Phylogenetic Analyses of the Water-Insoluble Glucan Synthesizing Glucosyltransferase Gene of *Streptococcus ratti*

Ryotaro Hirata

Nihon University Graduate School of Dentistry at Matsudo, Microbiology and Immunology, Matsudo, Chiba 271-8587, Japan

*Correspondence to:
Ryotaro Hirata
E-mail: mary07012@g.nihon-u.ac.jp

**Abstract**

The nucleotide sequence of the glucosyltransferase (GTF) gene, which encodes a GTF enzyme that synthesizes a water-insoluble glucan (WIG), was determined in *Streptococcus ratti* FA-1 (GTC 002457). The *gtf* of *S. ratti* consisted of 4,350 bp that encoded for a 1,449 amino acid protein with a molecular weight of approximately 159.9 kDa and was revealed to be a *gtfL* type gene, which were first found in *Streptococcus salivarius* ATCC 25975. The deduced 35 amino acid sequence of the N-terminal was thought to be a signal peptide required for the secretion of GTF-L in *S. ratti* as it showed high similarity to a known GTF-L from *S. salivarius*. In addition, three major functional domains of GTF: an N-terminal variable region, a conserved catalytic site for sucrase activity, and C-terminal YG repeating units for glucan binding were also found in GTF-L from *S. ratti*. The percentage homology of the GTF-L amino acid sequences from *S. ratti* and *S. salivarius* was 99.7%. Although GTF-L were classified into the WIG-synthesizing GTF group, phylogenetic analysis suggested that GTF-L were positioned outside of the WIG-synthesizing GTF group of mutants streptococci.

**Introduction**

Mutants streptococci have been implicated as the primary causative agent of dental caries in human and animals. Many factors associated with cariogenicity have been determined. These bacteria secrete several extracellular enzymes such as glucosyltransferases (GTF), which catalyze the formation of water-insoluble and water-soluble glucans (WIG and WSG, respectively) from dietary sucrose. The WIG adhere to smooth tooth surfaces and facilitate the aggregation of oral bacteria so WIG synthesizing GTF are one of the most important cariogenic factors in humans (1, 2).

Previously, various GTF enzymes of oral streptococci were purified, and their genes were cloned. So far, the nucleotide sequences of different GTF genes from *Streptococcus mutans*, *Streptococcus sobrinus*, *Streptococcus downei*, *Streptococcus salivarius*, *Streptococcus gordonii*, *Streptococcus oralis*, *Streptococcus sanguinis*, and *Streptococcus criceti* have been determined (3–5), and a phylogenetic tree of these GTF based on their amino acid sequences was constructed (5). In addition, two novel mutants streptococci, *Streptococcus orisuis* and *Streptococcus dentirousseti*, which were isolated from the oral cavities of pigs and bats, respectively, were reported (6, 7), and nucleotide sequence analyses of WIG synthesizing GTF genes from these strains was performed (5, 8).

*Streptococcus ratti* was originally isolated from the oral cavities of laboratory rats as a distinct serovar of *S. mutans* and has also been isolated from humans on rare occasions (9). Kawamura et al. reported that *S. ratti* diverged early in the evolutionary process of the mutants streptococcus group on the basis of a 16S rRNA gene tree (10). This suggestion led to a search for the cariogenic factor produced by the *gtf* gene of *S. ratti*. Indeed, WIG and WSG-synthesizing GTF enzymes have already been purified from the culture
supernatants of *S. ratti* BHT and characterized (11, 12). However, the genes encoding the GTF have not yet been cloned and analyzed.

In this study, sequence analyses of WIG synthesizing GTF genes from *S. ratti* is described, and the construction of a phylogenetic tree based on the deduced amino acid sequences of other streptococcal GTF was attempted.

**Materials and Methods**

*Bacterial strains and media*

*S. ratti* strain FA-1 (GTC 00245\(^7\)), which was isolated from rats by Fitzgerald (13); strain BHT, which was isolated from humans by Zinner (14); strain LB-2, which was isolated from humans by Gansser (15); and *S. salivarius* ATCC 25975 were used in this study. The strains were grown for 24 hours at 37 °C under anaerobic conditions on Bact™ Brain Heart Infusion (BHI; Becton, Dickinson, and Company NJ, USA) agar.

**Reconfirmation of BHT and LB-2**

Chromosomal DNA was extracted from the bacterial cultures using the Promega genome kit (Promega, Co., WI, USA) according to the manufacturer's instructions. To reconfirm the BHT and LB-2 strains, PCR analysis of the 16S rRNA gene was performed directly using a single colony and the following primer set: 27f and 1525r (16).

**PCR experiments**

The sequences of the primers used in this study are listed in Table 1 (17). The target regions of the *gft* genes were amplified using the KOD FX (Toyobo Co., Ltd., Japan) under the following conditions: (i) initial denaturation at 95 °C for 5 min; (ii) 30 cycles of amplification, denaturation at 95 °C for 30 sec, primer annealing at 60 °C for 30 sec, and extension at 72 °C for 30 sec to 4 min; and (iii) final extension at 72 °C for 10 min.

**Sequence analysis**

The molecular weights of the amplified PCR products were confirmed by agarose gel electrophoresis after staining with ethidium bromide, purification with a Suprec-PCR™ (Takara Bio Inc, Otsu, Japan), and sequencing with the ABI PRISM 3130 Genetic Analyzer using the Big Dye Terminator v1.1 cycle

<table>
<thead>
<tr>
<th>Target</th>
<th>Name</th>
<th>Primer sequence</th>
<th>Location*</th>
<th>Product size</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>gftB</em></td>
<td>GTFB-F</td>
<td>5'-actacatcctccggtgctgg-3'</td>
<td>793-814</td>
<td>517</td>
<td>(17)</td>
</tr>
<tr>
<td></td>
<td>GTFB-R</td>
<td>5'-cagtataggccgctgctctc-3'</td>
<td>1309-1288</td>
<td>368</td>
<td>This study</td>
</tr>
<tr>
<td><em>gftC</em></td>
<td>140F</td>
<td>5'-cggttcagaaacacgtgagctgtg-3'</td>
<td>5471-5494</td>
<td>712</td>
<td>(17)</td>
</tr>
<tr>
<td></td>
<td>526R</td>
<td>5'-ctggactaagttatagtcaagc-3'</td>
<td>5857-5838</td>
<td>438</td>
<td>This study</td>
</tr>
<tr>
<td><em>gftI</em></td>
<td>GTFI-F</td>
<td>5'-gataactacgcaagctgtcactg-3'</td>
<td>871-892</td>
<td>481</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>GTFI-R</td>
<td>5'-agctgcctgttgaagttaact-3'</td>
<td>1582-1561</td>
<td>481</td>
<td>This study</td>
</tr>
<tr>
<td><em>gff</em></td>
<td>144F</td>
<td>5'-tcagcggattaaacttaaggtg-3'</td>
<td>144-164</td>
<td>483</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>64R</td>
<td>5'-cgaagtaaaacgtgagcagca-3'</td>
<td>646-626</td>
<td>483</td>
<td>This study</td>
</tr>
<tr>
<td><em>gfl</em></td>
<td>3968F</td>
<td>5'-gcgcactgtaaggagaagaa-3'</td>
<td>3968-3987</td>
<td>481</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>4467R</td>
<td>5'-gcgcctccgtaaaagtaa-3'</td>
<td>4467-4448</td>
<td>481</td>
<td>This study</td>
</tr>
</tbody>
</table>

* The nucleotide sequences were obtained from GenBank (accession numbers: *M17361 (gftB and gftC), D90213 (gftI), M64111 (gff), and L35495 (gfl)*).
sequencing kit (Life Technologies, Ltd., USA). The sequence data were analyzed using the GENETYX® Ver. 7 genetic information processing software (Genetyx Co., Ltd., Japan). The closest known relatives of the gtf and 16S rRNA genes were determined by performing DDBJ/EMBL/GenBank database searches.

**Phylogenetic analysis**

Phylogenetic analysis was carried out using the Clustal W program (18). The topology of the tree was estimated (19), and a phylogenetic tree was constructed using the neighbor-joining method (20).

**Results**

**Reconfirmation of BHT and LB-2**

To confirm the identity of the BHT and LB-2 strains as *S. ratti*, the nucleotide sequences of their 16S rRNA genes were analyzed. The 16S rRNA gene sequences of both strains showed 100% homology with that of the *S. ratti* FA-1 strain.

**Detection of WIG-synthesizing gtf genes**

To investigate the presence of WIG synthesizing gtf genes in the *S. ratti* FA-1, BHT, and LB-2 strains, PCR experiments were performed using five sets of primers capable of detecting and identifying the *S. mutans* gtfB, gtfC, *S. sobrinus* gtfI, *S. salivarius* gtfJ, and gtfL genes (Table 1). As shown in Fig. 1, 517-bp gtfB fragments (A, lanes 1 to 3) and 481-bp gtfL fragments (C, lanes 1 to 3) were observed, while no fragments of gtfC, gtfI, or gtfJ were amplified, except for in the positive controls (B). These results suggested that gtfB-like and gtfL-like genes were present in the *S. ratti* strains. The amplified 517-bp and 481-bp fragments of *S. ratti* were purified, and the sequencing results revealed that they showed 100% homology to the corresponding region of the *S. mutans* gtfB gene and the *S. salivarius* gtfL gene, respectively.

**Sequence analysis**

To attempt amplification of the entire gtfB-coding region of *S. ratti*, inverse PCR was utilized to determine the unknown regions adjacent to the known sequence of the gtfB-like gene in *S. ratti* using an additional primer set, GTF-B-442F (21) and GTF-B-R (22). A 6.0-kb DNA fragment was amplified from *S. ratti* and subjected to nucleotide sequence determination. As a result, a 2.7-kb nucleotide sequence was found to correspond to parts of the *S. mutans* gtfB gene (1.1 kb downstream from the start codon of the open reading frame and 1.6 kb upstream of the gene). However, no full-length WIG producing gtf gene of the gtfB type was present in *S. ratti*.

To amplify the gtfL-like gene of *S. ratti*, other

![Fig. 1. Agarose gel electrophoresis of the PCR products of gtfB (A), gtfC, gtfI, gtfJ (B), and gtfL (C) produced using the primers listed in Table 1. Lanes 1 to 3, *S. ratti* FA-1, BHT, and LB-2, respectively; mu, *S. mutans* JCM5705; so, *S. sobrinus* OMZ176; sa, *S. salivarius* ATCC25975; M, 100-bp base ladder used as a size marker.](image)
primers were designed based on the *S. salivarius gtfL* sequence and used for PCR (Table 1). As shown in Fig. 2, 0.8-, 2.3-, 2.0-, and 0.3-kb fragments were obtained (lines a to d, respectively) and directly sequenced. Utilizing these fragments, the whole *gtfL*-like gene sequence of *S. rattii* was determined. As shown in Fig. 3 the determined nucleotide sequence of *S. rattii gtf* contains a 4,350-bp open reading frame, and the primary gene products are composed of 1,449 amino acids containing a signal sequence composed of the first 35 amino acids and a putative active site sequence (DGVRVDAVD). It was indicated that the GTF molecule of *S. rattii* possessed 15 YG repeats and one partial repeat, which were similar to those of *S. salivarius*, near its C-terminus (23). The deduced amino acid sequence of the *gtfL* gene coded for a preprotein with a molecular mass of ca. 159.9 kDa and a predicted secretory protein after cleavage of the signal peptide, which had a molecular mass of ca. 156.8 kDa. The percentage homology of the nucleotide sequence of the *gtf* gene from *S. rattii* was 99.7% compared with that of *S. salivarius* ATCC 25975; therefore, the WIG synthesizing *gtf* gene of *S. rattii* seems to correspond to *S. salivarius gtfL*. The percentage homology between the amino acid sequence of the GTF-L from *S. rattii* and that from *S. salivarius* was 99.7%.

**Phylogenetic analysis**

To study the relationship among the deduced amino acid sequences of the various GTF genes, a phylogenetic tree was constructed based on GTF including WIG- and WSG-producing enzymes from oral streptococci, as shown in Fig. 4. The WIG synthesizing GTF of mutants streptococci (GTF-B, GTF-C, and GTF-I) were positioned in the same group; on the other hand, the WIG synthesizing GTF of *S. salivarius* such as GTF-J and GTF-L were positioned outside of this group.

**Discussion**

Most mutants streptococci possess GTF enzymes, which catalyze WIG or WSG synthesis from sucrose. It has been proposed these GTF function as the principal virulence factors contributing to caries formation (1). In humans, *S. mutans* and *S. sobrinus* have the ability to secrete GTF-B, GTF-C, or GTF-I, which are different WIG-producing enzymes encoded by the *gtfB*, *gtfC*, and *gtfI* gene, respectively. In animals, mutants streptococci; i.e., *O. orisuis* and *D. dentifrouseti*, which were isolated from pig and fruit bats, respectively, demonstrated the ability to secrete GTF-I (5, 6).

*S. rattii*, which has been isolated from laboratory rats and humans, diverged at an early stage in the

---

Fig. 2. Restriction map of the *S. rattii gtfL* gene. The solid bar represents the 481-bp region amplified by the *gtfL*-3968F and *gtfL*^-4467R primers. B, *BamHI*; E, *EcoRI*; H, *HindIII*; K, *KpnI*; P, *PstI*. The solid lines (a to d) indicate the PCR products used to determine the nucleotide sequence of the *S. rattii gtfL* gene. The arrowheads denote the primer sets used for the PCR in this study (see Table 1). For details, see the text.
Fig. 3. Nucleotide sequence and deduced amino acid sequence of gftL. A potential ribosome binding site (RBS), and -10 and -35 promoter elements are indicated. The arrow shows a predicted signal sequence cleavage site. The putative active site is underlined. The dashes show YQ repeats in the gliadin binding domain.
evolution of the mutans group according to its 16S rRNA gene sequence (10). It was reported that one WIG synthesizing enzyme was purified from an S. ratti culture supernatant, and the isoelectric focusing and 2-dimensional electrophoresis patterns of the WIG synthases are clearly different from those of S. criceti, S. mutans, and S. sobrinus (11, 24, 25).

In this study, WIG–synthesizing GTF genes were detected using five sets of primers for the following genes: S. mutans gtfB, gtfC, S. sobrinus gtfI, S. salivarius gtfJ, and gtfL (Table I, Fig. I) because no WIG–synthesizing GTF gene from S. ratti has been cloned or reported.

As a result, 517–bp gtfB and 481–bp gtfL fragments were obtained. Initially, inverse PCR was carried out to determine the entire gtfB sequence because S. ratti is a member of the mutans streptococcal group. S. ratti did not possess a full–length WIG–producing gtf gene belonging to the gtfB group, although they did contain a section of the gtfB gene (located 1.1 kb downstream of the start codon of the open reading frame). The absence of the rest of the gtfB gene might have arisen by deletion or insertion. To determine the gtfL–like gene of S. ratti, four PCR products were obtained using originally designed primers, and the whole gtfL–like gene sequence of S. ratti was determined using these fragments (Fig. 2). The nucleotide sequence of the S. ratti gtf gene contains a 4,350–bp open reading frame, and the primary gene products are composed of 1,449 amino acids containing a signal sequence composed of the first 35 amino acids and a putative active site sequence. The percentage homology of the GTF–L amino acid sequences from S. ratti and S. salivarius was 99.7%, while those for S. ratti GTF–L were 47, 48, 44, and 47% compared to those of GTF–B, GTF–C, GTF–I, and GTF–J respectively. Furthermore, the GTF–L from S. salivarius produces a mixed–linkage insoluble glucan composed of α–1,3–linked and α–1,6–linked glucosyl residues (26). The WIG synthase of S. ratti also produces both α–1,3–linked and α–1,6–linked glucans (11). These findings suggest that the WIG–synthesizing GTF from S. ratti corresponds to the GTF–L from S. salivarius.
A phylogenetic tree was constructed based on the amino acid sequences of 20 GTF, including both WIG- and WSG-synthesizing GTF enzymes, from oral streptococci to study the relationships among GTF. WIG-synthesizing GTF from mutants streptococci such as GTF-I, GTF-B, and GTF-C were positioned in the same group; however, the GTF-L genes were more closely related to the WSG-synthesizing GTF cluster than the WIG-synthesizing GTF of mutants streptococci. When an additional GTF catalytic domain-based phylogenetic tree was constructed (23), a similar result was obtained (data not shown). The solubility of the WIG produced by GTF-L might be high compared to those of GTF-B, GTF-C, and GTF-I, as the WIG produced by GTF-L are thought to have relatively low side chains consisting of α-1,6-linked glucosyl residues attached to α-1,3-linked backbone.

The ultimate goal of this ongoing work is to identify the origin of cariogenic factors, especially WIG synthase in humans. In this study, the nucleotide sequence of the gtfL gene of S. ratti was determined. A phylogenetic tree constructed using GTF amino acid sequences suggested that GTF-L developed at an intermediate stage in the evolution of cariogenic factors in S. mutans.

Furthermore, streptococcus and other bacterial strains that produce WIG or WSG synthase have been isolated from the oral cavities of various animals and humans. However, some of the genes of these enzymes have not been analyzed and reported. It is suggested that if analyses of GTF genes are successful, the results might clarify the process by which cariogenic factors are obtained at a level beyond species or genera.

Acknowledgments

The author is grateful to Professor M. Hirasawa for his direction and helpful advice, and appreciate Professors M. Makimura, K. Shibutani, and M. Wada for reviewing the manuscript. I thank Doctors K. Takada and N. Kuwahara for their continuous and kind support. This study was supported in part by a Grant-in-Aid for SPSR from MECSST 2008–12.

References