

Short sequence-paper

Cloning and nucleotide sequence analysis of the *Streptococcus sobrinus* *gtfU* gene that produces a highly branched water-soluble glucan

Nobuhiro Hanada ^{a,*}, Kazuo Fukushima ^b, Yoshiaki Nomura ^a, Hidenobu Senpuku ^a, Mitsuo Hayakawa ^c, Hidehiko Mukasa ^d, Teruaki Shiroza ^c, Yoshimitsu Abiko ^c

^a Department of Oral Science, National Institute of Infectious Diseases, 1-23-1 Toyama, Sinjuku-ku, Tokyo 162-8640, Japan

^b Department of Microbiology, Nihon University School of Dentistry at Matsudo, Chiba 271-8587, Japan

^c Department of Biochemistry, Nihon University School of Dentistry at Matsudo, Chiba 271-8587, Japan

^d Department of Chemistry, National Defense Medical College, Tokorozawa, Saitama 359-8513, Japan

Received 5 June 2001; received in revised form 23 October 2001; accepted 2 November 2001

Abstract

Streptococcus sobrinus has four *gtf* genes, *gtfI*, *gtfS*, *gtfT*, and *gtfU*, on the chromosome. These genes correspond respectively to the enzymes GTF-I, GTF-S₁, GTF-S₂, and GTF-S₃. An *Escherichia coli* MD66 clone that contained the *S. sobrinus* *gtfU* gene was characterized. Immunological properties showed that the protein produced by the *E. coli* MD66 clone was similar to *S. sobrinus* GTF-S₁. Biological properties and a linkage analysis of the glucans by ¹³C NMR spectrometry revealed that the protein produced by the *E. coli* MD66 clone was GTF-S₁. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Glucosyltransferase

Streptococcus mutans and *Streptococcus sobrinus* are major pathogenic bacteria for dental caries. They have the ability to produce water-insoluble glucans (WIG) and water-soluble glucans (WSG) from dietary sucrose. These glucans are closely related to the maturation of bacterial biofilms on the tooth surface and to their virulence.

S. mutans has three *gtf* genes, *gtfB*, *gtfC*, and *gtfD*, on the chromosome [1]. These genes correspond to GTF enzymes GTF-I, GTF-S₁, and GTF-S.

S. sobrinus has four *gtf* genes, *gtfI*, *gtfS*, *gtfT*, and *gtfU*, on the chromosome [2]. These genes correspond respectively to the enzymes GTF-I, GTF-S₃, GTF-S₂, and GTF-S₁ [2]. GTF-I is an α-1,3-glucan synthesizing enzyme and synthesizes WIG that is activated by the WSG. The GTF-S₁ enzyme is an α-1,6-glucan and an α-1,3,6-glucan synthesizing enzyme and synthesizes WSG. GTF-S₁ is also activated by the addition of WSG. The properties of these enzymes were reported previously [3,4]. However, the genes of GTF-S₁ have not yet been identified. In this

study, we identify the *gtfU* gene and GTF-S₁ enzyme from the *S. sobrinus* strain B13N.

The genomic library was constructed by the chromosomal DNA from *S. sobrinus* B13N and was partially cleaved with *Sau3A*I and ligated with *Bam*HI-cut λL47.1 DNA [5]. Phage plaques that developed on the LB agar plates after in vitro packaging of the library into *Escherichia coli* C600 indicator cells were blotted onto nitrocellulose membranes. The membranes were incubated with antiserum raised against a crude GTF preparation from the B13N strain. They were then treated by goat anti-rabbit IgG conjugated with horseradish peroxidase and developed using 4-methoxy-1-naphthol. Several clones reacted with a monoclonal antibody (MAb B76) specific to the GTF-S₁ enzyme [2]. One of the positive clones, designated as λMDSM66, was selected. The DNA from λMDSM66 was digested with *Cla*I and ligated into the plasmid vector pACYC184 [6]. The ligation mixture was transformed into *E. coli* HB101. One of the chloramphenicol-resistant recombinant transformant MD66 cells was then selected on LB agar plates containing chloramphenicol (30 μg/ml) and was used for the following study.

E. coli MD66 cells containing plasmid pMD66 were subsequently tested for sucrose hydrolyzing ability. Reduction of the sugar-forming activity and WSG-forming ac-

Abbreviations: CCB staining, Coomassie brilliant blue staining; PAS staining, periodic acid Schiff staining

* Corresponding author. Fax: +81-3-5285-1172.

E-mail address: nhanada@nih.go.jp (N. Hanada).

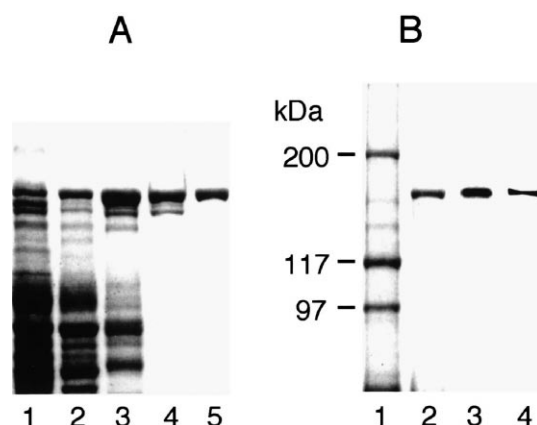


Fig. 1. SDS-PAGE, immunostaining and activity-staining analysis of the recombinant enzyme rGTF-S₁. (A) Protein staining (with CBB) of various rGTF-S₁ preparations in the purification process. Lane 1, cell-free extract; lane 2, ammonium sulfate fraction; lane 3, Butyl-Toyopearl 650 fraction; lane 4, DEAE Bio-Gel A fraction; lane 5, hydroxyapatite fraction. (B) Analysis of the purified enzyme (ca. 0.5 µg protein). Lane 1, CBB staining of size marker proteins; lane 2, CBB staining of the purified enzyme; lane 3, activity staining (activity is displayed after incubation with sucrose and subsequent PAS staining); lane 4, immunostaining with anti-GTF-S₁ serum.

tivity was assayed as previously reported [7,8]. One international unit (IU) of activity was defined as the amount of enzyme required to incorporate 1 µmol glucose from sucrose into glucans per minute under the standard assay. The MD66 cells evidently express a recombinant protein that possesses sucrase and WSG-forming activities.

Rabbit anti-GTF sera specific to GTF-I, GTF-S₁, GTF-S₂ and GTF-S₃ from the *S. sobrinus* B13N and rGTF-S₁ were prepared as described previously [9]. The extract of MD66 cells produced a single precipitation line with only the anti-GTF-S₁ serum in the Ouchterlony test [10]. When the enzyme localization of *E. coli* MD66 cells was examined, a large portion (74.3%) of sucrase activity was found in the soluble cytoplasmic fraction, compared with 13.7% in the periplasmic fraction and 12.0% in the membrane fraction. The recombinant GTF-S₁ (rGTF-S₁) enzyme was prepared from the soluble cytoplasmic and periplasmic fraction of MD66 cells. The recombinant enzyme rGTF-S₁ was then purified to homogeneity following ammonium sulfate precipitation, hydrophobic adsorption, anion exchange, and hydroxyapatite chromatographies.

rGTF-S₁ preparation migrated on a sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) gel as a 175 kDa protein band (Fig. 1A). The 175 kDa protein band possessed WSG-forming activity and was reactive to the anti-GTF-S₁ serum (Fig. 1B). The *pI* value and the specific activity (determined by WSG-forming assay) of the purified rGTF-S₁ enzyme were *pI* 4.0 and 13.3 IU/mg protein. In contrast, the molecular size, *pI* value, and

Table 1

Structures of glucans produced by GTF-S₁ and rGTF-S₁ measured by ¹³C NMR

Linkages	A	B
1,3- α -linked glucosyl residues	0	< 1
1,3,6- α -linked glucosyl residues	32.0	30.8
1,6- α -linked glucosyl residues	26.3	25.2
Terminal residues	41.7	43.9

(A) ¹³C NMR spectra of WSGs synthesized by B13N GTF-S₁ (%). (B) MD66 rGTF-S₁ enzymes (%).

specific activity of the GTF-S₁ enzyme purified from the B13N culture fluid were 176 kDa, *pI* 4.0, and 14.3 IU/mg protein. These data corresponded to the previous report [11]. Sucrase activities both of the rGTF-S₁ and GTF-S₁ enzymes were increased three times by addition of dextran T10 in the reaction mixtures.

A sequence analysis of plasmid pMD66 was carried out to determine the complete sequence of the gene (Fig. 2). The putative active site sequence (D-G-V-R-V-D-A-V-D) of pMD66 was identical to the previously published [12] active site peptide derived from GTF-S. GTF-I and GTF-S active-site peptides had similar but not identical sequences: GTF-I, D-S-I-R-V-D-A-V-D; GTF-S, D-G-V-R-V-D-A-V-D. Each has three aspartic acids (D) as potential sites of glucose conjugation [12].

The first 35 amino acids encoded by sequence MEKKL-HYKLHKVKKHWVTIAVASIGLVSLVGAGTV are similar to those encoded by the previously reported *S. sobrinus* *gtfT* gene, MERKLHYKLHKVKKQWVTIAVASAGLASIVGAGSV [13]. Since these amino acids contain the signal sequence of the GTF protein, we truncated the first 35 amino acids to estimate the *pI* value of rGTF-S₁ by computer (DNASIS, Hitachi, Japan). Computer analysis revealed that the total number of amino acids of rGTF-S₁ was 1519, the *pI* value was 4.35 and the molecular mass was 167 843. Values for the previously reported *gtfT* gene that encoded the GTF-S₃ protein were calculated to be 1433 amino acids, *pI* 5.22 and *M_r* 158 973, respectively. The *pI* value of the enzyme and its molecular mass both corresponded with the previous report [11].

A linkage analysis of the glucans by ¹³C NMR spectrometry was conducted to determine the structure of the glucans that were formed by GTF-S₁. The ¹³C NMR spectrum was recorded at 67.9 MHz with a JEOLJNM-GX270 spectrometer operated in the Fourier transform mode, with complete proton-decoupling, at 21°C.

Table 1 shows the result of ¹³C NMR spectra of glucans synthesized from sucrose by purified rGTF-S₁ and GTF-S₁ enzymes in the absence of dextran. As shown in Table 1, spectra A and B were similar, indicating that WSGs synthesized by both enzymes are highly branched 1,3,6-

Fig. 2. The complete sequence of the *gtfU* gene obtained from the *S. sobrinus* B13N. The first 35 amino acid peptide and putative active site sequence (D-G-V-R-V-D-A-V-D) is underlined.

9	18	27	36	45	54
5' ATG GAA AAA AAG CTA CAT TAT AAG CTT CAC AAG GTT AAA AAA CAT TGG GTT ACG					
M E K K L H Y K L H K V K K H W V T					
63	72	81	90	99	108
ATT GCG GTA GCC TCT ATT GGT TTA GTA AGC CTC GTT GGT GCT GGT ACT GTT TCC					
I A V A S I G L V S L V G A G T V S					
117	126	135	144	153	162
GCA GAG GAT AAG GTA GCT AAT GAT ACG ACG GCC CAA GCA ACA GTA GGA GAT GAT					
A E D K V A N D T T A Q A T V G V D					
171	180	189	198	207	216
ACT GGT CAG GAT CAG GCT ACC AAT GAC GCG AAT ACG AAC ACT ACA GAT ACT					
T G Q D Q A T T N D A N T N T T D T					
225	234	243	252	261	270
GAC ACG GCT GAC CAA TCA GCC AAT ACT AAT CAA GAT CAA GCA GGT TCT GAT CAA					
D T A D Q S A N T N Q D Q A G S D Q					
279	288	297	306	315	324
AGT AAC AAT CAA GAT CAG GCC AAG CAA GAT ACG GCC AAT ACC GAT CGA AAT CAG					
S N N Q D Q A K Q D T A N T D R N Q					
333	342	351	360	369	378
GCG GAT AAC AGT CAA ACT GAT AAT AAT CAA GCG ACT GAC CAA GCC ACT AGT CCA					
A D N S Q T D N N Q A T D Q A T S P					
387	396	405	414	423	432
GCG ACA GAT GGA ACC AGT GTC CAA CGA AGA GAT GCC AAC GTG GCA ACA GCA					
A T D G T S V Q R R D A A N V A T A					
441	450	459	468	477	486
GCA GAT CAA GAG GGA CAA ACA GCT CCT TCT GAA CAA GAA AAA TCA GCA GCC CTG					
A D Q E G Q T A P S E Q E K S A A L					
495	504	513	522	531	540
TCC CTT GAC AAT GTT AAG TTG ATT GAT GGG AAA TAC TAT TAT GTC CAA GCT GAT					
S L D N V K L I D G K Y Y Y V Q A D					
549	558	567	576	585	594
GGC TCT TAC AAG AAG AAT TTT GCC ATT ACT GTC AAC GGA CAA ATG CTC TAC TTT					
G S Y K K N F A I T V N G Q M L Y F					
603	612	621	630	639	648
GAT AGC GAT ACT GGT GCC CTT TCG TCA ACA TCA ACC TAT TCC TTC TCA CAA GGG					
D S D T G A L S S T S T Y S F S Q G					
657	666	675	684	693	702
ACA ACC AAC TTG GTT GAT GAC TTC TCC AGT CAT AAC AAG GCC TAC GAT TCA ACT					
T T N L V D D F S S H N K A Y D S T					
711	720	729	738	747	756
GCC AAA AGT TTT GAA TTG GTT AAT GGT TAT TTA ACA GCT AAC TCT TGG TAC CGT					
A K S F E L V N G Y L T A N S W Y R					
765	774	783	792	801	810
CCA GCT GGC ATT CTG CGC AAT GGT CAA ACT TGG GAA GCT TCA AAT GAA AAC GAC					
P A G I L R N G Q T W E A S N E N D					
819	828	837	846	855	864
CTG CGC CCT GTT TTG ATG AGC TGG TGG CCT GAC AAG GAT ACC CAA GTT GCT TAT					
L R P V L M S W W P D K D T Q V A Y					
873	882	891	900	909	918
GTC AAC TAC ATG AAT AAG TAC TTG AGT GCA AAT GAG ACA GAA GTC ACT AAT GAA					
V N Y M N K Y L S A N E T E V T N E					
927	936	945	954	963	972
ACA TCT CAG GTA GAT TTG AAT AAA GAA GCT CAA TCT ATT CAA ACC AAG AAT GAA					
T S Q V D L N K E A Q T S I Q T K I E					
981	990	999	1008	1017	1026
CAA AAG ATC ACC TCT GAT AAT AGT ACC CAA TGG TTA CGG ACA GCT ATG GAG GCC					
Q K I T S D N S T Q W L R T A M E A					
1035	1044	1053	1062	1071	1080
TTC GTT GCT GCT CAG CCT AAG TGG AAC ATG AGT ACT GAA AAC TTC AAT AAG GGT					
F V A A Q P K W N M S T E N F N K G					
1089	1098	1107	1116	1125	1134
GAC CAC CTG CAA GG; GGG GCT CTG CTC TAT ACC AAT TCA GAT TTG ACC CCT TGG					
D H L Q G G A L L Y T N S D L T P W					
1143	1152	1161	1170	1179	1188
GCA AAT TCT GAC TAC CGT CTG CTC AAC CGC ACC CCA ACT CAA CAA GAT GGT ACT					
A N S D Y R L L N R T P T Q Q D G T					
1197	1206	1215	1224	1233	1242
AAG AAA TAC TTT ACT GAA GGT GGT GAA GGG GGT TAT GAA TTC CTG TTG TCT AAT					
K K Y F T E G G E G G Y E F L L S N					
1251	1260	1269	1278	1287	1296
GAC GTT GAT AAC TCA AAC CCT GTC GTT CAA GCA GAA CAA CTG AAC CAA TTG CAC					
D V D N S N P V V Q A E Q L N Q L H					
1305	1314	1323	1332	1341	1350
TAC CTG ATG AAC TGG GGC GAT ATT GTC ATG GGA AAT AAG GAT GCC AAT TTT GAT					
Y L M N W G D I V M G D K D A N F D					
1359	1368	1377	1386	1395	1404
GGC GTT CGA GTC GAT GCG GTC GAT AAT GTC AAT GCC GAC TTG CTT CAA GTC TAC					
G V R V D A V D N V N A D L L Q V Y					
1413	1422	1431	1440	1449	1458
AGC AAT TAC TTC AAG GAC AAC TAT AAG GTA ACA GAT TCC GAA GCC AAT GCT TTA					
S N Y F K D N Y K V T D S E A N A L					
1467	1476	1485	1494	1503	1512
GCT CAT ATT TCT ATC CTT GAA GCT TGG TCA CTA AAT GAT AAC CAA TAT AAT GAA					
A H I S I L E A W S L N D N Q Y N E					
1521	1530	1539	1548	1557	1566
GAT ACA AAT GGT ACC GCC CTG TCT ATT GAT AAC TCA TCT CGT TTG ACC TCT CTA					
D T N G T A L S I D N S S R L T S L					
1575	1584	1593	1602	1611	1620
GCT GTT TTA ACC AAG CAA CCT GGT CAA CGG ATT GAC CTC TCA AAC TTG ATT AGT					
A V L T K Q P G Q R I D L S N L I S					
1629	1638	1647	1656	1665	1674
GAA TCG GTC AAT AAG GAG CGG GCT AAT GAT ACG GCC TAC GGC GAT ACT ATT CCG					
E S V N K E R A N D T A Y G D T I P					
1683	1692	1701	1710	1719	1728
ACC TAT TCC TTT GTT CGA GCT CAT GAC TCA GAA GTA CAA ACC GTT ACT GCT AAG					
T Y S F V R A H D S E V Q T V I A K					
1737	1746	1755	1764	1773	1782
ATT GTT AAG GAA AAG ATT GAT ACC AAT TCA GAT GGT TAT ACC TTT ACT CTT GAT					
I V K E K I D T N S D G Y T F T L D					
1791	1800	1809	1818	1827	1836
CAG TTA AAG GAT GCC TTC AAG ATT TAT AAT GAG GAT ATG GCT AAG GTT AAT AAG					
Q L K D A F K I Y N E D M A K V N K					
1845	1854	1863	1872	1881	1890
ACC TAT ACC CAT TAT AAT ATT CCG GCA GCC TAT GCG CTT TTG CTA AGC AAC ATG					
T Y T H Y N I P A A Y A L L L S N M					
1899	1908	1917	1926	1935	1944
GAA TCT GTC CCT CGA GTG TAT TAC GGT GAT CTT GAT ACC GAT GAC GGC CAG TAC					
E S V P R V Y Y G D L Y T D D G Q Y					
1953	1962	1971	1980	1989	1998
ATG GCT AAA AAA TCT CCT TAC TAT GAT GCT ATC GCA ACT ATG CTG CAA GGT CGC					
M A K K S P Y Y D A I A T M L Q G R					
2007	2016	2025	2034	2043	2052
ATA GCC TAT GTC TCA GGC GGT CAA AGT GAA GAA GTT CAT AAG GTT AAT GGG AAT					
I A Y V S G G Q S E E V H K V N G N					
2061	2070	2079	2088	2097	2106
AAC CAA ATC CTT TCA TCT GTC CGT TAC GGT CAA GAT CTC ATG TCT GCC GAT GAT					
N Q I L S S V R Y G Q D L M S A D D					
2115	2124	2133	2142	2151	2160
ACT CAG GGT ACC GAC CTT AGT CGG ACT TCT GGT CTA GTA ACT CTG GTC AGC AAT					
T Q G T D L S R T S G L V T L V S N					
2169	2178	2187	2196	2205	2214
GAT CCA AAC CTC GAC CTA GGC GGA GAC AGC CTT ACA GTC AAT ATG GGC CGA GCT					
D P N L D L G G D S L T V N M G R A					
2223	2232	2241	2250	2259	2268
CAT GCT AAC CAA GCC TAT CGT CCA TTG ATT TTA GGG ACT AAG GAT GGT GGT CAA					
H A N Q A Y R P L I L G T K D G V Q					
2277	2286	2295	2304	2313	2322
TCC TAT CTC AAG GAT TCT GAT ACC AAC ATT GTT AAA TAC ACT GAT GCC AAT GGT					
S Y L K D S D T N I V K Y T D A N G					
2331	2340	2349	2358	2367	2376
AAT TTA ACC TTC ACA GCC GAT GAT ATT AAG GGT TAC TCA ACC GTT GAT ATG AGT					
N L T F T A D D I K G Y S T V D M S					
2385	2394	2403	2412	2421	2430
GGT TAT TTG GCT GTT TGG GTG CCA GTT GGC GCT AAG GAT GGT CAA GAT GTG CGT					
G Y L A V W V P V G A K D G Q D V R					
2439	2448	2457	2466	2475	2484
GTT GCA GCA GAT ACC AAT CAA AAG GCA GAT GGT AAG TCC CTC AAG ACT CTA GCT					
V A A D T N Q K A D G K S L K T S A					

Fig. 2.

2493	2502	2511	2520	2529	2538
GCC CTT GAT TCT CAA GTC ATC TAT GAA GGC TTC TCA AAT TTC CAA GAC TTT GCA					
A L D S Q V I Y E G F S N F Q D F A					
2547	2556	2565	2574	2583	2592
AAT AAT GAT GCA GAT TAT A ^c C AAC AAG AAA ATT GCT GAA AAT GCC GAC TTC TTC					
N N D A D Y T N K K I A E N A D F F					
2601	2610	2619	2628	2637	2646
AAG AAA TTG GGT ATC ACT TCG TTT GAA ATG GCT CCA CAA TAC GTT TCA GCC ACA					
K K L G I T S F E M A P Q Y V S A T					
2655	2664	2673	2682	2691	2700
GAC GGT AGC TTT TTG GAT TCT ATC ATT CAA AAT GGT TAT GCC TTC TCA GAC CGC					
D G S F L D S I I Q N G Y A F S D R					
2709	2718	2727	2736	2745	2754
TAT GAC CTT GCG ATC AGC AAG AAC AAT AAA TAC GGT TCT AAG GAT GAT TTG GCT					
Y D L A M S K N N K Y G S K D D L A					
2763	2772	2781	2790	2799	2808
AAT GCC CTC AAG GCC CTC CAC GCT AAT GGT ATT CAA GCC ATT GCC GAC TGG GTA					
N A L K A L H A N G I Q A I A D W V					
2817	2826	2835	2844	2853	2862
CCA GAC CAA ATT TAT CAA TTA CCA GGT GAA GAA GTG GTA ACG GCT AAA CGG ACC					
P D Q I Y Q L P G E E Y V T A K R T					
2871	2880	2889	2898	2907	2916
AAT AGC TAT GGT AAT CCA ACC TTT GAT GCC TAC ATC AAT AAT GCC CTC TAT GCT					
N S Y G N P T F D A Y I N N A L Y A					
2925	2934	2943	2952	2961	2970
ACC AAT ACT AAG AGC AGC GGT AGT GAC TAC CAA GCT CAA TAT GGT GCC TTC					
T N T K S S G S D Y Q A Q Y G G A F					
2979	2988	2997	3006	3015	3024
TTG GAT GAG CTC AAG GCT AAA TAC CCA GAC ATG TTC ACC GTT AAC ATG ATT TCA					
L D E L K A K Y P D M F T V N M I S					
3033	3042	3051	3060	3069	3078
ACT GGT AAG CCA ATT GAT CCA TCA ACC AAG ATT AAA CAA TGG GAA GCT AAA TAC					
T G K P I D P S T K I K Q W E A K Y					
3087	3096	3105	3114	3123	3132
TTC AAT GGT ACC AAC GTC CTT GGC AAG GGT GCT GGT TAT GTC CTC AGT GAT GAT					
F N G T N V L G K G A G Y V L S D D					
3141	3150	3159	3168	3177	3186
GCA ACC GGT AAG TAC TTC ACC GTA AAT GAA AAT GGT GAC TTC CTA CCA GCC AGC					
A T G K Y F T V N E N G D F L P A S					
3195	3204	3213	3222	3231	3240
TTC ACC GGT GAC CAA AAT GCC AAG ACA GGC TTC TAC TAT GAT GGC ACT GGC ATG					
F T G D Q N A K T G F Y Y D G T G M					
3249	3258	3267	3276	3285	3294
GCT TAT TAC TCA ACT TCG GGT AAT AAG GCT GTC AAC AGC TTT ATC TAC GAA GGT					
A Y Y S T S G N K A V N S F I Y E G					
3303	3312	3321	3330	3339	3348
GGT CAC TAT TAT TAC TTC GAT AAA GAT GGT CAC ATG GTG ACT GGT AGC TAC AAG					
G H Y Y F D K D G H M V T G S Y K					
3357	3366	3375	3384	3393	3402
GCC GAA GAC GGT AAT GAT TAT TAC TTC TTG CCA AAT GGT ATT CAG ATG CGG GAT					
A E D G N D Y Y F L P N G I Q M R D					
3411	3420	3429	3438	3447	3456
GCC ATC TAT CAA GAT GCT CAA GGA AAT AGT TAC TAT TAC GGT CGG ACA GGT ATT					
A I Y Q D A Q G N S Y Y Y G R T G I					
3465	3474	3483	3492	3501	3510
CTT TAC AAG GGA GAC AAC TGG TAT CCA TTT GTA GAT CCT AAT AAT GCT AAC AAG					
L Y K G D N W Y P F V D P N N A N K					
3519	3528	3537	3546	3555	3564
ACG GTC TTC CGT TAC TTC GAT GCT AAT AAT GTC ATG GCT ATT GGC TAT AGA AAC					
T V F R Y F D A N N V M A I G Y R N					
3573	3582	3591	3600	3609	3618
ATG TAT GGT CAA ACC TAC TAC TTT GAT GAA AAT GGT TTC CAA GCT AAA GGC CAA					
M Y G Q T Y Y F D E N G F Q A K G Q					
3627	3636	3645	3654	3663	3672
CTC TTA ACT GAC GAT AAG GGT ACC CAT TAC TTC GAT GAA GAT AAT GGT GCC ATG					
L L T D D K G T H Y F D E D N G A M					
3681	3690	3699	3708	3717	3726
GCT AAG AAT AAA TTT GTC AAT GTT GGT GAT GAC TGG TAC TAC ATG GAT GGT AAC					
A K N K F V N V G D D W Y Y M D G N					
3735	3744	3753	3762	3771	3780
GGT AAT GCG GTC AAG GGA CAA TAC CCT GTC AAC AAC CAA ATC CTT TAC TTC AAT					
G N A V K G Q Y P V N N Q I L Y F N					
3789	3798	3807	3816	3825	3834
CCA GAA ACG GGT GTT CAG GTT AAG GGA CAA TTT ATT ACC GAT GCT CAA GGC CGG					
P E T G V Q V K G Q F I T D A Q G R					
3843	3852	3861	3870	3879	3888
ACC TCA TAC TAC GAT GCT AAT TCA GGT GCC CTC AAG TCC AGT GGT TTC TTC ACA					
T S Y Y D A N S G A L K S S G F F T					
3897	3906	3915	3924	3933	3942
CCA AAT GGT AGC GAC TGG TAC TAT GCT GAA AAC GGT TAT GTT AAA GGT TTC					
P N G S D W Y Y A E N G Y V Y K G F					
3951	3960	3969	3978	3987	3996
AAA CAA GTA GCT GAA AAC CAA GAT CAA TGG TAT TAC TTC GAC CAA ACT ACT GGT					
K Q V A E N Q D Q W Y Y G Q T T G					
4005	4014	4023	4032	4041	4050
AAG CAA GCC AAG GGA GCT GCC AAA GTT GAC GGA CGA GAC CTT TAC TTT AAC CCT					
K Q A K G A A K V D G R D L Y F N P					
4059	4068	4077	4086	4095	4104
GAT TCA GGT GTC CAA GTC AAG GGT GAC TTC GCA ACA GAC GAA TCT GGT AAT ACC					
D S G V Q V K G D F A T D E S G N T					
4113	4122	4131	4140	4149	4158
AGC TTC TAC CAT GGT GAT AAC GGT GAT AAG GTC GTC GGA GGT TTC TCA ACA ACC					
S F Y H G D N G D K V V G G F F T T					
4167	4176	4185	4194	4203	4212
GGT AAC AAT GCT TGG TAC TAC GCT GAT AAC AAT GGT AAT CTT GTC AAA GGC TTC					
G N N A W Y Y A D N N G N L V K G F					
4221	4230	4239	4248	4257	4266
CAA GAA ATA GAT GGC AAA TGG TAC CAC TTT GAC GAA GTA ACT GGC CAA CAA GCT					
Q E I D G K W Y H F D E V T G Q Q A					
4275	4284	4293	4302	4311	4320
AAG GGA GCA GCC TTG GTT AAT GGT CAA CAA CTC TAT TTC GAT GTA GAT TCT GGT					
K G A A L V N G Q Q L Y F D V D S G					
4329	4338	4347	4356	4365	4374
ATC CAA GTC AAG GGT GAC TTT GTC ACA GAT GGT CAA GGA AAT ACT TCC TAT TAT					
I Q V K G D F V T D G Q G N T S Y Y					
4383	4392	4401	4410	4419	4428
GAT GTC AAT TCT GGT GAT AAG AAG GTC AAT GGC TTC TCA ACT GGT GAT AAT					
D V N S S G D K K V N G F F T T G D N					
4437	4446	4455	4464	4473	4482
GCT TGG TAC TAT GCT GAT GGT CAG GGT AAT CTA GCC AAA GGT CGC AAG TCT ATT					
A W Y Y A D G Q G N L A K G R K S I					
4491	4500	4509	4518	4527	4536
GAT AAT CAA GAC CTC TAC TTT GAT CCT GCA ACA GGT AAG CAA GTT AAG GGG CAA					
D N Q D L Y F D P A T G K Q V K G Q					
4545	4554	4563	4572	4581	4590
CTC GTT TCT ATT GAT GGT CGC AAT TAT TAC TTC GAT AGT GGC TCT GGT AAT ATG					
L V S I D G R N Y Y F D S G S G N M					
4599	4608	4617	4626	4635	4644
GCC AAG AAC CGT TTT GTC CGC ATC GGT GAT CAA TGG ATT TAC TTC GGC AAC GAC					
A K N R F V R I G D Q W I Y F G N D					
4653	4662				
GGT GCC GCT ACC AAC CTA TAA3'					
G A A T N L					

Fig. 2 (continued)

α -linked glucosyl residues of nearly the same structure. The previously reported *gtfT* gene that encoded GTF-S₂ produced a small amount (5%) of branched 1,3,6- α -linked glucosyl residue [13]. The results of the linkage analysis completely corresponded with the previous report [14]. These results strongly suggest that the MD66 clone ob-

tained in this study has the native *S. sobrinus gtfU* gene and expresses an active GTF-S₁ enzyme.

References

- [1] N. Hanada, H.K. Kuramitsu, Infect. Immun. 57 (1989) 2079–2085.

- [2] A. Nanbu, M. Hayakawa, K. Takada, N. Shinozaki, Y. Abiko, K. Fukushima, *FEMS Immunol. Med. Microbiol.* 27 (2000) 9–15.
- [3] A. Shimamura, H. Tsumori, H. Mukasa, *FEBS Lett.* 157 (1983) 79–84.
- [4] Y. Yamashita, N. Hanada, T. Takehara, *J. Bacteriol.* 171 (1989) 6265–6270.
- [5] Y. Abiko, M. Hayakawa, S. Saito, H. Takiguchi, *Arch. Oral Biol.* 34 (1989) 571–575.
- [6] A.C. Chang, S.N. Cohen, *J. Bacteriol.* 134 (1978) 1141–1156.
- [7] Y. Namiki, I. Kantake, T. Ikeda, K. Fukushima, *IRCS Med. Sci.* 13 (1985) 501–502.
- [8] K. Fukushima, R. Motoda, K. Takada, T. Ikeda, *FEBS Lett.* 128 (1981) 213–216.
- [9] M. Hayakawa, K. Fukushima, Y. Abiko, T. Ikeda, H. Takiguchi, *Biochem. Mol. Biol. Int.* 31 (1993) 1167–1175.
- [10] O. Ouchterlony, *Prog. Allergy* 5 (1958) 1–9.
- [11] A. Shimamura, H. Tsumori, H. Mukasa, *Biochim. Biophys. Acta* 702 (1982) 72–80.
- [12] G. Mooser, S.A. Hefta, R.J. Paxton, J.E. Shively, T.D. Lee, *J. Biol. Chem.* 266 (1991) 8916–8922.
- [13] N. Hanada, Y. Isobe, Y. Aizawa, T. Katayama, S. Sato, M. Inoue, *Infect. Immun.* 61 (1993) 2096–2103.
- [14] A. Shimamura, *Carbohydr. Res.* 185 (1989) 239–248.