Molecular Cloning of Levan Fructotransferase Gene from *Arthrobacter nicotinovorans* GS-9 and Its Expression in *Escherichia coli*

Katsuichi Saito, Atsushi Yokota, and Fusao Tomita

Laboratory of Applied Microbiology, Department of Bioscience and Chemistry, Faculty of Agriculture, Hokkaido University, Sapporo 060, Japan

Received June 13, 1997

The gene encoding an extracellular levan fructotransferase, designated the *lff* gene, was cloned from the genomic DNA of *Arthrobacter nicotinovorans* GS-9, and expressed in *Escherichia coli*. It was found that a single open reading frame consisted of 1554 base pairs that encoded a polypeptide composed of a signal peptide of 33 amino acids and a mature protein of 484 amino acids (M, 53,152), and it was also found that a putative ribosome-binding site was present in the upstream from the ORF.

The primary structure had no significant similarity with those of inulin fructotransferases, but had low similarity to the catalytic regions of other fructosylhydrolases.

The expression of the *lff* gene was increased on a plasmid, pLFT-BB1, in which the *lff* gene was fused with z-peptide of the *lacZ* gene of pUC18. An *E. coli* transformant carrying pLFT-BB1 expressed six times as much activity of levan fructotransferase as that of the original strain, *A. nicotinovorans* GS-9.

**Key words:** levan; levan fructotransferase; DFA IV; *lff* gene; *Arthrobacter nicotinovorans* GS-9

Levan is a β-2,6-linked fructan used as a reserve carbohydrate of such monokotyledons as ryegrass and cocksfoot. Levan is also produced by several kinds of bacteria during their assimilation of sucrose through the action of levansucrase (EC 2.4.1.10), and it is possible to produce levan on a large scale by using these microorganisms. Thus, levan is an interesting starting material for the production of useful oligosaccharides such as DFA IV. DFA IV (di-o-fructose-2,6′,6,2″-dianhydride) is an oligosaccharide which is produced from levan by microbial enzymes. *i.e.*, levan fructotransferase (LFTase) and a kind of levanase.

We have reported the purification and properties of a novel type of LFTase from *Arthrobacter nicotinovorans* GS-9 for the effective production of DFA IV. However, it was considered that the yield of LFTase from *A. nicotinovorans* GS-9 was not enough for practical applications, and it would also be favorable for industrial use to introduce useful characteristics such as thermostability into the LFTase. Thus, cloning of the gene coding for LFTase (*lff* gene) and its expression in *E. coli* were done to obtain higher yields of LFTase and to characterize the enzyme.

In this paper, we describe cloning, sequencing, and expression in *E. coli* of the *lff* gene from *A. nicotinovorans* GS-9.

**Materials and Methods**

*N-Terminal and inner amino acid sequences of LFTase.* LFTase was purified from the culture supernatant of *A. nicotinovorans* GS-9 by the method described in our paper. The purified enzyme (1 nmol) was blotted onto a polyvinylidene difluoride (PVDF) membrane, and the amino acid sequence of the N-terminal region was analyzed with a gas-phase protein sequence (Model 477A-120A, Perkin-Elmer Applied Biosystems Division, Foster City, Calif., U.S.A.). For analysis of inner amino acid sequences, the purified enzyme (10 nmol) was denatured and pyridylethylated by the method of Cabins and Friedman, and digested with 90 pmol of lysyl endopeptidase from *Achromobacter lyciues* (Wako Pure Chemical Industries, Ltd., Osaka) at 37°C for 48 h in 0.1 M Tris–HCl buffer (pH 9.0)

containing 8.0 M urea. The digested peptide fragments were put on a reverse-phase HPLC column (Shodex C8P-50 4D, 4.6 × 150 mm, Showa Denko Co., Ltd., Tokyo) equilibrated with 0.1% (v/v) trifluoroacetic acid, and eluted with a linear gradient of 0 to 60% (v/v) acetonitrile at 50°C. The eluted peptides were fractionated by monitoring absorbance at 216 and 280 nm, and the sequences of the peptides were analyzed by using the same apparatus as mentioned above.

**DNA manipulations.** Recombinant DNA techniques were the standard methods described by Sambrook et al. and used as a host for introduction of plasmids, and transformation of this strain was done by the method of Hanahan.

**Synthesis of oligonucleotide probes.** On the basis of the internal partial amino acid sequences of LFTase, two oligonucleotide probes (probes 1 and 2) were synthesized with a DNA synthesizer (Model 381A, Perkin Elmer Applied Biosystems Division) as described in the Results and Discussion.

**Construction and screening of the genomic library of *A. nicotinovorans* GS-9.** *A. nicotinovorans* GS-9 was cultivated by the method described in our paper, and the genomic DNA was prepared as described by Sakurai et al. Partial digests of the genomic DNA with *Sma*I were fractionated by centrifugation on a gradient of 5 to 25% (w/v) NaCl at 37,000 rpm (SW 40T) rotor; Beckman, Fullerton, Calif., U.S.A.) for 4.5 h, and DNA of 10 to 20 kbp was recovered. The fragments were ligated with the BamHI site of JM113 arms and packaged with GigaPack III Gold packaging extract (Stratagene Cloning Systems, La Jolla, Calif., U.S.A.). Plating of the phages was done by using *E. coli* XL-1 Blue MRA (F2) cells as a host strain by the established methods, and the plaques that formed were transferred onto Hybrid-N+ membrane filters (Amersham International plc, U.K.). Plaque hybridization was done at 42°C for 18 h in a solution of 5 × SSC (1 × SSC is 0.15 M NaCl, and 0.015 M tri-sodium citrate), 5 × Denhardt solution (1 × Denhardt is 0.02% (w/v) Ficoll, 0.02% (w/v) polyvinylpyrrolidone, and 0.002% (w/v) bovine serum albumin, 0.1% (w/v) SDS, and 0.25 mg/ml of denatured salmon sperm DNA (Wako Pure Chemical Industries, Ltd.), and the probes 1 and/or 2 labeled with [γ-32P] ATP by using MEGALABEL labeling kit (Takara Shuzo Co., Ltd.). The filters were washed and autoradiographed to detect positive clones.

**Nucleotide sequencing.** The dyeoxy sequencing reaction was done by a Cy5 AutoRead sequencing kit (Pharmacia Biotech, Uppsala, Sweden)
with progressive oligonucleotide primers. These nucleotide sequences were analyzed with an ALF express DNA sequencer (Pharmacia Biotech). Computer analyses of the DNA sequence data and the deduced amino acid sequence were done with GENETYX-MAC (Software Development Co., Ltd., Tokyo). The resulting amino acid sequence was compared with other published sequences in the SWISS PROT database by Network Service. The nucleotide sequence data reported in this paper will appear in the DDBJ/EMBL/GerBank DNA databases under the accession number AB001984.

**Detection of lft gene expression in E. coli.** E. coli clones were cultivated overnight in a test tube containing 5 ml of LB medium (5 g/liter of yeast extract, 10 g/liter of Polypeptone (Nihon Pharmaceutical Co., Ltd., Tokyo), 10 g/liter of NaCl, pH 7.0) containing 100 μg/ml ampicillin at 37°C. Then 0.1 ml of these cultures were inoculated into the same medium with 1 mm IPTG, and cultivated at 37°C for 24 h. The culture broth were centrifuged (8000 g, 4°C, 15 min) to separate cells and culture supernatants. The cells were washed twice with 10 mm sodium phosphate buffer (pH 7.0), and resuspended in 5 ml of the same buffer. They were then disrupted by sonicating with a Sonifier (Model 250, Branson, Danbury, CT, U.S.A.), and the mixture was centrifuged (25,000 g, 4°C, 30 min). The supernatants were used as the cell-free extracts. The LFTase activities of the cell-free extracts and culture supernatants were assayed as described in our paper. The levan-degrading reaction was done in a reaction mixture containing 50 mm of sodium phosphate buffer (pH 6.0), levan (10 mg/ml), and 0.25 ml of the culture supernatant in a total volume of 0.5 ml at 37°C for 24 h. The reaction products were analyzed by TLC with the method B as described previously, and the main product was purified and analyzed by 13C-NMR by the method described in our paper. SDS-PAGE was done on 10% gel by the method of Laemmli.

**Results and Discussion**

**N-Terminal and inner amino acid sequences of LFTase and preparation of probes**

The N-terminal sequence was QASLRAIYHMTT. Five internal partial amino acid sequences were also identified. Two inner amino acid sequences, (IHWDATRNEWVC-IGRRAYAFTYSPNLRDWQWK) and (RLATAWMN-NWK), were selected for preparing probes 1 (5'-AA(C/T) GA(A/G) TGG GT(A/C/G/T) GT(G/T) GT(A/C/G/T) AT-3') and 2 (5'-AC(A/C/G/T) GC(A/C/G/T) TGG ATG AA(C/T) AA(G/T) TGG AA-3'), respectively. The N-terminal and inner amino acid sequences described above are underlined in Fig. 3.

**Cloning of lft gene**

The genomic library of A. nicotinovorans GS-9 was screened by plaque hybridization with probes 1 and 2. Among approximately 15,000 plaques, 17 positive plaques that hybridized with both probes were isolated, and the corresponding phage DNAs were extracted. All phage DNAs seemed to have similar regions of genomic DNA, because almost all phage DNAs had the same 3.6-kb BamHI fragments, which corresponded to the hybridized fragments of BamHI digested genomic DNA of A. nicotinovorans GS-9 with the same probes (data not shown). A phage containing the 3.6-kb BamHI fragment, lftp1T7, was selected, and digested with BamHI. Then the 3.6-kb BamHI fragment was cloned in pUC19. The plasmid thus obtained, pLFT-7C, was used for subsequent analysis.

LFTase activity of the supernatant of E. coli/pLFT-7C was measured, and the products after 24 h of reaction were analyzed as shown in Fig. 1. A product corresponding to DFA IV was detected as a main product by using the supernatant prepared from E. coli/pLFT-7C (lane 2) the same as that from A. nicotinovorans GS-9 (lane 1), and other minor products were also identical between them.

**Fig. 1. TLC Analysis of Reaction Products from Levan with Culture Supernatant of E. coli/pLFT-7C.**

The enzyme reaction was done as described in the text. S., partial-HCl hydrolysates of levan; S1, purified DFA IV; lane 1, purified LFTase from A. nicotinovorans GS-9; lane 2, culture supernatant of E. coli/pLFT-7C, lane 3, that of E. coli/pC19; F1, fructose; F2, levansamine; F3, levantarose; F4, levantarotose.

**Fig. 2. Restriction Map of the Cloned DNA Fragment in pLFT-7C and the Strategy for Nucleotide Sequencing.**

The cloned DNA fragment in a plasmid is represented by a box. The closed region shows the ORF of the lft gene. The arrows indicate the direction of the sequencing.

The main reaction product of E. coli/pLFT-7C was purified, analyzed by 13C-NMR, and identified as DFA IV (data not shown). Thus, it was considered that pLFT-7C was sure to include the lft gene, and LFTase was expressed in E. coli.

**Nucleotide sequence of lft gene**

A restriction map of pLFT-7C was constructed as shown in Fig. 2. Using appropriate restriction sites, the cloned DNA fragment was subcloned into pUC19. By using pLFT-7C and subcloned plasmids, nucleotide sequences of pLFT-7C were identified as indicated in Fig. 2. A single ORF was found between SmaI sites in the central position of 3679 bp BamHI fragment. The nucleotide sequence of the 1752-bp SmaI-SmaI DNA fragment of pLFT-7C is shown in Fig. 3. In the deduced amino acid sequence of this ORF, all the amino acid sequences analyzed above were found (underlined in Fig. 3). These data suggested that this ORF specified the lft gene. From the N-terminal amino acid sequence of the purified LFTase, this ORF was considered to encode a mature protein of 484 amino acids and a putative signal peptide of 33 amino acids. The molecular weight of the deduced mature protein (M, 53,152) agreed closely with that of the purified LFTase (M, 52,000 by SDS-PAGE). The GC content in the ORF was 62.2%. Upstream from
the ORF, a promoter sequence was not found by comparing this sequence with consensus sequences, but a putative ribosome-binding site was found (double-underlined in Fig. 3). The deduced primary structure of LFTase was compared with the other published sequences. No closely similar sequence was found by comparing with two kinds of inulin fructotransferase producing other types of difructose anhydrides from inulin.\(^8,12\) However, as shown in the Table, a part of the sequence had some similarity with the fructosylhydrolases such as a levanase from \textit{Bacillus subtilis},\(^13\) an inulinase from \textit{Kluveromycetes marxianus},\(^14\) and an invertase from yeast.\(^15\) The catalytic residues in these regions were reported as Asp and Glu (underlined in the Table), which acted as a nucleophile and a proton donor, respectively, in a double displacement catalytic mechanism of fructosylhydrolases.\(^16,17\) Thus, it was considered that Asp-52 and Glu-232 of LFTase were putative catalytic amino acid residues (boxed in Fig. 3) and levanan degradation was done by a similar catalytic mechanism. Differences of amino acid residues between LFTase and these fructosylhydrolases around the catalytic residues were supposed to cause the different catalytic reactions, \textit{i.e.}, transferase and hydrolase.

**Expression of \textit{Ift} gene in \textit{E. coli}**

As described above, \textit{E. coli}/pLFT-7C expressed the \textit{Ift} gene in the culture supernatant. To obtain more information on the \textit{Ift} gene and to increase the production of LFTase, the following experiments were done. The 3.6-kbp BamHI fragment of pLFT-7C was ligated to pUC19, and pLFT-14A was selected as a plasmid including \textit{Ift} gene so as to be the opposite direction from the \textit{lac} promoter. To construct pLFT-BB1 and pLFT-BB2, pLFT-7C was digested with \textit{BsrXI}, one of the recognition sites of which was in the gene encoding the signal peptide of LFTase, and the 1.8-kbp \textit{BsrXI} fragment, which included almost all of the \textit{Ift} gene, was treated with T4 DNA polymerase to generate blunt ends. The 1.8-kbp fragment was ligated into the \textit{SmaI} site of pUC18, and pLFT-BB1 and pLFT-BB2 were selected as plasmids carrying the \textit{Ift} gene in the opposite directions near the \textit{lac} promoter.

Figure 4 shows these plasmids and expression levels of LFTase in \textit{E. coli} JM109 carrying these plasmids. In both cases of \textit{E. coli}/pLFT-7C and \textit{E. coli}/pLFT-14A, the activities in cell-free extracts were 0.3 units/ml of culture broth, and the same amounts of the enzyme were present in culture supernatants. When \textit{A. nicotinovorans} GS-9 was cultured under the optimum conditions,\(^2\) LFTase production reached the level of 3.3 units/ml. Therefore, the total LFTase activity obtained by the culture of \textit{E. coli}/pLFT-7C was less than 15\% of that obtained from the culture of \textit{A. nicotinovorans} GS-9. Using pLFT-BB1, in which the \textit{Ift} gene was fused with \textit{a}-peptide of the \textit{lac} \textit{Z} gene of pUC18, good expression of LFTase was attained. As indicated in Fig. 4, the activities in cell-free extract and supernatant increased to 17.6 and 1.4 units/ml, respectively. The total activity was about 6-fold higher than that obtained with \textit{A. nicotinovorans} GS-9. However, \textit{E. coli}/pLFT-BB2 did not show any LFTase activity at all, while \textit{E. coli}/pLFT-14A could produce LFTase. This indicated that there was a promoter of the \textit{Ift} gene within about 800-bp upstream region from the initiation codon of the \textit{Ift} gene. Thus, it was considered that the promoter and ribosome-binding site of \textit{A. nicotinovorans} GS-9 were recognized in \textit{E. coli}, and \textit{E. coli} transformants carrying pLFT-7C and pLFT-14A produced LFTase under control of the promoter of the \textit{Ift}

---

**Table**

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Source</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>LFTase</td>
<td>\textit{A. nicotinovorans} GS-9</td>
<td>48-\textit{GCGD\textit{CG}F-58} 228-\textit{AGC\textit{CG}D\textit{C}-238}</td>
</tr>
<tr>
<td>Levanase</td>
<td>\textit{Bacillus subtilis}</td>
<td>45-\textit{GCGD\textit{CG}F-58} 223-\textit{AGC\textit{CG}D\textit{C}-229}</td>
</tr>
<tr>
<td>Inulinase</td>
<td>\textit{Kluveromycetes marxianus}</td>
<td>49-\textit{GCGD\textit{CG}F-58} 234-\textit{AGC\textit{CG}D\textit{C}-244}</td>
</tr>
<tr>
<td>Invertase</td>
<td>\textit{Saccharomyces cerevisiae}</td>
<td>38-\textit{GCGD\textit{CG}F-58} 219-\textit{AGC\textit{CG}D\textit{C}-229}</td>
</tr>
</tbody>
</table>

Identical amino acids in at least two sequences are indicated by inverted letters. The catalytic residues are underlined.
Molecular Cloning of Levan Fructotransferase Gene

LFTase activity
(units/ml of culture broth)

Cell-free extract
Supernatant

0.3
0.2
0.3
0.2
17.6
1.4
0
0
3.3

Fig. 4. LFTase Activities of E. coli Cells Carrying Constructed Plasmids.

The closed and hatched boxes show the lft gene encoding the mature protein and the signal peptide of LFTase, respectively. The lac promoter of pUC18 and pUC19 is shown by large arrows. The small arrow indicates the direction of transcription of the lft gene. The sequences are only for plFT-BB1, and the bold letters indicate the sequences from cloned DNA fragment. For ease of comparison, the activities of cell-free extracts were expressed by converting units per ml of culture broth.

pLFT-7C
pLFT-14A
pLFT-BB1
pLFT-BB2

A. nicotinovorans GS-9

Fig. 5. SDS-PAGE Analysis of the Cloned LFTase in Cell-free Extracts from E. coli Transformants.

Cell-free extracts corresponding to 30 μl of culture broths were put on, and the gel was stained with Coomassie brilliant blue. Lane M, marker proteins; Lane 1, purified LFTase from A. nicotinovorans GS-9; lane 2, cell-free extract of E. coli/pLFT-7C; lane 3, that of E. coli/pLFT-BB1; lane 4, that of E. coli/pUC18. Standard marker proteins were β-galactosidase (M, 116,000), fructose-6-phosphate kinase (M, 85,200), glutamate dehydrogenase (M, 55,600), akladose (M, 39,200), and trisphosphate isomerase (M, 26,600).

gene itself. A study of details of the upstream region of the lft gene is in progress.

The cell-free extracts from some E. coli transformants were analyzed by SDS–PAGE. As shown in Fig. 5, the cell-free extract from E. coli/pLFT-BB1 gave a strong band (lane 3) corresponding to LFTase from A. nicotinovorans GS-9 (lane 1). This band was transferred to a PVDF membrane and sequenced. The N-terminal amino acids sequence was AQLASLRAIYHMTPPSGWL, and was identical to the N-terminal sequence of LFTase from A. nicotinovorans GS-9 except for the existence of an Ala residue in front of the sequence. In spite of pLFT-BB1 being constructed so as to lack a part of the signal peptide and to fuse with α-peptide of lac Z gene, the signal peptide sequence of A. nicotinovorans GS-9 was recognized and digested in E. coli, and a mature protein was supposed to be secreted into the periplasmic space. This result is similar to our observation in inulin fructotransferase from Arthrobacter sp. H65-7.9)

We are currently studying recognition of the signal peptide sequence of Arthrobacter sp. in E. coli and searching for proteins with signal peptides.

Acknowledgments. The authors wish to thank GC-NMR Laboratory, Faculty of Agriculture, Hokkaido University for NMR spectroscopy. We are grateful to the Biopolymer Analysis Laboratory, Faculty of Agriculture, Hokkaido University for analysis of peptide sequences and synthesis of oligonucleotide probes.

References