 (Fig. 2A). The resulting plasmid, pUOX201, was used to transform *E. coli* JM109. Uricase activity of a clone harboring pUOX201 was 6.1 U/ml culture in the presence of 1 mM IPTG. Assuming that the specific activity of uricase is 25 U/mg protein (Ref. *20* and our data), the level of uricase expressed in *E. coli* represents about 20% of the soluble protein in the cells (Fig. 2B).

*Thermal Stability of Uricase—*Treatment at 60'C for 15 min caused a decrease of the uricase activity in the crude extract prepared from a suspension in culture medium of cells expressing the wild type uricase, and preincubation with 2-mercaptoethanol blocked the heat inactivation (Table II). Addition of 2-mercaptoethanol after heat treat-

C39S (lane 3), pUOX20lCl68S (lane 4), pUOX201C250S (lane 5), and pUOX20lC293S (lane 6). Lane M is the molecular weight marker (10 kDa Protein Ladder, GIBCO BRL). A  $10-\mu$ l aliquot of each crude extract was applied on 10-20% gradient SDS-PAGE (Daiichi Pure Chemicals, Tokyo) with Laemmli's buffer system *(26).* The protein band indicated by an arrow is the recombinant uricase.

ment restored about 90% of the initial activity (data not shown). The reduction of the enzyme activity was thought to result from the oxidation of cysteine residue(s). To confirm this hypothesis, we constructed mutants in which serine was substituted for each cysteine. None of the substitutions affected the producibility and solubility of uricase in mutants (Fig. 2B), and only the substitution of Ser for Cysl68 resulted in a decrease of the uricase activity (Table II). These results suggest that oxidation of Cysl68 is a cause of the thermal instability of *Candida* uricase.

# DISCUSSION

We have cloned and sequenced the uricase gene from a genomic DNA of *C. utilis,* which encodes a 34-kDa protein with 303 aa residues. A putative TATA box (TATAAAA) was found at position  $-99$  (Fig. 3) but no other consensus regulatory sequences were identified.

The genomic gene amplified by PCR could be expressed in *E. coli* under the control of a *lac* promoter, suggesting that this gene contains no introns. The recombinant uricase accumulated to a level of approximately 20% of soluble intracellular protein and had the same specific activity as uricase purified directly from C. *utilis.*

Compared with uricases from other sources, the amino acid sequence of *C. utilis* uricase possesses 49% identity to fungi, 33% to mammalians, 37% to fly, 33% to soybean, and 21% to *Bacillus.* Bairoch *(21)* and Legoux *et al. (8)* identified consensus patterns for eukaryotic uricase (motif 1 and motif 2 in Fig. 3). The sequences Val-Leu-Lys-Ser-Thr-Gly-Ser (154-160 in Fig. 1) and Ser-Pro-Ser-Val-Gln-Ala-Thr-Met-Phe (231-239 in Fig. 1) are slightly different from motif 1 (Val-Leu-Lys-Thr-Thr-Gln-Ser) and motif 2 (Ser-Pro-Ser-Val-Gln-Lys/His/Asn-Thr-Leu-Tyr), respectively. Recently. Yamamoto *et aL (10)* reported that the corresponding sequence of *Bacillus* sp. TB-90 was not highly homologous and that another highly conserved region was identified (region A in Fig. 3). This sequence, Asn-Ser-X-Val/Ile-Val/Ile-Ala/Pro-Thr-Asp-Ser/Thr-X-Lys-Asn, is found in the sequence of uricase from *C. utilis*

Fig. 3. A multiple alignment of uricases. The uricases from C. utilis, A. flavus (A38097), soybean (A25776), rat (A31774), and Bacillus (D49974) are aligned. The conserved residues among the five proteins are boxed. Motif 1 and motif 2 were proposed as consensus patterns for uricase by Bairoch  $(21)$  and Legoux  $(8)$ . respectively. Regions A and B were newly identified conserved sequences.

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 2, 2012



(54-65 in Fig. 1). In addition, there is another conserved sequence, Tyr/His-Gly-Lys-X-X-Val, in the amino terminal region of all uricases (region B in Fig. 3). A mutant with this region deleted had no uricase activity (data not shown). Thus, this sequence is important for the enzyme activity. Eukaryotic uricases are known to be copper-binding enzymes (22) and the copper-binding site has been proposed  $(2)$ . But uricases of C. *utilis* and A. *flavus* contain no copper  $(20, 23)$ , despite having a putative copper-binding site (His-Asp-His, 119-121 in Fig. 1 and Ref. 8). Yamamoto et al. reported that the copper-binding site was not found in the uricase of Bacillus sp. TB-90, which does not contain copper  $(10)$ . It is interesting to investigate how microbial uricases compensate functionally for the absence of copper in the active center.

Uricase of  $C$ . *utilis* is well characterized  $(20, 24, 25)$ . Nishimura et al. (20) reported that it contains three cysteines, of which one is located near the active site of the enzyme but is not necessary for uricase activity. It is known that subunits of *Candida* uricase can associate to form an enzyme with lower specific activity under aerobic conditions  $(25)$ . Based on our sequence data, the uricase of C. *utilis* IFO0988 contains four cysteines. We showed that the oxidation or substitution of Cys168 resulted in reduction of the uricase activity. One hypothesis is that a chemically reactive sulfhydryl group located on the surface of the molecule is spontaneously oxidized during purification and forms a disulfide bond linking the two molecules.

We thank Ms. K. Saito for technical assistance.

# **REFERENCES**

- 1. Reddy, P.G., Nemali, M.R., Reddy, M.K., Reddy, M.N., Yuan, P.M., Yuan, S., Laffler, T.G., Shiroz, T., Kuramitsu, H.K., Usuda, N., Chisholm, R.L., Rao, M.S., and Reddy, J.K. (1988) Isolation and sequence determination of a cDNA clone for rat peroxisomal urate oxidase: Liver-specific expression in the rat. Proc. Natl. Acad. Sci. USA 85, 9081-9085
- Wu, X., Lee, C.C., Muzny, D.M., and Caskey, C.T. (1989) Urate oxidase: Primary structure and evolutionary implications. Proc. Natl. Acad. Sci. USA 86, 9412-9416
- 3. Ito, M., Suzuki, M., and Takagi, Y. (1988) Nucleotide sequence of cDNA and predicted amino acid sequence of rat liver uricase. Eur. J. Biochem. 173, 459-463
- 4. Motojima, K., Kanaya, S., and Goto, S. (1988) Cloning and sequence analysis of cDNA for rat liver uricase. J. Biol. Chem. 263.16677-16681
- 5. Motojima, K. and Goto, S. (1989) Cloning of rabbit uricase cDNA reveals a conserved carboxy-terminal tripeptide in three species. Biochim. Biophys. Acta 1008, 116-118
- 6. Wallrath, L.L., Burnett, J.B., and Friedman, T.B. (1990) Molecular characterization of the Drosophila melanogaster urate oxidase gene, an ecdysone-repressible gene expressed only in the malpighian tubules. Mol. Cell. Biol. 10, 5114-5127
- 7. Nguyen, T., Zelechowska, M., Foster, V., Bergmann, H., and Verma, D.P.S. (1985) Primary structure of the soybean nodulin-35 gene encoding uricase II localized in the peroxisomes of uninfected cells of nodules. Proc. Natl. Acad. Sci. USA 82, 5040-5044
- 8. Legoux, R., Delpech, B., Dumont, X., Guillemot, J.-C., Ramond,

P., Shire, D., Caput, D., Ferrara, P., and Loison, G. (1992) Cloning and expression in *Escherichia coli* of the gene encoding *Aspergillus flavus* urate oxidase. *J. Biol. Chem.* **267,** 8565-8570

- 9. Oestreicher, N. and Scazzocchio, C. (1993) Sequence, regulation, and mutational analysis of the gene encoding urate oxidase in *Aspergillus nidulans. J. Biol. Chem.* **268,** 23382-23389
- 10. Yamamoto, K., Kojima, Y., Kikuchi, T., Shigyo, T., Sugihara, K., Takasio, M., and Emi, S. (1996) Nucleotide sequence of the uricase gene from *Bacillus* sp. TB-90. *J. Biochem.* **119,** 80-84
- 11. Roush, A.H. and Domnas, A.J. (1956) Induced biosynthesis of uricase in yeast. *Science* **124,** 125-126
- 12. Borck, J., Beggs, J.D., Brammer, W.J., Hopkins, A.S., and Murray, N.E. (1976) The construction *in vitro* of transducing derivatives of phage lambda. *Mol. Gen. Genet.* **146,** 199-207
- 13. Frischauf, A.M., Lehrach, H., Polstka, A., and Murray, N.M. (1983) Lambda replacement vectors carrying polylinker sequence. *J. Mol. Biol.* **170,** 827-842
- 14. Messing, J. (1985) Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mpl8 and pUC19 vectors. *Gene* 33, 103-119
- 15. Gross, E. and Witkop, B. (1962) Nonenzymatic cleavage of peptide bonds: The methionine residues in bovine pancreatic ribonuclease. *J. Biol. Chem.* **237,** 1856-1860
- 16. Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (eds.) (1987) *Current Protocols in Molecular Biology,* John Wiley & Sons, New York
- 17. Vieira, J. and Messing, J. (1987) Production of single-stranded plasmid DNA in *Methods in Emymology* (Wu, R. and Grossman,

L., eds.) Vol. 153, pp. 3-11, Academic Press, San Diego, CA

- 18. Koyama, Y. and Ohmori, H. (1996) Nucleotide sequence of the *Escherichia coli solA* gene encoding a sarcosine oxidase-like protein and characterization of its product. *Gene* (in press)
- 19. Kunkel, T.A., Roberts, J.D., and Zakour, R.A. (1987) Rapid and efficient site-specific mutagenesis without phenotypic selection in *Methods in Enzymology* (Wu, R. and Grossman, L., eds.) Vol. 154, pp. 367-382, Academic Press, San Diego, CA
- 20. Nishimura, H., Yosida, K., Yokota, Y., Matsushima, A., and Inada, Y. (1982) Physicochemical properties and states of sulfhydryl groups of uricase from *Candida utilis. J. Biochem.* **91,** 41-48
- 21. Bairoch, A. (1991) Prosite: A dictionary of sites and patterns in proteins. *Nucleic Acids Res.* **19,** 2241-2246
- 22. Mahler, H.R. (1963) Uricase in *The Enzymes* (Boyer, P.D., Lardy, H., and Myrback, K., eds.) Vol. 8, pp. 285-296, Academic Press, New York
- 23. Conley, T.G. and Priest, D.G. (1980) Thermodynamics and stoicheiometry of the binding of substrate analogues to uricase. *Biochem. J.* **187,** 727-732
- 24. Itaya, K., Yamamoto, T., and Fukumoto, J. (1967) Studies on yeast uricase: Purification and some enzymatic properties of yeast uricase. *Agric. BioL Chem.* 31, 1256-1264
- 25. Itaya, K., Fukumoto, J., and Yamamoto, T. (1971) Studies on urate oxidase of *Candida utilis:* Some physical and chemical properties of the purified enzyme. *Agric. Biol. Chem.* 35, 813- 821
- 26. Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227,** 680-685