The structural gene for NADP⁺-dependent serine dehydrogenase [EC 1.1.1.-] from Agrobacterium tumefaciens ICR 1600 was cloned into Escherichia coli cells and its complete DNA sequence was analyzed. The gene encodes a polypeptide containing 249 amino acid residues. The enzyme had high sequence similarity to short-chain alcohol dehydrogenases from bacteria and unknown proteins of Haemophilus influenzae, Escherichia coli, and Saccharomyces cerevisiae.

Key words: serine dehydrogenase; Agrobacterium tumefaciens; short-chain alcohol dehydrogenase; nucleotide sequence; amino acid sequence

Received November 7, 2001; Accepted December 19, 2001

Note

Cloning and Sequencing of the Serine Dehydrogenase Gene from Agrobacterium tumefaciens

Hisae FUJISAWA, Shinji NAGATA, Emran Kabir CHOWDHURY, Miki MATSUMOTO, and Haruo MISONO*

Department of Bioresources Science, Kochi University, Nankoku, Kochi 783-8502, Japan

We have purified and characterized a novel enzyme, NADP⁺-dependent serine dehydrogenase [EC 1.1.1.-] from Agrobacterium tumefaciens ICR 1600. The enzyme has a molecular mass of about 100 kDa and consists of four identical subunits. The enzyme catalyzes the oxidation of the 3-hydroxyl group of serine to form 3-aminomalonic semialdehyde, which is spontaneously decarboxylated into 2-aminoacetaldelyde. Although the 3-hydroxyisobutyrate dehydrogenase from Pseudomonas putida E23, which functions in L-valine metabolism, catalyzes the oxidation of the hydroxyl group of L-serine, the enzyme shows different properties from those of serine dehydrogenase. We cloned the structural gene for the enzyme from A. tumefaciens ICR 1600 into Escherichia coli JM109 to study its structural relationship with other proteins.

In this paper, we describe the cloning and nucleotide sequence of the serine dehydrogenase gene.

A. tumefaciens ICR 1600, which was used as a source of chromosomal DNA, was grown at 30°C for 20 h in Luria broth (LB)(1% peptone, 0.5% yeast extract, and 0.5% NaCl, pH 7.2) on a reciprocal shaker. Transformants were grown at 37°C for 20 h in 100 ml of LB containing ampicillin (50 μg/ml) and isopropyl-β-D-thiogalactopyranoside (120 μg/ml) on a reciprocal shaker. The enzyme assay was done as described in our previous paper.

Sense (N) and antisense (C) primers were designed from the N-terminal amino acid sequence and the amino acid sequence of the peptide P₃, obtained from the lysyl endopeptidase digest of the enzyme, respectively. Sequences were 5'-GGGAATTCGTAARGGTATHCA-3' (primer N) and 5'-GGGAGTTCGGAATTCATSCC-3' (primer C). PCR was done with Ampli Taq DNA polymerase (Perkin Elmer, U.S.A.). The reaction mixture (50 μl) consisted of 100 pmol of each of the primers, 0.5 μmol of Tris-HCl buffer (pH 8.3), 2.5 μmol of KCl, 0.15 μmol of MgCl₂, 10 nmol of each dNTP, 2.5 units of DNA polymerase (Ampli Taq), and 10 ng of the EcoRI-digested chromosomal DNA as a template. After 7 min of incubation at 94°C, the reaction mixture was heated at 94°C for 1 min, then cooled at 55°C for 1 min, and incubated at 72°C for 2 min. The programmed temperature shift was repeated 30 times. The nucleotide sequence of the amplified DNA fragment (504 bp) was analyzed using an Applied Biosystems 373A DNA sequencer with a DNA sequencing kit (ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit, Perkin Elmer, U.S.A.). From the sequence obtained, primers, N1 (5'-TCAAGGAGGCCTGGAAGG-3'), N2 (5'-TTTCGTCCTGTGCTGCG-3'), D1 (5'-GGTGGGATCGAGTTACC-3'), and D2 (5'-TCGGTGATATCCAACGCGA-3') were prepared. For the sequencing of unknown DNA regions at the 5'- and 3'-end of the structural gene for the enzyme, the cassette-ligation mediated PCR was done using a TaKaRa LA PCR in vitro cloning kit (Takara Shuzo, Kyoto, Japan). The cassette DNA containing a Sat3AI site was ligated to the BamHI-digested chromosomal DNA and the ligated fragments were used as the template DNA for PCR for the unknown 5'-end region. The product of the first PCR with the cassette primer C1 and the primer D1 was used as the template DNA for the second PCR with the cassette primer C2 and the primer D2. The DNA fragment containing the 5'-end region of the gene was amplified and se-
quenced. The DNA fragment containing the 3'-end region of the gene was amplified with the cassette DNA containing an EcoRI site, the cassette primer C1, the primer N1, the cassette primer C2, and the primer N2 and sequenced. Finally the structural gene for the enzyme was amplified by PCR with EX Taq DNA polymerase (Takara Shuzo, Kyoto, Japan) and with the primers EN, which contained a Shine-Dalgarno sequence and a BamHI site, and HC containing a HindIII site. Sequences were 5'-GGGGATCC AGGAAACAGACATGAGCGGAATCTCTTGATACAG-3' (an antisense primer HC) and 5'-GGGAAGCTTCAGCTTTCCCGATAGACCTGGAAA-TCTTGATCACAGTTTCCGGATAGACCTGGAAA-TCTTGATCACAGCTTTCCCGATAGACCTGGAAA-TCTTGATCACA

The DNA fragment was completely digested with both BamHI and HindIII and ligated into the BamHI-HindIII site of plasmid pUC18 (Takara Shuzo, Kyoto, Japan) with T4 DNA ligase. The plasmid was introduced into the competent E. coli JM109 cells. Transformants were selected on LB agar plates containing ampicillin (50 μg/ml), isopropyl-β-D-thiogalactopyranoside (120 μg/ml), and 5-bromo-4-chloro-3-indolyl-β-D-galactoside (100 μg/ml). A positive clone was selected from the transformants by measuring the enzyme activity of the cell extracts. We named the plasmid isolated from the positive clone pUSDH. The sequence of the structural gene for the enzyme in the plasmid was analyzed in both directions with an Applied Biosystems 373A DNA sequencer with a DNA Sequencing kit. The nucleotide sequence data will appear in the DDBJ/EMBL/GenBank nucleotide sequence databases under the accession number AB032322.

The activity of a cell extract of E. coli cells harboring pUSDH was 43.3 munits/mg of protein. Although E. coli JM 109 used as a host cell has the weak activity (1.3 munit/mg of protein), the activity of E. coli JM 109/pUSDH is 34-fold higher than that in the host cell.

The structural gene for the enzyme is not associated with other genes in an operon and encodes a polypeptide consisting of 249 amino acid residues. The first 23 predicted amino acids in the sequence except methionine and the C-terminal two amino acids were identical with N-terminal and C-terminal amino acid sequences, respectively, of the enzyme purified from A. tumefaciens. The amino acid sequences of 7 internal peptides isolated from the lysyl endopeptidase digest of the A. tumefaciens enzyme were also in good agreement with the predicted amino acid sequence (Fig. 1). The calculated molecular mass of this protein was 26.6 kDa, which agreed with the apparent subunit molecular mass of 25 kDa of the enzyme purified from A. tumefaciens. The initiator methionine, which is not observed in the N-terminal sequence of the enzyme from A. tumefaciens, may be removed by a post-translational modification in vivo.

The amino acid sequence of the enzyme was similar to those of short-chain alcohol dehydrogenases (Fig. 2) and its homologous proteins. The percentage of identical amino acids of the enzyme compared with glucose dehydrogenase from Bacillus subtilis,30 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase from E. coli,43–50 acetacetyl-CoA reductase from Alcaligenes eutrophus,51 7a-hydroxysteroid dehydrogenase from Eubacterium sp.,30 3-oxoacyl acyl-carrier protein reductase from E. coli52–58 were estimated to be 34, 29, 26, 26, and 25%, respectively. Those of the enzyme with the short-chain alcohol dehydrogenase homologues from Haemophilus influenzae,30 E. coli (YdfG),53 and Saccharomyces cerevisiae (YMR226c)54 were 51, 50, and 38%, respectively. A common GXXGXXG sequence, which is characteristic for a NAD(P)⁺-binding site55 is seen in the N-terminal region of these enzymes. Tyrosine and lysine residues, which are important residues in catalysis of short-chain alcohol dehydrogenases,55 are also conserved in these enzymes. The amino acid sequence of serine dehydrogenase, however, is not similar to that of 3-hydroxyisobutyrate dehydrogenase from P. aeruginosa, in which the catalytic residues are lysine and asparagine.140 These data

Fig. 1. Nucleotide Sequence of the Serine Dehydrogenase Gene and the Deduced Amino Acid Sequence.

The amino acid sequences determined by Edman degradation of the enzyme from A. tumefaciens and the peptides (P1–P8) isolated from the lysyl endopeptidase digest of the enzyme and the C-terminal two amino acids sequence, which is determined by the carboxypeptidase Y digestion method, are shown by a single underline. The putative –35 and –10 sequences are boxed. The putative Shine-Dalgarno (SD) sequence is indicated with an asterisk.
dependent serine dehydrogenase from A. tumefaciens belongs to the short-chain alcohol dehydrogenases.13) The reaction mechanism of the enzyme1) suggests that NADP+-dependent serine dehydrogenase from A. tumefaciens belongs to the short-chain alcohol dehydrogenases.13)

Alignment of the Amino Acid Sequences of Serine Dehydrogenase and Several Short-chain Alcohol Dehydrogenases.

The amino acid sequence of the enzyme from A. tumefaciens (SDH) was compared with those of glucose dehydrogenase from Bacillus subtilis (GDH),3) 3-oxoacyl acyl-carrier protein reductase from Alcaligenes eutrophus (ACR),6) 7α-hydroxysteroid dehydrogenase from Eubacterium sp. (HDH),3) and 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase from E. coli (BDH).4–5) Common residues in these dehydrogenases are shown by white letters in black boxes. The essential residues in catalysis and glycine residues in these dehydrogenases are shown by white letters in black boxes. The essential residues in catalysis and glycine residues in the NAD(P)⁺-binding site are indicated by asterisks.

and properties of the enzyme13) suggest that NADP⁺-dependent serine dehydrogenase from A. tumefaciens belongs to the short-chain alcohol dehydrogenase family. The reaction mechanism of the serine dehydrogenase is probably similar to that of short-chain alcohol dehydrogenases.13)

References


12) Skelton J., Churcher C. M., Barrell B. G., Rajandream M. A., and Walsh S. V., YMR226C (or YM9959.08C) appear in the EMBL, GenBank, and DDBJ databases under the accession number Z49939 (EMBL) and Q05016 (SWISS).

