

Title	Four gbpC Gene Homologues in Streptococcus sobrinus
Author(s) Alternative	Sato, Y; Ishikawa, A; Okamoto-Shibayama, K; Takada, K; Hirasawa, M
Journal	Journal of Oral Biosciences, 49(4): 303-308
URL	http://hdl.handle.net/10130/448
Right	

Short Communication

Four *gbpC* gene homologues in *Streptococcus sobrinus*.

Yutaka Sato¹⁾, Ayako Ishikawa¹⁾, Kazuko Okamoto-Shibayama¹⁾, Kazuko
Takada²⁾, and Masatomo Hirasawa²⁾.

¹⁾ Department of Biochemistry and Oral Health Science Center,
Tokyo Dental College

2-2, Masago 1-chome, Mihama-ku, Chiba City, 261-8502 Japan

²⁾ Department of Microbiology and Immunology, Nihon University
School of Dentistry at Matsudo

2-870-1, Sakaechonishi, Matsudo City, 271-8587 Japan

Corresponding author: Yutaka Sato (E-mail) yusato@tdc.ac.jp

(Phone)81-43-270-3750 (Fax)81-43-270-3752

Runing title: Four *gbpC* gene homologues in *S. sobrinus*

Field: Microbiology

Abstract

We previously identified the glucan-binding protein C gene, *gbpC*, solely involved in dextran-dependent aggregation (ddag) of *Streptococcus mutans*. Recently, we identified two *gbpC* gene homologues, *gbpC* and *dbl*, in *Streptococcus sobrinus* and suggested that the *dbl* gene was very likely responsible for ddag of this species. However, homology searches with the *gbpC* or *dbl* genes against the incomplete TIGR *S. sobrinus* ATCC 6715 database suggested that this strain may have other *gbpC* homologues. PCR-based chromosomal walking from the *gbpC* and *dbl* genes revealed two additional homologous genes designated as *gbpC2* and *dblB*. Proteins encoded by these genes exhibited alpha-1,6 glucan-binding activities. Therefore, *gbpC2* and *dblB* are also logical candidates responsible for the ddag phenotype.

Key words

Streptococcus sobrinus, glucan-binding protein C, *gbpC2* gene, *dblB* gene

Introduction

Streptococcus mutans and *Streptococcus sobrinus* are regarded as the two major principal etiologic agents of human dental caries ¹⁾. Several putative virulence factors involved in dental caries, e.g. surface protein antigens or glucosyltransferases (GTFs), are commonly found in both organisms. However, one clear difference in phenotype between these two organisms has been well-known since the 1980's and is that *S. sobrinus* exhibits more active dextran (α -1,6 glucan)-dependent aggregation (ddag) than *S. mutans* ²⁾. Therefore, we have used this phenotype as a simple convenient method to distinguish *S. sobrinus* from *S. mutans* as suggested previously ³⁾. We initially identified the *gbpC* gene ⁴⁾ involved in ddag of *S. mutans*. In addition, we recently identified two *gbpC* gene homologues in clinically isolated *S. sobrinus* strain 100-4 and suggested that the second homologue gene is very likely involved in ddag of this organism ⁵⁾. Ma *et al.* proposed to define glucan-binding proteins as glucan-binding lectins (Gbls) since they would confer the property of the aggregation on bacteria in the presence of exogenously added dextran ⁶⁾. Similar to this suggestion, and in order to avoid confusion between the two *gbpC* gene homologues within a species, we have proposed the designation Dbls (Dextran-binding lectins) as referring to a wall-anchored protein family characterized by dextran-binding activity and have designated this second homologue as the *dbl* gene as described previously ⁵⁾. Both glucan-binding protein C homologues (GbpC and Dbl) are classified as cell wall-anchored surface proteins in

Gram-positive bacteria ⁷⁾ possessing LPXTG motifs by which sortase mediates to tether the proteins to cell wall peptidoglycans. However, we could detect neither GbpC nor Dbl proteins in a wall-fraction prepared from *S. sobrinus* type strain ATCC6715 with Western blot analyses using anti-GbpC or anti-Dbl sera. However, both proteins were detected in concentrated culture supernatants with the same analyses. When homology searches with the *gbpC* or *dbl* genes against the incomplete TIGR (The Institute for Genomic Research) *S. sobrinus* 6715 database were carried out, multiple contigs containing parts of homologous sequence regions with 70-80% identities were detected in each search.

These results suggested that strain 6715 may have other cell-associated *gbpC* homologues involved in ddag of this strain, although the search results at TIGR provided no information concerning chromosomal locations of other *gbpC* homologues. Therefore, we provisionally carried out PCR-based chromosomal walking of both upstream and downstream regions from the *gbpC* and *dbl* genes.

Materials and Methods

S. sobrinus strain 6715 was maintained and cultured in Todd-Hewitt (TH) broth. Isolation of chromosomal DNA was carried out as previously described ⁴⁾. A genome walking library was constructed based on nucleotide sequence information obtained from the *gbpC* and *dbl* gene regions with the Universal GenomeWalker Kit (BD Biosciences Clontech, Palo Alto, CA) as described previously ⁸⁾. Purification of amplified fragments and sequencing analyses were

carried out as described previously ⁵⁾.

We used BLAST programs to search the *S. sobrinus* 6715 genome database at TIGR (<http://www.tigr.org>) for identical nucleotide sequences, and preliminary sequence data obtained from the TIGR website at

http://tigrblast.tigr.org/ufmg/index.cgi?database=s_sobrinus%7

Cseq were used to compare with our determined sequences.

Gene fragments expressing the extracellular domains of the *S. sobrinus* GbpC homologues without the signal sequences predicted by the SignalP program 3.0 ⁹⁾ were amplified by PCR and then subcloned into pBAD/His vectors (Invitrogen) as described previously.

His-tagged GbpC homologue proteins were overexpressed in *E. coli* Top10, purified with Ni-Sepharose 6 Fast Flow resin (GE Healthcare Biosciences) and used for glucan-binding assays as described previously ⁵⁾.

The nucleotide sequence data reported in this communication will appear in the DDBJ, EMBL and GenBank nucleotide sequence databases under accession numbers AB281278, AB281280, AB294108, and AB294109.

Results and Discussion

The *S. sobrinus* 6715 *gbpC* gene encoded an intact 621 amino acid protein, which was the same size as that in strain 100-4 ⁵⁾. Differences between the two homologous sequences of strain 6715 and 100-4 were detected only at 6 nucleotide positions (99.7% identity) and 2 of these are non-synonymous substitutions, which

resulted in amino acid substitutions at L94 to S and L427 to H in the strain 6715 sequence.

687 bp downstream from the *gbpC* termination codon, another open reading frame (ORF) quite similar to the *gbpC* gene was found on the same strand. The entire sequence of this ORF was determined as described above and was designated as the *gbpC2* gene encoding a 632 amino acid protein, which was slightly larger than the GbpC protein. Homology, allowing for gaps between the *gbpC* and *gbpC2* gene sequences, was 54.9%, which was slightly less than that between the *gbpC2* and *S. mutans gbpC* genes (57.2%). In order to avoid confusion between the *S. mutans* and *S. sobrinus gbpC* genes, we would like to rename *S. sobrinus gbpC* gene to *gbpC1*.

The *dbl* gene sequence in strain 6715 was also determined and encoded an intact 1188-amino acid protein that was 82aa-smaller than the strain 100-4 homologous protein. The *dbl* gene contains two direct repeats (DR) comprised of the 168nt and 78nt DR-units near the 5' and 3' ends, respectively. The numbers of repeating units in both repeats (DR-5' and DR-3') in strain 6715 were one unit-smaller than those in strain 100-4. The DR regions and difference of the DR numbers were schematically indicated by the Harr plot analysis (Fig. 1). Both DR-5' and DR-3' regions were separately amplified from chromosomal DNA of 8 strains of *S. sobrinus* isolated in our institutions and the amplicons, together with those from strains 6715 and 100-4 as controls, were subjected to agarose gel electrophoretic analysis to estimate their sizes. Only one strain from the 8 isolates exhibited the same sizes of

the DR-5' and DR-3' amplicons as those amplified from strain 6715 and the other 7 strains appear to contain the same repeating structures as strain 100-4 in both DR-5' and DR-3' regions (data not shown). The strain 6715-type *dbl* gene may not belong to major gene population in *S. sobrinus* species.

Similar to *gbpC1* and *gbpC2* gene arrangements, another open reading frame (ORF) similar to the *dbl* gene was found to start at 505bp downstream from the *dbl* gene termination codon on the same strand. Therefore, we propose that the previously reported *dbl* gene is renamed *dblA* and that the downstream ORF is designated as the *dblB* gene. The *dblB* gene encoded an 1425 amino acid protein, which was 337aa larger than the DblA protein. Homology, allowing gaps between the *dblA* and *dblB* gene sequences, was 55.8%.

BLAST searches with these 4 *gbpC* homologue sequences against the *S. sobrinus* 6715 genome database at TIGR on the date of June 20, 2007 did not hit any completely identical sequences. Detected homologous sequence regions on the database revealed still 70-80% identities in each search.

Pairwise comparisons and phylogenetic tree construction using the four *S. sobrinus* GbpC homologue proteins with *S. mutans* GbpC protein are shown in Table 1 and Fig. 2, respectively. These results indicated that the *S. sobrinus* GbpC2 protein was evolutionally more related to *S. mutans* GbpC protein than the *S. sobrinus* GbpC1 protein.

Comparisons of other features of wall-anchored proteins are summarized in Table 2. These features are well conserved among these

proteins. Purified recombinant proteins corresponding to the extracellular domains of these *S. sobrinus* GbpC homologue proteins were subjected to alpha-1,6 glucan-binding assays as described in Materials and Methods. All of the GbpC2, DblA, and DblB recombinant proteins derived from strain 6715 exhibited alpha-1,6 glucan-binding activities (Fig. 3), in contrast to the *S. sobrinus* GbpC (GbpC1) protein defective in such activity as recently reported⁵⁾.

We have identified the GbpC protein solely involved in ddag of *S. mutans*⁴⁾, although three other proteins designated as GbpA, GbpB, and GbpD¹⁰⁾ exhibiting glucan-binding activity are encoded by the genome of this organism. In addition, the genome project of this organism revealed the absence of other *gbpC* gene homologues¹¹⁾. The *S. sobrinus* genome contained at least four *gbpC* gene homologues presumed to be duplicated from their antecedent genes in contrast to *S. mutans* genome. Epidemiological studies have indicated that *S. sobrinus* is more frequently isolated from highly caries-susceptible patients compared to *S. mutans*¹²⁾. The cariogenic potential in rats of *S. sobrinus* is also greater than that of *S. mutans*¹³⁾. These results may be related to the genome structure of this organism, which is highly adapted to sucrose-rich environments where multiple glycosidic linkage types of glucans are abundantly synthesized by mutans streptococci or other streptococci. Proteins encoded by duplicated genes have often expanded a repertoire of functions. *S. mutans* GtfB/C/D¹⁴⁾ and *S. gordonii* SspA/B¹⁵⁾ proteins may be such examples. The relative

affinities of the three glucan-binding proteins for alpha-1,6 glucan were slightly different. The GbpC2 protein appeared to exhibit higher affinity in the lower glucan concentration range than the other two proteins and the GbpC (GbpC1) protein exhibited no alpha-1,6 glucan-binding activity but bound Sephadex resin as reported recently ⁵⁾. Each protein may bind a preferable type of glucans. We could detect more Dbl (DblA) protein from the strain 100-4 cell wall but detected neither GbpC (GbpC1) nor Dbl (DblA) proteins in a wall-fraction prepared from strain 6715 with Western blot analyses using anti-GbpC (GbpC1) or anti-Dbl (DblA) sera. Expression or the repertoire of each protein may be different from strain to strain. However, more precise kinetic analyses will be necessary to clarify this hypothesis and such works are in progress in our laboratory.

We suggested that the *dbl* (*dblA*) gene in *S. sobrinus*, but not the *gbpC* (*gbpC1*), is a logical candidate responsible for the ddag phenotype in a recent report ⁵⁾. However, GbpC2 and DblB also exhibited alpha-1,6 glucan-binding activities. Therefore, the *gbpC2* and *dblB* genes are also logical candidates contributing to the ddag phenotype in strain 6715. To resolve this question, it would be ideal to introduce gene-directed mutagenesis in strains of *S. sobrinus*.

Acknowledgments

Preliminary sequence data was obtained from The Institute for Genomic Research website at <http://www.tigr.org>.

This investigation was supported in part by Grants-in-Aid for Scientific Research (17591925) from the Research Fund of MEXT, and in part by a Grant from The Tsuchiya Foundation for Culture and Science to Y.S

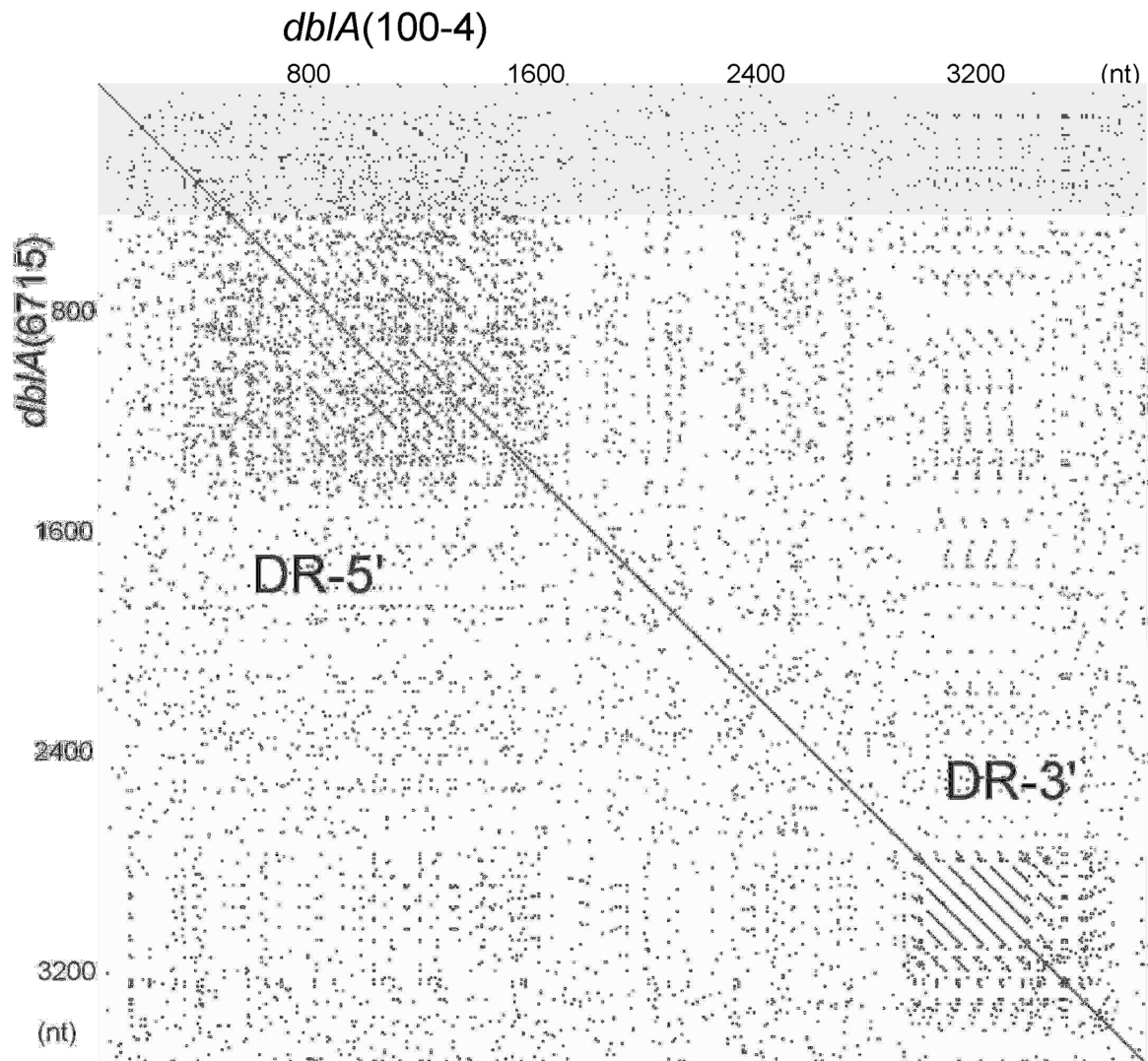


Fig. 1. The Harr plot analysis with the strains 100-4 and 6715 *dbIA* genes.

Unit size to compare was 10 nt and dot plot matching number was 8 nt.

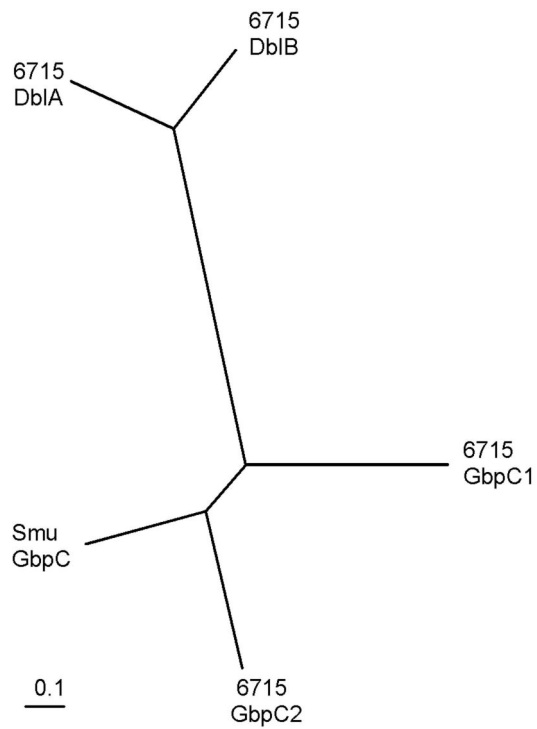


Fig. 2. Phylogenetic relationship among the *S. sobrinus* GbpC protein homologue sequences and *S. mutans* GbpC sequence.

Distances were calculated by the neighbor-joining (NJ) method using the DDBJ ClustalW program and the tree was drawn with the TreeView program ¹⁶⁾ using the calculated result.

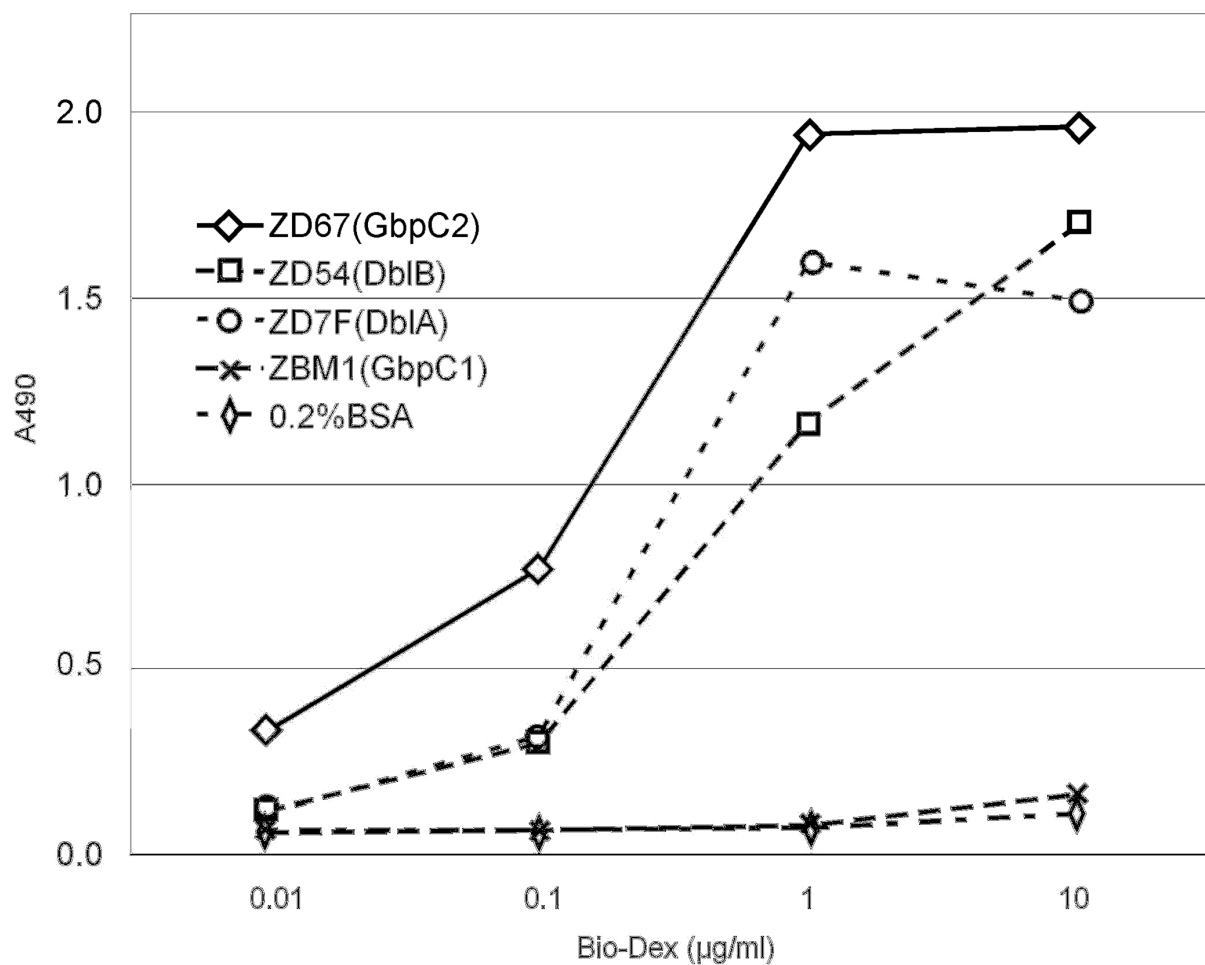


Fig. 3. Biotin-dextran binding by histidine-tagged derivatives of the *S. sobrinus* GbpC1, GbpC2, DblA, and DblB proteins.

GbpC1, GbpC2, DblA, and DblB proteins were purified as described in the text. BSA was used as a control. Higher A490 values represent higher binding activities. Wells of the HisSorb plates were coated with approximately 9 µg of protein. Each well was reacted with the indicated concentrations of biotin-dextran, washed, and detected as described previously. Data presented are averages from three independent determinations for each sample.

Table 1. Pairwise comparisons among the four *S. sobrinus* GbpC homologue proteins with the *S. mutans* GbpC protein

	(6715)GbpC1	(6715)GbpC2	(6715)DblA	(6715)DblB	SmuGbpC
(6715)GbpC1	100	39.7	35.4	27.1	42.8
(6715)GbpC2		100	27.4	26.4	45.5
(6715)DblA			100	47.3	43.8
(6715)DblB				100	29.9
SmuGbpC					100

Table 2. Comparison of features for wall-anchored proteins among the four *S. sobrinus* GbpC homologue proteins with the *S. mutans* GbpC protein

GbpC homologue proteins	Amino acid residue numbers	Putative signal peptide cleavage site*	LPXTG motif+ C-terminal residue numbers	Charged Tail
(6715)GbpC1	621	VLA-DE(28-29)	LPQTG+31aa	-K RKED
(6715)GbpC2	632	AFA-DD(28-29)	LPQTG+33aa	-R KMDENND
(6715)DblA	1188	VLA-DE(36-37)	LPQTG+31aa	-K KRQED
(6715)DblB	1425	VLA-DE(36-37)	LPQTG+32aa	-K KRKED
SmuGbpC	583	VFA-DE(26-27)	LPHTG+30aa	-R KKEN

*Predicted by the hidden Markov Model at SignalP website (<http://www.cbs.dtu.dk/services/SignalP/>).

References

1. Loesche, W.J.: Role of *Streptococcus mutans* in human dental caries. *Microbiol Rev* **50**: 353-380, 1986.
2. Drake, D., Taylor, K.G., Bleiweis, A.S. and Doyle, R.J.: Specificity of the glucan-binding lectin of *Streptococcus cricetus*. *Infect Immun* **56**: 1864-1872., 1988.
3. Wu-Yuan, C.D. and Gill, R.E.: An 87-kilodalton glucan-binding protein of *Streptococcus sobrinus* B13. *Infect Immun* **60**: 5291-5293, 1992.
4. Sato, Y., Yamamoto, Y. and Kizaki, H.: Cloning and sequence analysis of the *gbpC* gene encoding a novel glucan-binding protein of *Streptococcus mutans*. *Infect Immun* **65**: 668-675, 1997.
5. Kagami, A., Okamoto-Shibayama, K., Yamamoto, Y., Sato, Y. and Kizaki, H.: One of two *gbpC* gene homologues is involved in dextran-dependent aggregation of *Streptococcus sobrinus*. *Oral Microbiol Immunol* **in press**: 2007.
6. Ma, Y.S., Lassiter, M.O., Banas, J.A., Galperin, M.Y., Taylor, K.G. and Doyle, R.J.: Multiple glucan-binding proteins of *Streptococcus sobrinus*. *J Bacteriol* **178**: 1572-1577, 1996.
7. Navarre, W.W. and Schneewind, O.: Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol Mol Biol Rev* **63**: 174-229., 1999.
8. Okamoto-Shibayama, K., Sato, Y., Yamamoto, Y., Ohta, K. and Kizaki, H.: Identification of a glucan-binding protein C gene homologue in *Streptococcus macacae*. *Oral Microbiol Immunol* **21**: 32-41, 2006.
9. Nielsen, H., Engelbrecht, J., Brunak, S. and von Heijne, G.: Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* **10**: 1-6., 1997.
10. Banas, J.A., Russell, R.R. and Ferretti, J.J.: Sequence analysis of the gene for the glucan-binding protein of *Streptococcus mutans* Ingbritt. *Infect Immun* **58**: 667-673., 1990.
11. Ajdic, D., McShan, W.M., McLaughlin, R.E., Savic, G., Chang, J., Carson, M.B., Primeaux, C., Tian, R.Y., Kenton, S., Jia, H.G., Lin, S.P., Qian, Y.D., Li, S.L., Zhu, H., Najjar, F.,

- Lai, H.S., White, J., Roe, B.A. and Ferretti, J.J.: Genome sequence of *Streptococcus mutans* UA159, a cariogenic dental pathogen. Proc. Natl. Acad. Sci. USA **99**: 14434-14439, 2002.
12. Hirose, H., Hirose, K., Isogai, E., Miura, H. and Ueda, I.: Close association between *Streptococcus sobrinus* in the saliva of young children and smooth-surface caries increment. Caries Res **27**: 292-297, 1993.
 13. de Soet, J.J., van Loveren, C., Lammens, A.J., Pavicic, M.J., Homburg, C.H., ten Cate, J.M. and de Graaff, J.: Differences in cariogenicity between fresh isolates of *Streptococcus sobrinus* and *Streptococcus mutans*. Caries Res **25**: 116-122, 1991.
 14. Ooshima, T., Matsumura, M., Hoshino, T., Kawabata, S., Sobue, S. and Fujiwara, T.: Contributions of three glycosyltransferases to sucrose-dependent adherence of *Streptococcus mutans*. J Dent Res **80**: 1672-1677, 2001.
 15. Eglund, P.G., Du, L.D. and Kolenbrander, P.E.: Identification of independent *Streptococcus gordonii* SspA and SspB functions in coaggregation with *Actinomyces naeslundii*. Infect Immun **69**: 7512-7516, 2001.
 16. Page, R.D.M.: TREEVIEW: An application to display phylogenetic trees on personal computers. Computer Applications in the Biosciences **12**: 357-358, 1996.