Sequence Analysis of the *Streptococcus mutans* Ingbritt *dexA* Gene Encoding Extracellular Dextranase

Takeshi Igarashi*, Ayako Yamamoto, and Nobuichi Goto

Department of Oral Microbiology, Showa University School of Dentistry, Shinagawa-ku, Tokyo 142, Japan

Received May 31, 1995. Accepted July 26, 1995

**Abstract:** The complete nucleotide sequence (3,747 bp) of the dextranase gene (*dexA*) and flanking regions of the chromosome of *Streptococcus mutans* Ingbritt (serotype c) were determined. The open reading frame for *dexA* was 2,550 bp, ending with a stop codon TGA. A putative ribosome-binding site, promoter preceding the start codon, and potential stem-loop structure were identified. The presumed dextranase protein (*DexA*) consisting of 850 amino acids was estimated to have a molecular size of 94,536 Da and a pI of 4.79. The nucleotide sequence and the deduced amino acid sequences of *S. mutans dexA* exhibited homologies of 57.8% and 47.0%, respectively, to those of *Streptococcus sobrinus dex*. The homologous region of *dex* of *S. sobrinus* was in the N-terminal half. The C terminus of DexA consisted of a hexapeptide LPQTGD, followed by 7 charged amino acids, 21 amino acids with a strongly hydrophobic character, and a charged hexapeptide tail, which have been reported as a common structure of C termini of not only the surface-associated proteins of Gram-positive cocci but also the extracellular enzymes such as β-fructosidase of *S. mutans* and dextranase of *S. sobrinus*. The DexA protein had no significant homology with the glucosyltransferases, the glucan-binding protein, or the dextranase inhibitor of *mutans* streptococci.

**Key words:** Dextranase, *dexA* gene, *Streptococcus mutans*, Dextran

Mutans streptococci, especially *Streptococcus mutans* and *Streptococcus sobrinus*, are closely associated with dental caries. The formation of extracellular glucans from sucrose and the ability to produce acids from a variety of sugars are believed to be major cariogenic factors of these bacteria (13, 22). *S. mutans* is able to produce a water-soluble glucan (WSG), mainly α-1,6-linked glucan, and a water-insoluble glucan (WIG), mainly α-1,3-linked glucan (4, 13). Extracellular dextranase produced by *S. mutans* is an endo-enzyme which cleaves α-1,6-linkages in the glucans and produces isomaltosaccharides as its reaction products from the glucans (13, 18). Although the precise role(s) of extracellular dextranase in the virulence of *S. mutans* has not yet been clearly understood, the following observations have been reported. Dextranase can partially degrade WIG (11, 40), inhibit synthesis of WIG (32, 40), inhibit the adherence of the *S. mutans* cells to the smooth surface of the tooth (12, 32), and inhibit plaque formation and dental caries development (10, 27). Several hypotheses on dextranase function have been proposed: dextranase (i) may regulate WIG production by altering the ratio of α-1,6- to α-1,3-linkages in glucans (4, 21, 24, 41); and (ii) may act together with intracellular dextran glucosidase in a catabolic pathway which is responsible for acid production from glucans (26, 43). Experiments using a dextranase-deficient mutant obtained by insertional inactivation of the dextranase gene have shown that dextranase might play a role in controlling the amount and the nature of extracellular glucans, in adherence of *S. mutans*, and in the utilization of glucans as a carbohydrate source (5). In addition, a dextranase-negative mutant, which was obtained by chemical mutagenesis, was less cariogenic to experimental rats than the parent strain (37).

In our previous work, we purified and characterized the extracellular dextranase of *S. mutans* Ingbritt (serotype c) (18), and cloned the gene *dexA* in *Escherichia coli* cells (19). In this report, we describe the nucleotide sequence of *dexA* and the deduced amino acid sequence of the *dexA* gene product (DexA). A comparison of the nucleotide and amino acid sequences

---

*Address correspondence to Dr. Takeshi Igarashi, Department of Oral Microbiology, Showa University School of Dentistry, 1-5-8, Hatanodai, Shinagawa-ku, Tokyo 142, Japan.

Abbreviations: Dei, dextranase inhibitor; dex, *S. sobrinus* dextranase gene; Dex, *S. sobrinus* dextranase; DexA, *S. mutans* dextranase gene; DexA, *S. mutans* dextranase; FruA, exo-β-D-fructosidase; FTF, fructosyltransferase; GBP, glucan-binding protein; GPC-SP, Gram-positive coccal surface protein; GTF, glucosyltransferase; LB, Luria-Bertani; ORF, open reading frame; S.D., Shine-Dalgamo; SpaA, surface protein antigen A; SpaP, surface protein antigen I/II; WapA, wall-associated protein A; WIG, water-insoluble glucan; WSG, water-soluble glucan.
of dexA and DexA of *S. mutans* with those of the counterparts of other bacterial species is also discussed.

**Materials and Methods**

**Bacterial strains and plasmids.** *E. coli* strains DH5 (supE44 hsdR17 recA1 endA1 gyrA96 thi-1 relA1) (14) and JM109 (recA1 supE44 endA1 hsdR17 gyrA96 relA1 thi-(lac-proAB)F’(traD36 proAB + lacIq lacZ∆M15)) (44) were maintained on LB agar plates, grown in LB broth (23) at 37°C, and used as the host strains for transformation with recombinant plasmids. The recombinant plasmid pSD2 containing the *dexA* gene of *S. mutans* Ing Britt (serotype c) has been previously described (19). Charomid 9-36 (23) and pUC18 (44) were used for subcloning DNA fragments to be sequenced.

**Nucleotide sequencing and sequence analysis.** DNA fragments to be sequenced were subcloned into charomid 9-36 or pUC18. Sequential deletions of the insert fragments in the plasmids were performed by the exonuclease III and mungbean nuclease system (Takara Shuzo Co., Ltd., Kyoto, Japan). Plasmid DNA was isolated from the recombinant *E. coli* cells with the Wizard Prep DNA kit (Promega, Madison, Wis., U.S.A.). Double-stranded templates were used directly for sequencing. Sequencing was accomplished by the dideoxy chain-termination method (31) using Taq DNA polymerase supplied with the Taq Dye terminator sequencing kit (Perkin Elmer Applied Biosystems, Foster City, Calif., U.S.A.), and an ASTEC Program Temp Control System PC-700 (Sci-Media, Ltd., Tokyo). An automated DNA sequencer (Perkin Elmer Applied Biosystems Model 373A) was used to determine the nucleotide sequence. Sequence analysis was carried out with both the software Genetyx-Mac (ver. 7.2) and its CD-ROM databases (Software Development Co., Ltd., Tokyo).

**Nucleotide sequence accession number.** The *dexA* nucleotide sequence presented in this paper has been deposited in the DDBJ, EMBL, GSBD, and NCBI nucleotide sequence databases for the accession number D49430.

**Results**

**Nucleotide Sequence**

The nucleotide sequence of the entire insert in pSD2-del.4 (Fig. 1), which encoded an intact dextranase in the *E. coli* host (19), was determined. The complete sequence was confirmed in both directions and from at least three overlapping clones by using the deletion mutants. Restriction fragments of 0.4-kb *PstI-EcoRI*, 1.2-kb *EcoRI-SacI*, 0.8-kb *SacI-PstI*, and 0.9-kb *PstI-

![Fig. 1. Restriction map of pSD2 and the strategy used to sequence the 4-kb DNA fragment of pSD2-del.4 carrying the *dexA* gene of *S. mutans* Ing Britt (serotype c). The dextranase-coding region is depicted by a heavy arrow superimposed on the restriction map. The arrows below the map represent the extent and direction of the DNA sequencing which was performed with deletion mutants. Restriction sites: B, BamHI; E, EcoRI; H, HindIII; P, PstI; S, SacI.](image)

**Amino Acid Sequence**

The amino acid sequence deduced from the nucleotide sequence revealed that the product of *dexA*-ORF (DexA) is an acidic protein of 850 amino acids (Fig. 2) with a calculated molecular mass of 94,537 Da and a pI of 4.79. There were no cysteine residues found in the deduced amino acid sequence from the *dexA* gene (Fig. 2). This is consistent with the observations that most extracellular proteins from *mutans* streptococci appear to contain little or no cysteine (1, 3, 9, 30, 39).

The common motif for the C termini of Gram-positive cocal surface proteins (GPC-SP) such as streptococcal M protein, protein A, and fibronectin-binding protein of *Staphylococcus aureus*, and SpaP and WapA of *S. mutans* has been identified and partially characterized (8, 17, 20,
The common motif for the C-terminal portion of fructosyltransferase (FTF), FruA, glucan-binding protein could be detected between DexA and any of GTFs, amino acid sequences. Little or no sequence homologies cant homology with DexA in either the nucleotide or protein sequences. The dextranase of Arthrobacter CB-8 had no significant homology between these regions was 57.0%. S. sobrinus dextranase (Fig. 4). The homology of S. sobrinus dextranase (Fig. 4). The homology that homologous regions were found between positions 111 to 732 of S. mutans DexA and positions 169 to 800 of S. sobrinus dextranase (Fig. 4). The homology of S. mutans Ingbritt and S. sobrinus UAB66 (serotype g). We then compared the C-terminal portion of the extracellular enzymes to those of the surface proteins of mutans streptococci. As shown in Fig. 3, the C-terminal feature of the LPXTGX hexapeptide, membrane-spanning region, and charged tail were well conserved in all the extracellular and the surface proteins compared.

Homologies of DexA to Other Dextranases and Extracellular Enzymes in Oral Streptococci

The entire DNA and protein sequences of dexA and DexA were compared with those of the dextranase genes and their products of S. sobrinus UAB66 (serotype g) (42), which was a derivative of strain 6715, and Arthrobacter CB-8 (25). Significant homologies were found between the nucleotide sequences as well as the amino acid sequences in S. mutans Ingbritt and S. sobrinus UAB66. There was a 57.8% homology between the nucleotides of the genes and a 47.0% homology between the deduced amino acid sequences (data not shown). The alignment of the amino acid sequences revealed that homologous regions were found between positions 111 to 732 of S. mutans DexA and positions 169 to 800 of S. sobrinus dextranase (Fig. 4). The homology between these regions was 57.0%. S. sobrinus dextranase protein had a much longer heterologous C-terminal region than the S. mutans DexA protein (Fig. 4). The dextranase of Arthrobacter CB-8 had no significant homology with DexA in either the nucleotide or amino acid sequences. Little or no sequence homologies could be detected between DexA and any of GTFs, fructosyltransferase (FTF), FruA, glucan-binding protein and dextranase inhibitor (Dei)(data not shown)(1, 33, 38). The C-terminal portion of DexA had a number of common features with the cell wall-anchoring region of GPC-SP (Figs. 2 and 3), although dextranases has been found mainly in the culture suprnatant of S. mutans (13, 18, 41). The hexapeptide LPQTGD was at 35 to 40 residues from the C terminus of DexA. In GPC-SP, the highly conserved peptide LPXTGX is found at approximately 30 residues from the C termini of these proteins (8, 33). Following the LPQTGD peptide was a charged heptapeptide (NNETRSN) and then 21 highly hydrophobic amino acids. The protein terminated with a hexapeptide (KGRKND) carrying a net positive charge. These regions were structurally similar to the GPC-SP cell wall-anchoring region. The region preceding the LPQTGD sequence appeared to be enriched in Ser and Pro (Fig. 2), which seems to correspond to the cell wall-spanning region of the GPC-SP (8).

The common motif for the C-terminal portion of GPC-SP has also been reported in the extracellular polysaccharidases such as β-fructosidase (FruA) (3) of S. mutans GS5 (serotype c) and dextranase (42) of S. sobrinus UAB66 (serotype g). We then compared the C-terminal portion of the extracellular enzymes to those of the surface proteins of mutans streptococci. As shown in Fig. 3, the C-terminal feature of the LPXTGX hexapeptide, membrane-spanning region, and charged tail were well conserved in all the extracellular and the surface proteins compared.

Discussion

The present analysis of the nucleotide sequence of the dexA gene confirmed the restriction map previously proposed (19). The dexA sequence contained characteristic sequences for a prokaryotic ribosome-binding site (34) and promoter sequences (15, 16) immediately upstream of the putative initiation codon ATG. There was another ribosome-binding site AGG (positions 305 to 307) located 14 bp upstream from the proposed ATG initiation codon (Fig. 2). However, it appeared that the distance between this ribosome-binding site and the initiation codon was longer than observed for other S. mutans genes (1, 3, 7, 9, 20, 35, 36, 39, 42). Another candidate for the initiation codon was found at position 570 (Fig. 2), which was also accompanied with a putative ribosome-binding site AGGA (positions 554 to 558), but not any appropriate promoter sequences.

In the 321-bp fragment from the 5' end to the start codon, several stem-loop structures were present (Fig. 2). Two of them, located at positions 13 to 42 and at positions 184 to 214, were followed by a string of 5T's, suggesting that these sequences might be a set of transcriptional terminators of other genes upstream of dexA.

Although the secretory proteins of oral streptococci such as GTFs, FTF, Dei, and GBP (1, 9, 35, 36, 39) usually have a signal peptide at their N terminus, they have no cell wall-anchoring region at the C terminus. DexA is an unusual secretory protein as it has the putative cell wall-anchoring region at the C terminus as presented in Figs. 2 and 3. Schneewind et al (33) have demonstrated that the cell wall-anchoring region consisting of an LPXTGX peptide, a C-terminal hydrophobic region, and a charged tail of GPC-SP plays key roles in directing the protein to the cell surface. It is possible, therefore, that the C terminus of the DexA protein, i.e., the putative cell wall-anchoring region, also plays the same role as that of GPC-SP. It is uncertain, however, why the secretory DexA protein has such a C-terminal region and how this region functions in the DexA protein.

Two hypotheses could be proposed: one is that both the N- and the C-terminal regions are required for the secretion of the DexA protein, and another is that only the N-terminal region is required for secretion and the C-terminal region has another unknown function. Therefore it would be interesting to examine whether the putative cell wall-anchoring region on the C terminus of DexA is involved in DexA secretion. We are currently working
Fig. 2.
SEQUENCING OF S. MUTANS dexA GENE

It has been reported that GPC-SP usually has a Ser-, Thr-, Pro-, and Gly-rich wall-spanning region, which was 50 to 125 amino acids adjacent to the N-terminal side of the LPXTGX sequence. In the corresponding region of the DexA, there was quite a high content of Pro between positions 738 and 769, and the Ser content was high between positions 785 and 807. Therefore these regions are predicted to be involved in the wall-spanning function of DexA.

**Fig. 2.** Nucleotide sequence of the dexA gene and flanking regions of S. mutans Ingbritt with deduced amino acid sequence of the putative dextranase-precursor protein. The putative promoters (−35 and −10 sequences) and the predicted ribosome-binding site (Shine-Dalgarno [S.D.]) are underlined. The predicted ATG start codon is at position 322. The peptide LPQTGD is encoded by the nucleotides 2752 to 2769 (asterisks). The putative membrane-spanning domain is delineated by a dashed underline. The cell wall-anchor region is located immediately after the LPXTGX sequence and consists of seven charged residues followed by a 21-amino acid hydrophobic domain and a charged hexapeptide. The two facing arrows which are downstream of the stop codon at position 2872 (end) represent a potential stem-loop structure.
Fig. 3. Alignment of the amino acid sequences of the C-terminal regions of surface proteins and extracellular enzymes of mutans streptococci. The sequences are aligned along the LPXTGX motif common to all of the proteins (bold letters and boxed). The hydrophobic region found in all of those proteins is also boxed with a thinner line, to the right of which the charged tail is indicated by a bold letter. The cell wall-anchoring region consists of an LPXTGX motif, a C-terminal hydrophobic region, and a charged tail. Abbreviations: DexA, dextranase of S. mutans Ingbritt (serotype c) (this paper); Dex, dextranase of S. sobrinus UAB66 (serotype g) (GenBank accession no. M96978); FruA, β-fructosidase of S. mutans GS5 (serotype c) (GenBank accession no. L03387); SpaA, surface protein antigen of S. sobrinus MT3791 (serotype g) (GenBank accession no. D90354); SpaP, surface antigen I/II of S. mutans NG5 (serotype c) (GenBank accession no. X17390); and WapA, wall-associated protein A of S. mutans Ingbritt (serotype c) (GenBank accession no. M19346).

Fig. 4. Alignment of deduced amino acid sequences of the homologous regions of the S. mutans Ingbritt dextranase (DexA, this study) and the S. sobrinus UAB66 dextranase (Dex, [GenBank accession no. M96978]). The homologous regions in each gene are shown at the top of the figure. Identical amino acids are denoted by asterisks. Each number corresponds to the position of the residue in the DexA and Dex protein sequences. Hatched boxes at the C-terminal regions represent the Ser- and Thr-rich and cell wall-anchoring regions as shown in Fig. 3.
regions preceding the LPQTGD sequence might correspond to the wall-spanning region of the GPC-SP.

The molecular mass (94.5 kDa, this study) of DexA estimated from the deduced amino acid sequence was smaller than SDS-PAGE-estimated masses of the native dextranase (120 kDa) produced by S. mutans Ingbritt (19, 29) and the recombinant DexA (133 kDa) by E. coli cells (19). The same phenomenon was observed in S. sobrinus UAB66 dextranase, 143 kDa from the deduced molecular mass and 175 kDa from SDS-PAGE gels (2, 42). The molecular mass of the recombinant dextranase produced by E. coli cells is not clear because the complete dex gene of S. sobrinus UAB66 has not been cloned yet. Colby et al (5) reported that intact dextranase shows anomalous migration on SDS-PAGE and the exact size of dextranase can only be determined from the deduced amino acid sequence. As compared with the molecular masses from the deduced amino acid sequences, DexA (94.5 kDa) from S. mutans Ingbritt is smaller than that (143.3 kDa) from S. sobrinus UAB66. It is apparent that this difference (about 48.7 kDa) is due to the heterologous C terminus of the S. sobrinus dextranase, as shown in Fig. 4.

The deduced amino acid sequences of extracellular dextranase genes of S. mutans Ingbritt (dexA) and S. sobrinus UAB66 (dex) showed an extensive homology (Fig. 4). We have detected some active dextranases from the E. coli hosts carrying a pSD2-del.4-derived deletion mutant in which dexA had been deleted from the 3’ termini (data not shown). This implies that the C-terminus portion of DexA is not essential for the enzymatic activity. Wanda and Curtiss (42) also reported that while the N-terminal region of the dex gene from S. sobrinus was necessary to specify the Dex enzyme activity, the C terminus of the dex gene can be deleted to some extent without loss of enzymatic activity. These suggest that the homologous region as shown in Fig. 4 is essential for the enzymatic activity, and both the catalytic and the dextran-binding sites reside in this region. The proteins GTFs, GBP, and Dei, which have been known as glucan-binding proteins (1, 9, 28, 36, 39), have direct repeating units present in the dextran-binding domains which are the conserved sequence of about 23 amino acids in length (28). Although dextranases of both S. mutans and S. sobrinus also bind to dextran, the dextranases have no such repeats (Fig. 2 and ref. 42). The dextranases might interact with dextran at its catalytic site.

We thank Ms. Emiko Asaga for her technical assistance.

References