Oral Microbiology Immunology 2005: 20: 303–309
Printed in Singapore. All rights reserved

Oral Microbiology and Immunology

Molecular analysis of the genes involved in the biosynthesis of serotype specific polysaccharide in the novel serotype *k* strains of *Streptococcus mutans*

R. Nomura, K. Nakano, T. Ooshima
Department of Pediatric Dentistry, Osaka University Graduate School of Dentistry, Osaka, Japan

Nomura R, Nakano K, Ooshima T. Molecular analysis of the genes involved in the biosynthesis of serotype specific polysaccharide in the novel serotype k *strains of* Streptococcus mutans.
Oral Microbiol Immunol 2005: 20: 303–309. © Blackwell Munksgaard, 2005.

We previously reported the new serotype *k* of *Streptococcus mutans*, which, compared to serotypes *c*, *e*, and *f*, features a drastic reduction in the length of the glucose side chain linked to the rhamnose backbone of the serotype specific polysaccharide. The 5′ region of the *rgpF* gene of serotype *k* strains contains a distinctive nucleotide sequence, which suggests that an alteration of the *rgpF* gene in serotype *k* strains may explain the shortened glucose side chain. However, in the present study, expression of the *rgpF* gene of MT8148 (serotype *c*) in serotype *k* isolates was not found to lead to serotype conversion. Furthermore, mRNA expression of *rgpE*, known to be associated with glucose side chain formation, was not detected in any of the tested serotype *k* isolates with an RT-PCR method. The nucleotide alignment of all genes known to be involved in the biosynthesis of serotype specific polysaccharide in serotype *k* strains was shown to be quite similar to that of serotype *c* strains, as compared to serotype *e* and *f* strains, especially in the region downstream of *rgpF*. Our results indicate that the common characteristics of serotype *k* isolates may be caused by a lack of expression of the gene involved in glucose side chain formation.

Key words: *Streptococcus mutans*; serotype specific polysaccharide; glucose side chain; mRNA; reverse transcription-polymerase chain reaction

Dr. Takashi Ooshima, Department of Pediatric Dentistry, Osaka University Graduate School of Dentistry, 1–8 Yamada-oka, Suita, Osaka 565–0871, Japan
Tel.: + 81 6 6879 2961;
fax: + 81 6 6879 2965;
e-mail: ooshima@dent.osaka-u.ac.jp
Accepted for publication April 7, 2005

Streptococcus mutans is known to be a major causative bacterium of dental caries in humans and is occasionally isolated from the blood of patients with infective endocarditis (3, 17, 18). This pathogen was previously classified into *c*, *e*, and *f* serotypes, based on the serotype specific polysaccharide, and shown to have a backbone of rhamnose and a side chain of α- or β-linked glucosidic residue (5). Surveys of *S. mutans* serotypes in Japan have shown that serotype *c* is the most prevalent, with distribution ranging from 74% to 85% (4, 8, 11). However, in a recent study, none of four *S. mutans* blood isolates taken from patients with bacteremia after tooth extraction or with infective endocarditis were classified as serotype *c* (2). In addition, two strains could not be classified as serotype *c*, *e*, or *f*, and were shown to feature a drastic reduction of the glucose side chain linked to the rhamnose backbone in the serotype specific polysaccharide (2). Recently, we designated those strains as belonging to novel serotype *k* and reported that 2–5% of Japanese children possess *S. mutans* with this serotype (8, 9).

The genes involved in biosynthesis of the serotype specific polysaccharide of *S. mutans* have been cloned and sequenced, and four *rml* genes (*rmlA* through *rmlD*) have been shown to be related to the biosynthesis of the rhamnose backbone unit (15, 16). In addition, five genes (*rgpA*, *B*, *C*, *D*, and *F*) are known to participate in the formation of the rhamnose backbone, the *gluA* and *rgpE* genes have been shown to be related to the biosynthesis of glucose

Table 1. S. mutans strains used in the present study

Strain	Serotype	Features	Reference
TW295	*k*	Blood isolate from a 59-year-old male with bacteremia following a tooth extraction procedure	2
TW871	*k*	Blood isolate from a 45-year-old male with infective endocarditis complicated with subarachnoid hemorrhage	2
FT1	*k*	Oral isolate from a healthy 3-year-old girl	8
YT1	*k*	Oral isolate from a healthy 6-year-old boy	9
MT8148	*c*	Oral isolate from a healthy boy	7
NN2001	*c*	Oral isolate from a healthy 6-year-old boy	8
NN2004	*c*	Oral isolate from a healthy 11-year-old boy	8
NN2005	*c*	Oral isolate from a healthy 8-year-old boy	8
MT8148Fcon	*c*	MT8148 carrying *rgpF* replaced with that of MT8148 and *erm* cassette	This study
TW295Fcon	*k*	TW295 carrying *rgpF* replaced with that of MT8148 and *erm* cassette	This study
TW871Fcon	*k*	TW871 carrying *rgpF* replaced with that of MT8148 and *erm* cassette	This study
FT1Fcon	*k*	FT1 carrying *rgpF* replaced with that of MT8148 and *erm* cassette	This study
YT1Fcon	*k*	YT1 carrying *rgpF* replaced with that of MT8148 and *erm* cassette	This study

side chain formation (20, 21), and the *rgpH* and *rgpI* genes are thought to be regulators of its branching frequency (10).

Since serotype *k* strains have been isolated from the blood of patients with bacteremia or infective endocarditis (2), and were also shown to be less susceptible to phagocytosis by human polymorphonuclear leukocytes (8), we considered it important to identify subjects with serotype *k* strains in the oral cavity, as they may possibly carry a risk for infective endocarditis caused by *S. mutans*. Therefore, the genes involved in biosynthesis of the serotype specific polysaccharide were sequenced for use in a rapid PCR identification method (9). Those results revealed that the nucleotide sequences specific for serotype *K* were located in the 5' region of the *rgpF* gene, and that there were no prominent differences in the sequence of genes correlated to the serotype specific polysaccharide formation, except for *rgpf*, among the serotype *K* and other strains. In the present study, we performed a molecular characterization of *S. mutans* serotype *k* strains to investigate the etiology of the drastic shortening of the glucose side chain in the serotype specific polysaccharide.

Material and methods

Bacterial strains

Table 1 lists the *S. mutans* strains used in the present study, which included the serotype *k* blood isolates TW295 and TW871 (2), and oral isolates FT1 (8) and YT1 (9). In addition, MT8148 (serotype *c*) (7) and three serotype *c* oral isolates (NN2001, NN2004, and NN2005) (8) were selected as reference strains.

Predicted amino acid sequence of genes involved in the biosynthesis of serotype specific polysaccharide

Figure 1 shows the genes reported to be involved in biosynthesis of the serotype specific polysaccharide of *S. mutans*. The numbers appearing in Fig. 1 are based on the Oral Pathogen Sequence database (http://www.stdgen.lanl.gov/oragen/) funded by the National Institute of Dental and Craniofacial Research (NIDCR) of the National Institutes of Health, Bethesda, Maryland (1). The complete nucleotide alignments of the *rgp* genes (*rgpA* through ORF11; total 15,890 bp) were determined in our previous study for strains TW295 (*k*), TW871 (*k*), FT1 (*k*), YT1 (*k*), and MT8148 (*c*) (9). In the present study, the PCR products of the *rgpG*, *gluA*, and ORF12 genes of

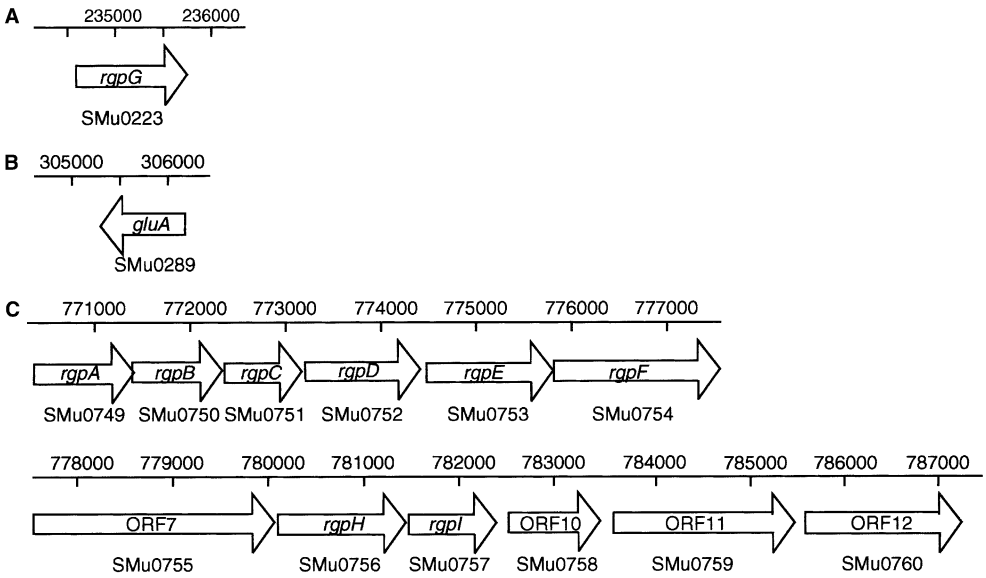

Fig. 1. Genes involved in biosynthesis of serotype specific polysaccharide of *S. mutans*. Numbers on the scale are based on the Oral Pathogen Sequence database (http://www.stdgen.lanl.gov/oragen/). The gene name is indicated inside the arrow, with the identification number from the database shown below. The *rgpG* (A) and *gluA* (B) genes were located separately, and 12 genes, *rgpA* through ORF12 (C), were located consecutively.

Table 2. Primers used in the RT-PCR assays

Gene	Primer	Sequence (5′-to 3′)	Product size (bp)
rgpG	rgpG-tF	CTC TTA GGG TGG GGG CAG	481
	rgpG-tR	ACC AAT CCC ATA GTG GTC	
	rgpG-bF	GAT CTT TGC TAT TGC TGG	478
	rgpG-bR	GCC AAT AAA AGC CAA TAG	
gluA	gluA-tF	GCC AAG TAA AAA AGT CAG AA	300
	gluA-tR	TGG AGA TTG ATT GCG GTA GT	
	gluA-bF	GAC ACG TCA ACT CAT GAA	479
	gluA-bR	CCG AGT TCT TTT GAG TCG	
rgpA	rgpA-tF	GGT GTT GAT TGT TTT ACT	400
	rgpA-tR	GTG ATA AAT CCG TGC CAT	
	rgpA-bF	TCA CCT TCA GGG CTG TTA	519
	rgpA-bR	AGC TCT TTG TCC ATA CTC	
rgpB	rgpB-tF	CTG ATG CGC GGA TTC GTT	325
	rgpB-tR	GTC CCG CCA GTA ACT GTA	
	rgpB-bF	CAT TGT CTA GCT AAG CAG	386
	rgpB-bR	GTA TGA AAG GCA CGA TTC	
rgpC	rgpC-tF	CTG ATT TTA AAC TTC GTT A	264
	rgpC-tR	AAA CAA TCG TCT GCT TCG	
	rgpC-bF	GCT TTT CTC AAC AGG GAT TG	284
	rgpC-bR	CAA GAT ATG GAA TGA CAG C	
rgpD	rgpD-tF	GCT AAG AAA TTT GCG GAG AT	361
	rgpD-tR	TCA CGG CCA GTC AGT TCT	
	rgpD-bF	TGA CAC CAA GAG ACT CCA TC	156
	rgpD-bR	GTT TCT TTC CTG AAC CAC	
rgpE	rgpE-tF	AAG GAA GCT GGA TTG GTG	252
	rgpE-tR	AAT AAT CAT CAC CAT CAC ATC TGG C	
	rgpE-bF	CCC GAT CCT CAA ATG ATC TA	271
	rgpE-bR	GTA GGA ATG GTC CAA CGG CG	
rgpF	rgpF-ctF	CGG GTA AGT TCC CAT GTC GTT	271
	rgpF-ctR	ACA TTT CCC AAA GAG GTC CAA	
	rgpF-ktF	CTA CCA ACT GAC ACA AAT GCG	248
	rgpF-ktR	CAT CCC ACA AAG GTC CAA AGC	
	rgpF-bF	GCA CCT GAG ATG AAT ACA	320
	rgpF-bR	GCT GAG TTA CCA CGT TCA	
ORF7	ORF7-tF	TGT CTC GGA AAA TGA CTC	339
	ORF7-tR	CAG CAT CAT TTT TTA AGC CG	
	ORF7-bF	TGT TGC TTC CAG TGG AAC	341
	ORF7-bR	CCA GGT AAA TGG TCT CCA TA	
rgpH	rgpH-tF	GAT GCA GGA GAT CAA ACG	305
	rgpH-tR	GTA GGA ATT AAA CCA GCA C	
	rgpH-bF	GGC CGT GGT TAT CTT GGT	185
	rgpH-bR	ATT TTC CTT CTT GCC ATC TC	
rgpI	rgpI-tF	CTT TCA ATA GTT ATT CCT TG	253
	rgpI-tR	CCA TCT ACT GCT TCC AAC CC	
	rgpI-bF	AGG GTT GGA AGC AGT AGA	237
	rgpI-bR	CGA AAG TCA CGC GCC CCA	
ORF10	ORF10-tF	CTT ACC GGA AGC TGA CAT	539
	ORF10-tR	TTC ACT GCC CTC GGG ACG	
	ORF10-bF	CGT CCC GAG GGC AGT GAA	340
	ORF10-bR	TAC GGC CAT TAC GCT CAT	
ORF11	ORF11-tF	GCT CTA ACT CAG CTG TTA	291
	ORF11-tR	AGC TAA AAT AGT CCA CGG	
	ORF11-bF	AGG GAT CAC AAA GAC CTA	404
	ORF11-bR	GTA CTA TTC CTC CCA TGC	
ORF12	ORF12-tF	GCT GGC TGT TAG TTC TTA	473
	ORF12-tR	CTG CCG TGT ACC AAT GAA	
	ORF12-bF	GCG GAC ATG TTT TGG CTA	387
	ORF12-bR	CCA AAT AGG AAG CCG TGG	

the serotype *k* oral and blood isolates, as well as MT8148, were directly cloned into a pGEM-T Easy vector (Promega, Madison, WI), for which the nucleotide sequences were determined using a dye-terminator reaction with a DNA Sequencing System (373–18 DNA sequencer, Applied Biosystems, Foster City, CA) and an ABI PRISM Cycle Sequencing kit, as described previously (9). Data analysis was performed with GENE WORKS software (IntelliGenetics, Mountain View, CA). The predicted amino acid sequences of these strains were compared using a multiple alignment analysis with CLUSTAL W from the DNA Data Bank of Japan (DDBJ; Mishima, Japan) (13). In addition, the *rgp* genes of strain Xc (*c*) (10, 11, 21) (GenBank accession no. AB010970, AB091254, and AB108684), for which the *rgp* genes were originally sequenced, and of strain UA159 (*c*) (GenBank accession no. AE014133) (1), for which the complete genome has been sequenced, were also compared as reference strains.

Construction and serologic characterization of replacement mutants

To examine the conversion of serotype *c* from serotype *k* strains by replacement of the *rgpF* gene from serotype *k* strains with the *rgpF* gene from the serotype *c* strain of MT8148, the replacement mutant strains TW295Fcon, TW871Fcon, FT1Fcon, and YT1Fcon were constructed as follows. The flanking region of the *rgpF* gene of MT8148 was amplified by PCR using LA-Taq polymerase (Takara Shuzo, Otsu, Japan), with the primers rgpE-5981F (5′-CCG ATC CTC AAA TGA TCT AT-3′) and ORF7–9059R (5′-GTT CAT TAC GCA TAC CTA CC-3′), which were constructed on the basis of the *rgpE*, *rgpF*, and ORF7 sequences of *S. mutans* strain Xc (21) (GenBank accession no. AB010970). The amplified fragment was then cloned into a pGEM-T Easy vector to generate pRN111. Next, the *Stu* I restriction site located just after the beginning of ORF7 was cleaved and ligated with the erythromycin resistance gene (*erm*) from recombinant plasmid pKN100, which carried an 830-bp fragment of *erm* from pVA838 (6) to yield pRN112. After linearization by digestion at the unique restriction site of *Nco* I, the plasmid was introduced into *S. mutans* strains TW295, TW871, FT1, and YT1 using the method of Tobian & Macrina (14). To confirm that there was no effect on the genes located downstream from the location of the *erm* insertion at the 5′ region of the initiation of the ORF7 gene, the mutant strain MT8148Fcon was also constructed using the method described above. These transformants were screened on mitis salivarius (MS) agar (Difco Laboratories, Detroit, MI) plates containing erythromycin (10 μg/ml). The appropriate replacement of *rgpF* of MT8148 and the *erm* cassette in each of the strains was confirmed by PCR amplification, as well as by sequencing analysis of the 5′-third region of the *rgpF* gene. The serotype of each tested strain was determined with an autoclaved polysaccharide extract and serotype specific antisera, as described previously (8).

Analysis of mRNA expression of genes involved in biosynthesis of serotype specific polysaccharide

To confirm the transcription of the genes involved in biosynthesis of the serotype specific polysaccharide in the tested strains, we utilized a reverse transcriptase-mediated PCR (RT-PCR) method. Total RNA was prepared from the tested strains using a FastPrep® Cell Disrupter (Model FP100A, Q-Bio gene, Carlsbad, CA) in combination with a Fast RNA® Pro Blue kit (Q-Bio gene), according to the manufacturer's instructions. SuperScript™ III Reverse Transcriptase (Invitrogen, Carlsbad, CA) was used to amplify cDNA synthesized from mRNA. The primers used to detect the transcription of each gene are listed in Table 2. Due to the differences in nucleotide alignment of the initial part of *rgpF* between the serotype *c* and *k* strains, serotype *c*- and serotype *k*-specific sets of primers (rgpF-ctF and rgpF-ctR, and rgpF-ktF and rgpF-ktR, respectively) were constructed. The cDNA was transcribed from 100 ng of total RNA at 55 °C for 60 min and heated at 70 °C for 15 min. Successive PCR assays were performed under the following condition: 30 cycles at 94 °C for 30 s, 50 °C for 30 s, and 72 °C for 30 s. Genomic DNA from each strain was used as a positive control and RNA samples from which reverse transcriptase had been omitted as a negative control.

Results

Comparison of predicted amino acid sequence of RgpF

Figure 2 shows the N-terminal one-fifth region of RgpF of the serotype *k* strains TW295, TW871, FT1, and YT1, as well as of the serotype *c* strains MT8148, Xc, and UA159. The serotype *k* specific sequences were located at the positions of residues 21–116 of the serotype *k* strains. The serotype *c* strains of Xc and UA159 showed a different amino acid sequence at around residues 130–175, as compared to MT8148. In the rest of the sequence, only three of the amino acids were different in the serotype *k* strains compared with the serotype *c* strains (data not shown).

Serologic characteristics of replacement mutants

To confirm that the serotype *k* specific amino acid alignment caused the reduction of the glucose side chain, replacement mutant strains MT8148Fcon, TW295Fcon, TW871Fcon, FT1Fcon, and YT1Fcon were constructed and then screened on MS agar containing erythromycin. The appropriate replacement of *rgpF* from MT8148 in each of the strains was confirmed by the presence of the *erm* cassette following PCR amplification (Fig. 3A), as well as sequencing analysis of the 5′-third region of the *rgpF* gene in each mutant strain. The autoclaved extracts from MT8148Fcon were reactive only with serotype *c* specific antiserum, while those from TW295Fcon, TW871Fcon, FT1Fcon, and YT1Fcon were reactive only with serotype *k* specific antiserum (Fig. 3B).

Expression of the genes involved in the biosynthesis of serotype specific polysaccharide

All of the tested genes shown in Fig. 1 were transcribed in MT8148, which was found to be similar to the three serotype *c* oral isolates NN2001, NN2004, and NN2005 (data not shown). Transcripts of the first half of *rgpE*, ORF7, and ORF10 were not detected in the serotype *k* strains TW295, TW871, FT1 and YT1, whereas

```
                                                                      60
MT8148 (c)      MKRLLLYVHFNKYNRVSSHVVYQLTQMRSLFSKVIFISNSQVADADVKMLREKHLIDDFI
Xc (c)          ............................................................
UA159 (c)       ............................................................
TW295 (k)       ..................Y.......P···R·V··T··HLSQENQDK··SQK·M···L
TW871 (k)       ..................Y.......P···R·V··T··HLSQENQDK··SQK·M···L
FT1 (k)         ..................Y.......P···R·V··T··HLSQENQDK··SQK·M···L
YT1 (k)         ..................Y.......P···R·V··T··HL·QE·QDK··NQN·M···L

                                                                     120
MT8148 (c)      QRQNSGFDFAAWRDGMVFVGFDELVTYDSVTTMNDTCFGPLWEMYSIYQEFETKTTVDFW
Xc (c)          ............................................................
UA159 (c)       ............................................................
TW295 (k)       ··E·I···········NHI···N·DS··I··V·········DVKEH·LSY·KQDEI···
TW871 (k)       ··E·I···········NHI···N·DS··I··V·········DVKEH·LSY·KQDEI···
FT1 (k)         ··E·I···········NHI···N·DS··I··V·········DVKEH·LSY·KQDEI···
YT1 (k)         ··E·I···········NHI···N·DS··I··V·········DVKDY·LSY·KQDE····

                                                                     180
MT8148 (c)      GLTNNRATKQFKEHIQSYFISFKKAVIQSEAFHNFWENIQNHADIQRVIDDYETQVTTTL
Xc (c)          ·········S·R···········AS·LR·T··RD·····KEYQ·V·K···Q···K·····
UA159 (c)       ·········S·R···········AS·LR·T··RD·····KEYQ·V·K···Q···K·····
TW295 (k)       ..................T.........................................
TW871 (k)       ..................T.........................................
FT1 (k)         ..................T.........................................
YT1 (k)         ..................T........N···D........T·V·················
```

Fig. 2. Multiple alignment of predicted amino acid sequence of the N-terminal one-third region of RgpF (total 583 amino acids) in serotype *k* strains compared to the reference strains. Only amino acid residues different from those of MT8148 are presented. Identical amino acids are indicated by dots.

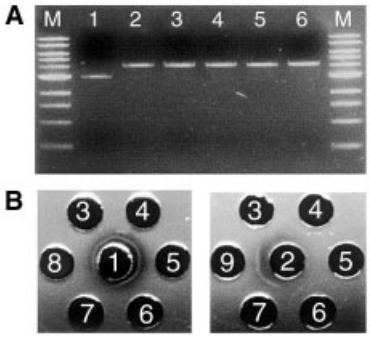

Fig. 3. Features of isogenic mutant strains following replacement of the *rgpF* gene with that of MT8148. A) Agarose gel electrophoresis of PCR products was performed to confirm insertion of the *erm* cassette into the mutant strains. Lane 1: MT8148. Lane 2: MT8148Fcon. Lane 3: TW295Fcon. Lane 4: TW871Fcon. Lane 5: FT1Fcon. Lane 6: YT1Fcon. M, molecular size marker (1-kb DNA ladder). B) Immunodiffusion between the autoclaved extracts of the parent and isogenic mutant strains. Wells 1 and 2 contained antiserum specific for serotypes *k* and *c*, respectively. The outer wells contained RR extracts from strains TW295Fcon (well 3), TW871Fcon (well 4), FT1Fcon (well 5), YT1Fcon (well 6), MT8148Fcon (well 7), TW295 (well 8), and MT8148 (well 9).

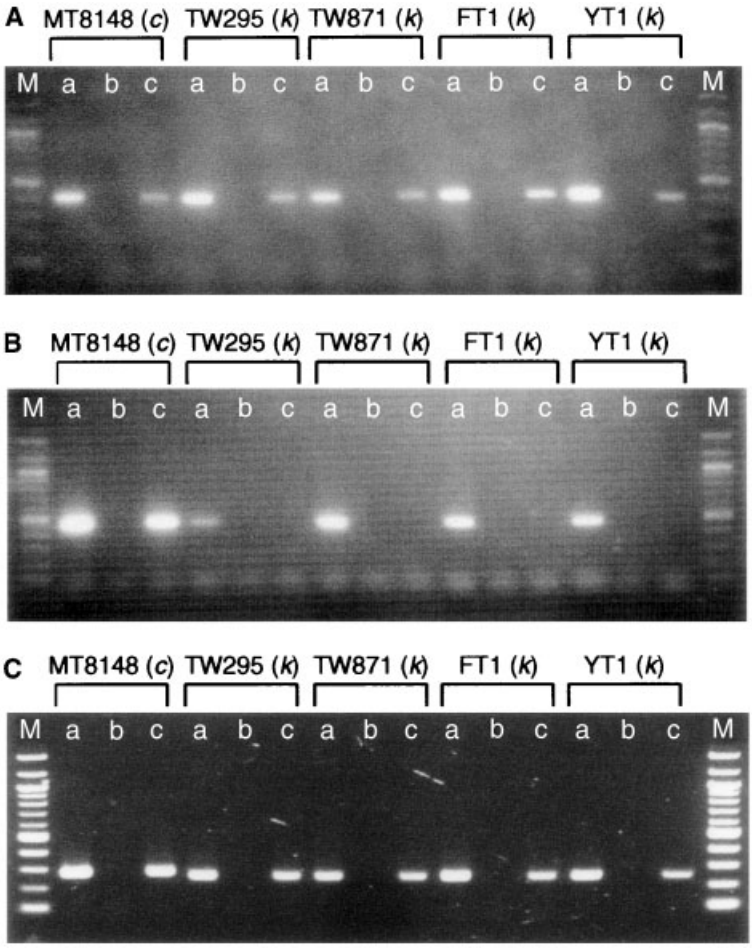

Fig. 4. Detection of transcription of genes involved in biosynthesis of serotype specific polysaccharide by RT-PCR. Three representative images are presented: A) first half of *rgpA,* B) first half of *rgpE,* C) first half of *rgpF.* In each image (a) represents genomic DNA, (b) the control without transcriptase, and (c) the RT-PCR result. M, molecular size marker (100-bp DNA ladder).

the other tested genes shown in Fig. 1 were transcribed in those serotype *k* strains. Figure 4A and C shows representative images of the detection of transcription of the first half of *rgpA* and *rgpF*, respectively, which demonstrated that those genes were transcribed in both MT8148 and the serotype *k* strains. Figure 4B shows a representative image of the detection of the transcription of the first half of *rgpE*, indicating that the gene was transcribed only in MT8148 and not in the serotype *k* strains.

Discussion

The process of biosynthesis of the serotype specific polysaccharide of *S. mutans* as well as the associated genes has been investigated (Fig. 5). The initial step in that process is considered to be the transfer of N-acetylglucosamine-1-phosphate to a lipid carrier by RgpG encoded by the *rgpG* gene (19). The enzymes encoded by 4 *rml* genes (*rmlA* through *rmlD*) catalyze the production of dTDP-L-rhamnose from D-glucose-1-phosphate (15, 16). RgpA, encoded by *rgpA*, is considered to be the first rhamnosyltransferase, and RgpB and RgpF, encoded by *rgpB* and *rgpF*, respectively, are speculated to alternate in elongating the rhamnose polymers (12). On the other hand, glucose side chain formation was shown to be initialized by the enzyme encoded by *gluA*, which catalyzes the conversion of D-glucose-1-phosphate and UTP into UDP-D-glucose (20). RgpE, which is encoded by *rgpE*, has been proposed to be the glucosyltransferase that produces glucose polymers (21). In addition, RgpH and RgpI, encoded by *rgpH* and *rgpI*, respectively, are considered to be involved in the regulation of the branching frequency of the glucose side chain in polysaccharide synthesis (10). The *rgpC* and *rgpD* genes, which encode the putative ABC transporter, may be associated with the export of polysaccharide to the cell surface (21).

In the present study, we attempted to determine the mechanism involved with the drastic reduction of the glucose side chain in the serotype specific polysaccharide of serotype *k* clinical strains. Several genes involved in the biosynthesis of the serotype specific polysaccharide of *S. mutans* have been identified (Fig. 1), among which the nucleotide alignment specific for serotype *k* strains has been found in the 5′ region of the *rgpF* gene (350 bp from the initial sequence) (9). We speculated that an alteration of the amino acid sequence of RgpF in serotype *k* strains, as compared to that of serotype *c* reference strains (Fig. 2) might be associated with the drastic reduction of the glucose side chain. RgpF has been reported to function in the transfer of the third rhamnose residue to the second rhamnose residue, and it was speculated that RgpB and RgpF alternate in elongating the rhamnose backbone (12). We considered that a variation of the *rgpF* gene compared to that of MT8148 may result in a short rhamnose backbone, causing a shortened attachment location of the glucose side chain. However, all of the mutant strains tested (TW295Fcon, TW871Fcon, FT1Fcon, and YT1Fcon), in which *rgpF* was replaced with that from MT8148, were shown to be serotype *k* (Fig. 3B),

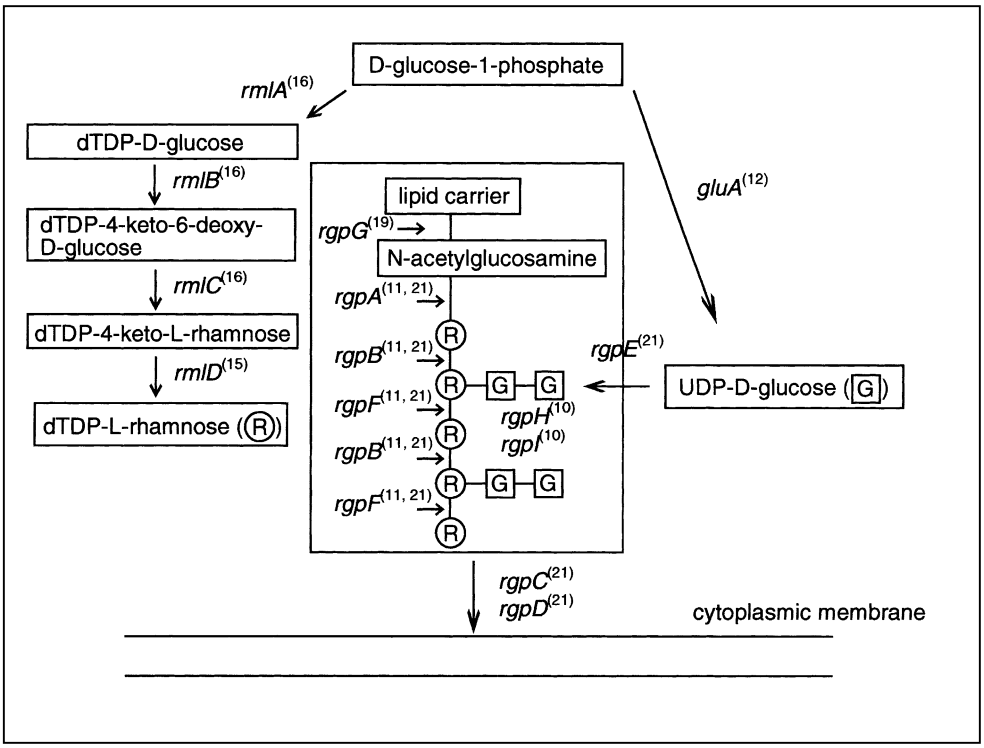

Fig. 5. Illustration of the pathway of biosynthesis of serotype specific polysaccharide. Reference numbers for each gene are shown in parentheses.

indicating that the prominent differences in nucleotide alignment of *rgpF* between serotype *k* strains as compared to MT8148, Xc, and UA159 are not associated with the drastic reduction in length of the glucose side chain.

During construction of the mutant strains with replacement of the *rgpF* gene of MT8148, the restriction site was found to be located just after the beginning of ORF7, which is possibly interrupted. Analysis of the sugar composition revealed that the strain with insertional inactivation of ORF7 had HPLC patterns very similar to those of the parental strain, as well as the same ratio of rhamnose to glucose (21). This indicates that inactivation of ORF7 does not lead to an alteration of the serotype specific polysaccharide. We concluded that the serotype *k* specific amino acid sequence in RgpF would not have an effect on its rhamnosyltransferase function.

The *gluA*, *rgpE*, *rgpH*, and *rgpI* genes have been shown to be correlated to glucose side chain formation in the serotype specific polysaccharide of *S. mutans* (10, 20, 21). However, in our previous study (9) there were no significant differences in the nucleotide alignment of those genes in the serotype *k* strains examined here. In contrast, the results of the present RT-PCR assays revealed that mRNA from the first half of the *rgpE* gene was not expressed in the serotype *k* strains (Fig. 4B). It was previously reported that insertional inactivation of the *rgpE* gene caused a reduction in the amount of glucose side chain of strain Xc (21), for which characterization of the serotype specific polysaccharide is consistent with serotype *k*, which we recently designated (8). The present results suggest that the tested wild type serotype *k* strains have a common feature of no expression of *rgpE* mRNA, which is regarded to be involved in biosynthesis of the glucose side chain of the serotype specific polysaccharide.

In our previous study, Southern blot analysis using the *rgpE* fragment as a probe showed that the serotype *k* strains TW295 and TW871 lacked a 2.7-kb band hybridized with the *rgpE* gene, indicating that the lack of an *rgpE* homolog in these strains may result in a glucose side chain defect. However, two bands were identified, not only in the serotype *c* strain MT8148, but also in the serotype *k* strains TW295 and TW871 (2). Our subsequent study showed that the *rgpE* gene was located in TW295 and TW871, and that its nucleotide alignment was nearly the same as of that in MT8148 (9).

It should also be noted that ORF10 mRNA was not expressed in the serotype *k* strains. In the Oral Pathogen Sequence database funded by NIDCR, ORF10 (SMu0758; 312 amino acids) is described as a possible glycosyltransferase. This was supported by the results of the BLAST search, which indicated that residues 4–309 are 63% similar to the glycosyltransferase in *Clostridium acetobutylicum* (GenBank accession no. AE007733). It is reasonable to speculate that ORF10 itself might be directly associated with glucose side chain formation, as well as being a possible regulator of the genes involved. However, Ozaki et al. (10) showed that inactivation of ORF10 in strain Xc did not cause a reduction of side chain formation, indicating that ORF10 is not associated with the side chain formation of the serotype specific polysaccharide.

The *rgpH* and *rgpI* genes were cloned and sequenced from the serotype *c* strain Xc, and were also found in the serotype *c* strain UA159, as SMu0756 for *rgpH* and SMu0757 for *rgpI*. Our previous study results indicated that the serotype *c* strain MT8148 also possesses the *rgpH* and *rgpI* genes, whose sequences are highly consistent with those of Xc and UA159 (9). The *rgpH* gene in the serotype *e* strain LM7 and serotype *f*

strain MT6219 was reported to have a low identity (less than 5%) with that of Xc, and the ORF3f gene of MT6219 (*f*) was shown to have a greater than 98% identity with *rgpI* of Xc, while the gene corresponding to *rgpI* was shown to be absent in LM7 (*e*) (11). Since serotype *c*, *e*, and *f* strains each possess a glucose side chain with a different linkage style (α1–2, β1–2, and α 1–3, respectively) (4), we speculated that *rgpH* and *rgpI* might be involved in the regulation of the branching frequency of the α1–2 glucose side chain. However, the nucleotide alignment of this region has been sequenced for only one serotype *e* and one serotype *f* strain up to this point; sequencing analyses of a much larger number of serotype *e* and *f* strains should reveal which genes are associated with regulation of the side chain formation of those strains. The serotype *c*, *e*, and *f* specific sequences were identified from the region downstream of the *rgpF* gene, with the primers used for the PCR identification method constructed from the *rgpH* gene (for serotype *c*) and its corresponding genes (for serotype *e* and *f*), and low identities were found among them (11). Our previous study findings indicate that one third of the 5′ region of *rgpF* in serotype *k* strains is different from that of serotype *c* strains, whereas the nucleotide alignment downstream of *rgpF* in serotype *k* strains is consistent with that of MT8148 (*c*), Xc (*c*), and UA159 (*c*) (9). Taken together, these results indicate that the progenitor of serotype *k* strains might be serotype *c* rather than serotype *e* or *f*.

Acknowledgments

This study was a part of 21st Century COE entitled 'Origination of Frontier BioDentistry' at Osaka University Graduate School of Dentistry supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan, and was supported by a Grant-in-Aid for Scientific Research (B) 16390605 from the Japan Society for Promotion of Science.

References

1. Ajdic D, McShan WM, McLaughlin RE, Savic G, Chang J, Carson MB, et al. Genome sequence of *Streptococcus mutans* UA159, a cariogenic dental pathogen. Proc Natl Acad Sci U S A 2002: **99**: 14434–14439.
2. Fujiwara T, Nakano K, Kawaguchi M, Ooshima T, Sobue S, Kawabata S et al. Biochemical and genetic characterization of serologically untypable *Streptococcus mutans* strains isolated from patients with bacteremia. Eur J Oral Sci 2001: **109**: 330–334.
3. Gauduchon V, Benito Y, Celard M, Mouren C, Delorme V, Philippe-Bert J, et al. Molecular diagnosis of recurrent *Streptococcus mutans* endocarditis by PCR amplification and sequencing. Clin Microbiol Infect 2001: **7**: 36–37.
4. Hirasawa M, Takada K. A new selective medium for *Streptococcus mutans* and the distribution of *S. mutans* and *S. sobrinus* and their serotypes in dental plaque. Caries Res 2003: **37**: 212–217.
5. Linzer R, Reddy MS, Levine MJ. Immunochemical aspects of serotype carbohydrate antigens of *Streptococcus mutans*. In: Hamada S, Michalek SM, Kiyono H, Manaker L, McGhee JR, eds. Molecular microbiology and immunology of Streptococcus mutans. Amsterdam: Elsevier, 1986: 29–38.
6. Macrina FL, Evans RP, Tobian JA, Hartley DL, Clewell DB, Jones KR. Novel shuttle plasmid vehicles for *Escherichia–Streptococcus* transgeneric cloning. Gene 1983: **25**: 145–150.
7. Minami T, Fujiwara T, Ooshima T, Nakajima Y, Hamada S. Interaction of structural isomers of sucrose in the reaction between sucrose and glucosyltransferases from mutans streptococci. Oral Microbiol Immunol 1990: **5**: 189–194.
8. Nakano K, Nomura R, Nakagawa I, Hamada S, Ooshima T. Demonstration of *Streptococcus mutans* with a cell wall polysaccharide specific to a new serotype, *k*, in the human oral cavity. J Clin Microbiol 2004: **42**: 198–202.
9. Nakano K, Nomura R, Shimizu N, Nakagawa I, Hamada S, Ooshima T. Development of a PCR method for rapid identification of new *Streptococcus mutans* serotype *k* strains. J Clin Microbiol 2004: **42**: 4925–4930.
10. Ozaki K, Shibata Y, Yamashita Y, Nakano Y, Tsuda H, Koga T. A novel mechanism for glucose side-chain formation in rhamnose-glucose polysaccharide synthesis. FEBS Lett 2002: **532**: 159–163.
11. Shibata Y, Ozaki K, Seki M, Kawato T, Tanaka H, Nakano Y et al. Analysis of loci required for determination of serotype antigenicity in *Streptococcus mutans* and its clinical utilization. J Clin Microbiol 2003: **41**: 4107–4112.
12. Shibata Y, Yamashita Y, Ozaki K, Nakano Y, Koga T. Expression and characterization of streptococcal *rgp* genes required for rhamnan synthesis in *Escherichia coli*. Infect Immun 2002: **70**: 2891–2898.
13. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994: **22**: 4673–4680.
14. Tobian JA, Macrina FL. Helper plasmid cloning in *Streptococcus sanguis*: cloning of tetracycline resistance determinant from the *Streptococcus mutans* chromosome. J Bacteriol 1982: **152**: 215–222.
15. Tsukioka Y, Yamashita Y, Nakano Y, Oho T, Koga T. Identification of a fourth gene involved in dTDP-rhamnose synthesis in *Streptococcus mutans*. J Bacteriol 1997: **179**: 4411–4414.
16. Tsukioka Y, Yamashita Y, Oho T, Nakano Y, Koga T. Biological function of the dTDP-rhamnose synthesis pathway in *Streptococcus mutans*. J Bacteriol 1997: **179**: 1126–1134.
17. Ullman RF, Miller SJ, Strampfer MJ, Cunha BA. *Streptococcus mutans* endocarditis: report of three cases and review of the literature. Heart Lung 1988: **17**: 209–212.
18. Vose JM, Smith PW, Henry M, Colan D. Recurrent *Streptococcus mutans* endocarditis. Am J Med 1987: **82**: 630–632.
19. Yamashita Y, Shibata Y, Nakano Y, Tsuda H, Kido N, Ohta M, et al. A novel gene required for rhamnose-glucose polysaccharide synthesis in *Streptococcus mutans*. J Bacteriol 1999: **181**: 6556–6559.
20. Yamashita Y, Tsukioka Y, Nakano Y, Tomihisa K, Oho T, Koga T. Biological functions of UDP-glucose synthesis in *Streptococcus mutans*. Microbiology 1998: **144**: 1235–1245.
21. Yamashita Y, Tsukioka Y, Tomihisa K, Nakano Y, Koga T. Genes involved in cell wall localization and side chain formation of rhamnose-glucose polysaccharide in *Streptococcus mutans*. J Bacteriol 1998: **180**: 5803–5807.