Note
Cloning and Sequencing of Trehalase Biosynthesis Genes from *Rhizobium* sp. M-11

Kazuhiko MARUTA, Kazuko HATTORI, Tetsuya NAKADA, Michio KUBOTA, Toshiyuki SUGIMOTO, and Masashi KURIMOTO

Hayashibara Biochemical Laboratories, Inc., 7-7 Amase-minami machi, Okayama 700, Japan
Received October 25, 1995

The two genes encoding maltoligosyl trehalase synthase and maltoligosyl trehalase trehalohydrolase, which are related to biosynthesis of α,α-trehalose, were cloned from *Rhizobium* sp. M-11. Sequence analysis showed that the synthase gene composed of 2316 bp was connected with the hydrolyase gene of 1788 bp by an overlap of one nucleotide. The deduced amino acid sequences of both enzymes have several regions common to amylolytic enzymes belonging to an α-2-amylace family.

Key words: trehalose; *Rhizobium*; gene cloning; nucleotide sequence; operon

Recently we have found several bacteria producing trehalose (α,α-trehalose) via a novel enzymatic system of maltoligosyl trehalase synthase (MTSase) and maltoligosyl trehalase trehalohydrolase (MTHase).1-3 MTSase catalyzes the conversion of maltohexa- to maltoligosyl trehaloses by forming α,α-T-glucosidic linkage by an intramolecular transglycosylation.4 MTHase hydrolyzes the products of the MTSase-catalyzed reaction on the α,α-glucosidic linkage between the maltoligosyl group and trehalose.5 Thus, trehalose is produced from maltodextrins by MTSase and MTHase in a joint operation. In this paper we present the nucleotide sequences of MTSase and MTHase genes from *Rhizobium* sp. M-11 (FERM BP-4130).2,3

Genomic DNA from *Rhizobium* sp. M-11 was partially digested with Sau3AI and was inserted into the BamHI site of the plasmid vector pBluescript II SK+ (Stratagene). *Escherichia coli* XL1-Blue was used as a host strain for constructing the genomic library. To isolate MTSase and MTHase genes from *Rhizobium* sp. M-11, colony hybridization was done by using four synthetic oligonucleotide probes. The MTSase-specific probes were 5'-CCGATTGTGTAAGGGTAAG-3' and 5'-ACGATTGTGTAAGGGTAAG-3' corresponding to the amino acid sequences of lysylendopeptidase-cleaved MTSase fragments, Pro-Glu-Trp-Glu-Lys and Thr-Glu-Pho-Trp-Asp, respectively, which were analyzed with a protein sequencer (model 473A, Applied Biosystems). The MTHase-specific probes were 5'-TGTGGACGATAGCTTGGTGGGAAGTG-3' and 5'-CCAATGGGGAAGTG-3' corresponding to Phe-Asp-Ile-Trp-Ala-Pro and Asp-Trp-Ala-Glu-Ala, respectively of the MTHase fragments. One clone (pBMTU1) positive for all of the MTSase- and MTHase-specific probes was obtained. The recombinant plasmid contained an insert of 8.3 kbp, of which the restriction map is shown in Fig. 1. The nucleotides were sequenced with a DNA sequencer (model 373A, Applied Biosystems) using the Taq DyeDeoxy terminator cycle sequencing kit (Applied Biosystems).

As shown in Fig. 2, sequence analysis found two open reading frames: ORF-1 (2316 bp, position 565 to 2880) and ORF-2 (1788 bp, position 2880 to 4667). The ORF-1 encodes 772 amino acid residues containing the N-terminal sequence of the MTSase protein and of their ten lysylendopeptidase-cleaved fragments as shown in Fig. 2. The deduced amino acid sequence has a calculated molecular mass of 84,881 Da, in agreement with the value of 82,000 Da by SDS-PAGE of MTSase protein.5 The ATG translational start codon of the ORF-1 (designated *treY*) is preceded by a putative ribosome binding sequence, 5'-GGGAGG-3'.

The N-terminal sequence of MTSase protein is identical to the deduced amino acid sequence of ORF-2 (designated *treZ*) from the 8th to the 30th residue. The amino acid sequences from five peptide fragments of MTHase were found in the deduced sequence. The molecular mass of the putative protein was calculated to be 65,388 Da after removal of the N-terminal seven residues. This value is in agreement with the 63,000 Da by SDS-PAGE of MTHase protein.3 The putative ribosome binding sequence, 5'-AGGAGG-3', precedes the ATG translational start codon. Thus, it is evident that ORF-1 (*treY*) and ORF-2 (*treZ*) are structural genes of MTSase and MTHase, respectively, of *Rhizobium* sp. M-11.

A palindromic sequence like a transcriptional termination signal (*ΔG* = -57 kcal/mol) was found 14 bp downstream from transcriptional termination codon of *treZ*. The 3'-end of *treZ* overlaps with the 5'-end of *treY* by one nucleotide. This overlap of the genes suggest that *treYZ* constitutes an operon, although promoters and transcriptional start points have not been identified. A Southern blot of genomic DNA from *Rhizobium* sp. M-11 digested with *HindIII* and probed with a *HindIII* fragment (2454 bp, position 2377 to 4830) of pBMTU1 gave one positive band of 2.5 kbp identical to the probe DNA. It shows that the *treYZ* is present in a single locus on the genomic DNA as a single copy.

When the amino acid sequence of MTSase was compared with that of MTHase, several regions were found to be homologous. Three of the homologous regions (boxed in Fig. 2) are common to α-amylolytic enzymes such as α-amylase, pullulanase.

![Fig. 1. Restriction Map of the Inserted DNA Fragments in pBMTU1.](image)

The thick arrows indicate the *treY* (solid line) and *treZ* (broken line). The thin arrows represent sequenced regions and directions.

**Abbreviations**: MTSase, maltoligosyl trehalose synthase; MTHase, maltoligosyl trehalose trehalohydrolase; *treY*, maltoligosyl trehalose synthase gene; *treZ*, maltoligosyl trehalose trehalohydrolase gene; ORF, open reading frame; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis.
Trehalose Biosynthesis Genes from *Rhizobium* sp. M-11

Fig. 2. Nucleotide Sequence of the *treY* and *treZ*

The deduced amino acid sequences of MTSase (*TreY*) and MTHase (*TreZ*) are shown below the nucleotide sequence. The amino acid sequences of MTSase, MTHase, and their hydrolase-peptidase-decarboxylase fragments are underlined. The homologous regions common to the *smyl* and *smyl* family are shown in boxes. Stop codons are marked by asterisks. The putative ribosome binding sequences are underlined and the transcriptional termination signal is marked by arrows below the nucleotide sequence.
cyclomaltodextrin glucanotransferases, amylase and branching en-
yzymes, which have been classified in an "z-amylase family". From
information on the three-dimensional structure8,9 and
site-directed mutagenesis,10,11 the amino acid residues in these re-
gions of the z-amylolytic enzymes have been proposed to be in-
volved in catalysis and substrate-binding. Therefore, we suggest
that MTSase and MTHase belong to the z-amylase family and
the conserved regions of our enzymes may participate in the en-
zymatic functions.

The DNA sequence presented in this report will appear in the
DDBJ, FMBL, and GenBank data bases with the accession
number D78001.

References
1) K. Maruta, T. Nakada, M. Kubota, H. Chaen, T. Sugimoto, M.
0606753A2 (Jul. 20, 1994).
4) T. Nakada, K. Maruta, K. Tsusaki, M. Kubota, H. Chaen, T.
6) M. H. Jespersen, A. E. MacGregor, R. M. Sierks, and B. Svensson,
7) T. Kuriki, H. Takata, S. Okada, and T. Imanaka, Denpun Kagaku,
8) Y. Matsuura, M. Kusunoki, W. Harada, and M. Kukudo, J.
Biochem., 95, 967 702 (1984).
9) M. Kubota, Y. Matsuura, S. Sakai, and Y. Katsube, Oyo Toshibsa
10) L. Holm, K. A. Koivula, M. P. Lehtovaara, A. Hemminki, and C.
11) K. Takase, T. Matsumoto, H. Mizuno, and K. Yamane, Biochim.