Cloning and Characterization of Zymomonas mobilis Genes Encoding Extracellular Levanasucrase and Invertase

Kiyoshi KYONO, Hideshi YANASE,* Kenzo TONOMURA,** Haruhiko KAWASAKI,*** and Takuo SAKAI
Department of Applied Biochemistry, Faculty of Agriculture, University of Osaka Prefecture, Gakuencho 1-1, Sakai, Osaka 593, Japan
Received September 5, 1994

The genes encoding the extracellular levanasucrase and invertase of Zymomonas mobilis have been cloned and sequenced. The levanasucrase gene, sucZ2, spans 1269 bp and encodes an M, 46,790 polypeptide, and the invertase gene, sucZ3, is of 1239 bp and encodes an M, 46,110 polypeptide. The 5'-terminal sequences of both genes corresponded to the N-terminal amino acid sequences of the secreted levanasucrase and invertase, implying that the secretion of both enzymes does not involve proteolytic processing of the N-terminals. Both enzyme molecules appear to carry no typical N-terminal secretion signal. Significant homology between sucZ2 and sucZ3 was observed, but both genes showed no homology to the gene encoding an intracellular invertase coexisting in Z. mobilis. Two genes, sucZ2 and sucZ3, are possibly placed in an operon because the expression of two genes were simultaneously controlled by the regulator gene zBlE, previously identified.

It is known that Zymomonas mobilis efficiently ferments glucose and fructose to produce theoretical yields of ethanol, but sucrose is a poorly fermentable substrate for this organism. We have found three kinds of sucrose-hydrolyzing enzymes, E1, E2, and E3, in Z. mobilis IFO 13756:1). E1 is an intracellular invertase, E2 is an extracellular levanasucrase, and E3 is an extracellular invertase. The activity levels of these enzymes in this organism are very low even if induced by sucrose. A spontaneous mutant of this organism, called strain Z6C, capable of fermenting sucrose vigorously, produced E2 and E3 constitutively in large quantities. 2) From this mutant, the enzymes E1, E2, and E3 have been purified and characterized. 2,3) E1 is a single polypeptide of 59 kDa, while E2 and E3 are composed of two homosubunits of 56 kDa and 58 kDa, respectively.

We are interested in the regulated production and the secretion of these enzymes in Z. mobilis. The E1 gene has been cloned and sequenced, 4) and recently a pair of genes that regulate the production and the secretion of E2 and E3 in Z. mobilis have been cloned and characterized. 5) This paper describes the cloning and characterization of the genes for E2 and E3 from Z. mobilis Z6C and the molecular relationships among E1, E2, E3, and other β-D-fructofuranosidases.

Materials and Methods

Bacterial strains and plasmids. Z. mobilis strain Z6C, a mutant of Z. mobilis IFO 13756 producing E2 and E3 constitutively, 2) was used as a source of the E2 and E3 genes. Escherichia coli HB101 (F-, hsdS20, recA13, ara-14, proA2, lacY1, galK2, rpsL20, xyl-5, mtl-1, supE44) and a cosmids vector pHC79 (6.4 kb, Ap+, Tc-) were used to construct a Z. mobilis genomic library. E. coli JM109 and MV1184, plasmid pUC119, and phage M13K07 used for DNA sequencing were obtained from Takara Shuzo Co., Ltd. (Kyoto).

Culture conditions. Z. mobilis Z6C was grown in RM medium consisting of 2% glucose, 1% yeast extract, and 0.2% KH₂PO₄. pH 6.0, under static conditions at 30°C. E. coli was cultured at 37°C in LB medium consisting of 1% peptone, 0.5% yeast extract, and 0.5% NaCl. pH 7.2. When required, the media had ampicillin (50 µg/ml) or chloramphenicol (50 µg/ml) added and were solidified by the addition of agar (1.5%).

DNA manipulations. The extraction and purification of plasmids, the cleavage and ligation of DNA, the transformation of E. coli with plasmids, Southern blotting of DNA fragments, and other general procedures for DNA manipulations were done as described by Maniatis et al. 7)

Construction of Z. mobilis genomic library. The Z. mobilis genomic DNA extracted by the method of Meade et al. 8) and purified by cesium chloride-ethidium bromide density-gradient centrifugation was partially digested with HindIII and ligated with HindIII-cleaved pHC79 DNA. The recombinant cosmid DNA was packaged with phage coat protein using a Lambda DNA in vitro Packaging Kit (Amersham) and Transfected into E. coli HB101 cells, which were spread and grown on nitrocellulose filters placed on LB plates containing ampicillin to construct the Z. mobilis genomic library.

Colonial hybridization. Library-replica filters were prepared by pressing

<table>
<thead>
<tr>
<th>13</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trp</td>
<td>Thr</td>
</tr>
</tbody>
</table>

Probes E2 5'- TGG ACC GTT GAT GCT ATG -3' C C C C

1 6

Met Phe Asn Phe Asn Ala

Probes E3 5'- ATG TTC AAT TTT AAT GC -3' T C C C

Fig. 1. Synthetic Oligonucleotide Probes for the E2 and E3 Genes.

Probe E2, a mixture of 32 oligonucleotides (21 bases), and probe E3, a mixture of 16 oligonucleotides (17 bases), were synthesized based on the N-terminal 13th to 19th amino acids of E2 and the 1st to 6th amino acids of E3, respectively. As to Thr, Arg, and Ala having 4 or more codons, two codons were chosen with reference to the average codon usage of Z. mobilis genes. 9)

* Present address: Department of Biotechnology, Faculty of Engineering, Tottori University, Tottori 680, Japan.
** Present address: Department of Food Technology, Faculty of Engineering, Fukuyama University, Fukuyama, Hiroshima 729-02, Japan.
*** To whom correspondence should be addressed.
† The nucleotide sequence data reported in this paper will appear in the DDBJ, EMBL, and GenBank Nucleotide Sequence Databases with the accession number D 17524.
new nitrocellulose filters softly over the colonies of the library plates and incubating at 37°C on new LB-plates. When the colonies grew up to the size of 0.5 mm, the filters were moved onto chloramphenicol plates and incubated for 12 h. The filters were treated sequentially with 0.5 M Tris-HCl (pH 7.5), and 1.5 M NaCl-1 M Tris-HCl, and baked at 80°C for 2 h. DNA hybridization was done as described by Duby\textsuperscript{9} using mixed oligonucleotides that were synthesized based on the N-terminal amino acid sequence of the enzyme E3\textsuperscript{11} or E3\textsuperscript{13} (Fig. 1) and labeled by [γ-32P] ATP (NEN Research Products).

Nucleotide sequencing. DNA fragments were subcloned into pUC119 and a series of deletion mutants of each subclone was constructed using a Kilo-Sequence Deletion Kit (Takara Shuzo Co., Ltd., Kyoto). Sequencing was done by the dyeoxy chain termination method\textsuperscript{10} using a Sequenase Version 2.0 Sequencing Kit (United States Biochemical Co.) and [α-32P] dCTP (NEN Research Products). The sequences were analyzed with the aid of computer programs of SDC-GENETYX (Software Development Co., Ltd., Tokyo).

Assay of sucrose-hydrolyzing activity. The enzyme activities of E. coli clones were assayed using cell-free extracts prepared by ultrasonic disruption. When necessary, the enzymes E1, E2, and E3 were separately assayed after DEAE-Toyopearl column chromatography.\textsuperscript{11} A reaction mixture (1 ml) consisting of 0.15 M sucrose, 0.1 M acetate buffer, pH 5.0, and an appropriate amount of enzyme was incubated at 30°C for 5 min, and the reducing sugar produced was measured by the method of Somogyi-Nelson.\textsuperscript{12} One unit of activity was defined as the amount of enzyme that produced reducing sugar equivalent to 1 μmol of glucose per min.

Results

Selection of E3 clones from the Z. mobilis genomic library

Of two hybridization probes for E2 and E3 genes (Fig. 1), the E3 probe was used in the first screening by colony hybridization. Among approximately 7000 colonies of the Z. mobilis genomic library, 26 positive clones were selected. They were classified into 5 groups based on the restriction patterns of recombinant plasmids they carried. Five representative clones had very poor but distinct sucrose-hydrolyzing activities (0.5 to 6 mU/ml) that were about one ten-thousandth of the activity of Z. mobilis Z6C (60 U/ml). The enzyme activity was not detected out of the E. coli cells.

One clone having the highest activity was chosen and its sucrose-hydrolyzing enzyme composition was analyzed by DEAE-Toyopearl column chromatography. The activity was eluted at two peaks corresponding to the positions of the Z. mobilis enzymes E2 and E3, but not E1. The ratios of the activities towards sucrose and raffinose are known to be 1:1 for E2 and 1:0.05 for E3. In this respect, too, the enzymes of two peaks agreed with E2 and E3, respectively. Therefore, it was concluded that this clone produced both E2 and E3 enzymes and its plasmid, termed pHc-ZS2, should contain two genes coding for E2 and E3, which could be faintly expressed in E. coli.

Subcloning of the genes for E2 and E3

Plasmid pHc-ZS2 carried an about 40-kb DNA insert. To curtail the insert, pHc-ZS2 was partially digested with HindIII, self-ligated, and used to transform E. coli JM109. One of transformants showing sucrose-hydrolyzing activity was verified to carry a shortened plasmid, termed pHc-ZS22, that contained a 6.6-kb insert. The column chromatography showed that this clone could produce both E2 and E3 enzymes. The 6.6-kb insert was digested with HindIII to generate three fragments of 0.2, 1.5, and 4.9 kb. The 4.9-kb fragment was subcloned into pHC79 to construct pHc-ZS23 and its reverse, pHc-ZS23R, both of which expressed E2 and E3 in E. coli.

To locate the gene-coding regions, various deletion mutants of pHc-ZS23 and pHc-ZS23R were constructed by Clal digestion. The 4.9-kb fragment was cleaved by Clal into three pieces, 0.9 kb, 2.2 kb, and 1.8 kb, as shown in Fig. 2. The deletion of the 0.9-kb segment from pHc-ZS23 did not arrest the E2 and E3 production (6 mU/ml), but the deletion of the 1.8-kb segment from pHc-ZS23R resulted in complete loss of the enzyme activity. The lack of the 2.2-kb segment strikingly reduced the enzyme activity, and

![Fig. 2. The 4.9-kb Fragment-Subclones pHc-ZS23 and pHc-ZS23R and their Deletion Derivatives.](image_url)

Deletion plasmids were constructed by partial digestion of pHc-ZS23 or pHc-ZS23R with Clal and religation. The insert DNA and the deleted regions are indicated by open boxes and broken lines, respectively. The sucrose hydrolyzing activity was measured using E. coli MV1184 as a host. A trace of activity is shown with ±. Abbreviations: H, HindIII; C, Clal.

![Fig. 3. Sequencing Strategies for the 2.2-kb and 1.8-kb Fragments and the Junction Region.](image_url)
Fig. 4. Nucleotide Sequence for the 2.2-kb Fragment Joined with the 1.8-kb Fragment and the Deduced Amino Acid Sequences from ORF-1 and ORF-2.

A Clp site in the middle of the sequence is the joining point of the 2.2-kb and 1.8-kb fragments. The sequences corresponding to the N-terminal amino acid sequences determined by protein sequencing are underlined. The putative promoters (~35 and ~10 sequences) recognized by Zymomonas RNA polymerase are boxed, and the putative E. coli σ70 promoter sequences and possible ribosome-binding sites are marked by broken lines.

the slightly produced enzyme was proved to be E3 by the analysis of substrate specificity. The E2 production needed both segments of 2.2-kb and 1.8-kb. These results suggested that the E3 gene was in the 1.8-kb segment and the E2 gene was laid across the 2.2-kb and 1.9-kb segments.

Southern hybridization also verified this assumption; the probe E3 hybridized with the 1.8-kb segment and the probe E2 did with the 2.2-kb segment, implying the terminal of the E2 gene resided in the 2.2-kb segment.

Nucleotide sequences of the 2.2-kb and 1.8-kb fragments

The 2.2-kb and 1.8-kb fragments were subcloned into pUC119 and nest deletion mutants of each subclone were constructed. The sequencing strategy for each fragment is shown in Fig. 3. The resulting 2.2-kb and 1.8-kb nucleotide sequences were joined and are given in Fig. 4. The sequence

NII-Electronic Library Service
around the Clal junction was confirmed as follows. The 4.9-kb fragment from pHG-ZS23 was subcloned into the HindIII site of pUC119, and then most of the 1.8-kb fragment was deleted by KpnI digestion. The region flanking the Clal site was sequenced from the KpnI site that located 45-bp apart from the Clal site (Fig. 3).

Identification of the E2 gene (sucZE2) and the E3 gene (sucZE3)

Two open reading frames (ORF) were found in the 4-kb sequence (Fig. 4). ORF-1, 1269 bp, starting at position 1124 and ending at position 2392, codes for a 423-amino acid protein with molecular weight 46,790. ORF-2 of 1239 bp starting at position 2551 and ending at position 3789 codes for a 413-amino acid protein with molecular weight 46,110. The deduced molecular weights are slightly smaller than those estimated for E2 and E3 by SDS-polyacrylamide gel electrophoresis, 56,000 and 58,000 respectively. However, N-terminal amino acid sequences deduced from ORF-1 and ORF-2 correspond completely to the first 19 amino acid residues of E2 and the first 21 residues of E3, respectively, which were analyzed by Edman degradation. Besides, the facts that ORF-1 is laid across the junction of the 2.2-kb and 1.8-kb fragments and ORF-2 is within the 1.8-kb fragment are just as our prediction based on the deletion analysis. Therefore, we have considered ORF-1 and ORF-2 to be the E2 gene and E3 gene, respectively, and designated the E2 gene as sucZE2 and the E3 gene as sucZE3.

The gene sucZE2 is preceded by a sequence AGGA, putative Shine-Dalgarno sequence, 9 bp upstream from the ATC initiation codon of the gene, and a sequence, AGAAAG, 16 bp upstream from the gene sucZE3, may constitute a potential ribosome-binding sequence.

We previously reported a sequence, A/G**A/C/C/AC/ATT/AG/TAT/C-17-18 bases-TAT/TGAT-14 bases-TAAAT, found about 360-bp and 200-bp upstream from the initiation codon of sucZE2. In the same region, a sequence TGT/GA-18 bases-TAAAT is observed, which is similar to a promoter sequence recognized by E. coli sigma RNA polymerase. On the other hand, upstream the sucZE3 gene no sequence like the Zymomonas promoter is found, while a sequence, CTGATT-17 bases-TATCAT, like an E. coli promoter is observed. It may be owing to this potential promoter that the deletion mutant containing only the 1.8-kb fragment (Fig. 2) slightly produced in E. coli.

The G+C contents of sucZE2 and sucZE3 are 47.4% and 47.2% respectively, which closely corresponds to the G+C content, 48.5%, of whole DNA of Z. mobilis.

Discussion

Since most of the E2 and E3 enzymes produced by Z. mobilis are secreted out of the cells, we presumed that the E2 and E3 genes should have S'-terminal nucleotide sequences coding for an N-terminal export signal. However, the S'-terminal sequences of sucZE2 and sucZE3 correspond to the N-terminal amino acid sequences of the extracellular E2 and E3 enzymes. This implies that the E2 and E3 secretion processes don't involve proteolytic processing of the N-terminal. E2 and E3 molecules don't carry the typical N-terminal signal peptide usually found on proteins translocated across the cytoplasmic membrane. The E2 and E3 secretion seems to be done through some other secretion mechanism. We have reported an essential gene, zils, involved in the E2 and E3 secretion. The secretion of proteins lacking N-terminal signal sequences have been reported, but the mechanism of such secretion has not been clear. Recently, Kenny et al. reported that the secretion of haemolysin from E. coli was dependent on a C-terminal secretion signal. As reference data, the hydrophathy plots of the predicted amino acid sequences of E2 and E3 are given in Fig. 5.

Comparison of the nucleotide sequences of sucZE2 and sucZE3 found high homology between levansucrase E2 and invertase E3. The similarity of the whole sequence was 68.1% on the nucleotide level and 65.3% on the amino acid level (Fig. 6A). We are interested in such structural similarity in spite of the functional differences of E2 and E3, though the relationships between their structures and functions remain to be studied.

On the other hand, sucZE2 and sucZE3 showed no homology to the E1 gene coding for intracellular invertase E1 coexisting in Z. mobilis. Finding no molecular relationship between two invertases E1 and E3 surprised us and suggested the diversity of invertase. The genes sucZE2 and sucZE3 showed significant homology with Bacillus subtilis extracellular levansucrase gene, SucB, but no homology with the B. subtilis intracellular sucrose gene, sacA, extracellular levansucrase gene, sacC, E. coli intracellular sucrose hydrolase gene, rafD, or Saccharomyces cerevisiae extracellular invertase gene, suc2. It has been previously reported that the E1 gene has homology with sacA, sacC, rafD, and suc2. This suggests at least two genealogies of β-D-fructofuranosidase.
Fig. 6. Comparison of the Amino Acid Sequences of Levensucrase E2 and Invertase E3 of *Z. mobilis* and Levensucrase of *B. subtilis.*

(A) Amino acid alignment of E2 (sucE2) and E3 (sucE3). (B) Alignment of the most homologous regions among E2, E3 and *B. subtilis* levensucrase (sucB). Identical and similar amino acid residues are indicated by asterisks and dots, respectively. Gaps are introduced to optimize the amino acid alignment.

References