# A 40-Kilodalton Cell Wall Protein-Coding Sequence Upstream of the sr Gene of Streptococcus mutans OMZ175 (Serotype f)

JOËLLE A. OGIER,\* MARIE SCHÖLLER, YVES LEPOIVRE, SOPHIE GANGLOFF, RIDHA M'ZOUGHI, AND JEAN-PAUL KLEIN

Institut National de la Santé et de la Recherche Médicale Unité 157, Faculté de Chirurgie Dentaire, 1 Place de l'Hôpital, Strasbourg 67000, France

Received 9 October 1990/Accepted 4 February 1991

Streptococcus mutans surface proteins may be important in immunization against dental caries. We report the existence of an open reading frame of 1,005 bp that lies 1,162 bases upstream of the S. mutans OMZ175 sr gene and that encodes a cell wall-associated protein. This open reading frame codes for 335 amino acid residues. The first 18-amino acid region is predominantly hydrophobic and resembles a signal peptide, and the hydrophobic C-terminal region may function as an anchor to the bacterial cell wall. On the basis of the predicted antigenic determinants of the deduced amino acid sequence, a 16-residue synthetic peptide corresponding to the middle hydrophilic coiled region was synthesized. Antibodies raised against this synthetic peptide reacted with a protein with an apparent  $M_r$  of 40,000 that was identified by Western immunoblotting in a cell wall extract from S. mutans OMZ175. The high reactivity in an enzyme-linked immunosorbent assay of the antibodies with whole S. mutans OMZ175 cells showed that this protein was located on the bacterial cell surface. Furthermore, the antipeptide immunoglobulin G recognized an identical determinant on the cell surface of other members of the S. mutans group. However, the function of this protein is not yet known.

Streptococcus mutans, the primary etiologic agent of the initiation of dental caries, possesses a number of cell surface and extracellular proteins which are implicated in the attachment of these bacteria to salivary glycoproteins in the tooth pellicle (15). These proteins are highly immunogenic and are protective in immunization studies (2). The virulence of S. mutans has been investigated by a genetic approach, and several genes specifying cell surface and extracellular proteins have been cloned. Extracellular proteins recently characterized are (i) glucosyltransferases (31), fructosyltransferases (27), and a glucan-binding protein (26), which are implicated in sucrose-dependent adherence; (ii) antigen I/II (12), a protein of  $M_r$  185,000 present in all members of the S. mutans group except S. rattus and involved in sucroseindependent adherence; and (iii) antigen A (5), a protein of  $M_r$  29,000 found in S. mutans and S. rattus and the role of which is not yet known. Animal studies have demonstrated that protection against dental caries could be obtained after immunization with purified antigen I/II or related proteins (13) and with antigen A (25). However, some side effects, such as the induction of heart-reactive antibodies (9), the production of rheumatoid factors (24), and antigen mimicry with human immunoglobulin G (IgG) (33), have hampered the development of a vaccine against S. mutans, although it was recently shown that the induction of heart-reactive antibodies was not due to antigenic similarity between antigen I/II and components of the human heart (34).

Therefore, the characterization of other S. mutans cell surface proteins is of particular importance in the development of a safe vaccine. Previously, we reported the cloning (19) and nucleotide sequence (20) of the sr gene of S. mutans OMZ175 (serotype f), which encodes the SR protein, an antigen I/II-related protein (12) implicated in the interactions of S. mutans with salivary glycoproteins. The present communication describes an open reading frame (ORF) lying

1,162 bases upstream of the sr gene. The results presented here indicate that this ORF encodes a cell wall protein of  $M_r$ 40,000 containing an antigenic determinant present in members of the S. mutans group. However, the role of this cell wall protein in S. mutans virulence is still not known.

### MATERIALS AND METHODS

**Bacterial strains and plasmids.** S. mutans OMZ70 (serotype c), B-2 (serotype e), and OMZ175 (serotype f); S. cricetus E49 (serotype a); S. rattus OMZ51 (serotype b); S. sobrinus OMZ176 (serotype d) and 6715 (serotype g); and S. downei MFe 28 (serotype h) have been described elsewhere (1, 28). The strains were grown in brain heart infusion broth (Difco Laboratories, Detroit, Mich.). Escherichia coli NM522 (7) was used as the host for plasmids pTZ18R and pTZ19R (Pharmacia LKB, Uppsala, Sweden). Recombinant plasmid pHBsr-1 has been described elsewhere (19).

Antigens and CE. Cell wall extracted antigens (WEA) and extracellular soluble antigens (ESA) were prepared from exponential-phase cultures of *S. mutans* OMZ175 as previously described (28). Cellular extracts (CE) were prepared from *E. coli* HB101 cells transformed by plasmid pHBsr-1, and recombinant SR protein was prepared from CE as described by Ogier et al. (19).

DNA cloning and analysis procedures. Routine DNA manipulation procedures such as miniscale plasmid preparation, restriction endonuclease analysis, purification of DNA fragments from agarose gels, and subcloning were performed as described by Maniatis et al. (16). Restriction endonucleases and T4 DNA ligase were obtained from Appligene (Strasbourg, France).

**DNA sequencing and computer analysis.** Restriction fragments covering the 5'-flanking region of the *sr* gene were generated from pHBsr-1 and subcloned in plasmids pTZ18R and pTZ19R. Single-stranded templates were produced from the pTZ plasmids in accordance with the instructions of the manufacturer. Series of deletions extending from the *Bam*HI

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



FIG. 1. Restriction map of the sr gene and its 5'-flanking region. The thick black bars indicate the ORFs. The open box represents the probe used in the hybridization experiment. Restriction sites: B, BamHI; E, EcoRI; H, HindIII; P, PstI; S, SacI; X, XbaI.

and *PstI* restriction sites, respectively localized 2,983 bases and 1,692 bases upstream from the start of the *sr* gene, were produced by digestion with exonuclease III (8). The nucleotide sequence was determined on both strands by the dideoxy chain termination method with T7 DNA polymerase (Pharmacia), [<sup>35</sup>S]dATP (>37 TBq/mmol; Amersham, Les Ulis, France), and the universal reverse primer or appropriate specific synthetic oligonucleotide primers (Appligene). The software package PC Gene (IntelliGenetics, Geneva, Switzerland) was used for the analysis of the sequence data and the prediction of the secondary structure of the protein and epitopes. The hydrophobicity and amphiphilicity properties of the deduced amino acid sequence were analyzed by the method of Kyte and Doolittle (10).

Southern blotting and hybridization procedure. Streptococcal DNA was recovered by the protocol of Robeson et al. (22). Each genomic DNA (20 µg) was BamHI-PstI digested to completion, and fragments were electrophoresed on an 0.8% agarose gel. The Southern blotting protocol (29) was slightly modified: DNA was transferred under vacuum onto nylon membranes (Amersham) and bound to the nylon under UV light at 254 nm. The BamHI-PstI restriction fragment covering the 5' part of ORF 2 (see below) was used to produce a probe labeled by nick translation with biotin-7dATP (BRL, Cergy-Pontoise, France) in accordance with the instructions of the manufacturer. Hybridizations were performed under standard conditions (16) in the following solution: 6× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)-5× Denhardt's reagent-0.5% sodium dodecyl sulfate (SDS)-100 µg of denatured salmon sperm DNA per ml. Detection was done with a streptavidin-alkaline phosphatase (AP) conjugate (BlueGene system; BRL).

**Peptide synthesis and conjugation to carrier protein.** On the basis of the predicted antigenic determinants of the deduced amino acid sequence, a 16-residue peptide (SNPQVNVETD SSEKDE) corresponding to the middle region (amino acids 167 to 182) of the protein was synthesized (Neosystem, Strasbourg, France) by a solid-phase method (17). The composition of the peptide was confirmed by amino acid analysis, and the purity (>85.9%) was checked by high-pressure liquid chromatography. Tyrosine was added to the carboxy terminus of the peptide to couple the peptide to ovalbumin through bisdiazobenzidine. The coupling ratio reached 12 mol of peptide per mol of ovalbumin. The peptide was also conjugated to bovine serum albumin (BSA) by glutaraldehyde, with a coupling ratio of 8 mol of peptide per mol of BSA.

Immunization and preparation of antisera. Antipeptide serum was raised in two rabbits by subcutaneous injections of equal mixtures of peptide conjugated to ovalbumin (100  $\mu$ g of peptide in phosphate-buffered saline [PBS]) and Freund incomplete adjuvant at 15-day intervals. At least 400  $\mu$ g of peptide was injected, and blood was collected 10 days after the last injection. Anti-SR protein serum was prepared as previously described (33). The IgG fraction of each antiserum was purified by chromatography on DEAE-Trisacryl M essentially by the method of Corthier et al. (3). Control preimmune sera were subjected to the same purification procedure. AP-antirabbit IgG was from Miles (Puteaux, France).

ELISA. The antibody recognition of the expressed protein was checked by an enzyme-linked immunosorbent assay (ELISA) as previously described (18, 33). In brief, microtiter plates (Nunc, Roskilde, Denmark) were coated with 50  $\mu$ l of *S. mutans* whole cells (10<sup>9</sup> bacteria per ml), WEA, ESA, CE (10  $\mu$ g/ml), or BSA-peptide (1  $\mu$ g/ml). After the remaining binding sites were blocked with 0.5% gelatin in PBS containing 0.05% Tween 20, the plates were incubated with serial dilutions of antipeptide IgG (1 h, 37°C). Antibody binding was detected with AP-antirabbit IgG and by incubation with a substrate. The  $A_{405}$  was read with a Dynatech MR 5000 spectrophotometer.

**Electrophoresis and immunoblotting.** Proteins were separated by SDS-polyacrylamide gel electrophoresis (PAGE) on 15% gels with the discontinuous buffer system of Laemmli (11). Proteins were detected by staining with Coomassie blue. In Western immunoblotting experiments, proteins were transferred from SDS-PAGE gels to nitrocellulose sheets by the procedure of Towbin et al. (30) and detected by incubation with antipeptide IgG and with AP-antirabbit IgG and an enzyme substrate.

### RESULTS

Nucleotide sequence and analysis. We previously described the construction of plasmid pHBsr-1, which contains the sr gene of S. mutans OMZ175 on a 9-kbp BamHI-HindIII insert (Fig. 1), as well as the sequencing strategy used to determine the 4,667 bp of the coding sequence of the sr gene (19, 20). We now determined the nucleotide sequence of the upstream region. Analysis of this nucleotide sequence revealed the existence of three ORFs (Fig. 1). The first ORF (ORF1), at the 5' end, extends from an undetermined initiation codon to the termination codon TAA (position 656) and is probably incomplete (Fig. 2). The second ORF (ORF2), starting at the ATG codon (position 816) which is preceded by possible promoter sequences (-35 and -10) and a potential ribosome-binding site, extends to the termination codon TAA (position 1821) (Fig. 2). The third ORF (ORF3) extends from base 1733 to base 2743, with a putative initiation codon ATG at position 2207. Analysis of ORF3 was done on the basis of this potential initiation codon. So defined, ORF3 has the potential to encode a peptide of 179 amino acids, mainly hydrophobic. Because of the poor probability of the existence of a secretory signal sequence in ORF3, we gave

	10	20	30	40	50	60		730	740	750	760	770	780
GGATCCI	CECCACCEA	ATCGTATTG	сстттете	GEATTETCTCA	GGTGAATTT	GAACG	ATAGTA		TCAAAAAAT	ATATTACET	-35 AAGTATT <u>GCT</u>	ANTATTTCT	TTTGTGTTT
υP	кнк	DRI	AFV	RIVS	GEF	ER						-	
	70	80	90	100	110	120		790	800	810	820	830	840
CEECATO	I		I	1		1	_10	1	1	RBSI	820	830	840
G M	S V N	L A R	T G K	SVKL	S N V	T Q	TCAATA	TAGGTGAA	MAAAAGAAAA	TGAAGGAAG	ATTATGAATC	AAAAAATAG1 2 K I V	V I
	130	140	150	160	170	180		850	860	870	880	890	900
GTŢTATO	GCAGAATCT	GTGAGAATO	i STAGAAAAT	; GCTGTCGCTGG1	: IGATATTATO	: GGAGT	CGTCAT	: TTTACATO	: TTAGGTGC1	: CATTCATTT	TCAAAGGCAG		TGATAGGA
FΜ	AES	REN	VEN	AVAG	DII	GΥ	SS	FYM	LGA	H SAF	SKA	У У Н К	DR
										T	T		
	190	200	210	220	230	240		910	920	930	940	950	960
TTACGA	ACAGGAACT.	TATCAGGTTO	GTGATACC	TTAACTGTTGG	TAAAAATAAA	ATTTGA	GTGTGA	AACTTAT	AAAAGAAT	GATATTAAT	CATCAAGCAC	AACGTTTTT	TATTCGTA
ΥD	1 6 1	Y U V	врт	LTVG	кпк	FE	SV	K L M	KRI	DIN	HQA	9 R F 9	SIR
	250	340	270	-									
	1	280	2/0	280	290	300		970 1	980	990	1000	1010	1020
ATTTGA	SCCACTGCCG	ACCITTACAC	CAGAGCTC	TTTATGAAAGT	TCTGCTAA	AATGT	AATATG	CATTTGG	GCTGCATC	TGTTTTAATI	GGCTGTGTCT	TTTTTCTAG	STACCCAAA
• -		• • •		F II K V	энк		КҮ	AFG	AAS	VLI	6 C V	FFLO	5 T Q
	310	320	330	340	350	360		1070	1040	1050		1070	1000
TATGOA				1				1030	1040	1030	1080	10/0	1080
MK	Q K S	FHK	G I E	Q L V Q	AGAAGGTGC	I Q	ATGTTT	CTGCACA	AGAGCAGGG	AACTCAATT(	CCAGCAAGTG	AAAACGCAG	TTGTGAACG
	370	380	390	400	410	420		1090	1100	1110	1120	1130	1140
GCTTTA	I TACCAACTAT	: CAAACTGGT(	: SAATATATG	: CTTGGAGCAGTI	: CGGTCAGCTI	COAGTT	TECOT			1	1	1	1
LY	TNY	QTG	EYM	LGAV	6 Q L	QF	V A	E N S	V A I	S Q A	V A D	к а а	T Q T
	430	440	450	460	470	480		1150	1160	1170	1180	1190	1200
TGAAGT	TTTCAAACAC	CGCATGGAA	ATGAATAC	AATGCGGAGGT	TATCATGAC	ACCCAT	CTCTA	I CAGAAAC	I ACCCCAAGT	TGAAGTTGA		STAASSTAA	ATECTCTE
ΕV	FKH	RME	NEY	NAEV	т м т	PM	ΤL	тет	PQV	EVE	ЕКЕ	SKV	N A P
	490	500	510	520	530	540		1210	1220	1230	1240	1250	1260
GGGTAA	GAAAACGGTG	CGTTGGATT	AGGAGGAA	GATCTTGATGA	ACGTATGTC	TTCCAG	CTTTAA	ATGTCGA	TGACAAAGG	TGCAAAATC	CAAAGAAGATG	TGAACCCTA	CTATTTCAA
6 K	KIV	RWI	KEE	DLDE	RMS	S S	A L	NVD	DKG	AKS	кер	VNP	TIS
	FFA	F.( A											
	1	1	570	580	590	600 1		1270	1280	1290	1300	1310	1320
CCGCAA	TATATTGGCT	AAAGACCGC	TTAACAAG	CCCGTTTTCCT	CTTTCAAAA	TGACTT	AGACA	SCAAGTGA	AGTGGAAGC	TTCTGCAGT	AACTECTACTE	ATACTAAAA	ATTCAAATC
			FNK	F V F L	run	U F	KI	ASE	VEA	SAV	TAT	DTK	N 5 N
	610	620	630	640	650	660		1330	1340	1750	1740	1770	1700
TTOTOT				1	1	:		1000	1340	1330	1380	13/0	1360
S L	H W F	A D K	Y P D	I V L E	E K M	GTAACA -	CACAAO P G	STCAATGT	TGAAACTGA	CTCAAGTGA	K D E	VATAAAATGG NKM	TCACCTCGG
							2	•	•			•••	· · •
	670	680	690	700	710	720		1390	1400	1410	1420	1430	1440
ATGATT	I TGTTTAAATC	: TTATAAATA	ACGATCAAC	CATTICTION	: TT66T6TTT	: TTTTAT	CTCCA			1	TEAFAAAFOC	TOGOGOGO	ATCTTATCO
							A P	AKE	T E A	E Q N	EKA	VRE	NLM

FIG. 2. Nucleotide sequence of the upstream region of the *sr* gene. Numbering starts at the *Bam*HI site (5' end). The deduced amino acid sequence specified by the ORFs is given below the nucleotide sequence. Possible -35 and -10 promoter regions and a putative ribosome-binding site (RBS) are underlined. The vertical arrows show the putative cleavage sites for the removal of the signal peptide. The underlined ATG (position 2983) is the initiation codon of the gene *sr*.

priority to the study of ORF2, encoding a presumptive transmembrane exported protein.

**Characteristics of ORF2.** Computer analysis of the nucleotide sequence revealed the absence of internal repeats. ORF2 is transcribed with the same polarity as the *sr* gene and has the potential to encode a 335-amino-acid polypeptide (Fig. 2). To identify this putative gene product, we screened the EMBL Swiss-Prot protein sequence data bank for homologous sequences, but none was found. Examination of the deduced amino acid sequence revealed that the putative protein is hydrophilic, with hydrophobic regions at the N and C termini (Fig. 3). The hydrophobic N-terminal region, consistent with the -3, -1 rule (32), contains two potential signal peptide cleavage sites at amino acids 18 and 19, followed by a phenylalanine residue, and at amino acids 22 and 23. The hydrophobic C-terminal region does not contain any predicted antigenic determinant and may function as an anchor to the bacterial cell wall, although it does not display the features of a cell wall-spanning region described for other gram-positive organisms (21). The deduced molecular weight of the ORF2-encoded polypeptide is 37,141.

Characterization of the expression product. Our goal was to study the expression product of ORF2 by using the following three steps: (i) prediction analysis of the secondary structure of the putative protein, (ii) antibody production against a potential epitope by use of a synthetic peptide, and (iii) detection of the expression product in recombinant *E. coli* and *S. mutans* cells.

......

. . . .

\_\_\_\_

		14	υ,		-				14/0	,		1 46	Υ.		1.	470			1200			Z	230			224			22	ю. -		Z	260		4	2/0		2	280
-	GAC	-	БĊТ	AAG	6C ]	бтс	TCA	TT	CCAT	СБ	CAAF	56C/	AAT'	TATE	5ТТ	ττċ	-00	600		06	-		، حمد	CCTI	5TA	مم	י מרמ	محم	5674	ά.	<b>Δ</b> ΤΤ/	AGC/	ATTE	icc/	ATA	محف	SCAG	CCA	455
2	R	G	A	ĸ	A	v	S	I	P	S	0	6	-N	Ŷ	v	Ę	a	E	т	140	2	P	E	P	- V	E	T	T	6	I	T	s	1	A	N	KS	3 S	0	6
-							_	-		-	-	-		-			_	-	-		-	•	-	•	•	-	•	•	-	-	-	-	-	••				-	-
		15	10		1	520			1530	>		15	40		1	550			1560			23	290			230	o		231	0		2	320		- 2	2320		2	340
			1			:						_	1			:			:				:				:			1			1			1			1
Ст	CCTI	STA	-	AAT	GCf	ACCC	AGT	ATG	TCCA	- GC	CCA	ACCI	CAA	ттт	AAC	TTT	GAT	'AAA	GGAG	A	TT	6AT(	STT	TTG	ATT	ACT	AAT	GCT	TCC/	<b>ACC</b>	ACT	CAA	GAC/	TA	YAA	AGG.	TTT	AGT	ecc
•	P	•	ĸ	N	A	A	S	74	5	5	P	т	Q	F	N	F	D	ĸ	6		F	D	v	L	I	т	N	A	S	S	Т	•	D	I	ĸ	E١	V L	. V	P
			70		,	590			150	•			~						1430			~					~					-	700					-	
		10	ĩ			1000			1370	ĩ		10	~~			910			1020			2.	350			د∠∠	÷		23	·•		- 2	380		•	2370		-	400
AT	AAG	6TT	TTT	TAT	GA.	raai	бтт	TTA	600	SCS	GAT	666	ĊĂT	-	TEE	ידד	-	TAT	тата	6		TGG	тга 1	<u></u>	<b>6</b> 00	-	660	-	5000	та:	ATT	ΔΤΤ	таат	TAT	-001	-	ста	ACA	wee'
D	ĸ	v	F	Y	D	N	v	Ľ	E	A	D	G	н	<u> </u>		ī	s	~~~~	0.01		511	100	5	F	0	N	6	0	n	5	<b>.</b>		- GG	<b>~</b>	0	Δ	TK		3 6
								-	-	•••	-	-	••	-		-	-		•		•	-	5	•.			0			-	•	•		•	-		• •	•	
		16	30			1640	)		165	0		16	60		1	670	)		1680			2	410	,		242	0		24	30		2	440		:	2450		1	2460
			1							1			1			:			:				:				:			:			1			1			:
CT	TAC	AGT	'6G1	TAT	CG	TCGO	TAT	6CT	CCT	ATT	6CT	6TG	IACA	ATT	GA4	GAA	TTE	5440	SCAAA	5	GAA	GGC	GTT	TAT	AAG	GTG	GCC	GTT	AAG	БТС	AGT	GAC	CAT	<b>AAA</b>	AAT	<b>SACA</b>	6760	5TA	<b>NCTA</b>
S	Y	S	6	1	R	R	Y	A	P	1	A	v	т	I	Е	E	L	ĸ	Q		Ε	6	v	Y	к	v	Α	v	κ	v	S	D	н	κ	N	D	s e	5 P	Y P
		• •	~							~												_					~		-	-		-							-
		16	140			1700			1/1	Ŷ.		1/	20		1	1/30	,		1/40			2	4/0	·		246			24	70		4	500			2310		-	2320
مم	500	AT1	re'r	TCAS	:	<u>^</u>	TTA	cre			SC0	-	ידאי	CAC			-	0.C.04	-raca	т	ممר	ATT	مم	·CTT	· T Δ1	-		·CTT	GTA	مרז	661	600	TTA	ممح	GTT	STT6	-	-	
ĸ	F	- T			~~~	N		P			6	T	- T-	4		T	<b>K</b>		0	•	N		-		10		6	~	507	Ť	6	F		ĸ	<u> </u>	0.10	5 (	5 5	< T
~	-	•	•				-			-		•	•			•	~		-			•		-	•	•		-	•	•		-	-		•	•		•••	• •
		17	750			176	D		177	o		17	780		:	1790	>		1800			2	530	)		254	0		25	50		2	:560			2570	•	:	2580
			:				:			1			1			1	1		:				:				:			:			:			:			:
GC	CTTA		AT	GAAI	5CT	AAA	CTG1	ICT/	AGTC	:C64	÷CCC	:001	L L L L	rcg1	ITT.	TAC/	÷AC	GGA	GATCA	. A	ACG	5ACA	GTA	<b>\GA</b> f	AGCI	CCCI	raa1	rag/	GTC	AA1	ICTI	ICC4	<b>IGCA</b>	CAA	GGA	ACTT	ATG	TTT	TCAC
S	L	к	M	ĸ	L	. N	С	L	v	R	P	N	S	R	F	т	т	E	1		т	т	v	E	Α	P	N	R	v	Ν	L	Р	A	Q	6	т	Y١	V 1	FΤ
		10	810			182	0		183	so		16	840			185	0		1860	)		2	294	2		260			26	10			620	)		2630	,	,	2040
~			1				i 			1									; 					i Fran			•	TEA				TCI	I ACT			1 7007	TTA		
	5111 E		1011	GHI	-		11144	31-11-11	50.56	31411	913141		LAHA	1 (56)		HISL	IAI	616	TUCTA	1	N		101	- GHC	501	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	N	5			T		- e	0	T T	0	E	T 1	FN
•		•		•			_															~	•	-	•			-			•			•	•	-	•	•	• ••
		1	870	,		188	0		189	70		1	900			191	0		1920	•																			
			1				1			1		-	1				1					2	2654	0		26	50		26	70		:	2680	•		2690	>		2700
C	AGTI	56T	ATC	CGT	CG1	TAT	GTT	6TT:	ATTE	<b>SGA</b>	AAG	CTT	ACG	ACA	CAA	CCC	тст	'CCA	ATTGA	۱			1	:			:			1			:			1	1		1
																				7	TAA	AGG4	AGA	AAG'	TAT	TTA	CTA	TGA	CAGT	AT	CTT	5AA'	TGCT	'GA1	'GG4	CAT	CAAT	GGA	TTAG
																					ĸ	G	Ε	S	I	Y	۰Y	D	S	I	L	N	Α	D	6	н	Q	H	I S
		1	930	•		194	0		19	50		1	960			197	0		1980	)																			
							:			1			1				:			_																			
-	HUI		6 I A	ITCA	66	IACT	ATT	510	ATC	:AA	AATI	AAAA	AC6	GCT	CAA	CAA	TTC	:GAT	GTTGI	r			271	<u>0</u>		27	20		27	730			2740	>		275	0		2760
																								1				-		1	<b>*</b> **	<b>*</b> **		rees		ACT	i ^^^^		 •••••
			<b></b>			200	<b>•</b>		20	••		2	020			201	~		2040	、 '	CIA	ICG	110	CIA	CAG	166	IAI	ILG D	105	. 16	IH!	141		600		44911			BIIH
		•	110			200	ĩ		20	ĩ		-	.020			200	ĩ		2040		T	R	3	•				R	R	T	•				-				
С	ATT	тст	AAT	GCT	тс	AAGO	AAT	CAA	GGC	AŤA	-	GAG	GTA	TTA	GTG	<b>ICCA</b>	ĠТІ	TGE	ITCAG	à			277	0		27	80		27	790			2800	>		281	0		2820
																								:			1			1			- 1	1			:		1
																				(	GGA	TGA	CAA	AAT	CCT	GAC	TTT	TTT	GTG	CTT	TAG	AAT	TAA	rgt'	rGG	TAA	AGTE	ITGE	SAGTT
		2	:050	>		206	0		20	70		2	080	,		209	o		2100	0																			
							1			1			- 1				1		1	1									_	-							~		
G	CAA	AAC	:666	SCAE	GA	TGAC	ATT	GTC	:T66	TAT	CAA	6CA	ACT	'AAA	CAF	661	GA	NGG(	GTTT	<b>A</b>			283	0		28	40		2	820	,		286			28/	Ŷ		2880
																					T	607	<b>6</b> 60		ΔT/		i Anna		600	í TÆŦ	677	TAT	AAT	TTG.	<b>ATT</b>		כפדי		, ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
		-					•		-	-		-								•		301							3744		3.1								
		2	4119			213	20		21	30		2	:140	,		215	NO .		216																				
	-005		-00-	i Pett	-	CC74	1	-	~~~	1		~~~	1			-	1	TC.T/		1 T			289	0		29	00		2	910	•		292	0		293	0		2940
					HH	001	,1413 T	GHC	ا الد			rafi T	HIS	.001	-	.141	LH	(	20110	•				1			1			:				1			:		1
																					ATT	TCA	AGC	CAAP	AA/	ATT1	GAC	CAAP	TCA	AAT	CAP	TTA	TAT	TAC	AAT	TTT	TAA	CGT/	ATATI
		2	2170	D		210	30		21	90		2	200	•		221	0		222	0																			
			1	1			:		-	1		_	1			-	1		_	:			70-	50		~			_	<b>07</b> /			200	~					
٦	TAT	TAT	CT	TTTE	<b>iGA</b>	TAA	r661	GAA	CAA	AGA	GGA	6TC	:666	<b>IGCA</b>	AC/	ATE	iac'	TGA	56T66	A			273	, ,		~	100		2	770			278	~					
																м	T	E	V 1	E	AC	-		• • • • •	ГТТ	560		TTT4	<b>1</b> 110	<u>66</u>		664	-	TTT					
																																			- 19				

. . . .

.....

FIG. 2-Continued.

In accordance with the secondary structure prediction analysis, (14), we selected the middle hydrophilic coiled region ranging from residue 167 to residue 182 for modeling of the peptide. After synthesis, the 16-residue peptide was conjugated to ovalbumin and used for antibody production in two rabbits. The antipeptide response was measured in an ELISA with the peptide conjugated to a different carrier protein, BSA. The peptide elicited high titers of antibodies in both rabbits. Figure 4 shows the titration curve of antipeptide IgG. We used antipeptide IgG to detect the putative expression product of ORF2 by Western blotting and ELISA. Analysis of CE from *E. coli* HB101 cells carrying recombinant plasmid pHBsr-1 by Western blotting revealed that the expression product of ORF2 migrated as a single band with an apparent  $M_r$  of 40,000 (Fig. 5B). This  $M_r$  is close to that estimated from the deduced amino acid se-

quence. Antipeptide IgG showed strong reactivity in ELISA with CE from transformed E. coli (Fig. 4). To determine whether this protein was present in S. mutans, we analyzed various cellular fractions by ELISA. All of the activity was found associated with whole S. mutans cells and WEA but not with ESA, suggesting a cell surface localization of this protein (Fig. 4). Western blot analysis of S. mutans WEA and ESA revealed that antipeptide IgG reacted strongly with one protein present in WEA but not in ESA (Fig. 5B). This protein had a molecular weight of 40,000 and migrated like the protein detected in transformed E. coli. The presence of this protein in the other streptococci of the S. mutans group was determined by a whole-cell ELISA with antipeptide IgG. The results for heterologous binding revealed the presence of the antigenic determinant of the peptide in all of the streptococci (Table 1). On the other hand, we used a



FIG. 3. Hydropathy plot of the ORF2 potential expression product. The x axis represents the amino acid number. Hydrophobic and hydrophilic domains are, respectively, above and below the dotted line.

DNA probe covering the 5' portion of ORF2 (amino acids 1 to 157), with the exception of the peptide-encoding region, in the hybridization analysis. Under high-stringency conditions, the probe hybridized only with the homologous restriction fragment of S. mutans OMZ175 chromosomal DNA and with chromosomal fragments derived from S. mutans strains from serotypes c and e, indicating the absence of close homology with the other streptococcal genes (Fig. 6) or at least a weaker conservation of the 5' portion of the gene.

## DISCUSSION

Several approaches, including biochemical, immunological, and genetic methods, have been used to characterize *S. mutans* surface proteins because of their possible importance as potential antigens in immunization against dental caries. Here we report the existence of an ORF that lies upstream of the *S. mutans* OMZ175 sr gene and that encodes a cell-associated protein. From the DNA sequence analysis, we deduced that the encoded protein was composed of 335 amino acids and was possibly shortened by a signal peptide. To study the putative expression product of this ORF, we elicited antibodies against a synthetic 16-residue peptide derived from the deduced amino acid sequence and corresponding to the middle hydrophobic region of high predicted flexibility. Antipeptide IgG was used to determine, identify, and localize the authentic streptococcal protein. The antibodies reacted with a 40-kDa protein present in both CE from recombinant E. coli and WEA from S. mutans, showing that the expression of this S. mutans protein was efficient in E. coli. The  $M_r$  deduced from the ORF sequence is 37,100, similar to the  $M_r$  of the gene product in E. coli and to the  $M_r$ of the protein identified in S. mutans, suggesting the absence of any posttranslational maturation process. Antipeptide IgG reacted with the streptococcal protein not only in immunoblotting, possibly indicating a denatured state, but also in ELISA, in which the structure of the protein is likely to be more native. Furthermore, the high reactivity of the antibodies with whole S. mutans serotype f cells confirmed the fact that the protein is exported at the cell surface and that the hydrophobic C-terminal region probably anchors the protein to the cell wall. Antipeptide IgG recognized an identical determinant on the cell surface of other streptococci of the S. mutans group. However, that genetic diversity exists within these streptococci is evidenced by the hybridization



FIG. 4. Dose-dependent binding of antipeptide IgG to peptide conjugated to BSA ( $\blacksquare$ ), CE from *E. coli* HB101 ( $\bullet$ ), and WEA ( $\times$ ) and ESA ( $\Box$ ) from *S. mutans* OMZ175. Values obtained with control preimmune IgG are deducted.



FIG. 5. (A) Coomassie blue-stained SDS-PAGE gel of CE from transformed (lane 1) and control (lane 2) *E. coli* HB101 and of *S. mutans* OMZ175 WEA (lane 3) and ESA (lane 4). Forty micrograms of protein per fraction was run. (B) Western blot of the gel incubated with antipeptide IgG and with AP-antirabbit IgG. Relative molecular mass markers are indicated at the left in kilodaltons (k).

 
 TABLE 1. Binding of antipeptide IgG to streptococcal whole cells as measured by ELISA<sup>a</sup>

Strain (serotype)	A <sub>405</sub>
S. cricetus E49 (a)	
S. rattus OMZ51 (b)	
S. mutans OMZ70 (c)	
S. sobrinus OMZ176 (d)	
S. mutans B-2 (e)	
S. mutans OMZ175 (f)	
S. sobrinus 6715 (g)	
S. downei MFe 28 (h)	

<sup>*a*</sup> Antipeptide IgG was used at a concentration of 50  $\mu$ g/ml, and the results were read after 1 h at 405 nm. A<sub>405</sub> values represent the read values minus values obtained with preimmune serum.

experiments, which revealed a poor degree of homology in DNAs from the different serotypes. The S. mutans OMZ175 DNA fragment used as a probe did not appear to have sequence homology with the DNAs of the strains of S. cricetus, S. rattus, S. sobrinus, and S. downei used in this study. A similar finding was recently reported by Goldschmidt and Curtiss concerning the antigen I component of the S. sobrinus SpaA protein (6). Although the antigen I-encoding fragment of S. sobrinus DNA did not appear to have sequence homology with the chromosomes of other members of the S. mutans group, the immunodeterminant of antigen I was shown to cross-react with cell surface proteins from S. mutans of various serotypes. Russell described a cell wall antigen (antigen III) present in S. mutans serotypes c, e, and f and S. rattus serotype b and having an  $M_r$  of 39,000, similar to the  $M_r$  of our protein (23). Recently, the gene specifying a precursor form of antigen A (wapA), a cell wall-associated antigen from S. mutans serotype c that is similar to antigen III, has been cloned (4) and sequenced (5). A comparison of the amino acid sequence reported here and the sequence of the wapA-encoded protein as well as those of other S. mutans proteins failed to reveal any homology, demonstrating that our 40-kDa protein is different from antigen III or A. In summary, by using antipeptide IgG, we identified a new S. mutans cell wall protein. The function of



FIG. 6. Southern hybridization. Lanes a, b, c, d, e, f, g, and h indicate the patterns of hybridization of the biotin-7-dATP probe to the *Bam*HI-*Pst*I-restricted chromosomal DNAs (20  $\mu$ g each) of S. *mutans* strains from the corresponding serotypes.

this 40-kDa protein remains unknown, and further studies will be necessary to define the role, if any, of this protein in the virulence of *S. mutans* and to determine whether interactions occur with other cell wall components.

#### REFERENCES

- 1. Beighton, D., R. R. B. Russell, and H. Hayday. 1981. The isolation and characterization of *Streptococcus mutans* sero-type h from dental plaque of monkeys (Macaca fascicularis). J. Gen. Microbiol. 124:271–278.
- Colman, G., and B. Cohen. 1979. Immunisation of monkeys (Macaca fascicularis) with *Streptococcus mutans*, p. 214–215. *In M. T. Parker (ed.)*, Pathogenic streptococci. Reedbooks Ltd., Bracknell, United Kingdom.
- 3. Corthier, G., E. Boschetti, and J. Charley-Poulain. 1984. Improved method for IgG purification from various animal species by ion exchange chromatography. Immunol. Methods 66:75.
- 4. Dao, M. L., C. Chavez, Y. Hirachi, and J. J. Ferretti. 1989. Molecular cloning of the *Streptococcus mutans* gene specifying antigen A. Infect. Immun. 57:3372–3376.
- 5. Ferretti, J. J., R. R. B. Russell, and M. L. Dao. 1989. Sequence analysis of the wall-associated protein precursor of *Streptococcus mutans* antigen A. Mol. Microbiol. 3:469–478.
- 6. Goldschmidt, R. M., and R. Curtiss III. 1990. Cross-reactivity between the immunodominant determinant of the antigen I component of *Streptococcus sobrinus* SpaA protein and surface antigens from other members of the *Streptococcus mutans* group. Infect. Immun. 58:2276–2282.
- Gough, J. A., and N. E. Murray. 1983. Sequence diversity among related genes for recognition of specific targets in DNA molecules. J. Mol. Biol. 166:1–19.
- Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28:351– 359.
- Hughes, M., S. M. McHardy, A. J. Sheppard, and N. C. Woods. 1980. Evidence for an immunological relationship between *Streptococcus mutans* and human cardiac tissue. Infect. Immun. 27:576-588.
- Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157:105-132.
- 11. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) 227:680-685.
- Lee, S. F., A. Progulske-Fox, and A. S. Bleiweis. 1988. Molecular cloning and expression of a *Streptococcus mutans* major surface protein antigen, P1 (I/II), in *Escherichia coli*. Infect. Immun. 56:2114-2119.
- Lehner, T., M. W. Russell, J. Caldwell, and R. Smith. 1981. Immunization with purified protein antigen from *Streptococcus mutans* against dental caries in rhesus monkeys. Infect. Immun. 34:407-415.
- Levin, J., B. Robson, and J. Garnier. 1986. An algorithm for secondary structure determination in proteins based on sequence similarity. FEBS Lett. 205:303-308.
- 15. Loesche, W. J. 1986. Role of *Streptococcus mutans* in human dental decay. Microbiol. Rev. 50:353-380.
- 16. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Merrifield, R. B. 1963. Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. J. Am. Chem. Soc. 85:2149-2154.
- Ogier, J., J. P. Klein, M. Scholler, and R. M. Frank. 1986. Studies of *Streptococcus mutans* interactions with saliva glycoproteins by an enzyme linked immunosorbent assay. J. Microbiol. Methods 5:157-165.
- 19. Ogier, J. A., A. Pini, P. Sommer, and J. P. Klein. 1989. Purification and characterization of the expression product of the sr gene of *Streptococcus mutans* OMZ 175. Microb. Pathog. 6:175-182.
- Ogier, J. A., M. Scholler, Y. Lepoivre, A. Pini, P. Sommer, and J. P. Klein. 1990. Complete nucleotide sequence of the sr gene

from Streptococcus mutans OMZ 175. FEMS Microbiol. Lett. 68:223-228.

- Pancholi, V., and V. A. Fischetti. 1988. Isolation and characterization of the cell-associated region of group A streptococcal M6 protein. J. Bacteriol. 170:2618-2624.
- Robeson, J. P., R. G. Barletta, and R. Curtiss III. 1983. Expression of a Streptococcus mutans glucosyltransferase gene in Escherichia coli. J. Bacteriol. 153:211-221.
- Russell, M. W. 1979. Purification and properties of a protein surface antigen of *Streptococcus mutans*. Microbios 25:7–18.
- Russell, M. W. 1987. Analysis of heart-reactive antibodies induced in rabbits by immunization with *Streptococcus mutans*. J. Oral Pathol. 16:234-240.
- 25. Russell, R. R. B., D. Beighton, and B. Cohen. 1982. Immunization of monkeys (Macaca fascicularis) with antigens purified from *Streptococcus mutans*. Br. Dent. J. 152:81-84.
- Russell, R. R. B., D. Coleman, and G. Dougan. 1985. Expression of a gene for glucan-binding protein from *Streptococcus mutans* in *Escherichia coli*. J. Gen. Microbiol. 131:295–299.
- Sato, S., and H. K. Kuramitsu. 1986. Isolation and characterization of a fructosyltransferase gene from *Streptococcus mutans* GS-5. Infect. Immun. 52:166–170.
- 28. Schöller, M., J. P. Klein, and R. M. Frank. 1981. Common antigens of streptococcal and nonstreptococcal oral bacteria:

immunochemical studies of extracellular and cell-wall-associated antigens from *Streptococcus sanguis*, *Streptococcus mutans*, *Lactobacillus salivarius*, and *Actinomyces viscosus*. Infect. Immun. **31**:52-60.

- Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503-517.
- Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA 76:4350-4354.
- 31. Ueda, S., T. Shiroza, and H. K. Kuramitsu. 1988. Sequence analysis of the *gtfC* gene from *Streptococcus mutans* GS-5. Gene 69:101-109.
- 32. Von Heijne, G. 1983. Patterns of amino acids near signalsequence cleavage sites. Eur. J. Biochem. 133:17-21.
- Wachsmann, D., F. Ackermans, C. Vincenzotto, M. Schöller, H. Bazin, J. A. Ogier, and J. P. Klein. 1989. Human IgG and *Streptococcus mutans* SR protein contain cross-reactive epitopes. J. Immunol. 143:4257–4263.
- Wu, H., and M. W. Russell. 1990. Immunological cross-reactivity between *Streptococcus mutans* and human heart tissue examined by cross-immunization experiments. Infect. Immun. 58:3545-3552.