INFECTION AND IMMUNITY, Aug. 1988, p. 1956–1960
0019-9567/88/081956-05$02.00/0
Copyright © 1988, American Society for Microbiology
Vol. 56, No. 8

Sequence Analysis of the *Streptococcus mutans scrB* Gene

YUTAKA SATO AND HOWARD K. KURAMITSU*

Department of Microbiology-Immunology, Northwestern University Medical-Dental Schools, 303 East Chicago Avenue, Chicago, Illinois 60611

Received 11 February 1988/Accepted 18 April 1988

The complete nucleotide sequence of the *Streptococcus mutans* GS-5 *scrB* gene coding for sucrose-6-phosphate hydrolase activity was determined. A potential ribosome-binding site as well as promoter sequences were identified upstream from the gene. The deduced amino acid sequence of the enzyme suggested a molecular weight of 51,750, which is similar to that estimated for the enzyme isolated from strain GS-5. The enzyme is slightly acidic, with a pI of 5.9, and is a relatively hydrophilic protein. The nucleotide and amino acid sequences of the enzyme showed significant homology with those of the *sacA* protein from *Bacillus subtilis*. In addition, a region of amino acid homology with the *S. mutans* fructosyltransferase and *B. subtilis* levansucrase proteins was also detected.

It is well recognized that *Streptococcus mutans* plays an important role in the etiology of human dental caries (12). This organism displays a wide variety of sucrose-metabolizing enzymes which appear to be responsible for the important role of dietary sucrose in cariogenesis. Among these are glucosyltransferases, which catalyze the formation of water-insoluble glucan implicated in colonization of the organisms; fructosyltransferases, which synthesize fructans, which may act as a reserve source of carbohydrate in dental plaque; invertases, which hydrolyze sucrose to glucose and fructose; and the sucrose transport systems (12).

Although much study has been concerned with the synthesis of extracellular glucans by mutans streptococci, most of the sucrose metabolized by this organism is degraded by cellular enzymes (25). At pH levels which normally occur in dental plaque, most of the sucrose is transported into the cells by means of a sucrose phosphotransferase system (23). In addition, evidence for a non-phosphotransferase sucrose transport system has also been obtained (21). Sucrose is converted by the former transport system into sucrose 6-phosphate, which is cleaved to glucose 6-phosphate and fructose by the sucrose-6-phosphate hydrolase (Suc-6-PH) characterized previously in this organism (4).

Several intracellular invertases have been purified from *S. mutans* (10, 24). Subsequent investigations have suggested that these enzymes are primarily concerned with sucrose 6-phosphate hydrolysis (4). More recently, the genes coding for these enzymes have been isolated in *Escherichia coli* (8, 13, 16). The Suc-6-PH expressed by these clones had a molecular size of approximately 58 kilodaltons (kDa) (8, 13, 16). However, the comparable enzyme purified from *S. mutans* strains had a somewhat smaller molecular size, approximately 48 kDa (10, 24). These results suggested different types of posttranslational modification of the enzyme in *E. coli* and *S. mutans*. To examine this possibility, we determined the nucleotide sequence of the *scrB* gene coding for Suc-6-PH activity in *S. mutans* GS-5. The present results suggest that the direct translational product of the gene has the molecular size previously predicted for the enzyme extracted from strain GS-5 cells (10).

* Corresponding author.

MATERIALS AND METHODS

Plasmids. Plasmid pMH613 containing the *scrB* gene has been previously described (8). A 2.0-kilobase (kb) *Pvu*II fragment from pMH613B2 was isolated and ligated to *Sma*I-*Hin*cII-digested pUC18 to produce a chimeric plasmid with a single *Bam*HI site. The ligation mixture was transformed into *E. coli* JM83, and transformants were selected on LB (Luria broth) agar plates containing ampicillin (40 μg/ml). Transformants containing the intact *scrB* gene were identified on MacConkey sucrose agar plates (8), and the fragment orientation was determined following *Bam*HI-*Hin*dIII digestion. The resultant plasmids, pPV5 and pPV7, which contained the *scrB* fragment oriented in both directions relative to the vector, were used to sequence the gene.

DNA manipulations. DNA isolation, endonuclease restriction, ligation, and transformation of competent *E. coli* cells were carried out as previously described (1).

Nucleotide sequencing. Nucleotide sequencing was done by the dideoxy chain termination method (17) with single-stranded M13mp18 and M13mp19 bacteriophage DNAs, the 20-mer universal sequencing primer (New England BioLabs, Inc., Beverly, Mass.), and [α-^{35}S]dATP (600 μCi/mmol; Amersham Corp., Arlington Heights, Ill.). DNA fragments to be sequenced were isolated on agarose gels (14) following restriction endonuclease digestion of pPV5 or pPV7. Each fragment was introduced into phage M13mp18 or M13mp19 (Bethesda Research Laboratories, Inc., Gaithersburg, Md.) or, alternatively, into phage M13tg130 (Amersham Corp.). However, the RNA polymerase sense strand containing the *scrB* promoter region was sequenced with double-stranded pPV7 DNA and the reverse sequencing primer (New England BioLabs). Both strands encompassing the *scrB* gene were entirely sequenced.

Sequence analysis. The nucleotide sequences were analyzed with the Pustell sequence analysis programs (International Biotechnologies, Inc., New Haven, Conn.).

Enzyme analysis. Preparation of cell extracts, assays for Suc-6-PH activity, electrophoretic analysis of the enzymes, and Western blot (immunoblot) analysis were carried out as previously described (8).

RESULTS

Sequencing strategy. Recent results from our laboratory (8) indicated that the *scrB* gene was expressed in plasmid

TABLE 1. Codon usage in the *scrB* gene

Codon	Amino acid	No. of residues	Codon	Amino acid	No. of residues	Codon	Amino acid	No. of residues	Codon	Amino acid	No. of residues
TTT	Phe	24	TCT	Ser	11	TAT	Tyr	23	TGT	Cys	2
TTC	Phe	3	TCC	Ser	2	TAC	Tyr	6	TGC	Cys	2
TTA	Leu	8	TCA	Ser	5	TAA		1	TGA		0
TTG	Leu	5	TCG	Ser	3	TAG		0	TGG	Trp	6
CTT	Leu	10	CCT	Pro	8	CAT	His	8	CGT	Arg	6
CTC	Leu	6	CCC	Pro	3	CAC	His	3	CGC	Arg	4
CTA	Leu	4	CCA	Pro	10	CAA	Gln	17	CGA	Arg	1
CTG	Leu	2	CCG	Pro	0	CAG	Gln	4	CGG	Arg	0
ATT	Ile	18	ACT	Thr	13	AAT	Asn	25	AGT	Ser	8
ATC	Ile	7	ACC	Thr	4	AAC	Asn	3	AGC	Ser	5
ATA	Ile	3	ACA	Thr	7	AAA	Lys	26	AGA	Arg	1
ATG	Met	4	ACG	Thr	1	AAG	Lys	5	AGG	Arg	0
GTT	Val	14	GCT	Ala	12	GAT	Asp	18	GGT	Gly	9
GTC	Val	9	GCC	Ala	3	GAC	Asp	9	GGC	Gly	6
GTA	Val	2	GCA	Ala	7	GAA	Glu	25	GGA	Gly	15
GTG	Val	1	GCG	Ala	0	GAG	Glu	6	GGG	Gly	2

pMH613. When this plasmid was introduced into *E. coli* HK730 (11), sucrase-positive transformants were identified, and one of these yielded plasmid pMH613B2. This plasmid, which was approximately 700 base pairs larger than pMH613, was subsequently shown to harbor an IS*1* element downstream from the *scrB* gene (unpublished results). Based on the restriction map of pMH613B2, it was possible to isolate the intact *scrB* gene on a 2.0-kb *Pvu*II fragment in plasmid vector pUC18, producing plasmids pPV5 and pPV7 (the *Pvu*II fragment oriented in both directions relative to the vector). Both plasmids expressed high sucrase activity in *E. coli*. The fortuitous presence of an IS*1* *Pvu*II site downstream from the *scrB* gene allowed convenient isolation of the intact gene.

Utilizing a detailed restriction map of the *scrB* gene, we isolated a series of overlapping fragments encompassing the entire gene. These fragments were isolated after appropriate restriction endonuclease digestion of plasmids followed by agarose gel electrophoresis and ligation into M13mp18 and M13mp19 bacteriophages for nucleotide sequencing. The 2.0-kb *Pvu*II insert was sequenced in its entirety on both strands.

It was not possible to isolate the RNA polymerase sense strand of the *Pvu*II-*Eco*RV fragment corresponding to the amino-terminal sequences of the *scrB* gene in the M13 bacteriophages. This result resembles recent results in our laboratory in that both strands of the fragments containing the promoter regions of the *S. mutans gtfB* (20) and *ftf* (19) genes also could not be isolated in these bacteriophages. Therefore, this region of the *scrB* gene was sequenced with double-stranded pPV7 DNA and the universal reverse sequencing primer.

Nucleotide sequence of the *scrB* gene. The sequencing data indicated that only one open reading frame capable of coding for the Suc-6-PH protein (4) could be identified on the *Pvu*II DNA fragment (Fig. 1). This open reading frame begins with an ATG initiating codon (position 232) and ends at the TAA termination codon (position 1594). It is preceded by a potential Shine-Dalgarno sequence, AGGAG, 10 base pairs upstream from the potential initiating codon. The region upstream from the ribosome-binding site is A+T rich, typical of the promoter regions of gram-positive bacteria (6, 9, 18). One potential promoterlike sequence, TAGAAA-N_{17}-TAGTAT (positions 171 to 199), could be identified in this region. Several other potential −10 and −35 sequences could also be recognized in this region. In addition, no sequences typical of termination sequences were identified downstream from the termination codon.

Codon usage in the *scrB* gene. As predicted from the relatively low G+C content of 36 to 38 mol% for *S. mutans* chromosomal DNA (5), the third base positions of the codons utilized for *scrB* gene expression were A+T rich (75%) (Table 1). A comparison of the codon utilization for individual amino acids indicated that the *scrB* gene was more similar to the *S. mutans* GS-5 *ftf* gene (19) than to the *gtfB* gene from the same organism (20).

Amino acid composition of the *scrB* gene product. The deduced amino acid sequence of the *scrB* gene indicated a molecular weight of 51,750 for the Suc-6-PH protein. This value is close to that estimated for the enzyme purified from strain GS-5, 48,000 (11), but lower than that estimated for the enzyme expressed in *E. coli*, 58,000 (8). The amino acid composition of the protein (Table 1) suggested that the protein was slightly acidic (58 acidic versus 43 basic amino acid residues), a result which was also indicated by the estimated pI of 5.9 for the protein. This value approximates that of pI 5.1 estimated for the enzyme purified from strain GS-5 (15). No signal peptide sequence was evident at the amino terminus of the protein, in keeping with the cytoplasmic location of the enzyme in strain GS-5 (15). Four cysteine residues were also found in the protein, but evidence for the existence of disulfide bonds in the mature protein has not yet been obtained.

A Kyte-Doolittle hydrophobicity plot (11a) of the enzyme (Fig. 2) indicated that the protein was relatively hydrophilic, with no extensive regions of hydrophobicity.

Homology of Suc-6-PH with other proteins. It was of interest to compare the nucleotide and amino acid sequences of the *scrB* gene with those of another protein exhibiting Suc-6-PH activity, the *Bacillis subtilis sacA* protein (7). A comparison of both amino acid and nucleotide sequences between the two genes (Fig. 3) revealed several regions of homology. Further examination of the regions of homology revealed one region of high amino acid homology (16 of 22 identical amino acids) (Fig. 4). Interestingly, nine of these amino acid sequences were also present in two other enzymes catalyzing the transfer of the fructose moiety of

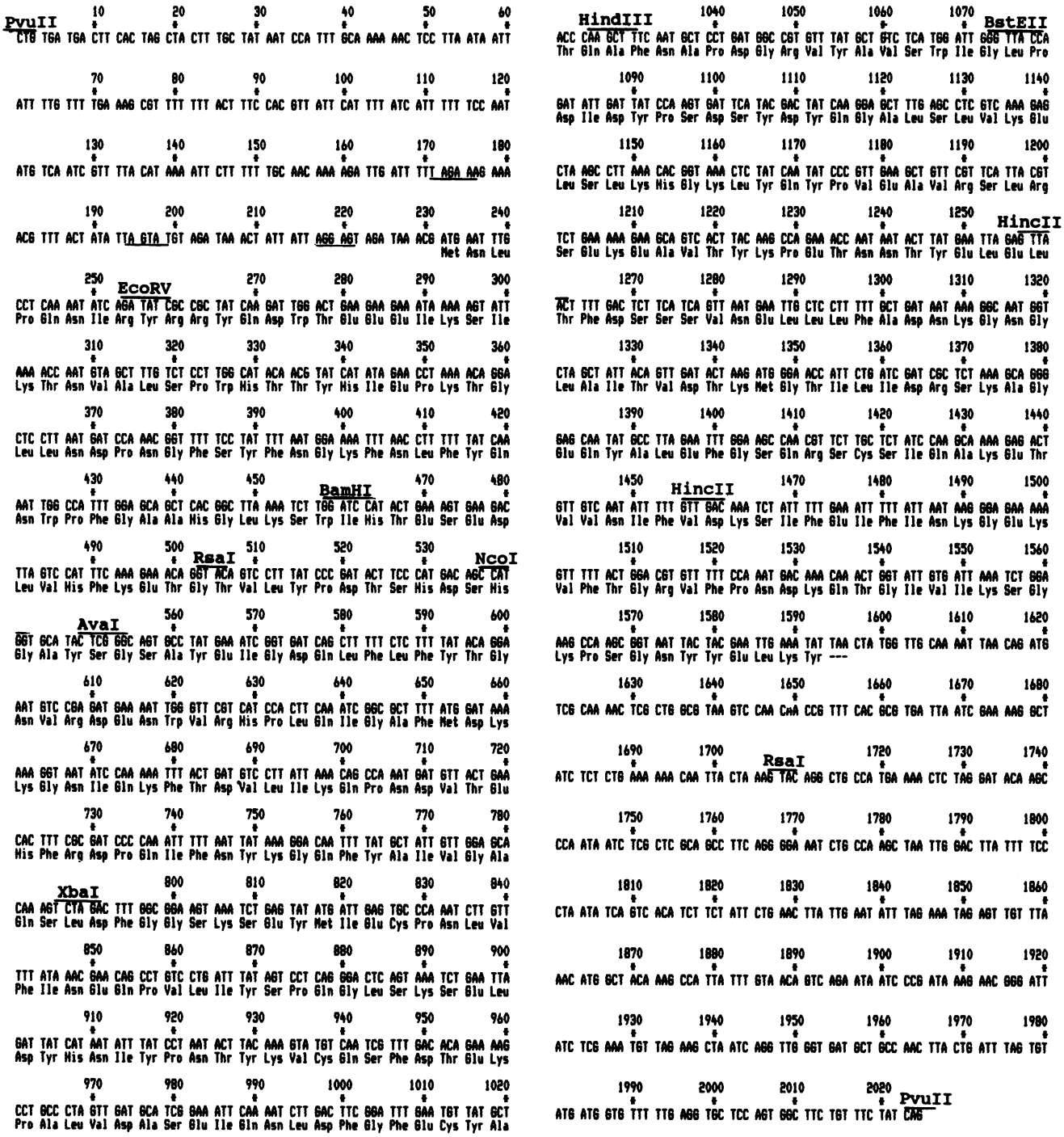

FIG. 1. Nucleotide sequence of the 2.0-kb *Pvu*II fragment and the deduced amino acid sequence. Positions 171 and 194, representing possible −35 and −10 promoter regions, and position 217, representing a putative ribosome-binding site, are underlined.

sucrose, the *S. mutans* GS-5 fructosyltransferase (19) and *B. subtilis* levansucrase (22) proteins.

DISCUSSION

The nucleotide sequence of the *scrB* gene suggested that the immediate translation product is a protein of 51.8 kDa. However, previous results indicated that the Suc-6-PH expressed by the cloned gene in *E. coli* was a protein of approximately 58 kDa (8). In the present investigation, the Suc-6-PH expressed in plasmids pPV5 and pPV7 was also 58 kDa (data not shown). However, the enzyme extracted from strain GS-5 is a 48-kDa protein (10). Since the molecular weight of the enzyme expressed in strain GS-5 was near that predicted from the sequence of the *scrB* gene, it is possible that the cloned gene product migrates at a slower-than-normal rate or that the enzyme undergoes unusual processing in *E. coli*. Furthermore, these data indicate that, contrary to previous suggestions (8, 13), no extensive posttranslational modification of Suc-6-PH occurs in *S. mutans*.

A comparison of the nucleotide and amino acid sequences

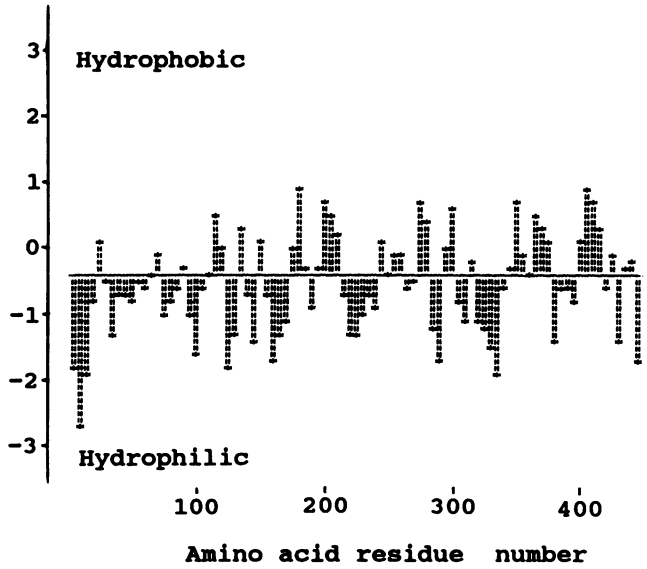

FIG. 2. Hydrophobicity plot of the *scrB* gene product.

of the *scrB* gene with the corresponding *sacA* sequences in *B. subtilis* revealed several regions of homology (Fig. 3). It was also of interest that an amino acid sequence of the *scrB* gene product shared homology with similar sequences in the *B. subtilis sacA* and *sacB* genes as well as the *ftf* gene from *S. mutans*. These regions, designated as the "sucrose box" (Fig. 4, boxed regions), may be important for enzymes catalyzing the transfer of fructose from sucrose. In addition, similar sequences could not be detected in the *S. mutans gtfB* gene coding for an enzyme which catalyzes the transfer of glucose from sucrose to either the enzyme or water (20). Therefore, these sequences may be regions subject to site-directed mutagenesis to identify the functional domains of the enzymes.

Previous results have suggested that the yeast *SUC2* gene (26) codes for the expression of both intracellular and extracellular invertase activities. Both proteins share the same reading frame in the gene but are initiated from different positions of the gene. In addition, more recent evidence (7) suggests that the *B. subtilis sacA* gene shares partial homology with the yeast *SUC2* gene. However, no sequence homology can be detected between the *scrB* gene and the yeast *SUCA* gene. In addition, an examination of the sequence of the former gene (or flanking regions) did not reveal the presence of an initiation codon which might allow for the expression of a protein larger than 51.8 kDa.

A hydrophobicity plot of the *scrB* gene product indicated that Suc-6-PH is a relatively hydrophilic protein (Fig. 2). This property is compatible with the cytoplasmic location of the enzyme in *S. mutans* (15). The previous observation (2) that antibody against purified Suc-6-PH reacts with the strain GS-5 membrane suggests the existence of a membrane protein sharing antigenic similarity with Suc-6-PH. Further investigation is required to examine this possibility.

The nucleotide sequence immediately upstream from the *scrB* gene indicates the presence of both a ribosome-binding site and putative promoter sequences (Fig. 1). This region was relatively A+T rich, and such sequences were also identified upstream from previously cloned *S. mutans* (19, 20) and other streptococcal (6, 9, 18) genes. This region contained several potential −10 and −35 sequences, but the actual transcription start site needs to be identified by mRNA characterization.

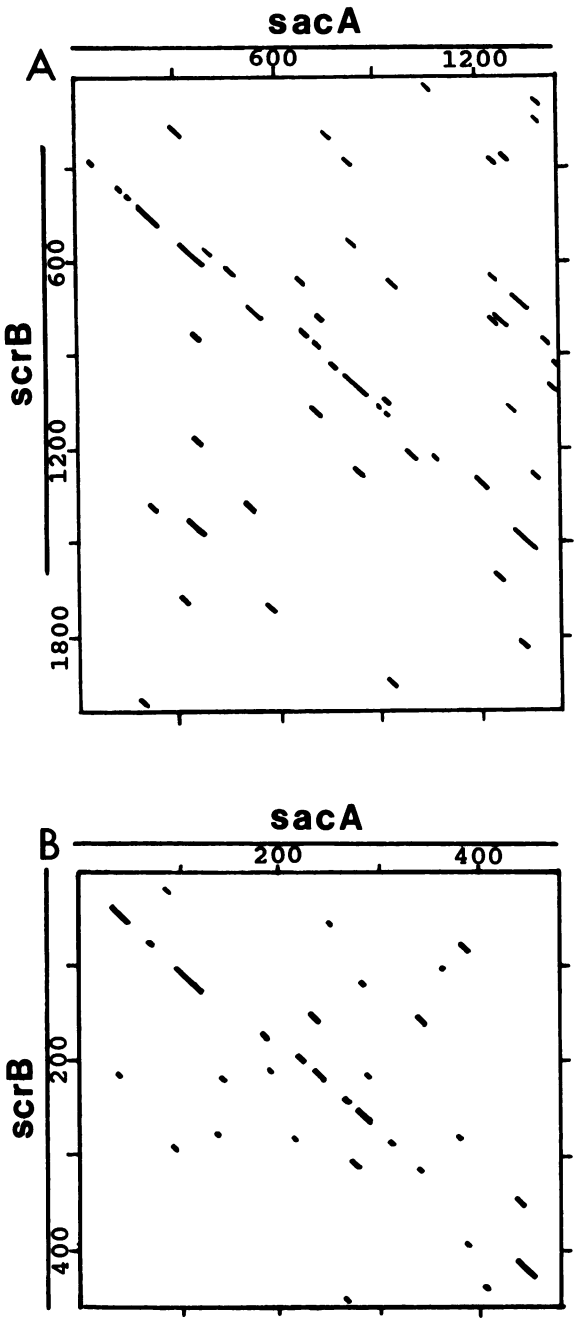

FIG. 3. Homology between the *scrB* and *sacA* genes. (A) Nucleotide sequence homology matrix. The 2.0-kb *Pvu*II fragment containing *scrB* and the protein-coding region of *sacA* were compared. Each point represents greater than 50% homology with a 41-nucleotide window. (B) Amino acid sequence homology matrix. Each point represents greater than 50% homology with a 7-amino-acid window.

The presence of a strong promoter immediately upstream from the *scrB* gene was suggested previously (13) and further supported by the inability to isolate the RNA polymerase sense strand from this region in the M13mp18 and M13mp19 bacteriophages. A similar situation was also encountered in the promoter regions of the *S. mutans gtfB* (20) and *ftf* (19) genes. In all three situations, high-level expression of the gene products was observed in *E. coli*. It is possible that the orientation of these promoters in the opposite direction

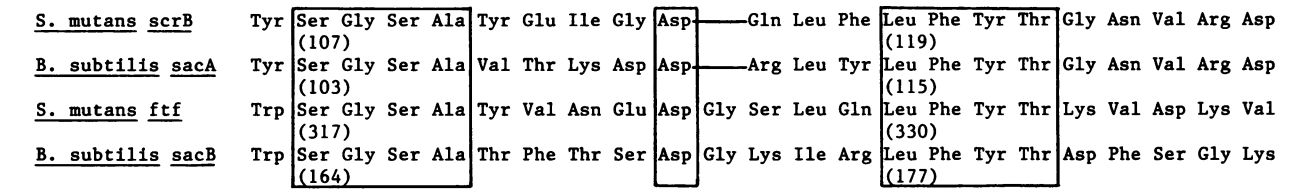

FIG. 4. Comparison of amino acid sequences among four gene products catalyzing the transfer of fructose from sucrose: *scrB* and *sacA*, Suc-6-PHs; *ftf*, fructosyltransferase; *sacB*, levansucrase. The homologous region of nine amino acids is boxed.

relative to M13 transcription is lethal to the single-stranded bacteriophage. The presence of a strong promoter upstream from the *scrB* gene was also suggested by the observations that the *Pvu*II fragment containing the gene could be cloned in both directions relative to plasmid vector pUC18 and that there were high levels of sucrase activity in both plasmid constructs.

Since the Suc-6-PH enzyme is involved in the transport and conversion of sucrose into a metabolizable form, it would not be surprising to find that the *scrB* gene is closely linked to other genes involved in the phosphoenolpyruvate-dependent sucrose phosphotransferase system of *S. mutans*. Furthermore, it has recently been observed that the comparable genes coding for lactose transport and hydrolysis in *Staphylococcus aureus* constitute an operonlike structure (3). Therefore, it will be of interest to determine the nucleotide sequences flanking the *scrB* gene to search for genes involved in the sucrose phosphotransferase system. Such an approach is currently under way in our laboratory.

ACKNOWLEDGMENTS

We thank Teruaki Shiroza for advice in the sequencing experiments.

This investigation was supported in part by Public Health Service grant DE-03258 from the National Institutes of Health.

LITERATURE CITED

1. **Aoki, H., T. Shiroza, M. Hayakawa, S. Sato, and H. K. Kuramitsu.** 1986. Cloning of a *Streptococcus mutans* glucosyltransferase gene coding for insoluble glucan synthesis. Infect. Immun. **53**:587–594.
2. **Bozzola, J. J., H. K. Kuramitsu, and M. T. Maynard.** 1981. Localization of *Streptococcus mutans* GS-5 glucosyltransferases and intracellular invertase. Infect. Immun. **32**:830–839.
3. **Breidt, F., Jr., W. Hengstenberg, U. Finkeldei, and G. C. Stewart.** 1987. Identification of the genes for the lactose-specific components of the phosphotransferase system in the *lac* operon of *Staphylococcus aureus*. J. Biol. Chem. **262**:16444–16449.
4. **Chassy, B. M., and E. V. Porter.** 1979. Initial characterization of sucrose-6-phosphate hydrolase from *Streptococcus mutans* and its apparent identity with intracellular invertase. Biochem. Biophys. Res. Commun. **89**:307–314.
5. **Coykendall, A. L.** 1974. Four types of *Streptococcus mutans* based on their genetic, antigenic, and biochemical characteristics. J. Gen. Microbiol. **83**:327–338.
6. **Fahnestock, S. R., P. Alexander, J. Nagle, and D. Filpula.** 1986. Gene for an immunoglobulin-binding protein from a group G streptococcus. J. Bacteriol. **167**:870–880.
7. **Fouet, A., A. Klier, and G. Rapoport.** 1986. Nucleotide sequence of the sucrase gene of *Bacillus subtilis*. Gene **45**:221–225.
8. **Hayakawa, M., H. Aoki, and H. K. Kuramitsu.** 1986. Isolation and characterization of the sucrose-6-phosphate hydrolase gene from *Streptococcus mutans*. Infect. Immun. **53**:582–586.
9. **Hollingshead, S. K., V. A. Fischetti, and J. R. Scott.** 1986. Complete nucleotide sequence of type 6 M protein of the group A *Streptococcus*. J. Biol. Chem. **261**:1677–1686.
10. **Kuramitsu, H. K.** 1973. Characterization of invertase activity from cariogenic *Streptococcus mutans*. J. Bacteriol. **115**:1003–1010.
11. **Kuramitsu, H. K.** 1987. Utilization of a mini-mu transposon to construct defined mutants in *Streptococcus mutans*. Mol. Microbiol. **1**:229–231.

11a. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.

12. **Loesche, W. J.** 1986. Role of *Streptococcus mutans* in human dental decay. Microbiol. Rev. **50**:353–380.
13. **Lunsford, R. D., and F. L. Macrina.** 1986. Molecular cloning and characterization of *scrB*, the structural gene for the *Streptococcus mutans* phosphoenolpyruvate-dependent sucrose phosphotransferase system sucrose-6-phosphate hydrolase. J. Bacteriol. **166**:426–434.
14. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
15. **Maynard, M. T., and H. K. Kuramitsu.** 1979. Purification and antigenic properties of intracellular invertase from *Streptococcus mutans*. Infect. Immun. **23**:873–883.
16. **Russell, R. R. B., P. Morrissey, and G. Dougan.** 1985. Cloning of sucrase genes from *Streptococcus mutans* in bacteriophage lambda. FEMS Microbiol. Lett. **30**:37–41.
17. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
18. **Shaw, J. H., and D. B. Clewell.** 1985. Complete nucleotide sequence of macrolide-lincosamide-streptogramin B-resistant transposon Tn*917* in *Streptococcus faecalis*. J. Bacteriol. **164**:782–796.
19. **Shiroza, T., and H. K. Kuramitsu.** 1988. Sequence analysis of the *Streptococcus mutans* fructosyltransferase gene and flanking regions. J. Bacteriol. **170**:810–816.
20. **Shiroza, T., S. Ueda, and H. K. Kuramitsu.** 1987. Sequence analysis of the *gtfB* gene from *Streptococcus mutans*. J. Bacteriol. **169**:4263–4270.
21. **Slee, A. M., and J. M. Tanzer.** 1982. Sucrose transport by *Streptococcus mutans*. Evidence for multiple transport systems. Biochim. Biophys. Acta **692**:415–424.
22. **Steinmetz, M., D. LeCoq, S. Aymerich, H. 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.
23. **St. Martin, E., and C. L. Wittenberger.** 1979. Characterization of a phosphoenolpyruvate-dependent sucrose phosphotransferase system in *Streptococcus mutans*. Infect. Immun. **24**:865–868.
24. **Tanzer, J. M., A. T. Brown, and M. F. McInerney.** 1973. Identification, preliminary characterization, and evidence for regulation of invertase in *Streptococcus mutans*. J. Bacteriol. **116**:192–202.
25. **Tanzer, J. M., B. M. Chassy, and M. I. Krichevsky.** 1972. Sucrose metabolism by *Streptococcus mutans* SL-1. Biochim. Biophys. Acta **261**:379–387.
26. **Taussig, R., and M. Carlson.** 1983. Nucleotide sequence of the yeast *SUC2* gene for invertase. Nucleic Acids Res. **11**:1943–1954.