JOURNAL OF BACTERIOLOGY, Dec. 1990, p. 6727–6735
0021-9193/90/126727-09$02.00/0
Copyright © 1990, American Society for Microbiology
Vol. 172, No. 12

Cloning and Sequencing of the *sacA* Gene: Characterization of a Sucrase from *Zymomonas mobilis*

P. GUNASEKARAN,† T. KARUNAKARAN,† B. CAMI, A. G. MUKUNDAN,† L. PREZIOSI, AND J. BARATTI*

Laboratoire de Chimie Bactérienne, Centre National de la Recherche Scientifique, 13277 Marseille Cedex 9, France

Received 19 March 1990/Accepted 12 September 1990

The *Zymomonas mobilis* gene (*sacA*) encoding a protein with sucrase activity has been cloned in *Escherichia coli* and its nucleotide sequence has been determined. Potential ribosome-binding site and promoter sequences were identified in the region upstream of the gene which were homologous to *E. coli* and *Z. mobilis* consensus sequences. Extracts from *E. coli* cells, containing the *sacA* gene, displayed a sucrose-hydrolyzing activity. However, no transfructosylation activity (exchange reaction or levan formation) could be detected. This sucrase activity was different from that observed with the purified extracellular protein B46 from *Z. mobilis*. These two proteins showed different electrophoretic mobilities and molecular masses and shared no immunological similarity. Thus, the product of *sacA* (a polypeptide of 58.4-kDa molecular mass) is a new sucrase from *Z. mobilis*. The amino acid sequence, deduced from the nucleotide sequence of *sacA*, showed strong homologies with the sucrases from *Bacillus subtilis*, *Salmonella typhimurium*, and *Vibrio alginolyticus*.

The ethanologenic gram-negative bacterium *Zymomonas mobilis* can grow only on glucose, fructose, or sucrose and metabolizes these sugars with the production of ethanol and carbon dioxide as main fermentation products (1, 22, 41). Carbohydrate metabolism in *Z. mobilis* has been reviewed recently (43). The monosaccharides glucose and fructose are transported inside the cell by a facilitated diffusion system mediated by a carrier (12), phosphorylated by a specific kinase, and metabolized through the Entner-Doudoroff pathway. The disaccharide, sucrose, is first hydrolyzed to liberate glucose and fructose in the culture medium, and these sugars enter the cell by using the transport system described above. The number and nature of the enzymes involved in sucrose catabolism are not clearly known in *Z. mobilis*.

Sucrose metabolism has been intensively studied in *Bacillus subtilis* (15). Three saccharolytic enzymes are present: an intracellular sucrase (*sacA* gene), an extracellular levansucrase (*sacB* gene), and a levanase (*sacC* gene). All enzymes act as β-D-fructofuranosidases; in addition, levansucrase catalyzes the formation of levan, a high-molecular-weight polymer of fructose. The nucleotide sequences of *sacA*, *sacB*, and *sacC* genes have been determined (13, 21, 39). A strong homology of the N-terminal protein sequences of sucrase, levanase, and yeast invertase (*SUC2* gene) was observed, while no similarity with levansucrase could be detected (21).

Levan formation during growth of *Z. mobilis* on sucrose is well known and the presence of levansucrase is generally well accepted (22, 41). Furthermore, it has been demonstrated that levan formation is cell linked (25), while a high saccharolytic activity was detected in culture medium (28, 33). These results raised the question of the existence, in addition to levansucrase, of a second enzyme, a sucrase, which may be liberated in culture medium during cell growth (41). More recently, two other polymers has been characterized: a cell-linked, high-molecular-mass polysaccharide, identified as an α-fructofuranosyl-(2-1)-β-fructofuranosyl-(2-6)-polymer (2); and extracellular, low-molecular-mass oligomers made of one glucose molecule linked to two or three fructose units (43). Whether these polysaccharides are formed by levansucrase, sucrase, or another enzyme is not clear.

In an attempt to better understand sucrose metabolism, we have cloned and sequenced the *sacA* gene from *Z. mobilis*. The product of this gene was characterized as a saccharolytic enzyme (EC 3.2.1.26) and showed high homologies with other well-known sucrases.

MATERIALS AND METHODS

Bacterial strains and plasmids. *Escherichia coli* RR1 ΔM15 [*leu pro thi rpsL hsdR hsdM lacZ ΔM15*(F′ *lacI^q lacZ ΔM15 pro$^+$*)] and XL1 Blue [endA1 *hsdR17* (r^- m^+) *supE44 thi-1* λ^- *recA1 gyrA96 relA1 Δlac*(F′ *proAB lacI^q ΔM15* Tn*10*)] and *Z. mobilis* ZM1 (ATCC 10988) were used. The cloning vector pUC19 (Pharmacia) was used for construction of the genomic library, and plasmid Bluescript KS (Stratagene) and helper phage R408 (Pharmacia) were used for DNA sequencing.

Growth conditions. *Z. mobilis* ZM1 was grown on RM medium containing the following (per liter): 100 g of glucose, 10 g of yeast extract, 2 g of KH_2PO_4, 0.5 g of $MgSO_4 \cdot 7H_2O$, 0.1 g of $(NH_4)_2SO_4$. Inocula (5% in volume) were grown for 18 h in stationary flasks. Cultures were run at 30°C in a 2-liter fermentor containing 1.5 liters of medium with mild agitation (200 rpm), and pH was controlled at 5.5 by addition of 2 N KOH.

E. coli strains were grown on LB, 2×TY (20), or M63 medium (40). Cells were grown at 37°C under agitation and harvested in the exponential growth phase.

Cloning the gene encoding sucrase. Chromosomal DNA from *Z. mobilis* was prepared by the method of Byun et al. (6) and partially digested with *Sau3A*. Fragments of 5 to 10 kb were isolated by centrifugation on a sucrose gradient (5 to 40%, wt/vol) and were inserted into the dephosphorylated *Bam*HI site of pUC19. The ligation mixture was used to transform *E. coli* RR1, and transformants were selected on LB agar containing ampicillin (50 μg/ml), isopropyl-β-D-thiogalactopyranoside (40 mg/liter), and X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside; 20 mg/liter).

* Corresponding author.
† Present address: Department of Microbiology, School of Biological Sciences, Madurai Kamaraj University, Madurai 625021, India.

Transformants were replicated onto M63 agar plates (40) supplemented with glycerol (10 g/liter), sucrose (10 g/liter), and ampicillin (50 μg/ml). After 48 h at 37°C, GOD-perid reagent (Boehringer, Mannheim, Federal Republic of Germany) was sprayed onto the plates. The appearance of a green halo zone around the clone indicated a sucrose-hydrolyzing activity.

Minicell experiment. *E. coli* DS410 was transformed with plasmid pZS11 and a transformant was used for minicell preparation, using the method of Komai et al. (16). The minicells, purified from the sucrose gradient, were suspended in 0.4 ml of Davis medium containing all amino acids except methionine and shaken vigorously at 30°C for 10 min. A 20-μCi amount of [^{35}S]methionine was added, and the mixture was incubated at 37°C for 10 min. Then 0.5 ml of stop buffer was added and the mixture was spun at 6,000 rpm for 10 min to collect the minicells, which were washed twice with electrophoresis buffer.

DNA sequencing. The gene coding for sucrase was sequenced by the dideoxy method of Sanger et al. (33) with the Kilobase sequencing kit (Bethesda Research Laboratories) and [^{35}S]ATP (Amersham). Nested deletions were obtained with exonucleases III and VII (44). The method of Russel et al. (32) was used to prepare single-stranded plasmid DNA. Sequencing was performed for the entire length of both strands, and all ends of the DNA fragments which were sequenced overlapped one another. The sequence data obtained were analyzed with a computer, using the Microgenie program (29) from Beckmann Instruments Inc. to arrange the overlapping sequence determinations as contiguous units and to perform homology comparisons.

DNA manipulations. Rapid preparation of plasmid DNA from *E. coli* strains was carried out by using the modified alkaline lysis method of Birnboim and Doly (4), and DNA was further purified by CsCl-ethidium bromide gradient centrifugation. The DNA fragments were separated by agarose gel electrophoresis, recovered by electroelution, and purified by passing through an Elutip-d column (Schleicher & Schuell, Dassel, RFA). All cloning experiments were done by using the standard methods described in Maniatis et al. (20).

Determination of sucrase activity. The sucrase-positive *E. coli* clones were grown in 100 ml of LB medium containing 50 μg of ampicillin per ml for 18 h at 37°C, and cells were harvested by centrifugation (7,000 × *g*, 5 min, 4°C). The pellet was washed with 10 mM Tris hydrochloride buffer (pH 7.0) containing 1 mM EDTA and resuspended in 10 ml of the same buffer. Lysozyme was added at a final concentration of 1 mg/ml, and the mixture was incubated for 30 min at room temperature. Ultrasonic disruption of cells was carried out with a Branson Sonicator (three 15-s periods at 40 W, with 45-s intervals between periods). Extracts were clarified by centrifugation (100,000 × *g*, 1 h, 4°C) and supernatant was used as enzyme source. Proteins were determined by the method of Lowry et al. (18).

Sucrase activity was measured by the rate of liberation of reducing sugars from sucrose in the following mixture: 0.25 ml of enzyme solution, 0.25 ml of 0.1 M sodium acetate buffer (pH 5.0) containing 0.1 M sucrose. The reducing sugars liberated after 10 min at 30°C were estimated by the Somogyi-Nelson method (38). One unit is defined as the amount of enzyme hydrolyzing 1 μmol of sucrose per min.

Determination of exchange activity. The exchange reaction between [^{14}C]glucose and sucrose was tested by the method of Chambert and Gonzy-Treboul (7). The enzyme solution was incubated at 30°C in a mixture (0.1 ml) containing 50 mM glucose, 50 mM sucrose, and 185 kBq of [^{14}C]glucose (Amersham) per ml in 10 mM sodium acetate buffer, pH 5.0. Aliquots of 10 μl were taken every 10 min and treated with 100 μl of 70% (vol/vol) ethanol at 65°C. The liquid was evaporated under vacuum overnight. The samples were then resuspended in 70% ethanol and applied to a Whatman 1M paper. Elution was performed by using the solvent butanol-acetic acid-water (4:1:1, vol/vol/vol). The spots were detected by autoradiography, using Kodak X-Omat AR films. For quantification, the spots were cut from the paper chromatogram and radioactivity was assayed by liquid scintillation. One unit is defined as the amount of enzyme incorporating 1 μmol of [^{14}C]glucose into sucrose per min.

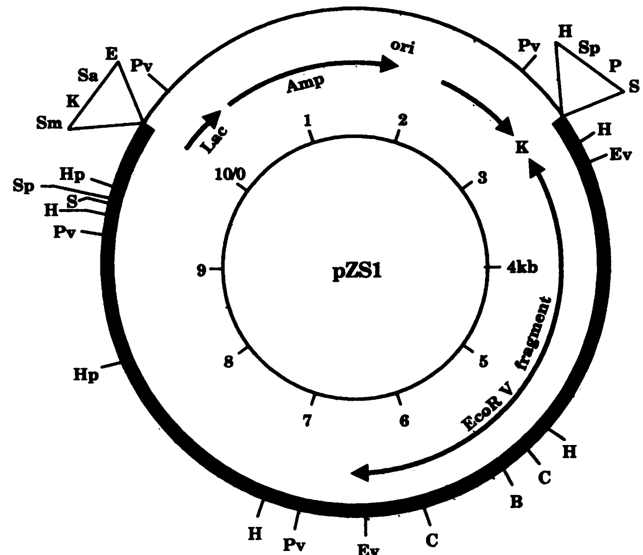

FIG. 1. Circular restriction map of plasmid pZS1. Restriction sites are shown by letters: H, *Hin*dIII; C, *Cla*I; Ev, *Eco*RV; Sp, *Sph*I; P, *Pst*I; S, *Sal*I; Pv, *Pvu*II; B, *Bam*HI; Hp, *Hpa*I; Sm, *Sma*I; K, *Kpn*I; Sa, *Sac*I.

Assay of levan formation. Analysis of sucrose hydrolysis, oligomer, and levan formation was performed at 30°C in a mixture containing 0.1 M [^{14}C]sucrose (740 kBq/ml) in 10 mM sodium acetate buffer (pH 5.0) containing 10 g of *B. subtilis* levans (average molecular mass, 15,000 daltons) per ml and the enzyme solution. Aliquots of 10 μl were taken every 30 min (for 2 h) and treated as above for the exchange reaction. The following reaction products were separated by paper chromatography: unreacted sucrose, glucose, fructose, oligomers containing three or four hexoses, and levans.

RESULTS

Cloning and localization of sucrase gene. Of 1,200 Amp^r transformants of *E. coli* RR1, 3 were sucrase positive on agar plates. Plasmid DNA analysis of the recombinant derivatives of pUC19 carrying inserts of chromosomal DNA from *Z. mobilis* ZM1 revealed the presence of a 7-kb insert in pZS1 and 10-kb inserts in pZS2 and pZS3. Partial restriction analysis of these three fragments indicated that they had overlapping regions. One plasmid, pZS1 (Fig. 1), was studied further.

Deletion of a *Cla*I fragment in pZS1 (approximate size, 1.4 kb) abolished the enzyme activity, indicating that the *Cla*I region is part of the sucrase gene (Fig. 2). Subcloning of the

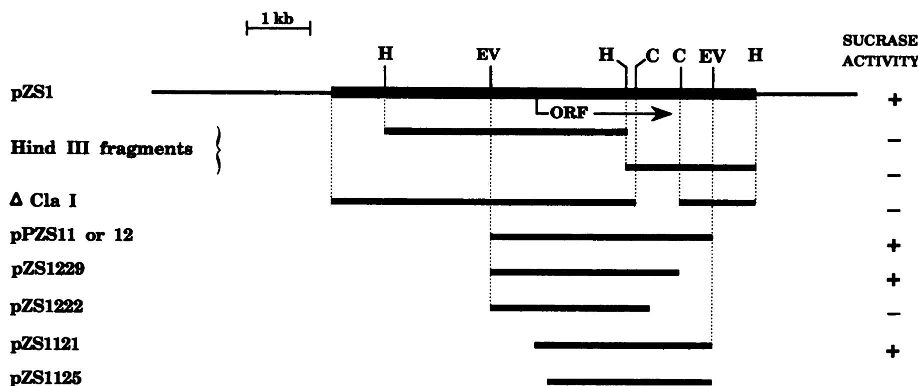

FIG. 2. Linear maps of plasmids: deletion plasmids (sucrase activity of clones indicated). Restriction sites are shown by letters: H, *Hin*dIII; C, *Cla*I; E, *Eco*RV. The most probable open reading frame (ORF) is boxed.

two *Hin*dIII fragments (5.6 and 3.6 kb) from pZS1 did not retain sucrase activity. However, the activity was retained on a 3.2-kb *Eco*RV fragment which was subcloned in both orientations with respect to the *lac* promoter at the *Sma*I site of pUC19 to produce pZS11 and pZS12 (Fig. 2). Plasmids pZS11 and pZS12 were subcloned in Bluescript, and deletion derivatives were generated to identify the coding region. Approximately 0.8 kb was deleted from the left side of the *Eco*RV fragment, resulting in plasmid pZS1121 (Fig. 2), without affecting activity. Deletion of 0.8 kb from the right side, resulting in plasmid pZS1229 (Fig. 2), did not abolish the activity. Further deletions on both sides (plasmids pZS1125 and pZS1222) resulted in loss of sucrase activity. Therefore, it was concluded that the approximate size of the coding region was 1.6 kb.

A Southern blot experiment was run with DNA from plasmid pZS11 as a labeled probe and total chromosomal DNA from *Z. mobilis* after digestion by *Hin*dIII. As expected, two bands were visible (data not shown) at 2.0 and 3.5 kb, which confirmed that the cloned gene was of *Z. mobilis* origin.

Analysis of the product of *sacA* in *E. coli*. The minicell-producing strain of *E. coli* DS410 was transformed by plasmid pZS11 as described in Materials and Methods. After purification of the minicells and incorporation of [^{35}S]methionine, the polypeptide synthesized were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and autoradiography (5). One major protein band of around 56-kDa molecular mass was detected. This protein is likely the product of the *sacA* gene, and the molecular mass agrees reasonably with that (58.4 kDa) deduced from the nucleotide sequence (see below). Thus, expression of the *Z. mobilis sacA* gene was observed in *E. coli*.

Expression and localization of *Z. mobilis* sucrase gene in *E. coli*. *E. coli* strains carrying plasmids pZS1, pZS11, pZS12, and pZS1121, were grown in either LB or M63 medium containing 50 μg of ampicillin per ml, and cell extracts were prepared as described in Materials and Methods. Sucrase activity was detected in all clones (Table 1). Maximum activity was found with plasmid pZS1121. Identical activity for clones with plasmids pZS11 and pZS12 (containing the 3.2-kb *Eco*RV fragment in both orientations with respect to the *lac* promoter) suggested that the sucrase gene was expressed from its own promoter. Addition of isopropyl-β-D-thiogalactopyranoside (40 μg/ml) to the culture medium showed no effect on sucrase activity.

The localization of sucrase activity in *E. coli*(pZS1121) was investigated by using cells from the late exponential growth phase. Cells were separated from the culture medium by low-speed centrifugation (7,000 × *g*, 5 min, 4°C). The sucrase activity was totally (>99.9%) recovered in the cell fraction (specific activity, 0.043 IU/mg), as assayed after ultrasonic disruption of cells. High-speed centrifugation (100,000 × *g*, 1 h, 4°C) of the disrupted cell fraction resulted in the recovery of >90% of sucrase activity in the supernatant. Thus, the sucrase activity was cell bound but not associated with a membrane fraction of *E. coli* cells.

To differentiate further between a periplasmic or cytoplasmic localization, an osmotic shock was run on whole cells (24). Only 16% of the sucrase activity was recovered in the shock fluids; the remainder was found in the shocked cells. A soluble, cytoplasmic localization for the product of *sacA* in *E. coli* was deduced from these experiments.

The kinetics of sucrase activity production by *E. coli* (pZS1121) was investigated on LB medium. No activity was detectable during the exponential growth phase. Activity appeared in the late growth phase, being maximum in early stationary phase. These results suggested the existence of a regulation of the expression of *sacA* in *E. coli* such as repression by nutrients contained in LB medium.

Characterization of sucrase activity. The product of the *sacA* gene was studied in *E. coli* XL1(pZS1121). To determine whether this enzyme is a levansucrase, the exchange reaction described for the *B. subtilis* levansucrase (7) was used. By incubation of [^{14}C]glucose with unlabeled sucrose, the formation of [^{14}C]sucrose appeared according to the reaction:

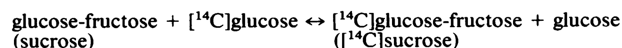

TABLE 1. Expression of sucrase activity in *E. coli*

E. coli strain	Plasmid	Sucrase activity (U/mg of protein)[a]	
		LB	M63
RR1	pUC19	<0.001	<0.001
RR1	pZS1	0.11	0.06
RR1	pZS11	0.17	0.08
RR1	pZS12	0.18	0.09
XL1	pZS1121	0.36	0.10

[a] Cells were grown on LB or M63 medium as described in Materials and Methods.

TABLE 2. Comparison of reaction products after incubation with [^{14}C]sucrose of the product of the *sacA* gene in *E. coli* and purified protein B46 from *Z. mobilis*

Reaction product	*E. coli* XL1(pZS1121)		Protein B46[a]	
	Rate[b]	%	Rate[b]	%
Sucrose	124.4	100	2,766	100
Glucose	58.6	47.8	1,366	49.4
Fructose	55.8	45.5	1,187	43.0
Oligomer 1[c]	4.9	4.0	200	7.2
Oligomer 2[c]	3.3	2.7	9.1	0.3
Levans	$<5 \times 10^{-3}$	<0.004	3.6	0.18

[a] Protein B46 is an extracellular sucrase with fructosyltransferase activity (27).

[b] Expressed as nanomoles of hexose reacting per minute per milliliter of reaction mixture. For sucrose, the reaction rate is consumption, while it is apparition for all of the other products.

[c] Oligomers 1 and 2 contained one glucose and two or three fructose residues, respectively.

Cellular extracts of *E. coli* XL1(pZS1121) showed sucrose-hydrolyzing activity (0.11 IU/mg), but no exchange activity could be detected ($<3 \times 10^{-5}$ IU/mg). Thus, the ratio of exchange to sucrase activity was <0.0003. By comparison, protein B46, a saccharolytic enzyme purified from the culture medium of *Z. mobilis* (27), showed a value 230-fold higher for the same ratio (Preziosi et al., unpublished results). These results suggested that the product of *sacA* displayed a different enzymatic activity than protein B46, the former being a sucrase and the latter being a sucrase with a transfructosylation activity.

Confirmation of the sucrase nature of the product of the *sacA* gene was obtained by analysis of the reaction products formed in vitro by incubation with [^{14}C]sucrose. The results (Table 2) are expressed as reaction rates. With cell extracts from *E. coli* XL1(pZS1121), the main reaction products were glucose and fructose. Although low-molecular-mass oligomers were present, levan formation could not be detected. The pattern was significantly different for the purified protein B46 of *Z. mobilis*: (i) high amounts of glucose and fructose were found as well as fructo-oligomers; (ii) levan formation was demonstrated at a low but reproducible level. Levan formation was never detected with cell extracts from *E. coli* XL1(pZS1121) with the same amount of sucrase units in the incubation mixture.

The product of *sacA* is a sucrase. Given the differences found in exchange activity and levan formation, it was concluded that the product of *sacA* was a sucraselike enzyme (EC 3.2.1.2.6) and was different from protein B46 purified from *Z. mobilis*. However, both enzymes were able to catalyze the formation of fructo-oligomers from fructose (Table 2). Further evidence for the nonidentity of these two proteins was seen by comparison of migration after gel electrophoresis. Sucrase activity of cell extracts from *E. coli* XL1(pZS1121) was detected after gel electrophoresis in native conditions. In the conditions used, the active band migrated with an R_f of 0.08, while an R_f of 0.38 was observed for protein B46. For an unknown reason, sucrase activity could not be detected, after sodium dodecyl sulfate-polyacrylamide gel electrophoresis, with *E. coli* XL1(pZS1121) extracts. An antiserum, raised against protein B46 (27), showed no reaction with *E. coli*(pZS121) extracts after native or sodium dodecyl sulfate-gel electrophoresis and immunoblotting in conditions in which a positive spot was observed for protein B46 (results not shown).

The product of *sacA* is present in *Z. mobilis*. Native gel electrophoresis was used to detect sucrase activity in *Z. mobilis* ZM1 extracts (extracellular and cellular fractions). The culture medium was concentrated by ultrafiltration and dialyzed before use, while the cell fraction was ultrasonically disrupted and centrifuged. Sucrase activity was detected on native gels and showed two bands corresponding to proteins B46 and A51 (28) in the extracellular fraction and a weak band, in the soluble cellular fraction, at the same level as the product of *sacA* in *E. coli* (results not shown). This experiment confirmed the presence of the product of *sacA* in the cellular fraction of *Z. mobilis*.

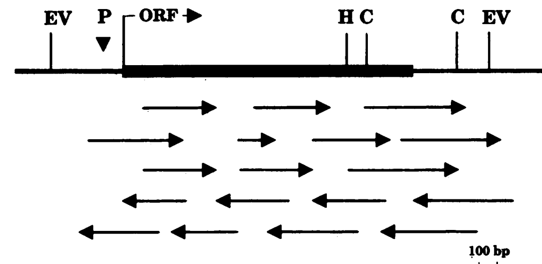

FIG. 3. Sequencing strategy for the *sacA* gene. ORF, Open reading frame; P, promoter; H, *Hin*dIII; C, *Cla*I; EV, *Eco*RV.

Nucleotide sequence of *sacA* gene. The sequencing strategy for the *sacA* gene is shown in Fig. 3, and the nucleotide sequence of the 2,091-bp fragment containing the sucrase gene is shown in Fig. 4. Two open reading frames were found starting with an ATG codon at positions 93 and 212 and both ending with a TAA codon at position 1626. Several arguments favored ATG 93 as the initiation codon. (i) ATG 93 is preceded, 8 bp upstream, by a putative ribosome-binding site (see below); this sequence is not present for ATG 212. (ii) The region near ATG 93 was sequenced from a clone containing pZS1121 which exhibited a positive result for sucrase activity, while a clone containing pZS1125 (lacking this region, but containing the region with ATG 212) showed no sucrase activity (Fig. 2).

Although uniquivocal proof is missing, it is likely that ATG at position 93 is the real initiation codon. The open reading frame would be 1,533 bp long and would encode a polypeptide chain of 511 residues with a calculated molecular mass of 58,360 Da. The G+C content of the *sacA* gene was 43.3%, a value lower than the reported 48.5% for the whole DNA of *Z. mobilis* (22).

The coding region of *sacA* is preceded by the sequence AAAGGCA, a probable ribosome-binding site (37) located 8 bp upstream from the start codon. This sequence matches three bases, AGG, with the *E. coli* consensus sequence (36) and *adhB* (10) and *pgk* (8) genes from *Z. mobilis*. The spacing between the ATG and the ribosome-binding sequence was 8 bases, similar to the average for *E. coli* (7 ± 1) (10) and identical to the average for *Z. mobilis* (8 ± 2) (26). The region upstream from the translational start was rich in A+T (65 versus 56% for the *phoC* gene) and contained poly(A) (one with seven, one with five, and two with four bases) and poly(T) (one with five and two with three bases) regions. The sequence TATAAT matches five bases of the *E. coli* consensus for the −10 region and three bases with the proposed *Z. mobilis* consensus (26). A similar homology was observed with the sequence TTGTCTTTGGTC, which matches seven bases with the *E. coli* consensus for the −35 region (26). The putative −10 and −35 regions of the *sacA* gene showed more similarity with those of *phoC* than any other well-expressed

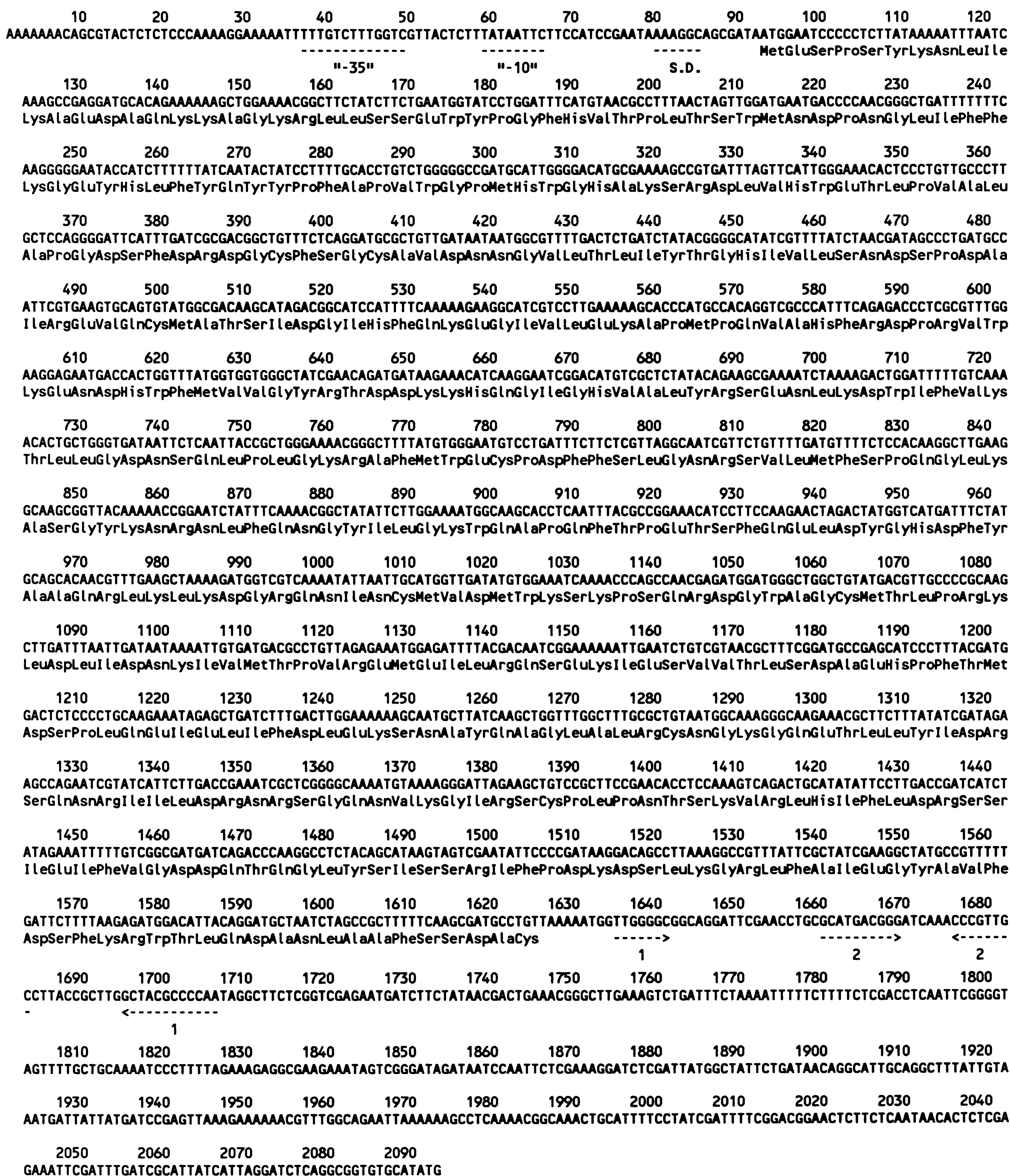

FIG. 4. Nucleotide sequence of the 2.1-kb fragment containing the *sacA* gene. Putative ribosome-binding site (S.D.), −10, and −35 regions are underlined. Convergent arrows indicate two regions of dyad symmetry. The translated amino acid sequence is also shown.

gene in *Z. mobilis*. The sequence TTGTCTT of the −35 region was identical in *sacA* and *phoC*, and four bases of the TATAA sequence from the −10 region of *sacA* were identical in *phoC*. However, unlike *phoC* and three other *Z. mobilis* genes, *pdc* (9), *adhB* (10), and *gap* (11), the coding region of *sacA* was not preceded by a long, untranslated leader sequence. The exact role of this untranslated sequence in four *Z. mobilis* genes is not known (26).

The sequences in regions 1635–1644 and 1660–1669 were highly complementary to those at positions 1695–1706 and 1676–1683, respectively. A possible secondary structure for mRNA could be speculated, from computer analysis, with a two stems–two loops pattern with predicted energies of −26.6 and −15.8 kcal/mol (ca. 111.2 and 66.0 kJ/mol), respectively. This structure seemed characteristic of sites recognized by RNase III and might be involved in mRNA processing (30, 31).

Codon usage and bias. Table 3 shows a comparison of codon usage for *sacA*, the average of five proteins of *Z. mobilis*, and the average of 52 proteins of *E. coli* (26). The pattern of codon usage in *sacA* showed differences. All codons were used with the exception of AGG. Sixteen codons showed a frequency of <1% (seven were also poorly used in *E. coli* and in *Z. mobilis* combined). Frequencies of >3% were observed for six codons in *sacA*; of them, two were also highly used in *E. coli* or *Z. mobilis* combined. However, codon bias was very low in *sacA* compared with *Z. mobilis* combined. The average codon bias index (3) was much lower for *sacA* (0.17) than for the average of *Z. mobilis* (0.58) when calculated for all amino acids. In this respect, *sacA* showed more similarity with *phoC* than with any other highly expressed *Z. mobilis* gene (26). It has been suggested that codon abundance might reflect the relative abundance of tRNA species with a subsequent limitation of expression for genes with low codon bias (14). In addition, the two codons (CAT and TGT) clearly dominant in *sacA* were different from the four dominant codons in *Z. mobilis* combined (ATC, TAC, GAA, and TGC).

Comparison of sucrase with other β-fructosidases. The deduced amino acid sequence of the *sacA* gene from *Z. mobilis* was compared with those of the gram-positive *B. subtilis* levansucrase, levanase, and sucrase (13, 21, 40), the gram-positive *Streptococcus mutans* sucrase (34), the gram-negative *Vibrio alginolyticus* (35) and *Salmonella typhimurium* (*scrB* gene from plasmid pUR400; K. Jahreis, unpublished results) sucrases, and the *Saccharomyces cerevisiae* invertase encoded by the *SUC2* gene (42). The best alignment of the *Z. mobilis* sucrase with that of other related proteins showed the following percentages of homology with respect to identical amino acids: *B. subtilis* sucrase, 34.4%; *Salmonella typhimurium* sucrase, 34.0%; *V. alginolyticus* sucrase, 33.5%; *Streptococcus mutans* sucrase, 29.8%; *Saccharomyces cerevisiae* invertase, 29.1%; *B. subtilis* levanase, 26.4%; and *B. subtilis* levansucrase, 18.0%.

Furthermore, maximal homology was found in the NH_2 terminal region of these seven proteins, which showed six well-conserved regions (A to F) with identical or similar amino acids (Fig. 5). Boxes A, C, and F had the highest number of identical amino acids, with box A containing almost the same amino acids in all seven proteins. Boxes B and D contained no identical but a high proportion of similar amino acids, while box E contained both identical and similar amino acids. These regions might be important for expression of enzyme activity.

In addition, the amino acid sequence Met-Trp-Glu-Cys-Pro-Asp around cysteine 231 in *Z. mobilis* sucrase (MWECPD in

TABLE 3. Comparison of translated codon usage

Amino acid	Codon	Frequency (mol%)		
		sacA	*Z. mobilis* combined[a]	*E. coli* combined[b]
Phe	TTT	3.5	0.7	1.3
	TTC	2.0	2.5	2.2
Leu	TTA	2.0	0.6	0.7
	TTG	1.6	1.5	0.9
	CTT	2.5	2.3	0.8
	CTC	0.4	1.7	0.8
	CTA	1.4	<0.1	0.2
	CTG	1.6	3.5	6.8
Ile	ATT	2.3	1.7	2.2
	ATC	2.0	3.3	3.7
	ATA	1.2	<0.1	0.2
Met	ATG	2.7	2.6	2.8
Val	GTT	2.0	5.0	2.9
	GTC	1.6	2.2	1.2
	GTA	0.6	<0.1	1.8
	GTG	0.8	0.6	2.2
Ser	TCT	2.2	1.1	1.3
	TCC	0.6	1.4	1.5
	TCA	1.0	0.6	0.4
	TCG	0.8	0.5	0.6
	AGT	0.6	0.3	0.3
	AGC	2.3	1.7	1.4
Pro	CCT	2.0	0.8	0.5
	CCC	1.6	0.6	0.3
	CCA	0.6	0.5	0.7
	CCG	1.0	2.7	2.5
Thr	ACT	0.4	0.8	1.1
	ACC	0.4	3.7	2.4
	ACA	1.2	0.2	0.3
	ACG	1.6	1.2	0.8
Ala	GCT	2.3	7.4	2.6
	GCC	1.6	3.8	2.2
	GCA	1.4	2.6	2.3
	GCG	0.4	1.2	3.2
Tyr	TAT	2.3	1.5	1.0
	TAC	1.0	0.4	1.5
His	CAT	2.5	1.5	0.7
	CAC	0.2	1.3	1.2
Gln	CAA	3.5	0.7	1.0
	CAG	1.2	1.7	3.2
Asn	AAT	3.5	1.7	1.0
	AAC	1.0	2.5	2.8
Lys	AAA	4.7	4.2	4.1
	AAG	1.6	2.4	1.3
Asp	GAT	4.9	4.2	2.5
	GAC	2.3	2.4	3.0
Glu	GAA	3.7	5.4	4.9
	GAG	1.0	<0.1	1.8
Cys	TGT	1.4	0.2	0.4
	TGC	0.4	1.6	0.5
Trp	TGG	2.5	1.8	0.7
Arg	CGT	1.4	1.8	3.1
	CGC	1.0	1.7	2.0
	CGA	1.2	0.2	0.2
	CGG	0.6	0.2	0.2
	AGA	1.4	0.2	<0.1
	AGG	0	0	<0.1
Gly	GGT	1.0	5.1	3.8
	GGC	2.7	2.6	3.1
	GGA	1.8	0.4	0.4
	GGG	1.6	<0.1	0.6

[a] Average of *Z. mobilis* including *gap* (11), *pgk* (8), *adh* (9), *pdc* (9), and *phoC* (26).
[b] Average of 52 proteins from *E. coli* (26).

```
                                                            A             B
ZS  MESPS----- --YKNLIKAE -DAQKKAGKR LLSSEWYPGF HUTPLTSW MN DPNG LIFFKG --EY HLFYQ -
BS  MTAHD----- ---QELRRRA YEEUEKKEPI ANSDPHRQHF HIMPPUGL LN DPNG UIYWKG --SY HUFFQ -
PU  MSLPS----- ---RL--PAI LQAUMQGQPR ALADSHYPRW HHAPUTGL MN DPNG FIEFAG --RY HLFYQ -
UB  MSLNNRWTUE QRYRRLEQIP QCDIEEMTLS RQQDKGFPSF HIAPKFGL LN DPNG LCYFNG --EH HIFYK -
SM  MNLPQ----N IRYRRYQDWT EEETKSIKTN UALSPWHTTY HIEPKTGL LN DPNG FSYFNG --KF NLFYQ N
YI  M--------- ---------S M-TNETSD-- ------RPLU HFTPNKGW MN DPNG LWYDEK DAKW HLYFQ -
BL  MKKRLI-QUM IMFTLLLTMA F-SADAADSS YYDEDYRPQY HFTPEANW MN DPNG MUYYAG E--Y HLFYQ -

                                                      C
ZS  YYPFAPUWGP -MHWGHAKSR DLUHWETLPU ALAPGDSFDR DGC FSGCA UD NNGULTLIYT GHIULSNDSP
BS  WQPFQTGHGA -KFWGHYTTQ DUUNWKREEI ALAPSDWFDK NGC YSGSA UT KDDRLYLFYT GNURDQDGNR
PU  WNPLACDHTF -KCWAHWSSI DLLHWQHEPI ALMPDEEYDR NGC YSGSA UD NNGTLTLCYT GNUKFAEGGR
UB  WTPUGPUHGM -KYWYHLSTK DFIHFTDHGU GLHPDQDYDS HGU YSGGA LU ENNQULLFFT GNKRDQNWNR
SM  W-PFGAAHGL -KSWIHTESE DLUHFKETGT ULYPDTSHDS HGA YSGSA YE IGDQLFLFYT GNURDENWUR
YI  YNPNDTUWGT PLFWGHATSD DLTNWEDQPI AIAPKR--ND SGA FSGSM UU DYNNTSGFFN DTIDPRQRCU
BL  YHPYGLQWG- PMHWGHAUSK DLUTWEHLPU ALYPD---EK GTI FSGSA UU DKNNTSGFQT GKEKPL---U

                                                D
ZS  DAI------R EUQCMATSID GIHFQKEGIU LEKAPMPQUA HFRDPRU-W - KENDHWFMUU GYRTDDKKHQ
BS  ETY------- --QCLAUSDD GLSFEKKGUU ARLPEAILTA HFSRSEU-W - EHEGTWYMUI GAQTENLK--
PU  TAW------- --QCLATENA DGTFRKIGPU LPLPEGY-TG HURDPKU-W - RHEDLWYMUL GAQDRQKR--
UB  IP-------- -TQCFATMDS DGSIEK--HG UUIENEHYTE HFRDPKU-W - KKGDDYLMUU GAQTKTEH--
SM  HP-------- -LQIGAFMDK KGNIQKFTDU LIKQPNDUTE HFRDPQI-F - NYKGQFYAIU GAQSLDFG--
YI  AIWTYNTPES EEQYISYSLD GGYTFTEYQK NPULA-ANST QFRDPKUFW Y EPSQKWIMTA AKSQDYK---
BL  AIYTQDREGH QUQSIAYSND KGRTWTKYAG NPUIPNPGKK DFRDPKUFW Y EKEKKWUMUL AAGD--R---

                                             E ↓
ZS  GIGHUALYRS ENLKDWIFUK TLLGDNSQLP LGKRA FMWEC PDF FSLGNRS U------LMF SPQGLKASG-
BS  --GQAULFAS DNLTEW-RFL GPITGAGFNG LDDFG YMWEC PDL FSLQGSD U------LIU SPQGLEADG-
PU  --GKULLFSS ADLHQWTSMM GEIAGHGING LDDUG YMWEC PDL FPLGDQH I------LIC CPQGIAREE-
UB  --GSMALYQS KDLKTWQH-K GPIKTK---- FSDLG YMWEC PDF FEINGQS U------MLF SPQGUSSSNP
SM  --GSKS---- ---------- ---------- ----E YMIEC PNL UFINEQP U------LIY SPQGLSKSE-
YI  ----IEIYSS KNLKQWTYAS EFGQDQGSHG ----- YQYEC PGL IEUPTEQ DPSKSYWUMF ISINPGAPAG
BL  ----ILIYTS KNLKQWTYAS EFGQDQGSHG ----- GUWEC PDL FELPUDG NPNQKKWUMQ USUGNGAUSG

                                            F
ZS  YKNRNLFQNG YILGK-WQ-- APQFTPETSF QEL DYGHDFY A AQRLKLKDG RQNINCMUDM WKSKPSQRD-
BS  FRYQNUYQSG YFUGR-LDYN KPELK-HGEF TEL DQGFDFY A PQTLEDDQG RR----ILFA WMAUPDQDEG
PU  ECYLNTYPAU WMAGE-FDYA AGAFR-HGEL HEL DAGFEFY A PQTMLTSDG RR----LLUG WMGUPEGEEM
UB  YDFKNIYSUA YIUGDQLNLE SMTLENHQDI LQP DYGFDFY A PQTYLDESG RR----ILIA WIGLPEID--
SM  LDYHNIYPNT YKUCQSFDTE KPALUDASEI QNL DFGFECY A TQAFNAPDG RU----YAUS WIGLPDID--
YI  GSFNQYFUGS FN-GTHFE-- --AFDNQSRU U-- DFGKDYY A LQTFFNTDP TYGSA-LGIA WASNWEYSAF
BL  GSGMQYFUGD FD-GTHFK-- --NENPPNKU LWT DYGRDFY A AUSWSDIPS TDSRR-LWLG WMSNWQYAND
```

FIG. 5. Comparison of amino acid sequences of *Z. mobilis* sucrase (ZS), *B. subtilis* sucrase (BS) and levanase (BL), yeast invertase (YI), *Streptococcus mutans* sucrase (SM), *Salmonella typhimurium* (plasmid pUR400) sucrase (PU), and *V. alginolyticus* sucrase (VB). Well-conserved regions are boxed. Arrow indicates the putative cysteine residue of the active site. X's and squares indicated identical and similar amino acids, respectively.

box E) was also found in *B. subtilis* sucrase and levanase and in *Salmonella typhimurium* and *V. alginolyticus* sucrases and was very similar to that found in yeast invertase and *Streptococcus mutans* sucrase. This sequence might contain the catalytic sulfhydryl group as postulated by Martin et al. (21).

DISCUSSION

The *sacA* gene from *Z. mobilis* was cloned in *E. coli* and its nucleotide sequence was determined. The open reading frame could encode a polypeptide of 511 amino acids with a calculated molecular mass of 58,360 Da. This is in agreement

with the value determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the polypeptide synthesized by minicells. There was no apparent signal sequence in the N-terminal part of the product of *sacA*, and the hydrophobicity plot showed no special region rich in hydrophobic residues. These observations together with the detection of sucrase activity after gel electrophoresis in cellular extracts strongly suggested that the product of *sacA* is an intracellular protein in *Z. mobilis* similar to the corresponding sucrase in *B. subtilis* (17).

The *sacA* gene was expressed in *E. coli* as shown by minicell experiment and detection of sucrase activity in *E. coli* extracts. A putative ribosome-binding site and −10 and −35 regions were identified by comparison with those of other *Z. mobilis* genes. These regions showed homology with the corresponding *E. coli* consensus, which may explain the expression of *sacA* in *E. coli* from its own promoter.

The saccharolytic activity of the product of *sacA* was studied in cell extracts of *E. coli* carrying plasmid pZS1121. The activity was intracellular in this bacterium, and no levansucrase-like activity could be detected by either the exchange reaction (7) or levan formation. Therefore, it was concluded that *sacA* encoded a true sucrase which was different from levansucrase (19) and from protein B46 purified from the culture medium (Preziosi et al., unpublished results). The gene was called *sacA* by analogy with the *B. subtilis* sucrase gene. The protein sequence of the sucrase from *Z. mobilis* showed a high homology with other well-characterized sucrases.

The exact role of this new enzyme in sucrose metabolism by *Z. mobilis* is not clear. It is well known that sucrose is hydrolyzed by an extracellular saccharolytic enzyme and that glucose and fructose enter the cell by a specific transport system. Thus, there is no need for an intracellular sucrase. In *B. subtilis* the sucrase showed a higher affinity for sucrose-phosphate than for sucrose. It was concluded that this sucrase acts more as a sucrose-phosphate hydrolase than a true sucrase (17). Assuming a similar function for the *Z. mobilis* sucrase, the presence of an active transport for sucrose (through a phosphotransferase-like system) would be necessary. No evidence for existence of such a transport system in *Z. mobilis* has been published.

Further work is needed to better understand the specific functions of levansucrase, protein B46, and sucrase in sucrose metabolism.

ACKNOWLEDGMENTS

This work was supported by a grant from the Communauté Economique Européenne (DGXII, EN3B/B1/194F) and the Indo-French Centre for the Promotion of Advanced Research (New Delhi).

We are grateful to F. Titgemeyer and K. Jahreis for communication of the sequence of the *scrB* gene from *S. typhimurium* before publication. N. Aït-Abdelkhader is acknowledged for running detection experiments of sucrase activity in *Z. mobilis* ZM1.

LITERATURE CITED

1. **Baratti, J., and J. Bu'Lock.** 1986. *Zymomonas mobilis*, a bacterium for ethanol production. Biotechnol. Adv. **4**:95–115.
2. **Barrow, K. D., J. G. Collins, P. L. Rogers, and G. M. Smith.** 1984. The structure of a novel polysaccharide isolated from *Zymomonas mobilis* determined by nuclear magnetic resonance spectroscopy. Eur. J. Biochem. **145**:173–179.
3. **Bennetzen, J. L., and B. D. Hall.** 1982. Codon selection in yeast. J. Biol. Chem. **257**:3026–3031.
4. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
5. **Bonner, W., and R. Laskey.** 1974. A film detection method for tritium labelled protein and nucleic acid in polyacrylamide gels. Eur. J. Biochem. **46**:83–88.
6. **Byun, M. O. K., J. B. Kaper, and L. O. Ingram.** 1986. Construction of a new vector for the expression of foreign genes in *Zymomonas mobilis*. J. Ind. Microbiol. **1**:9–15.
7. **Chambert, R., and G. Gonzy-Treboul.** 1976. Levansucrase of *Bacillus subtilis*: kinetic and thermodynamic aspects of transfructosylation processes. Eur. J. Biochem. **62**:55–64.
8. **Conway, T., and L. O. Ingram.** 1988. Phosphoglycerate kinase gene from *Zymomonas mobilis*: cloning, sequencing, and localization within the gap operon. J. Bacteriol. **170**:1926–1933.
9. **Conway, T., Y. A. Osman, J. L. Konnan, E. M. Hoffman, and L. O. Ingram.** 1987. Promoter and nucleotide sequences of the *Zymomonas mobilis* pyruvate decarboxylase. J. Bacteriol. **169**:949–954.
10. **Conway, T., G. W. Sewell, Y. A. Osman, and L. O. Ingram.** 1987. Cloning and sequencing of the alcohol dehydrogenase II gene from *Zymomonas mobilis*. J. Bacteriol. **169**:2591–2597.
11. **Conway, T., G. W. Sewell, and L. O. Ingram.** 1987. Glyceraldehyde-3-phosphate dehydrogenase gene from *Zymomonas mobilis*: cloning, sequencing, and identification of promoter region. J. Bacteriol. **169**:5653–5662.
12. **Di Marco, A. A., and A. H. Romano.** 1985. D-Glucose transport system of *Zymomonas mobilis*. Appl. Environ. Microbiol. **49**:151–157.
13. **Fouet, A., A. Klier, and G. Rapoport.** 1986. Nucleotide sequence of the sucrase gene of *Bacillus subtilis*. Gene **45**:221–225.
14. **Hoekema, A., R. A. Kastelein, M. Vasser, and H. A. De Boer.** 1987. Codon replacement in the PGK1 gene of *Saccharomyces cerevisiae*: experimental approach to study the role of biased codon usage in gene expression. Mol. Cell. Biol. **7**:2914–2924.
15. **Klier, A. F., and G. Rapoport.** 1988. Genetics and regulation of carbohydrate catabolism in *Bacillus*. Annu. Rev. Microbiol. **42**:65–95.
16. **Komai, N., T. Nishizawa, Y. Hayakawa, T. Murotsu, and K. Matsubara.** 1982. Detection and mapping of six minF encoded proteins by cloning and analysis of dissected minF segments. Mol. Gen. Genet. **186**:193–230.
17. **Kunst, F., M. Pascal, J. A. Lepesant, J. Walle, and R. Dedonder.** 1974. Purification on some properties of an endocellular sucrase from a constitutive mutant of *Bacillus subtilis* Marburg 168. Eur. J. Biochem. **42**:611–620.
18. **Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall.** 1951. Protein measurement with the Folin phenol reagent. J. Biol. Chem. **193**:265–275.
19. **Lyness, E. W., and H. W. Doelle.** 1983. Levansucrase from *Zymomonas mobilis*. Biotechnol. Lett. **5**:345–350.
20. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. **Martin, I., M. Débarbouillé, E. Ferrari, A. Klier, and G. Rapoport.** 1987. Characterization of the levanase gene of *Bacillus subtilis* which shows homology to yeast invertase. Mol. Gen. Genet. **208**:177–184.
22. **Montenecourt, B. S.** 1985. *Zymomonas*, a unique genus of bacteria, p. 261–289. *In* A. L. Demain and N. A. Solomon (ed.), Biology of industrial microorganisms. Benjamin-Cummings Publishing Co., Menlo Park, Calif.
23. **Mortatte, M. P. L., H. H. Sato, and Y. K. Park.** 1983. Induction and some characteristics of sucrose hydrolyzing enzyme from *Zymomonas mobilis*. Biotechnol. Lett. **5**:229–232.
24. **Nossal, N. G., and L. A. Heppel.** 1966. The release of enzyme by osmotic shock from E. coli in exponential phase. J. Biol. Chem. **241**:3055–3062.
25. **Park, Y. K., M. P. L. Mortatte, and H. H. Sato.** 1983. Study of levan formation during fermentation of *Zymomonas mobilis* on sucrose. Biotechnol. Lett. **5**:515–518.
26. **Pond, J. L., C. K. Eddy, K. F. Mackenzie, T. Conway, D. J. Borecky, and L. O. Ingram.** 1989. Cloning, sequencing, and

characterization of the principal acid phosphatase, the *phoC*+ product, from *Zymomonas mobilis*. J. Bacteriol. **171**:767–774.

27. **Preziosi, L., G. P. F. Michel, and J. Baratti.** 1990. Characterization of sucrose-hydrolyzing enzymes of *Zymomonas mobilis*. Arch. Microbiol. **153**:181–186.
28. **Preziosi, L., G. P. F. Michel, and J. Baratti.** 1990. Sucrose metabolism in *Zymomonas mobilis*: production and localization of sucrase and levansucrase activities. Can. J. Microbiol. **36**: 159–163.
29. **Queen, C., and L. J. Korn.** 1984. A comprehensive sequence analysis program for the IBM personal computer. Nucleic Acids Res. **12**:581–599.
30. **Regnier, P., and C. Portier.** 1986. Initiation, attenuation and RNase III processing of transcripts from the *Escherichia coli* operon encoding ribosomal protein S15 and polynucleotide phosphorylase. J. Mol. Biol. **187**:23–32.
31. **Robertson, H. D.** 1982. *Escherichia coli* ribonuclease III cleavage sites. Cell **30**:669–672.
32. **Russel, M., S. Ridd, and M. R. Kelly.** 1986. An improved filamentous helper phage for generating single stranded plasmid DNA. Gene **45**:333–338.
33. **Sanger, F., S. Nicklan, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
34. **Sato, Y., and H. Kuramitsu.** 1988. Sequence analysis of the *Streptococcus mutans scrB* gene. Infect. Immun. **56**:1956–1960.
35. **Scholle, R., S. Robb, F. Robb, and D. Woods.** 1989. Nucleotide sequence and analysis of the *Vibrio alginolyticus* sucrase gene (*scrB*). Gene **80**:49–56.
36. **Shinagawa, H., K. Makino, M. Amemrua, and A. Kakata.** 1987. Structure and function of the regulatory genes for the phosphate response in *Escherichia coli*, p. 20–25. *In* A. Torriani-Gorini, F. G. Rothman, S. Silver, A. Wright, and E. Yagil (ed.), Phosphate metabolism and cellular regulation in microorganisms. American Society for Microbiology, Washington, D.C.
37. **Shine, J., and L. Dalgarno.** 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71**:1342–1346.
38. **Somogyi, M. J.** 1952. Notes on sugar determination. J. Biol. Chem. **195**:19–23.
39. **Steinmetz, M., D. Le Coq, S. Aymerich, G. Gonzy-Treboul, and P. Gay.** 1985. The DNA sequence of the gene for the secreted *Bacillus subtilis* enzyme levansucrase and its genetic control sites. Mol. Gen. Genet. **200**:220–228.
40. **Steinmetz, M., D. Le Coq, H. B. Djemia, and P. Gay.** 1983. Analyse génétique de *sacB*, gène de structure d'une enzyme sécrétée, la levane-saccharase de *Bacillus subtilis*. Mol. Gen. Genet. **191**:138–144.
41. **Swings, J., and J. De Ley.** 1977. The biology of *Zymomonas*. Bacteriol. Rev. **41**:1–46.
42. **Taussing, R., and M. Carlson.** 1983. Nucleotide sequence of the yeast SUC2 gene for invertase. Nucleic Acids Res. **11**:1943–1954.
43. **Viikari, L.** 1988. Carbohydrate metabolism in *Zymomonas*. Crit. Rev. Biotechnol. **7**:237–261.
44. **Yanish-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.