# 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-6phosphate 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).

## MATERIALS AND METHODS

**Plasmids.** Plasmid pMH613 containing the *scrB* gene has been previously described (8). A 2.0-kilobase (kb) *PvuII* fragment from pMH613B2 was isolated and ligated to *SmaI-Hinc*II-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  $\mu$ g/ml). Transformants containing the intact *scrB* gene were identified on MacConkey sucrose agar plates (8), and the fragment orientation was determined following *Bam*HI-*Hind*III 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 singlestranded M13mp18 and M13mp19 bacteriophage DNAs, the 20-mer universal sequencing primer (New England Bio-Labs, Inc., Beverly, Mass.), and  $[\alpha^{-3^5}S]dATP$  (600  $\mu$ Ci/ 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

<sup>\*</sup> Corresponding author.

Codon	Amino acid	No. of residues	Codon	Amino acid	No. of residues	Codon	Amino acid	No. of residues	Codon	Amino acid	No. of residues
ттт	Phe	24	ТСТ	Ser	11	TAT	Tyr	23	TGT	Cys	2
TTC	Phe	3	тсс	Ser	2	TAC	Tyr	6	TGC	Cys	2
TTA	Leu	8	TCA	Ser	5		•	1	TGA		0
TTG	Leu	5	TCG	Ser	3	TAG		0	TGG	Тгр	6
CTT	Leu	10	ССТ	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	ī	GCG	Ala	0	GAG	Glu	6	GGG	Gly	2

TABLE 1. Codon usage in the scrB gene

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*I* 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 *PvuII* fragment in plasmid vector pUC18, producing plasmids pPV5 and pPV7 (the *PvuII* fragment oriented in both directions relative to the vector). Both plasmids expressed high sucrase activity in *E. coli*. The fortuitous presence of an IS*I PvuII* 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 PvuII insert was sequenced in its entirety on both strands.

It was not possible to isolate the RNA polymerase sense strand of the PvuII-EcoRV 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 PvuIIDNA 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<sub>17</sub>-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

INFECT. IMMUN.

1958	SATO	AND	KUR.	AN	ИT	SI	J

2 <u>vu</u> C19	II Tga	T6/	10 • •	CAC	TAG	20 CTA	CTT	TGC	30 TAT	) Aat	CCA	TTT	40 # 6ca	AAA	AAC	50 TCC	TTA	ATA	60 ATT	ACC Thr	H CĀ 61	ind A 601 A Ala	III TTC Phe	AAT Asn	GCT Ala	1040 CCT Pro	GAT Asp	66C 61y	1050 E CGT GTT Arg Val	TAT Tyr	106 SCT S Ala V	) TC TC al Se	A 166 r Trp	1070 ATT Ile	Вз 666 61 у	tE] TTA Leu	II CCA Pro
			70			80			90	)			100			110			120			1	090			1100			1110		112	3		1130		1	140
ATT	TTG	ITI	TGA	AAG	CGT	ŧΪΤ	TTT	ACT	TTČ	CAC	6TT	ATT	ĊAT	TTT	ATC	AŤT	TTT	TCC	AAT	GA1 Asp	AT 11	t GA1 e Asp	TAT Tyr	CCA Pro	AGT Ser	GAT Asp	TCA Ser	TAC Tyr	GAC TAT Asp Tyr	CAA Gln	GBA B Siy A	CT TT la Lei	6 A60 u Ser	CTC Leu	GTC   Val	AAA ( Lys (	686 61u
			130 #			140 #			15(	)			160			170			180			1	150			1160			1170 ±		119	)		1190		13	200
ATG	TCA	ATC	: <b>6</b> tt	TTA	CAT	AAA	ATT	CTT	ш	TGC	AAC	AAA	aga	TTG	ATT	ш	AGA	<u>AA</u> G	AAA	CTA Leu	AB	ic CT1 Ir Leu	AAA Lys	CAC His	66T 61y	AAA Lys	CTC Leu	TAT Tyr	CAA TAT Gln Tyr	Pro	GTT G Val G	MA 6C Lu Al	T GT1 a Val	CGT Arg	TCA Ser [	TTA ( Leu (	CGT Arg
			190 #			200			210	)		ł	220 *			230			240			1	210			1220			1230		124	)		1250	H	lin	cII
ACG	Ш	ACT	ATA	TTA	GTA	101	aga	TAA	ACT	ATT	ATT	<u>A66</u>	AGT	aga	TAA	ACS	ATG Net	AAT Asn	TT6 Leu	TC1 Ser	6A 61	u Lys	644 61u	6CA Ala	GTC Val	ACT Thr	TAC Tyr	AAG Lys	CCA GAN Pro Glu	ACC Thr	AAT A Asn A	AT AC' sn Thi	T TAI r Tyr	61u	TTA ( Leu (	SAG Glu (	TTA Leu
			250 +	Ec	OR	v			27( *	)		ł	280 #			290			300 *	_		1	270			1280			1290		130	)		1310		1	320
CCT Pro	CAA Gln	AAT Asn	ATC	AGA Arg	TAT Tyr	ČGC Arg	CGC Arg	TAT Tyr	CAA 61n	GAT Asp	166 Trp	ACT Thr	6AA 6lu	688 61 u	684 61 u	ATA Ile	AAA Lys	AGT Ser	ATT Ile	ACT Thr	Ph	t GAC e Asp	TCT Ser	TCA Ser	TCA Ser	6TT Val	AAT Asn	GAA Glu	TT& CTO Leu Leu	CTT Leu	TTT G Phe A	la Asi	t AA1 p Asi	AAA Lys	66C ( 61y (	AAT ( Asn (	86T 61 y
			310			320			33( *	)		:	340 #			350			360			1	330			1340			1350		136	)		1370		1	380
AAA Lys	ACC Thr	AAT Asn	6TA Val	GCT Ala	TTG Leu	TCT Ser	CCT Pro	TGG Trp	CAT His	ACA Thr	AC6 Thr	TAT Tyr	CAT His	ATA Ile	6AA 61u	CCT Pro	AAA Lys	ACA Thr	66A 61y	CTA Leu	GC Al	a lle	ACA Thr	STT Val	GAT Asp	ACT Thr	AAG Lys	AT6 Net	66A ACC 61y Thr	ATT Ile	CTS A Leu I	ic ga le Asi	T CGC p Arg	TCT Ser	AAA (	SCA   Ala	866 61 y
			370 •			380			390	)		4	100			410			420			1	390		:	1400			1410		142	)		1430		1 <sup>/</sup>	440
CTC Leu	CTT Leu	AAT Asn	GAT Asp	CCA Pro	AAC Asn	66T 61y	TTT Phe	TCC Ser	TAT Tyr	TTT Phe	AAT Asn	66A 61y	AAA Lys	TTT Phe	AAC Asn	CTT Leu	TTT Phe	TAT Tyr	CAA Gin	6A6 61u	CA 61	A TAT n Tyr	GCC Ala	TTA Leu	688 61u	TTT Phe	61y	AGC Ser	CAA CO1 Gln Arg	TCT Ser	TGC TO Cys Se	IT ATI	C CAA e Glm	GCA Ala	AAA I Lys I	686 ( 51u	ACT Thr
			430			440			450	)		Ba	mH	I		470			480			1	450	H	lin		[		1470		148	)		1490		ľ	500
AAT Asn	TGG Trp	CCA Pro	TTT	66A 61 y	GCA Ala	GĈT Ala	CAC His	66C 61 y	TTĂ Leu	AAA Lys	TCT Ser	T86 Trp	ATC 11e	CAT His	ACT Thr	60A 61u	AGT Ser	688 61 u	GAČ Asp	GTT Val	8T Va	C AAT 1 Asn	ATT Ile	TTT Phe	GTT Val	GAC Asp	AAA Lys :	TCT Ser	ATT TTT Ile Phe	684 61u	ATT TI Ile Pl	IT AT	t AA1 e Ast	AAG Lys	66A ( 61y (	6AA / 61u /	AAA Lys
			490			500	Rs	aT	510	)		1	520			530		N	COT			1	510		1	1520			1530		154	)		1550		Ľ	560
TTA Leu	GTC Val	CAT His	TTC Phe	AAA Lys	SAA Glu	ACA Thr	66T 61y	ACA Thr	GTĈ Val	CTT Leu	TAT Tyr	CCC Pro	GAT Asp	ACT Thr	TCC Ser	CAT His	GAC Asp	AGC Ser	CAT His	GTT Val	TT Ph	T ACT e Thr	66A 61y	CGT Arg	GTT Val	TTT Phe	CCA Pro	AAT Asn	GAC AAA Asp Lys	CAA 61n	ACT G	IT AT	r ere e Val	ATT	AAA I	TCT   Ser	66Å 61y
		,	va	I		560			570	)		:	580			590			600			1	570		:	1580			1590		160	)		1610		1	620
GGT Gly	GCA Ala	TAČ Tyr	TC6 Ser	66C 61y	AGT Ser	9ČC Ala	TAT Tyr	6AA Glu	ATC 11e	66T 61y	GAT Asp	CAG Gln	CTT Leu	TTT Phe	CTC Leu	TTT Phe	TAT Tyr	ACA Thr	66A 61y	AAG Lys	CC Pr	A AGC o Ser	66T 61y	AAT Asn	TAC Tyr	TAC Tyr	GAA Glu i	TTG Leu	AAA TAT Lys Tyr	TAA	CTA T	<b>16 TT</b>	6 CAA	AAT	TAA	CAG (	ATS
			610			650			630	)		6	540			650			660			1	630		1	1640			1650		166	)		1670		1	680
AAT Asn	GTC Val	CGA Arg	SAT Asp	688 61 u	AAT Asn	TGG Trp	6TT Val	CGT Arg	CAT His	CCA Pto	CTT Leu	CAA 61 n	ĂTC 11e	86C 61 y	OCT Ala	TTT Phe	AT6 Net	GAT Asp	AAA Lys	TCG	CA	a aac	ŤCG	CT6	606	TĂA	GTC (	CAA	CHĂ CCE	TTT	CAC 6	:8 T6/	A TTA	ATC	GAA /	MAG #	6CT
			670 •			680			690	)			700			710			720			1	690		1	1700		Re	La T		1720	)		1730		1	740
AAA Lys	66T 61 y	AAT Asn	ATC Ile	CAA Gin	AAA Lys	TTT Phe	ACT Thr	GAT Asp	6TČ Val	CTT Leu	ATT Ile	AAA Lys	CAG Gln	ССА Рте	AAT Asn	GÄT Asp	6TT Val	ACT Thr	644 61u	ATC	TC	T CTO	ÅAA	AAA	CAA	TTA	CTA (	AAG	TAC AGE	CTE	CCA TE	ja aai	A CTC	TÂG	GAT (	nca (	AGC
			730 *			740			750	)			760			770			780			1	750		1	1760			1770		178	)		1790		1	B00
CAC His	TTT Phe	CGC Arg	GAT Asp	CCC Pro	CAA Gln	ATT Ile	TTT Phe	AAT Asn	TAT Tyr	AAA Lys	66A 61 y	CAA 61n	TTT Phe	TAT Tyr	GCT Ala	ATT Ile	6TT Val	66A 61y	GCĂ Ala	CCA	AT	A ATC	TCG	CTC	6CA	339	TTC	A66	66A AAT	CT6	CCA A	)c tai	A TTE	GAC	TTA	TTT 1	TCC
	,	Cba	I			800			810	)		(	950			830			840			1	810		1	1820			1830		184(	)		1850		1	860
CAA G1n	AGT Ser	CTA	GAC Asp	TTT Phe	66C 61y	66A 61y	AGT Ser	AAA Lys	TCT Ser	686 61u	TAT Tyr	ATS Net	ATT 11e	6A6 61 u	TOC Cys	CCA Pro	AAT Asn	CTT Leu	GTT Val	CTA	AT	A TCA	<b>G</b> TC	ACA	TCT	TOT	ATT (	CTG	NAC TTA	TTG	AAT AT	it taf	s aaa	TAG	AGT	ret 1	TTA
			850			860			870	)		(	380			890			900			1	870		1	1880			1890		190	)		1910		1	920
TTT Phe	ATA Ile	AAC Asn	6AA 61u	CAG 61n	CCT Pro	GTC Val	CTS Leu	ATT Ile	TAT Tyr	AGT Ser	CCT Pro	CAG Gln	66A 61y	CTC Leu	AGT Ser	AĂA Lys	TCT Ser	688 61 u	TTA Leu	AAC	AT	6 <b>6</b> CT	ACA	<b>AA</b> 6	CCA	TTA	TTT (	GTA	ACĂ STC	AGA	ATA AT	ic cci	6 ATA	AAG	AAC (	<b>366</b> (	ATT
			910			920			930	)		9	740			950			960			1	930		1	1940			1950		196	)		1970		1	980
GAT Asp	TAT Tyr	CAT His	AAT Asn	ATT Ile	TAT Tyr	CCT Pro	AAT Asn	ACT Thr	TAC Tyr	AAA Lys	6TA Val	TØT Cys	CAA Gln	TCS Ser	TTT Phe	GAC Asp	ACA Thr	688 61 u	AAG Lys	ATC	TC	G AAA	Tet	TAG	aag	CTA	ATC (	A66	TT8 66T	GAT	ect e	ic aai	C TTA	cte	ATT ?	TAG 1	TGT
			970			<b>98</b> 0			990	)		10	000		1	010			1020			1	990		i	2000			2010		202	) _		-			
CCT Pro	OCC Ala	CTA Leu	GTT Val	GAT Asp	GCA Ala	TCS Ser	888 61u	ATT Ile	CAA Gln	AAT Asn	CTT Leu	GAC Asp	TTC Phe	66A 61y	TTT Phe	6AA Glu	TØT Cys	TAT Tyr	GCT Ala	ATG	AT	6 6TB	m	TTB	<b>AG</b> 6	TEC	TCC (	AGT	<b>66C</b> TTC	TOT	TTC T	IT CM		L			

FIG. 1. Nucleotide sequence of the 2.0-kb PvuII 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-thannormal 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



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 sitedirected 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 PvuII fragment containing *scrB* and the protein-coding region of *sacA* were compared. Each point represents greater than 50% homology with a 41nucleotide window. (B) Amino acid sequence homology matrix. Each point represents greater than 50% homology with a 7-aminoacid 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

S. mutans scrB	Tyr Ser Gly Ser Ala	Tyr Glu Ile Gly As	sp-Gln Leu Phe	Leu Phe Tyr Thr	Gly Asn Val Arg Asp
	(107)			(119)	
<u>B. subtilis</u> <u>sacA</u>	Tyr Ser Gly Ser Ala	Val Thr Lys Asp As	spArg Leu Tyr	Leu Phe Tyr Thr	Gly Asn Val Arg Asp
	(103)			(115)	T
S. mutans ftf	Trp Ser Gly Ser Ala	Tyr Val Asn Glu As	spicity Ser Leu Gin	Leu Phe Tyr Thr	Lys Val Asp Lys Val
P subtilis seeP		The Dho The Con As		(330) Tou Pho Tur Thr	Aco Pho Sor Cly Lyc
B. SUDCIIIS SACE	(164)	Int the Int Set As	sploty Lys the Arg	(177)	Asp file Set Giy Lys
	(164)			(177)	

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 *PvuII* 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 phosphoenolpyruvatedependent 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

- 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.
- 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.
- 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.
- Coykendall, A. L. 1974. Four types of *Streptococcus mutans* based on their genetic, antigenic, and biochemical characteristics. J. Gen. Microbiol. 83:327-338.
- 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.
- 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.

- 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.
- 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 Tn917 in *Streptococcus faecalis*. J. Bacteriol. 164: 782-796.
- Shiroza, T., and H. K. Kuramitsu. 1988. Sequence analysis of the *Streptococcus mutans* fructosyltransferase gene and flanking regions. J. Bacteriol. 170:810–816.
- Shiroza, T., S. Ueda, and H. K. Kuramitsu. 1987. Sequence analysis of the *gtfB* gene from *Streptococcus mutans*. J. Bacteriol. 169:4263-4270.
- Slee, A. M., and J. M. Tanzer. 1982. Sucrose transport by *Streptococcus mutans*. Evidence for multiple transport systems. Biochim. Biophys. Acta 692:415–424.
- 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.
- St. Martin, E., and C. L. Wittenberger. 1979. Characterization of a phosphoenolpyruvate-dependent sucrose phosphotransferase system in *Streptococcus mutans*. Infect. Immun. 24:865– 868.
- 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.
- Tanzer, J. M., B. M. Chassy, and M. I. Krichevsky. 1972. Sucrose metabolism by *Streptococcus mutans* SL-1. Biochim. Biophys. Acta 261:379–387.
- Taussig, R., and M. Carlson. 1983. Nucleotide sequence of the yeast SUC2 gene for invertase. Nucleic Acids Res. 11:1943– 1954.