Oral Microbiology Immunology 2007: 22: 240–247
Printed in Singapore. All rights reserved

ORAL MICROBIOLOGY
AND IMMUNOLOGY

One of two *gbpC* gene homologues is involved in dextran-dependent aggregation of *Streptococcus sobrinus*

A. Kagami, K. Okamoto-Shibayama, Y. Yamamoto, Y. Sato, H. Kizaki
Department of Biochemistry and Oral Health Science Centre, Tokyo Dental College, Mihama-ku, Chiba City, Japan

Kagami A, Okamoto-Shibayama K, Yamamoto Y, Sato Y, Kizaki H. One of two gbpC *gene homologues is involved in dextran-dependent aggregation of* Streptococcus sobrinus. *Oral Microbiol Immunol 2007: 22: 240–247.* © 2007 The Authors. Journal compilation © 2007 Blackwell Munksgaard.

Introduction: *Streptococcus sobrinus* exhibits marked dextran-dependent aggregation mediated by glucan-binding proteins (GBPs). In contrast to *Streptococcus mutans*, in which the *gbpC* gene responsible for dextran-dependent aggregation of this organism has been characterized, genes encoding the *S. sobrinus* GBPs have not yet been identified.
Methods: Recently, we identified the *gbpC* gene homologue from *Streptococcus macacae* using polymerase chain reaction primers based on the conserved regions of the *gbpC* sequence exhibiting intraspecies variations. This method was applied to amplify a *S. sobrinus* homologue.
Results: Unexpectedly, two *gbpC* gene homologues were identified in *S. sobrinus* strain 100-4. One homologue, named *gbpC*, was more similar to the *S. mutans gbpC* gene than the other and was approximately half the molecular size of its homologue with similar regions interrupted by several non-similar stretches. However, the dextran-binding activity of the protein expressed from *gbpC* in *Escherichia coli* was not detected in contrast to the other homologue, a protein designated as Dbl, expressing this activity. The *gbpC* gene was shown to be intact on the chromosome of strain OMZ176, which does not exhibit dextran-dependent aggregation, while the *dbl* gene of this strain contained a single adenine nucleotide insertion at approximately one-third the distance from the 5′-end. The insertion mutation in the *dbl* gene resulted in translation of a premature protein missing its LPXTG sequence signature sequence of the wall-anchored proteins.
Conclusion: These results suggest that the *dbl* gene is very likely responsible for *S. sobrinus* dextran-dependent aggregation.

Key words: aggregation; *dbl* gene; *gbpC* homologues; glucan-binding; *Streptococcus sobrinus*

Yutaka Sato, Department of Biochemistry and Oral Health Science Centre, Tokyo Dental College, 2-2, Masago 1-chome, Mihama-ku, Chiba City, 261
Japan
Tel.: +81 43 270 3750;
fax: +81 43 270 3752;
e-mail: yusato@tdc.ac.jp
Accepted for publication October 11, 2006

Streptococcus sobrinus and *Streptococcus mutans*, two species of the mutans streptococci, are established as two major pathogens associated with human dental caries; they reside in oral biofilms (dental plaque) and attached to tooth surfaces (11). Many genes encoding putative virulence factors involved in dental caries, including surface protein antigens or glucosyltransferases, have been identified from both *S. sobrinus* and *S. mutans* in studies carried out during the past two decades (10). These studies confirmed the glucan-binding properties of glucosyltransferases and led to the isolation of non-glucosyltransferase glucan-binding proteins (GBPs) (1). In contrast to glucosyltransferases, non-glucosyltransferase GBP genes responsible for dextran-dependent aggregation in *S. sobrinus* have not yet been identified, although five proposed GBPs have been reported (24), with two of them, GBP2 (26) and GBP4 (12), as candidate proteins for this phenomenon. Dextran-dependent aggregation was originally reported for *S. sobrinus* and *Streptococcus criceti* strains and has been regarded as specific for these two species (6). Epidemiological studies have indicated that *S. sobrinus* is more frequently isolated from highly caries-susceptible patients than *S. mutans* (7). In addition, the cariogenic potential in rats of *S. sobrinus* is greater than that of

S. mutans (5). Therefore, it is important to characterize the genes responsible for dextran-dependent aggregation in *S. sobrinus. Streptococcus sanguinis* also expresses glucosyltransferase activity, being able to synthesize glucans that are rich in α-1,6 linkages (9). Colonization by the sanguinis group of streptococci on human tooth surfaces precedes that by the mutans streptococci (3). Therefore, when cells of mutans streptococci colonize the dental plaque biofilm, cell-associated GBPs of the mutans streptococci are likely to be important in allowing these organisms to adhere (13, 22) to preformed plaque, mediated by glucans that were synthesized following colonization by the antecedent sanguinis group.

Four GBPs of *S. mutans* have been designated in the order of their discovery and the genes responsible for these GBPs were designated as *gbpA*, *gbpB*, *gbpC* and *gbpD* (2, 4, 14, 21, 23). Among the *S. mutans* GBPs, only the GbpC protein was associated with the cell wall by covalent linkage mediated by the extracellular sortase enzyme (17), as a wall-anchored protein (an LPXTG-motif protein). This wall-associated GbpC was regarded as one of the virulence factors of this organism (16) and was solely involved in dextran (α-1,6 glucan)-dependent aggregation (21). Therefore, the present consensus regarding aggregation is that *S. sobrinus* aggregates more readily than *S. mutans* and the conditions for aggregation of *S. mutans* are more limited than those for *S. sobrinus*, e.g. cells grown under certain defined stress conditions (1).

Our earlier Southern hybridization analysis with a *S. mutans gbpC* gene fragment probe against the chromosomal DNA of the other six species of mutans streptococci revealed no cross-hybridization of the probe with *S. sobrinus* chromosomal DNA (unpublished results), suggesting that *S. sobrinus* contains no *gbpC* gene homologues, although this might have resulted from the limitations of Southern hybridization analysis when dealing with species with different G + C contents. In contrast to the results with *S. sobrinus*, the same analysis suggested that *Streptococcus macacae* might contain a *gbpC* gene homologue. Recently, we identified a *gbpC* gene homologue in *S. macacae*, and subsequent Southern hybridization analysis unexpectedly exhibited cross-hybridization of the *S. macacae gbpC* gene homologue probe with all other mutans streptococcal DNA (19).

In this communication, we describe the presence of two *gbpC* gene homologues on the *S. sobrinus* chromosome. Furthermore, one, but not the other, is probably involved in the dextran-dependent aggregation of this species.

Materials and methods

Bacterial strains and plasmids

S. sobrinus strain 100-4 (serotype *d/g*, a clinical isolate of our laboratory) was used for the identification and cloning of the *gbpC* gene homologues; the serotype of this strain was confirmed with antiserum kindly provided by K. Fukushima, Nihon University, School of Dentistry at Matsudo, Japan. *S. sobrinus* strain OMZ176 (serotype *d*) was provided by K. Fukushima 25 years ago and has been kept frozen in our laboratory; it was used as a reference for sequence analysis of the homologues. The streptococci were maintained and cultured in Todd–Hewitt broth/agar plates. *Escherichia coli* strain TOP10, obtained from a commercial supplier (Invitrogen, Carlsbad, CA), was used as a host for plasmid pBAD/His and its derivatives.

Amplification of *gbpC* homologues from the *S. sobrinus* chromosome by the polymerase chain reaction

The polymerase chain reaction (PCR) primers used in this study are listed in Table 1. Primers YF2 and YR2 were the same as those used in the previous report (19) to amplify an internal fragment of the *S. macacae gbpC* homologue as an initial amplicon. This initial PCR amplification was carried out with high-fidelity *PfuTurbo* hot start DNA polymerase (Stratagene Corp., La Jolla, CA). A genome-walking library was constructed based on nucleotide sequence information obtained from the initial amplicons with the Universal GenomeWalker Kit (BD Biosciences Clontech, Palo Alto, CA) as described previously (19).

Nucleotide sequencing and sequence analysis

The amplified fragments were purified with a QIAquick PCR Purification Kit (Qiagen GmbH, Hilden, Germany) or with a QIAquick Gel Extraction Kit (Qiagen).

Table 1. Primers used in this study

ATGAGTCTCTCATCGTTAAAYCAYTGGAC	(YF2)
GTCATAGCACCTGCACCRTARTA	(YR2)
gbpC	
ATAGTGATGGTGAAAATGGTT	(SOF1)
GTTAAGCAAACAGCCGATAACA	(SOF2)
CTGAACTGACACAAGCCGTT	(SOF5)
GCCAATAAGGCACAAGCC	(SOF6)
CTCTGACACTTCTTGGAGGCA	(SOF7)
AGAAGTTGTACGTTTTTCACCATC	(SIR1)
TGGTGAGTAGATAGCACCATCAGCAT	(SOR2)
GAGTCTTCGTTAGGCGCAT	(SOR3)
CATCTGCTTCTTCAACTGTATCG	(SOR4)
ATCTCGAGGATGAAGCTACAGACACAAGTGC	(SOFFuXho)
ATGAATTCATGTTTGTGGCAGTACTTGAGC	(SORFuEco)
dbl	
GATGCCAATACGCATGAAGCA	(SsadF1)
GAGCCGCCTAAGATGACAA	(SsadF2)
CAGTGCCACCTAAGACAATTACA	(SsadF3)
TGACAAGGCTAAGTATGATGCTAA	(SsadF4)
AGGCTGCCTATGATAAGGCT	(SsadF5)
GAACAGCAAAATAATTTGGCTATT	(SsadF6)
AACCAATCAGCTGCTGATTCT	(SsadF7)
ACCACCGAAGTCAGCATAACCA	(SsadR1)
GACACTTTCAACGACAATGTTCT	(SsadR2)
TGTTTGGGCTCAGTGGGT	(SsadR3)
CATAAGGGCTTTCAGAAACGA	(SsadR4)
CAACTTCGTACTTAGCCAAGTCTT	(SsadR5)
TTGCAGCTGCATTTGCTT	(SsadR6)
TACCGGCACCACCGTAATA	(SsadR7)
AGACTATCTTAGTCTTCTTGGCGT	(SsadR8)
GCGCTTCAGGAGTCAGATAA	(SsadR9)
CTTCATCAGCTGCTTGGACA	(SsadR10)
GTACTTAGCCAAGTTTTGGT	(SsadR11)
TTGTACTTAGCCAATGCTTG	(SsadR12)
ATCTCGAGTGACGAACAAACCAACCAATCAG	(SsadFFuXh)
ATCTGCAGCCAGTTTGTGGCAAGCTTTCTT	(SsadRFuPs)

The corresponding regions were directly sequenced with a BigDye Terminator Cycle Sequencing FS Ready Reaction Kit using the primers listed in Table 1 and an ABI PRISM Genetic Analyzer 3130Avant (Applied Biosystems, Foster City, CA). The nucleotide sequences of both strands of the 2-kilobase (kb) and 4-kb fragments encompassing the two *S. sobrinus gbpC* gene homologues were then determined. We used BLAST programs to search the *S. sobrinus* 6715-genome database at The Institute for Genomic Research (TIGR) (http://www.tigr.org) for the identical nucleotide sequences, and preliminary sequence data were obtained from the TIGR website at http://tigrblast.tigr.org/ufmg/index.cgi?database=s_sobrinus%7Cseq. BLAST, FASTA and the CLUSTALW programs at the International DNA databases [European Molecular Biology Laboratory (EMBL), GenBank and DNA Data Bank of Japan (DDBJ)] were also used for similar amino acid sequence analysis and for multiple alignment of homologous sequences. The other sequence analyses were carried out with DNASIS-MAC (Hitachi Software Engineering, Yokohama, Japan) and GENETYX-MAC (Genetyx corporation, Tokyo, Japan).

Cloning and expression of *S. sobrinus* GbpC homologue proteins

Gene fragments of 1.7 and 3.6 kb expressing the extracellular domains of the two *S. sobrinus* GbpC homologues without the signal sequences predicted by the SIGNALP program 3.0 (http://www.cbs.dtu.dk/services/SignalP/) (18) were amplified by PCR using the forward and reverse primer sets (SOFFuXho and SORFuEco; SsadFFuXh and SsadRFuPs) (TABLE 1). PCR amplifications were carried out with high-fidelity DNA polymerases [either *PfuTurbo* hot start DNA polymerase (Stratagene Corp) or KOD-Plus-DNA polymerase (TOYOBO, Osaka, Japan)]. The amplified fragments were then ligated in-frame to the 5′ histidine-tag region of expression vectors pBAD/HisA or pBAD/HisB (Invitrogen). Transformation with *E. coli* strain TOP10, selection and analyses of the resulting clones, induction of homologue proteins, and the preparation of crude extracts by ultrasonication were carried out as described previously (19). Clones for the two GbpC proteins, designated respectively as ZBM1 and ZCQ1, were used throughout these experiments.

Protein sample preparation, sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE) and Western blot analysis

The two *S. sobrinus* GbpC proteins were overexpressed from *E. coli* clones ZBM1 and ZCQ1, respectively. Following disruption of the cells by sonication, His-tagged proteins were purified with Ni-Sepharose 6 Fast Flow resin (Amersham Biosciences Corp., Piscataway, NJ) according to the supplier's recommendation. The ZBM1 cells disrupted by sonication were also subjected to incubation overnight with 10 ml Sephadex G100 resin equilibrated with phosphate-buffered saline and the resin was collected from a glass-filter. Following washes with phosphate-buffered saline on the glass-filter, proteins bound to the Sephadex resin were eluted with 15 ml 3 M guanidine–HCl. Eluates were then dialysed, lyophilized and rehydrated with Milli-Q water.

Preparation of wall protein extracts from *S. sobrinus* strains was carried out as described previously (15) with a slight modification. A method to prepare concentrated *S. sobrinus* GBPs from culture supernatants was simplified from that reported by Smith et al. (24). Briefly, *S. sobrinus* cells were grown overnight in a screw-capped bottle containing 200 ml Todd–Hewitt broth supplemented with 0.3 weight/volume% glucose. After neutralization, culture supernatants were obtained by centrifugation, and then incubated for 4 h with 8 ml Sephadex G100 resin equilibrated with phosphate-buffered saline. Following washes with phosphate-buffered saline on a glass-filter, GBPs were eluted, dialysed, lyophilized and rehydrated with 200 µl Milli-Q water as described above. Proteins in the supernatant of a *S. sobrinus* strain OMZ176 BTR broth (20) culture were also conventionally precipitated with 65% saturated ammonium sulphate and were concentrated 1000-fold. These streptococcal sample extracts and the proteins from the induced *E. coli* cells harbouring the overexpression plasmids or vector controls described above were mixed with SDS sample buffer and frozen until electrophoretic analysis using the Laemmli-gel system.

Following SDS–PAGE, the separated proteins were electrophoretically transferred (Horizblot AE6675, ATTO, Tokyo, Japan) to a polyvinylidene fluoride membrane (Immobilon-P^{TM}, Millipore Corporation, Bedford, MA) and Western blot analysis was carried out as described previously (19) with either Penta-His HRP Conjugate antibody (Qiagen), anti-GBP2 kindly provided by D. J. Smith (Forsyth Institute), or anti-Dbl(ZCQ1) serum, which was prepared using the His-tagged protein purified from ZCQ1 cells described above injected into a female Japanese white rabbit (Operon Biotechnologies Inc., Tokyo, Japan).

Glucan-binding assay and detection of dextran-dependent aggregation

Glucan-binding assays of the two GbpC homologue proteins respectively purified from induced ZBM1 and ZCQ1 cells were carried out as described previously (19) using biotin-dextran (dextran, biotin, molecular weight 70,000, lysine fixable; Molecular Probes Europe BV, the Netherlands), streptavidin biotinylated horseradish peroxidase complex (Amersham Biosciences Corp.), and the Ni-NTA-coated 96-well HisSorb Plates (Qiagen) to immobilize the His-tagged proteins. Absorbance at 490 nm, resulting from the horseradish peroxidase reaction with its substrate, was measured in a microplate reader (BioRad MICROPLATE READER model 3550).

Dextran-dependent aggregation was observed visually as described previously (19).

Nucleotide sequence accession number

The nucleotide sequence data reported in this communication will appear in the DDBJ, EMBL and GenBank nucleotide sequence databases under accession numbers AB237533, AB237534, AB237535 and AB241126.

Results

Amplification of a *gbpC* gene homologue in *S. sobrinus* strain 100-4

Southern hybridization analysis of *S. sobrinus* chromosomal DNA probed with a *S. macacae gbpC* gene homologue fragment in our recent report (19) suggested that *S. sobrinus* also contains a homologous gene. We initially attempted PCR amplification using *S. sobrinus* chromosomal DNA as template with two primer sets with which the *S. macacae gbpC* partial fragments could be amplified. Amplification with one primer set was negative but PCR with the other primer set (YF2 and YR2) revealed an approximately 0.3-kb major band and a 0.7-kb faint band (Fig. 1A) following agarose gel electrophoresis of the amplification mixture. The 0.3-kb fragment (YF2–YR2) was excised from the gel, purified and directly

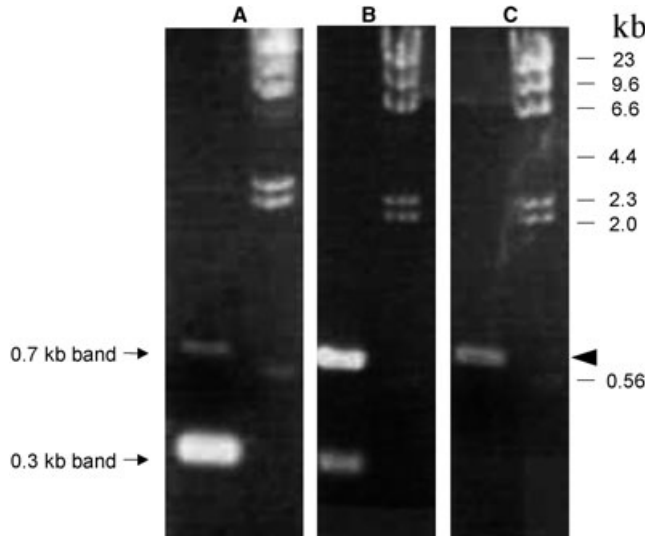

Fig. 1. Electropherograms of PCR amplification products using primers YF2 and YR2. The initial attempt to amplify gbpC homologues using *Streptococcus sobrinus* chromosomal DNA as template revealed multiple-sized products (A). A 0.3-kb band was excised, purified and directly sequenced. This fragment was part of the *S. sobrinus gbpC* gene. A 0.7-kb band was subjected to two rounds of an isolation procedure including excision from the gel and purification for use as a template in the subsequent PCR as described in the text (B, C). Finally, a 0.7-kb single band (arrow head) was obtained and sequenced. This 0.7-kb fragment was part of the second *gbpC* homologue designated as *dbl*.

sequenced. A 266-base-pair (bp) nucleotide sequence was determined and had 62.0% and 66.2% homologies with the corresponding 260-bp sequences within the *S. mutans gbpC* gene and the *S. macacae* homologue, respectively. Therefore, we carried out PCR-based chromosomal walking, as described in the Materials and methods, to obtain upstream and downstream regions. Nucleotide sequencing of these regions revealed a 1863-bp open reading frame beginning with an ATG and terminating with a TGA codon. This open reading frame encodes a putative 621-amino-acid protein with a calculated molecular mass of 66 kDa, which is similar to that (67 kDa) of the *S. macacae* GbpC protein. A FASTA homology search at DDBJ with the amino acid sequence deduced from the *S. sobrinus gbpC* gene homologue showed the *S. macacae* and *S. mutans* GbpC protein sequences as the highest (Opt. 1553) and second highest (Opt. 1383) similar sequences, respectively. However, the overall per cent identity of the sequences between *S. sobrinus* and *S. macacae* (43.1% identity in a 638-amino-acid overlap) was slightly lower than that between *S. sobrinus* and *S. mutans* (44.1% identity in a 630-amino-acid overlap). A phylogenetic tree constructed with these three sequences, which was similar to that constructed with 16S ribosomal RNA sequences of three species (data not shown), indicated that the *S. sobrinus gbpC* sequence is genetically more distant than the other two. A homology search against the TIGR Microbial database for the unfinished *S. sobrinus* (strain ATCC6715) sequence showed a 1793-bp nucleotide sequence region with 78% identity. This region probably corresponds to the homologue of *gbpC* gene in strain ATCC6715. However, a 1401-bp sequence with 76% identity was also shown in the same search.

Detection and amplification of a second *gbpC* gene homologue in *S. sobrinus* strain 100-4

We have described above that we also observed faint amplified bands in the agarose gel electropherogram of the same sample from which the ZBM1 *gbpC* gene homologue was identified. Since we predicted that these faint bands might be a part of another homologue, the 0.7-kb fragment (Fig. 1A) was excised from the gel, purified and used as a template for subsequent PCR with the same primer set. This amplification still revealed two (0.7 and 0.3 kb) bands (Fig. 1B) and an additional round of excision, purification as a template, and amplification was performed (Fig. 1C). The resultant amplification product was sequenced and subjected to a homology search. As a result, a 689-bp nucleotide sequence encoding a 229-amino-acid sequence was detected and the C-terminal two-thirds (amino acids 88–229) of the amino acid sequence was similar to a part (amino acids 288–428) of the *S. mutans* GbpC protein sequence with a 34% identity and 50% similarity. Therefore, we carried out PCR-based chromosomal walking and sequencing as described above. Nucleotide sequencing of the region unexpectedly revealed a long open reading frame of 3810 bp, beginning with an ATG and terminating with a TAA codon. This open reading frame encodes a 1270-amino-acid protein with a calculated molecular mass of 139 kDa. A BLAST homology search at GenBank with the amino acid sequence deduced from this *gbpC* gene homologue showed the *S. mutans* and *S. macacae* GbpC protein sequences as the most similar sequences but arranged in an interrupted manner. Therefore, this open reading frame encodes a previously unidentified and novel gene and was designated as *dbl* (dextran-binding lectin) to avoid confusion with the *gbpC* homologue characterized above. The signature sequence (-LPXTG-) for wall-anchored proteins of gram-positive bacteria followed by a charged tail sequence was detected in Dbl as LPQTG in the C-terminal region (1235–1239). Similar regions of the *S. sobrinus* Dbl protein corresponding to regions of the *S. mutans*, *S. macacae* and *S. sobrinus* GbpC proteins are shown in Fig. 2. A group of surface protein antigens, but not GBPs, was detected with a FASTA homology search as similar proteins to the Dbl protein. *Streptococcus downei* antigen I/II and the *Streptococcus gordonii* SspA protein are also included in Fig. 2. Two potential candidates for a Pribnow box, TTGGCT-N17-TATAAT and TGGCTA-N16-TATAAT, were detected 72 bp upstream of the *dbl* coding sequence. A potential ribosome-binding site (GGAGG) was found 10 bp upstream of the ATG initiation codon of the *dbl* coding sequence and an inverted repeat was detected 50 bp downstream of the TAA termination codon with – 130·2 kJ of free energy per mol (25). This suggests that the *dbl* gene in *S. sobrinus* is monocistronic, similar to the *S. mutans* and *S. macacae gbpC* genes. A homology search with the 3810-bp *dbl* coding region against the TIGR *S. sobrinus* (ATCC6715) Microbial database showed a 2708-bp nucleotide sequence region with 74% identity.

Glucan-binding assays of proteins encoded by the *S. sobrinus gbpC* and *dbl* genes

The purified *S. sobrinus* GbpC and Dbl proteins expressed in *E. coli* (clones

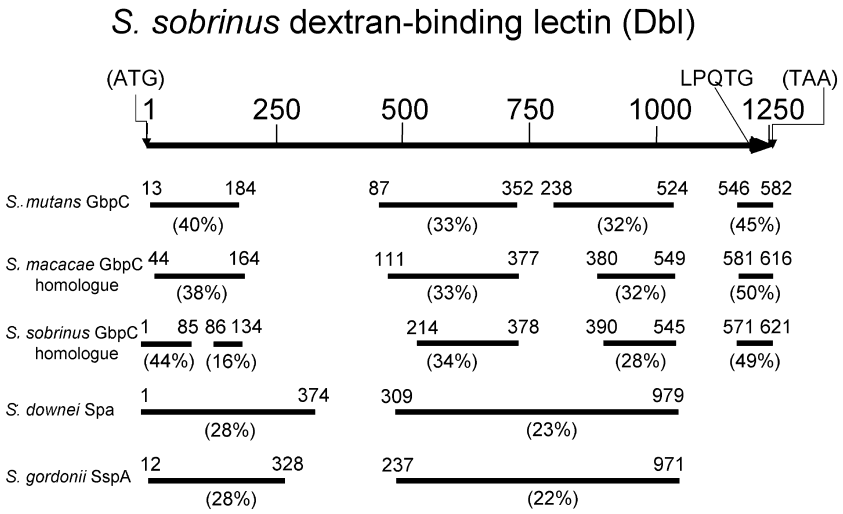

Fig. 2. Similar regions of the *Streptococcus sobrinus* Dbl protein sequence compared with three GbpC proteins from mutans streptococci and two streptococcal surface protein antigens. Similar regions of the *S. sobrinus* Dbl protein are depicted as solid interrupted lines. Numbers in parentheses are per cent similarities of the *S. sobrinus* Dbl partial region to corresponding protein regions and the numbers on both ends of the solid lines represent amino acid numbers of those corresponding to the similar regions.

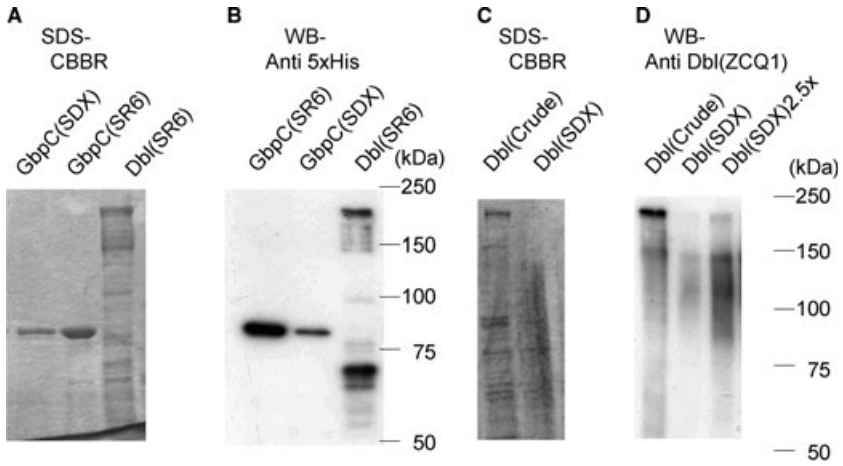

Fig. 3. The results of the SDS–PAGE and Western blotting of GbpC and Dbl proteins purified from induced *Escherichia coli* cells with Penta-His HRP Conjugate antibody. (A, C) SDS–PAGE electropherogram stained with Coomassie Brilliant Blue R (CBBR). (B) The Western blot analysis with Penta-His HRP Conjugate antibody. (D) The Western blot analysis with anti-Dbl(ZCQ1) serum. Abbreviations as GbpC(SR6) and Dbl(SR6) respectively, represent GbpC proteins purified with Ni-Sepharose 6 Fast Flow resin. GbpC(SDX) and Dbl(SDX) respectively, represent GbpC and Dbl proteins isolated with Sephadex resin. Dbl (Crude) represents induce *E. coli* ZCQ1 crude extract. Amounts of GbpC(SR6), GbpC(SDX), and Dbl(SR6) proteins applied to the gels for CBBR staining (A) were 867, 733 and 1670 ng, respectively, and 1/100 amounts of those proteins were used for the Western blot analysis (B). Amounts of Dbl (Crude) and Dbl(SDX) proteins applied to the gels for CBBR staining (C) were 3 and 2.9 μg, respectively, and amounts of Dbl (Crude), Dbl(SDX) and, Dbl(SDX)2.5x proteins for the Western blot analysis (D) were 30, 29 and 72.5 ng.

ZBM1 and ZCQ1, respectively) were prepared as described in the Materials and methods. The approximate calculated sizes (70 and 140 kDa) of the predicted extracellular domains without the signal sequences were confirmed with SDS–PAGE and Coomassie Brilliant Blue R staining as well as Western blot analysis using Penta-His HRP Conjugate antibody (Fig. 3A,B). Both the GbpC and Dbl proteins migrated larger (80 and 200 kDa) than their calculated molecular sizes, as does the *S. mutans* GbpC protein. The glucan-binding assays for the GbpC and Dbl proteins were then performed, using bovine serum albumin as a negative

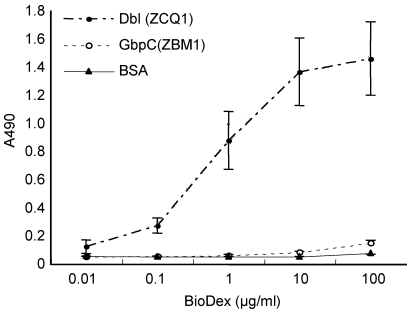

Fig. 4. Biotin-dextran binding by histidine-tagged derivatives of *Streptococcus sobrinus* two GbpC homologue proteins. GbpC and Dbl proteins were purified as described in the text. Bovine serum albumin was used as a control. Higher values for absorbance at 490 nm (A490) equate to higher binding activities. Wells of HisSorb plates were coated with approximately 7.8 μg GbpC or 3.0 μg Dbl proteins. Numbers of GbpC protein molecules per well used for coating were approximately five-fold more than those of Dbl protein molecules. Each well was reacted with the indicated concentrations of biotin–dextran, washed and detected as described in the text. Data presented are averages and standard deviations from two duplicated independent determinations for each sample.

control. The GbpC protein unexpectedly exhibited no α-1,6 glucan-binding activity. In contrast, the Dbl protein exhibited significant glucan-binding activity dependent on dextran concentrations (Fig. 4). We also determined whether or not the GbpC protein expressed α-1,3 glucan-binding activity as described previously (19) because *S. sobrinus* synthesizes water-insoluble glucan containing relatively long chains consisting of α-1,3 glycoside linkages. However, the results were uniformly negative (data not shown). In spite of negligible α-1,6 and α-1,3 glucan-binding activities for the GbpC protein, when overexpressed ZBM1 cells disrupted by sonication were subjected to incubation with Sephadex resin, the GbpC protein was recovered from the resin as a His-tagged protein (Fig. 3A,B). In this respect, the GbpC protein behaved as a GBP. In contrast, the His-tagged Dbl protein was not efficiently recovered as Sephadex-binding protein (Fig. 3C,D) when overexpressed ZCQ1 cells were subjected to the same procedure as described above for ZBM1 cells.

Sequence analyses of the *gbpC* and *dbl* genes in *S. sobrinus* strain OMZ176

Strain 100-4 was selected for these experiments because this strain exhibited the highest and most stable dextran-dependent aggregation phenotype in our culture

collection. In contrast to strain 100-4, OMZ176 did not exhibit dextran aggregation following growth under any of the growth conditions that were examined. Therefore, we expected that strain OMZ176 might have undergone a mutation in a gene responsible for the aggregation phenotype. The *gbpC* and *dbl* genes in strain OMZ176 were amplified and sequenced with primers listed in Table 1 for comparison to the corresponding genes of strain 100-4. The OMZ176 *gbpC* gene homologue encoded an intact 621-amino-acid protein, as did the 100-4 homologue. Differences in these two homologous sequences only occurred in five nucleotides (99.7% identity) and three of those were non-synonymous, which resulted in amino acid substitutions at L93 to S, N102 to D, and E571 to G in the OMZ176 sequence. In contrast to the *gbpC* sequences, a frame-shift mutation was found in the *dbl* sequence of OMZ176 resulting from a single adenine nucleotide insertion into the heptad adenine nucleotides located between the 1723 and 1729 nucleotide region. This resulted in a premature termination of Dbl protein translation at amino acid 585.

Western blot analysis of GBPs expressed in *S. sobrinus*

There are problems with transforming *S. sobrinus* that preclude the performance of gene-knockout experiments. Therefore, we employed *S. sobrinus* strain OMZ176, which does not exhibit dextran-dependent aggregation. This strain was not expected to express a Dbl protein tethered to its cell wall because a truncation of this protein resulted in the loss of its LPXTG motif. To confirm this, wall extracts from both strains 100-4 and OMZ176 and a culture supernatant concentrated with ammonium sulphate from strain OMZ176 were examined by Western blot analysis using anti-Dbl(ZCQ1) serum. A strong positive band around 170 kDa was observed in a wall extract from strain 100-4 in contrast to no corresponding band in strain OMZ176 (Fig. 5A). A predicted truncated Dbl protein with a calculated molecular size of 61.4 kDa from OMZ176 was detected in a culture supernatant concentrated with ammonium sulphate as an approximately 65-kDa band. These results support the conclusion that the Dbl protein is involved in dextran-dependent aggregation of this species. We also examined expression of the GBP2 protein in culture supernatants concentrated with Sephadex resin from both strains, because the *S. sobrinus* GBP2 (87 kDa GBP) was suggested to be involved in dextran-dependent aggregation of this species (26). Anti-GBP2 serum reacted with an approximately 85-kDa protein in culture supernatants with the same intensities for the two strains (Fig. 5B). This result suggests that GBP2 is not involved in dextran-dependent aggregation.

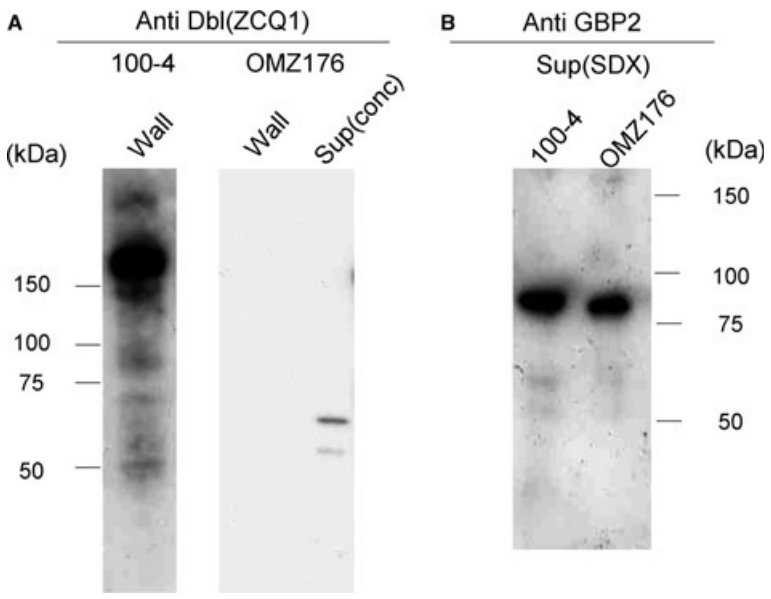

Fig. 5. Western blot analyses of proteins from strains 100-4 and OMZ176 using anti-Dbl(ZCQ1) (A) and anti-GBP2 (B) sera. Abbreviations as wall and sup respectively represent wall extracts and concentrated culture supernatants prepared using Sephadex resin (SDX) or using ammonium sulphate (conc) as described in the text. Amounts of sample proteins applied to the gels were 200 ng wall extracts and 100 ng supernatants from each strain analysed with anti-Dbl(ZCQ1) (A) or anti-GBP2 (B) sera.

Discussion

Glucosyltransferases from mutans streptococci synthesize glucans consisting of α-1,3 and α-1,6 glycosidic linkages and are also members of the GBP family. However, one of the non-glucosyltransferase GBPs, GbpC protein, was found to be solely involved in dextran-dependent aggregation in *S. mutans* (21). *S. sobrinus* exhibits robust dextran-dependent aggregation. Therefore, the presence of a causative gene for this phenomenon has long been predicted in this species. Based on our recent Southern hybridization results suggesting that *S. sobrinus* might have a *gbpC* homologue, we carried out PCR amplification and identified two. One, designated as the *dbl* gene, encoded a protein exhibiting α-1,6 glucan-binding activity, while the other, designated as the *gbpC* gene, did not (Fig. 4). The *S. sobrinus dbl* gene encoded a wall-anchored protein possessing a LPXTG motif (Fig. 2) by which the extracellular sortase enzyme tethers the protein to cell wall peptidoglycans. Igarashi et al. (8) demonstrated that an *S. mutans* sortase enzyme mediated the cell-wall-anchoring of the GbpC protein and that a sortase mutant of this organism lost the dextran-dependent aggregation phenotype. Therefore, with respect to cell association and dextran-binding, the *S. sobrinus dbl* gene provides the necessary conditions for a candidate gene responsible for dextran-dependent aggregation of this organism. In addition, the *dbl* gene in strain OMZ176 contained a single adenine nucleotide insertional mutation, resulting in translation of a premature protein missing the LPXTG motif and, as a consequence, this strain did not exhibit dextran-dependent aggregation. Therefore, the *dbl* gene is very likely responsible for this phenotype, although we cannot exclude the possibility that another gene involved in dextran-mediated aggregation was inactivated in strain OMZ176.

In contrast to the GbpC and Dbl proteins covalently linked to the cell wall, the GBP2 protein was mainly isolated from the culture supernatant (24, 26). Therefore, GBP2 apparently lacks an apparent mechanism for associating with cells for involvement in dextran-dependent aggregation. Our data indicated that GBP2 was detected at similar levels in the supernatants of both aggregation-positive and aggregation-negative strains (100-4 and OMZ176). Furthermore, Ma et al. (12)

also reported that both spontaneous and ethyl methane sulphonate-induced *S. sobrinus* mutants deficient in glucan-induced aggregation retained the approximately 87-kDa protein corresponding to GBP2. However, GBP2 was reported to be missing in a spontaneous aggregation negative mutant that was isolated (26). These contradictory findings will be finally resolved following completion of the *S. sobrinus* genome project currently in progress at TIGR, as well as further genetic analyses.

The Dbl protein showed similarity to the *S. mutans*, *S. macacae* and *S. sobrinus* GbpC proteins, although some gaps were present in the homologous regions. This protein exhibited similarity also to a group of surface protein antigens including *S. downei* antigen I/II and the *S. gordonii* SspA protein with a FASTA homology search (Fig. 2). The *dbl* gene may be the evolutionary product of recombination of antigen I/II and *gbpC* genes yielding the chimeric gene structure.

The greater divergence of the *S. sobrinus* GbpC protein sequence compared to the divergence of the other two homologues in *S. mutans* and *S. macacae* may suggest that the *S. sobrinus* GbpC protein possibly evolved another function. The *S. sobrinus* GbpC protein bound to neither α-1,3 nor α-1,6 glucans but it did bind to Sephadex resin composed of cross-linked α-1,6 glucan. This may indicate that the *S. sobrinus* GbpC protein has no affinity for linear glucan molecules but does have affinity for branched glucans, e.g. those consisting of α-1,3 and α-1,6 glycosidic linkages, while the Dbl protein may preferably bind to linear glucan molecules. It is therefore of interest to determine the physiological function of this unique character of the *S. sobrinus* GbpC protein. This protein may play a role in the attachment of this organism into *in vivo* plaque biofilms, where branched insoluble glucans are present, having been synthesized by antecedent mutans streptococci cells in the presence of sucrose. A glucan-binding lectin with a molecular size of 60 kDa isolated by a Sephadex affinity column method (12) has been designated GBP4 (24). It was reported that this protein appeared to have been responsible for *S. sobrinus* glucan-induced aggregation (12). The *gbpC* gene may encode GBP4 because of the similarities of molecular size and the amino acid contents of these two proteins. However, a question as to whether or not the *gbpC* homologue encodes GBP4 still remains to be answered.

Homology searches with the *gbpC* and *dbl* genes against the incomplete TIGR *S. sobrinus* (ATCC6715) database showed 78% identical 1814-bp and 74% identical 2708-bp nucleotide sequence regions, respectively. These identical values were much lower than those for the *gbpC* gene homologues of strains 100-4 and OMZ176 (99.7%), or for those in the *gtfI* genes retrieved from the GenBank databases between strains ATCC6715 (accession no. D90213) and OMZ176 (accession no. D13858) (98.2%). Therefore, our targeted regions of the TIGR database may contain incomplete sequence information. Alternatively, the *S. sobrinus* 6715 genome may contain several *gbpC* gene homologues.

In conclusion, we have identified two *gbpC* homologues in *S. sobrinus.* Candidate proteins responsible for the dextran-dependent aggregation phenotype of *S. sobrinus* were reported a decade ago. However, no genes responsible for this phenotype have been identified until now. We now suggest that the novel *dbl* gene in *S. sobrinus*, but not the *gbpC* homologue, is a logical candidate responsible for the dextran-induced aggregation phenotype.

Acknowledgments

We thank D. J. Smith (Forsyth Institute, Boston, MA) for providing anti-GBP2 serum; K. Fukushima (Nihon University School of Dentistry at Matsudo, Matsudo, Japan) for providing *S. sobrinus* serotype-specific antiserum and strain OMZ176; and T. Iijima for his technical assistance in sequencing the OMZ176 *gbpC* gene. Preliminary sequence data were obtained from The Institute for Genomic Research website at http://www.tigr.org. We also thank H. K. Kuramitsu (State University of New York at Buffalo, NY) for critically reading the manuscript. This investigation was supported in part by 'High-Tech Research Centre' Project for Private Universities: matching fund subsidy from MEXT (Ministry of Education, Culture, Sports, Science and Technology), 2001–2005, 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.

References

1. Banas JA, Vickerman MM. Glucan-binding proteins of the oral streptococci. Crit Rev Oral Biol Med 2003: **14**: 89–99.
2. Banas JA, Russell RR, Ferretti JJ. Sequence analysis of the gene for the glucan-binding protein of *Streptococcus mutans* Ingbritt. Infect Immun 1990: **58**: 667–673.
3. Caufield PW, Dasanayake AP, Li Y, Pan Y, Hsu J, Hardin JM. Natural history of *Streptococcus sanguinis* in the oral cavity of infants: evidence for a discrete window of infectivity. Infect Immun 2000: **68**: 4018–4023.
4. Chia JS, Chang LY, Shun CT, Chang YY, Tsay YG, Chen JY. A 60-kilodalton immunodominant glycoprotein is essential for cell wall integrity and the maintenance of cell shape in *Streptococcus mutans*. Infect Immun 2001: **69**: 6987–6998.
5. de Soet JJ, van Loveren C, Lammens AJ et al. Differences in cariogenicity between fresh isolates of *Streptococcus sobrinus* and *Streptococcus mutans*. Caries Res 1991: **25**: 116–122.
6. Drake D, Taylor KG, Bleiweis AS, Doyle RJ. Specificity of the glucan-binding lectin of *Streptococcus cricetus*. Infect Immun 1988: **56**: 1864–1872.
7. Hirose H, Hirose K, Isogai E, Miura H, Ueda I. Close association between *Streptococcus sobrinus* in the saliva of young children and smooth-surface caries increment. Caries Res 1993: **27**: 292–297.
8. Igarashi T, Asaga E, Sato Y, Goto N. Inactivation of *srtA* gene of *Streptococcus mutans* inhibits dextran-dependent aggregation by glucan-binding protein C. Oral Microbiol Immunol 2004: **19**: 57–60.
9. Kopec LK, Vacca Smith AM, Wunder D, Ng-Evans L, Bowen WH. Properties of *Streptococcus sanguinis* glucans formed under various conditions. Caries Res 2001: **35**: 67–74.
10. Kuramitsu HK. *Streptococcus mutans*: molecular genetic analysis. In: Fischetti VA, Novick RP, Ferretti JJ, Portnoy DA, Rood JI, ed. Gram-positive pathogens. Washington, DC: American Society for Microbiology, 2000: 280–286.
11. Loesche WJ. Role of *Streptococcus mutans* in human dental caries. Microbiol Rev 1986: **50**: 353–380.
12. Ma YS, Lassiter MO, Banas JA, Galperin MY, Taylor KG, Doyle RJ. Multiple glucan-binding proteins of *Streptococcus sobrinus*. J Bacteriol 1996: **178**: 1572–1577.
13. Matsumoto M, Fujita K, Ooshima T. Binding of glucan-binding protein C to GTFD-synthesized soluble glucan in sucrose-dependent adhesion of *Streptococcus mutans*. Oral Microbiol Immunol 2006: **21**: 42–46.
14. Mattos-Graner RO, Jin S, King WF, Chen T, Smith DJ, Duncan MJ. Cloning of the *Streptococcus mutans* gene encoding glucan binding protein B and analysis of genetic diversity and protein production in clinical isolates. Infect Immun 2001: **69**: 6931–6941.
15. Murakami Y, Nakano Y, Yamashita Y, Koga T. Identification of a frameshift mutation resulting in premature termination and loss of cell wall anchoring of the PAc antigen of *Streptococcus mutans* GS-5. Infect Immun 1997: **65**: 794–797.
16. Nakano K, Matsumura M, Kawaguchi M et al. Attenuation of glucan-binding protein C reduces the cariogenicity of *Streptococcus mutans*: analysis of strains isolated from

human blood. J Dent Res 2002: **81**: 376–379.
17. Navarre WW, Schneewind O. Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. Microbiol Mol Biol Rev 1999: **63**: 174–229.
18. Nielsen H, Engelbrecht J, Brunak S, von Heijne G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 1997: **10**: 1–6.
19. Okamoto-Shibayama K, Sato Y, Yamamoto Y, Ohta K, Kizaki H. Identification of a glucan-binding protein C gene homologue in *Streptococcus macacae*. Oral Microbiol Immunol 2006: **21**: 32–41.
20. Sato Y, Yamamoto Y, Kizaki H, Kuramitsu HK. Cloning of the phospho-beta-galactosidase gene in *Escherichia coli* from lactose-negative mutants of *Streptococcus mutans* isolated following random mutagenesis with plasmid pVA891 clone banks. FEMS Microbiol Lett 1992: **70**: 219–224.
21. Sato Y, Yamamoto Y, Kizaki H. Cloning and sequence analysis of the *gbpC* gene encoding a novel glucan-binding protein of *Streptococcus mutans*. Infect Immun 1997: **65**: 668–675.
22. Sato Y, Senpuku H, Okamoto K, Hanada N, Kizaki H. *Streptococcus mutans* binding to solid phase dextran mediated by the glucan-binding protein C. Oral Microbiol Immunol 2002: **17**: 252–256.
23. Shah DS, Russell RR. A novel glucan-binding protein with lipase activity from the oral pathogen *Streptococcus mutans*. Microbiology 2004: **150**: 1947–1956.
24. Smith DJ, King WF, Wu CD, Shen BI, Taubman MA. Structural and antigenic characteristics of *Streptococcus sobrinus* glucan binding proteins. Infect Immun 1998: **66**: 5565–5569.
25. Tinoco I, Jr, Borer PN, Dengler B et al. Improved estimation of secondary structure in ribonucleic acids. Nat New Biol 1973: **246**: 40–41.
26. Wu-Yuan CD, Gill RE. An 87-kilodalton glucan-binding protein of *Streptococcus sobrinus* B13. Infect Immun 1992: **60**: 5291–5293.