INFECTION AND IMMUNITY, May 2005, p. 3147–3151 Vol. 73, No. 5
0019-9567/05/$08.00+0 doi:10.1128/IAI.73.5.3147–3151.2005


# A Hypothetical Protein of *Streptococcus mutans* Is Critical for Biofilm Formation

Thomas A. Brown, Jr., Sang-Joon Ahn, Roslyn N. Frank, Yi-Ywan M. Chen, José A. Lemos, and Robert A. Burne*

*Department of Oral Biology, University of Florida College of Dentistry, 1600 SW Archer Rd., Gainesville, Florida 32610-0424*

Received 30 July 2004/Returned for modification 11 October 2004/Accepted 4 January 2005

**Inactivation of the Smu0630 gene of *Streptococcus mutans* resulted in dramatic decreases in biofilm formation, regardless of the carbohydrate source. The Smu0630 protein contained numerous interesting features, including a possible signal sequence and two conserved regions of repeated sequences. Smu0630 may represent a potential target for novel therapeutics.**

*Streptococcus mutans* exists almost exclusively in oral biofilms and is considered the primary etiological agent of human dental caries. Establishment and persistence in the human oral cavity present significant challenges to oral organisms because of the plethora of immune and nonimmune defenses that the host brings to bear against the organisms. Mechanical forces, antimicrobial systems, and rapid and severe changes in nutrient availability, pH, and other environmental conditions mandate that oral bacteria have multiple complex pathways by which they adhere, accumulate, and tolerate significant stresses if they are to persist and to initiate diseases. The ability to survive these environmental challenges and to establish themselves as members of stable oral biofilm communities is the primary requirement for the persistence and cariogenicity of *S. mutans*.

Adhesion is the initial step in the formation of biofilm communities. SpaP is a major surface antigen of *S. mutans* that has been shown to be important in saliva-mediated aggregation and adherence (11). Another primary mechanism for adherence of *S. mutans* is the production of glucan homopolymers from sucrose via glucosyltransferases (GTFs) (10, 17). The GTFs, in concert with glucan binding proteins (GBPs), contribute in major ways to initial adherence and to the formation of biofilms (4).

Beyond initial adherence, it appears that a variety of genes are required for the proper maturation of biofilms formed by *S. mutans* and other oral streptococci. By the use of specific- and random-mutagenesis strategies, many different types of genes that are required for these organisms to transition from adherent microcolonies to complex, three-dimensional biofilms have been identified. These include those for intercellular communication systems and environmental sensing systems, components of the general stress response pathway involved in protein repair and turnover, global regulators of carbohydrate metabolism, and adhesion-promoting genes (for examples, see references 3, 5, 12, 13, and 18). Interestingly, many of the genes identified to affect biofilm formation affect the expression of a large panel of genes, many of which are either unidentified or have no known function. Characterization of these genes of unknown function is generally recognized as an essential step toward fully understanding the biology of the host organisms and for establishing potential targets for novel and broadly effective therapeutics against *S. mutans* and other pathogens.

The *S. mutans* UA159 genome has been sequenced (1), and an annotated genome is available from the Oral Pathogens Sequence database (http://www.stdgen.lanl.gov/oragen). There are a large number of hypothetical proteins of as-yet-unknown function present in the annotated genome. We hypothesized that among these proteins, there are surface proteins that play an important role in adhesion and maturation of biofilms. We used a variety of computer algorithms to identify candidate proteins that could be necessary for biofilm formation. Specifically, the database was searched for proteins with a predicted signal sequence or transmembrane domains. From this group of hypothetical proteins, we selected those with molecular masses of greater than 20 kDa and pI values ranging from 4.0 to 10.0. From this subset of genes, we selected four open reading frames (ORFs) that fit the search criteria and had the ProDom database or the Pfam database provide hits that suggested which proteins could play a role in promoting adherence through adhesin presentation or facilitating binding to carbohydrates or proteins. The first ORF, Smu0176, was 39.8 kDa and had a predicted pI of 7.3 and a predicted signal sequence. Smu1484 was identified as a putative *ppiA* homologue because of its similarity to the same gene in *Streptococcus pyogenes* and in *Streptococcus pneumoniae*. This gene encodes a presumptive peptidyl-prolyl isomerase with a molecular mass of 28.6 kDa and a predicted pI of 10.7 that was identified as a possible lipoprotein. Smu1543 was identified to have three transmembrane domains and a cleavable signal sequence. The predicted mass and pI were 49.7 kDa and 4.2, respectively, and the protein was flagged as a possible hemolysin. One locus, designated Smu0630, was of interest due to its size (107 kDa), slightly acidic pI (5.70), and similarity to two ORFs that are annotated as predicted glucan-binding proteins (Smu0018 and Smu0760).

The Smu0630, Smu0176, Smu1543, and Smu1484 genes

* Corresponding author. Mailing address: Department of Oral Biology, University of Florida College of Dentistry, 1600 SW Archer Rd., Gainesville, FL 32610-0424. Phone: (352) 392-4370. Fax: (352) 392-7357. E-mail: rburne@dental.ufl.edu.

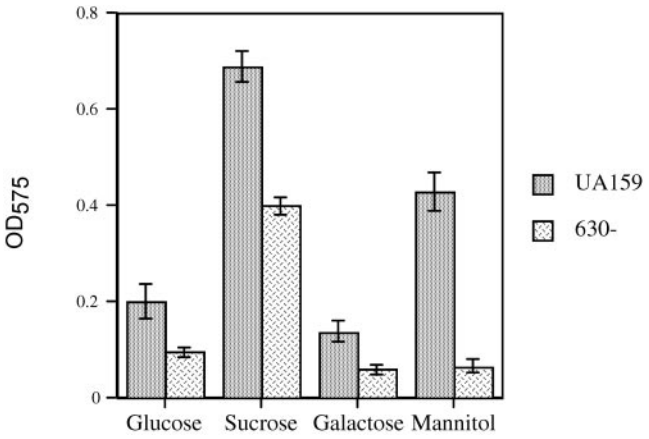


FIG. 1. Biofilm assays of cells grown in BM supplemented with different sugars on the surfaces of 96-well microtiter plates. Biofilms were grown and quantified as described in the text. Results shown are averages of three separate experiments, and error bars indicate the standard deviations. Differences were statistically significant as assessed by the *t* test: glucose, $P = 0.0006$; sucrose, $P = 0.0224$; galactose, $P = 0.0004$; mannitol, $P = 0.0018$. $OD_{575}$, optical density at 575 nm; 630-, Smu0630.

were amplified by PCR from the chromosomal DNA of strain UA159 (6) and insertionally inactivated by a polar (ΩKm) kanamycin resistance gene (15) using natural transformation of *S. mutans* (16). The mutant strains showed no differences in growth rate in brain heart infusion broth compared to the wild-type strain (data not shown). However, the Smu0630 mutant strain exhibited extensive clumping in the bottom of the test tubes during growth, in contrast to the wild-type strain, which produced a confluent culture. Additionally, the Smu0630 mutant formed longer chains than did the parent during planktonic or biofilm growth.

We assessed the biofilm-forming capacity of the mutants in relation to UA159 in BM (13) supplemented with different sugars. Biofilm assays were performed as previously described (5) with minor modifications. Biofilms that had been grown on the surfaces of 96-well polystyrene microtiter plates for 24 h were used in this assay. The stained biofilms were eluted once with 200 μl of 80% ethanol–20% acetone and diluted to a final volume of 1 ml with water before being read at 575 nm. The ORF Smu0176, Smu1484, and Smu1543 mutants formed biofilms as well as the parental strain under the conditions tested (data not shown). Thus, no further experiments were conducted using these three mutants. The results of the biofilm experiments using strain UA159 and the Smu0630 mutant are shown in Fig. 1. Under the conditions tested, the ability of the mutant to form biofilms was greatly reduced compared to the wild type. This trend was evident for BM containing glucose, sucrose, mannitol, or galactose as the primary carbohydrate source. In some ways, this finding was particularly surprising for sucrose. To the best of our knowledge, and with the exception of strains with defects in the GTFs or GBPs, mutants of *S. mutans* that show defects in biofilm formation in BM with glucose generally retain the ability to form biofilms efficiently in the presence of sucrose.

In addition to a microtiter plate assay, we investigated biofilms grown on the surfaces of hydroxylapatite (HA) disks using scanning electron microscopy (SEM). Biofilms were allowed to form on HA disks for 24 h and were processed as previously described (12). All images were produced at the University of Florida EM core laboratory (Biotechnology Program, University of Florida, Gainesville, FL). The results were consistent with those of the microtiter plate biofilm assays. The wild-type strain formed abundant biofilms, whereas the Smu0630 mutant was severely impaired in biofilm formation in BM supplemented with either glucose or sucrose (Fig. 2), although the defect was more apparent when glucose was the carbohydrate source. SEM of biofilms grown in sucrose revealed that the Smu0630 mutant formed more biofilms than it did in glucose, although it was still impaired compared to the parent.

The Smu0630 gene region is displayed in Fig. 3. The Smu0630 gene was inactivated 590 base pairs from the start of the gene. A gene coding for a hypothetical protein (Smu0631) is located 24 bp downstream from the stop codon of Smu0630. Smu0631 has a predicted molecular mass of 21.5 kDa and a pI of 9.10 and is annotated as a possible lipoprotein. Due to the very small intergenic region between Smu0630 and Smu0631, it is likely that the two genes are cotranscribed. For this reason, an Smu0630 mutant was constructed using a nonpolar kanamycin resistance gene (NP-Km) (9). In addition, the Smu0630 mutation was complemented in both the polar and nonpolar Smu0630 mutants by expressing Smu0630 in *trans* on plasmid pMSP3535 as was previously done for complementation of the *S. mutans relA* gene (12). The complementing plasmid contained the intact Smu0630 gene, including the putative promoter region of Smu0630. Biofilm assays with microtiter plates were repeated using the polar Smu0630 mutant (Smu630P), the nonpolar Smu630 mutant (Smu630NP), and the Smu630-complemented strains for the nonpolar insertion (SAB40) and the polar insertion (SAB41). In BM medium with glucose and sucrose, the nonpolar mutant showed reductions in biofilm formation similar to those of the polar mutant (Fig. 4). The expression of Smu0630 in both mutant strains restored the capacity to form biofilms (Fig. 4). Similar results were obtained with mannitol and galactose as the primary carbohydrate source, and the strains no longer formed clumps in the bottoms of the test tubes or formed long chains (data not shown). Given that the behavior of the polar and nonpolar mutants was the same and that complementation of both mutants by Smu630 alone was sufficient to restore the noted phenotypes, it is reasonable to conclude that the loss of Smu630 is responsible for the biofilm formation defect. However, these results do not exclude the possibility that there is some association between the products of these two ORFs.

Using available protein computer analysis tools, several interesting features of the Smu0630 protein were discovered (Fig. 5). PSORT analysis indicated a predicted signal peptide cleavable at an alanine residue at position 18. With a ProDom search, amino acid residues 774 to 886 placed Smu0630 in glycoside hydrolase family 25. The only representative enzyme of this family of glycosylhydrolases is lysozyme, suggesting the potential for a role in cell wall binding. Alignment of Smu0630 with two other ORFs predicted as glucan binding proteins (Smu0018 and Smu0760) revealed that the similarities are based primarily on conservations in the N-terminal region of the protein. Moreover, Smu0630 does not contain the YG

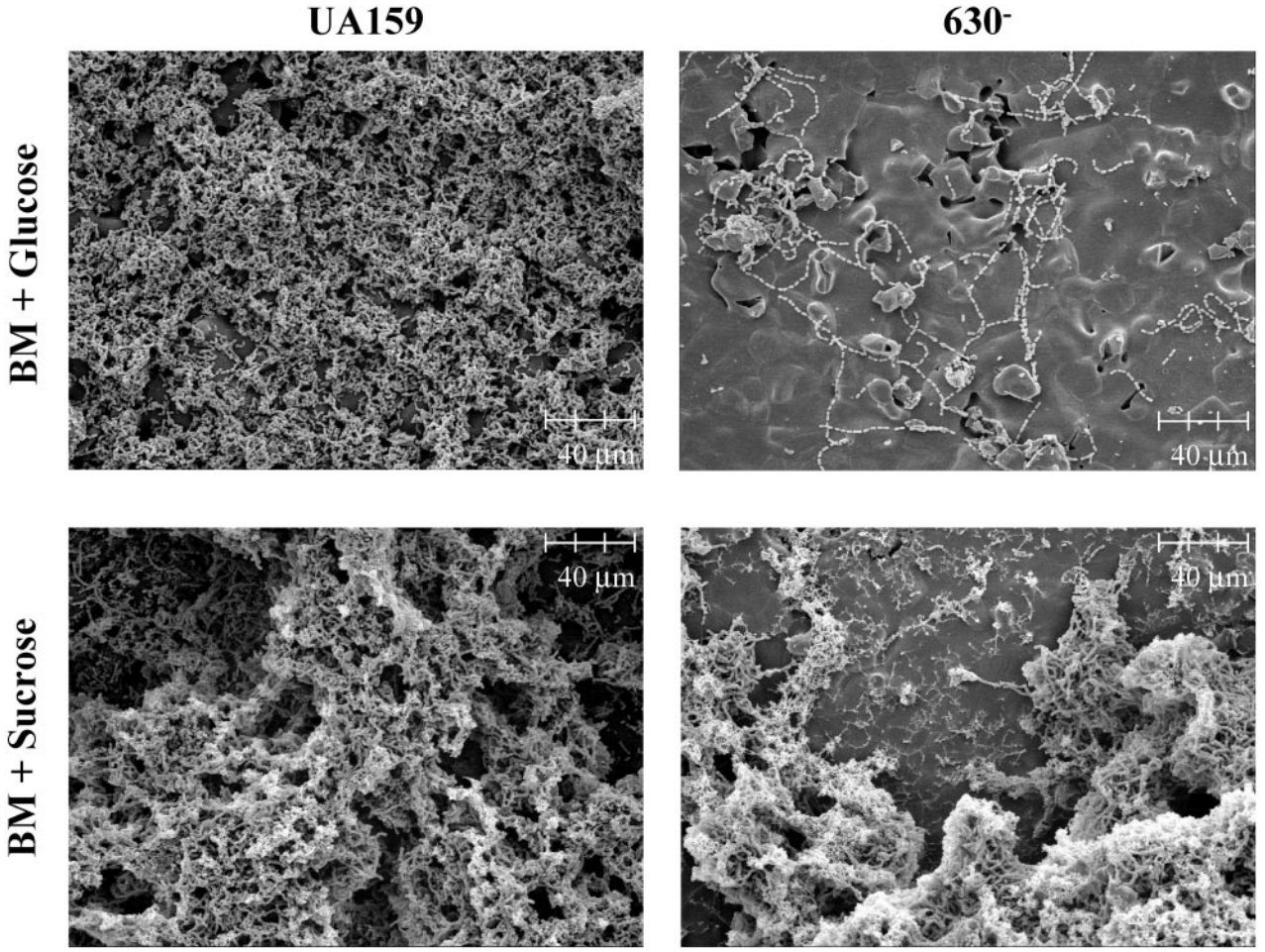


FIG. 2. SEM of biofilms grown on the surfaces of HA disks in BM supplemented with glucose or sucrose. Biofilms were grown as described in the text. All images shown were taken at ×1,500 magnification. The selected images were chosen as the best representatives of the amount of biofilm on the surface of the HA disk. 630-, Smu0630.

repeats present in true glucan binding proteins (7). Also of note was the 28% similarity of amino acid residues 12 to 186 of Smu0630 to a streptococcal hemagglutinin domain.

One of the most notable features of the protein is the occurrence of conserved repeated domains localized from amino acid positions 208 to 347 and positions 502 to 64. Using MEME analysis (2), we discovered that the central portion of the protein contains a series of repeated amino acid sequences. Three separate conserved sequences are arranged in a specific order with identical amounts of space between the sequences in these two regions (Fig. 5B). Sequence I is 20 amino acids long and is followed by a nonconserved region of 10 amino acids and then by the 31-amino-acid sequence III. After a nonconserved gap of 30 amino acids, another sequence I is repeated and then come 4 amino acids, followed by the 25-amino-acid sequence II. The orientation of sequences and the spacing between them are the same for the two regions, which start at amino acids 208 and 502.

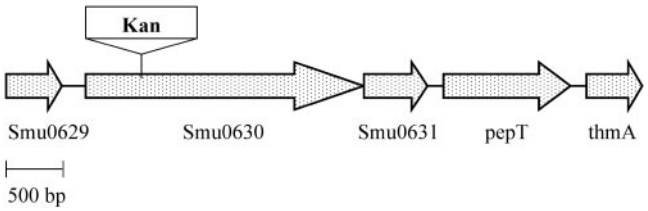


FIG. 3. Schematic diagram of the Smu0630 region of the *S. mutans* UA159 chromosome showing surrounding genes and the site of inactivation by a kanamycin resistance gene. Smu0629 codes for a conserved hypothetical protein showing no significant homologies. Smu0630 is a hypothetical protein showing some homology to the GBPs GbpB and GbpC. Smu0631 is a hypothetical protein showing no significant homologies. PepT is identified as peptidase T. ThmA is identified as a pore-forming protein.

By performance of BLAST searches using one of the I-III-I-II regions, a total of four proteins, in addition to Smu0630, were found to contain these motifs; two were from *S. mutans*, one was from *Streptococcus gordonii*, and one was from *Streptococcus agalactiae* (Table 1). It is important to note that the Bsp protein of *S. agalactiae* contains only the I-III region and is not followed by a I and II repeat. The amounts of space between the repeats of the I-III-I-II region are similar in all of the proteins. In order to derive a consensus sequence, the I, II, and III repeats were aligned (Fig. 5B). The most conserved of

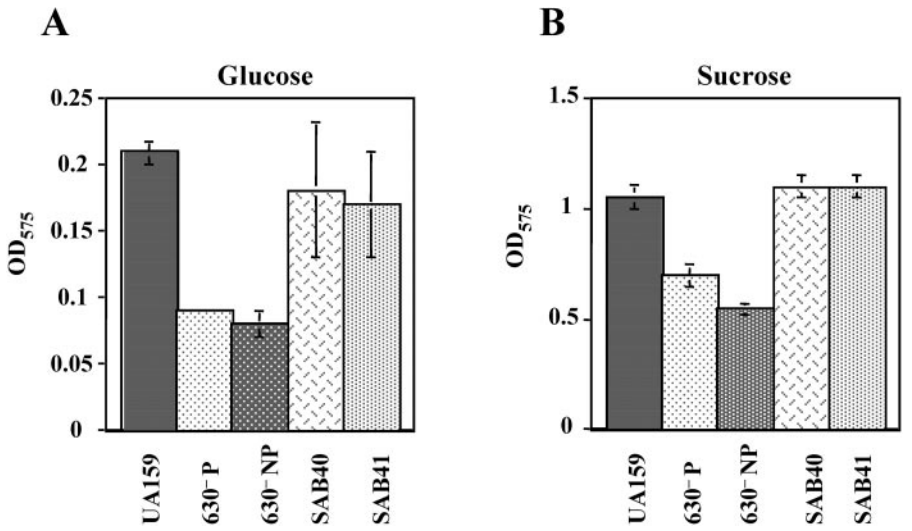


FIG. 4. Biofilm assays of cells grown in BM supplemented with glucose (A) or sucrose (B) on the surfaces of 96-well microtiter plates. Bars: UA159, wild type; 630-P, Smu0630 polar mutant; 630-NP, Smu0630 nonpolar mutant; SAB40, complemented 630-NP; and SAB41, complemented 630-P. Results shown are averages of three separate experiments, and error bars indicate the standard deviations. $OD_{575}$, optical density at 575 nm.

the three repeats is I, with an average of 13 identical amino acids out of the 20-amino-acid sequence. With the exception of repeat III, the sizes of the repeats are constant in all five proteins. In Smu0555, ORF O, and Bsp, repeat III is two amino acids larger than in Smu0630 and Smu0760.

In conclusion, a previously uncharacterized protein of *S. mutans* that is important in both sucrose-dependent and sucrose-independent biofilm formation has been identified and contains conserved, reiterated sequences. The importance of the I-III-I-II motif to the function of the protein is unknown

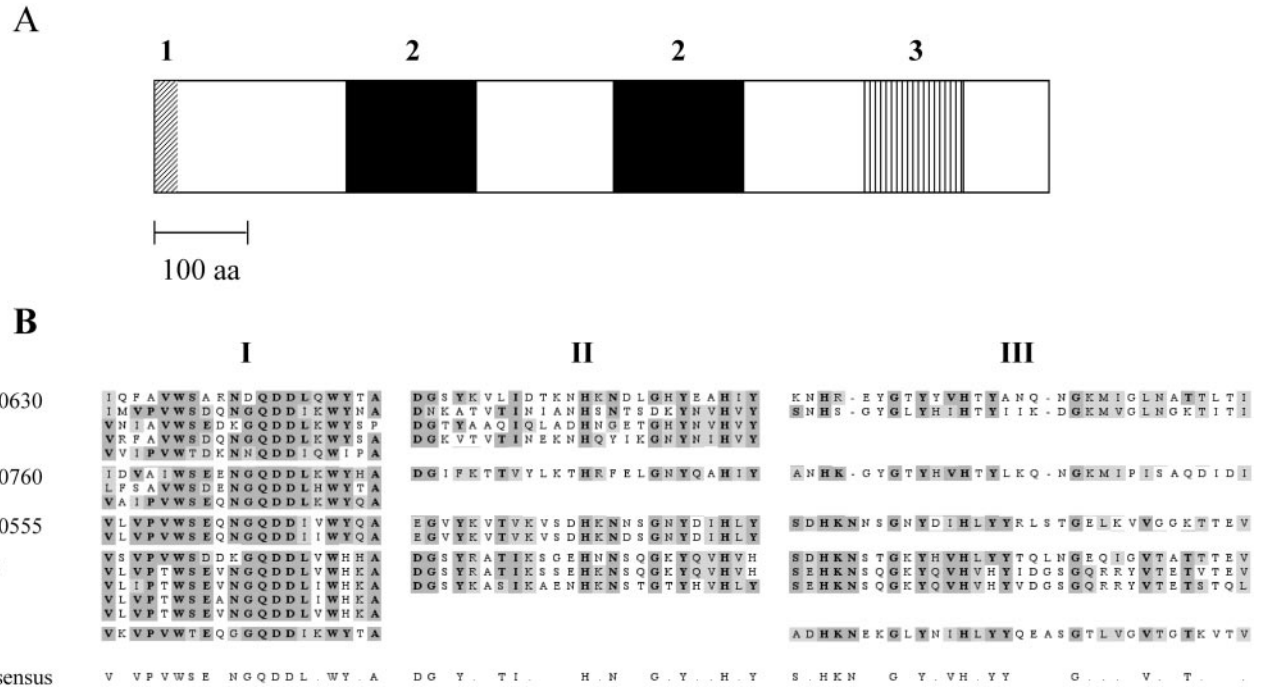


FIG. 5. (A) Diagram of Smu0630 indicating regions of interest. Region 1 is a cleavable signal peptide at an alanine residue at position 18. The shaded regions indicated by region 2 represent a series of repeats at amino acid positions 208 to −347 and 502 to −641. Region 3 spans amino acid residues (aa) 774 to 886 and places the Smu0630 protein in glycoside hydrolase family 25 (lysozyme). (B) ClustalW alignment of the I, II, and III repeats present in five genes in *Streptococcus* species. The regions from each protein were aligned in order to derive a consensus sequence for the region. Dark shading with bold print indicates an identical residue, light shading with normal print indicates a similar residue, and no shading indicates no similarities. Dashes are inserted in the alignment to fill in gaps. In the consensus sequence line, a period indicates that a single amino acid cannot be assigned to that position, but there is similarity in the types of amino acids found at that position.

TABLE 1. Summary of proteins containing the I-III-I-II repeat motif, including accession number and function

| Organism | Gene | Accession no. | Function of protein | No. of repeats |
|---|---|---|---|---|
| *S. mutans* UA159 | Smu0630 | NP_721116.1 | Hypothetical | 2 |
| *S. mutans* UA159 | Smu0760 | BAC75703.1 | Hypothetical | 1 |
| *S. mutans* UA159 | Smu0555 | NP_721041.1 | Putative cell wall protein precursor | 1 |
| *S. gordonii* 38 | *orfO* | AAN64579.1 | Putative *N*-acetylmuramidase | 3 |
| *S. agalactiae* NEM316 | *bsp* | AJ305309.1 | Group B *Streptococcus* secreted protein | NA[a] |

[a] NA, not applicable. The *bsp* gene contains only the I-III region of the I-III-I-II repeat.

but is worthy of further investigation since repeated regions in surface proteins of gram-positive pathogens are of primary importance in function, structural integrity, and immunogenicity (3, 4, 8, 14). It is also of interest that these motifs found in the Smu630 protein appear exclusively in streptococcal species. Should these regions prove critical to adherence and biofilm maturation, the repeat region may prove to be a desirable target for therapeutics that disrupt biofilm formation.

This work was supported by NIH-NIDCR award DE13239.

## REFERENCES


1. **Ajdic, D., W. M. McShan, R. E. McLaughlin, G. Savic, J. Chang, M. B. Carson, C. Primeaux, R. Tian, S. Kenton, H. Jia, S. Lin, Y. Qian, S. Li, H. Zhu, F. Najar, H. Lai, J. White, B. A. Roe, and J. J. Ferretti.** 2002. Genome sequence of *Streptococcus mutans* UA159, a cariogenic dental pathogen. Proc. Natl. Acad. Sci. USA **99:**14434–14439.
2. **Bailey, T. L., and M. Gribskov.** 1998. Combining evidence using p-values: application to sequence homology searches. Bioinformatics **14:**48–54.
3. **Banas, J. A.** 2004. Virulence properties of *Streptococcus mutans.* Front. Biosci. **9:**1267–1277.
4. **Banas, J. A., and M. M. Vickerman.** 2003. Glucan-binding proteins of the oral streptococci. Crit. Rev. Oral Biol. Med. **14:**89–99.
5. **Bhagwat, S. P., J. Nary, and R. A. Burne.** 2001. Effects of mutating putative two-component systems on biofilm formation by *Streptococcus mutans* UA159. FEMS Microbiol. Lett. **205:**225–230.
6. **Burne, R. A., K. Schilling, W. H. Bowen, and R. E. Yasbin.** 1987. Expression, purification, and characterization of an exo-β-D-fructosidase of *Streptococcus mutans.* J. Bacteriol. **169:**4507–4517.
7. **Giffard, P. M., and N. A. Jacques.** 1994. Definition of a fundamental repeating unit in streptococcal glucosyltransferase glucan-binding regions and related sequences. J. Dent. Res. **73:**1133–1141.
8. **Koga, T., T. Oho, Y. Shimazaki, and Y. Nakano.** 2002. Immunization against dental caries. Vaccine **20:**2027–2044.
9. **Kremer, B. H. A., M. van der Kraan, P. J. Crowley, I. R. Hamilton, L. J. Brady, and A. S. Bleiweis.** 2001. Characterization of the *sat* operon in *Streptococcus mutans*: evidence for a role of Ffh in acid tolerance. J. Bacteriol. **183:**2543–2552.
10. **Kuramitsu, H. K.** 1993. Virulence factors of mutans streptococci: role of molecular genetics. Crit. Rev. Oral. Biol. Med. **4:**159–176.
11. **Lee, S. F., A. Progulske-Fox, G. W. Erdos, D. A. Piacentini, G. Y. Ayakawa, P. J. Crowley, and A. S. Bleiweis.** 1989. Construction and characterization of isogenic mutants of *Streptococcus mutans* deficient in major surface protein antigen P1 (I/II). Infect. Immun. **57:**3306–3313.
12. **Lemos, J. A. C., T. A. Brown, Jr., and R. A. Burne.** 2004. Effects of RelA on key virulence properties of planktonic and biofilm populations of *Streptococcus mutans.* Infect. Immun. **72:**1431–1440.
13. **Loo, C. Y., D. A. Corliss, and N. Ganeshkumar.** 2000. *Streptococcus gordonii* biofilm formation: identification of genes that code for biofilm phenotypes. J. Bacteriol. **182:**1374–1382.
14. **Patti, J. M., B. L. Allen, M. J. McGavin, and M. Hook.** 1994. MSCRAMM-mediated adherence of microorganisms to host tissues. Annu. Rev. Microbiol. **48:**585–617.
15. **Perez-Casal, J., M. G. Caparon, and J. R. Scott.** 1991. Mry, a *trans*-acting positive regulator of the M protein gene of *Streptococcus pyogenes* with similarity to the receptor proteins of two-component regulatory systems. J. Bacteriol. **173:**2617–2624.
16. **Perry, D., and H. K. Kuramitsu.** 1981. Genetic transformation of *Streptococcus mutans.* Infect. Immun. **32:**1295–1297.
17. **Yamashita, Y., W. H. Bowen, R. A. Burne, and H. K. Kuramitsu.** 1993. Role of the *Streptococcus mutans gtf* genes in caries induction in the specific-pathogen-free rat model. Infect. Immun. **61:**3811–3817.
18. **Wen, Z. T., and R. A. Burne.** 2002. Functional genomics approach to identifying genes required for biofilm development by *Streptococcus mutans.* Appl. Environ. Microbiol. **68:**1196–1203.


*Editor:* V. J. DiRita