ELSEVIER

FEMS Microbiology Letters 192 (2000) 53–57

FEMS MICROBIOLOGY Letters

www.fems-microbiology.org

# Location of repeat elements in glucansucrases of *Leuconostoc* and *Streptococcus* species

Štefan Janeček [a,*], Birte Svensson [b], Roy R.B. Russell [c]

[a] *Institute of Microbiology, Slovak Academy of Sciences, Štefánikova 3, SK-81434 Bratislava, Slovak Republic*
[b] *Department of Chemistry, Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-2500 Copenhagen, Valby, Denmark*
[c] *Oral Biology, Dental School, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4BW, UK*

Received 22 August 2000; accepted 31 August 2000

**Abstract**

Glucosyltransferases of oral streptococci, dextransucrases and alternansucrase of *Leuconostoc mesenteroides*, collectively referred to as glucansucrases, are large extracellular enzymes that synthesise glucans with a variety of structures and properties. A characteristic of all these glucansucrases is the possession of a C-terminal domain consisting of a series of tandem amino acid repeats. These repeat units are thought to interact with glucan but closely resemble the cell wall binding domain motif found in choline binding proteins in *Streptococcus pneumoniae* and surface-located proteins in a range of other bacteria. Analysis of dextransucrase and alternansucrase sequences has now shown that they also contain these repeat motifs in the N-terminal region, raising questions about their evolutionary origin and functional importance. 

*Keywords:* *Streptococcus*; *Leuconostoc*; Glucansucrase; Glucosyltransferase; Dextransucrase; Alternansucrase

## 1. Introduction

Glucansucrases (EC 2.4.5.1) are extracellular enzymes produced by species of oral streptococci, *Leuconostoc mesenteroides* and some lactobacilli that synthesise α-linked glucans from sucrose. The glucans can differ in their size, types of linkage between glucose units and degree of branching and hence have different physical properties, particularly with regard to solubility, adhesiveness and susceptibility to breakdown by human or bacterial enzymes [1,2]. There is thus considerable interest in understanding their catalytic mechanism to gain insight into the process of dental disease and to explore the biotechnological applications of defined or novel glucans. In addition, their modular construction, their sharing of certain sequence motifs with other proteins and the fact that they are members of the α-amylase family raise intriguing questions about their evolution [3,4].

The first glucansucrase genes to be cloned were those for glucosyltransferases (GTFs) from oral streptococci and a basic pattern of organisation was identified [5]. GTFs are of high molecular mass, c. 160 kDa, and possess a signal peptide followed by a highly variable stretch of around 150 amino acids and a well-conserved catalytic core domain of c. 900 residues [6]. The catalytic core domain includes a region with the secondary structure of the classical $(\beta/\alpha)_8$-barrel fold of α-amylases, though with the component elements arranged in a circularly permutated form [3]. The C-terminal one-third of the molecule is capable of binding glucans and consists of a series of related but non-identical tandem repeats, each of 20–30 amino acids [7]. These repeats share sequence motifs with repeats found in proteins of other bacteria or bacteriophages [5,8]. As further streptococcal GTF sequences became available, it was found that all fell into this pattern, as do dextransucrases (DSRs) from *Leuconostoc mesenteroides* [2]. Recently, the sequence of alternansucrase (ASR) has been determined [9]. This enzyme produces a glucan with alternating α1,3 and α1,6 linkages and has a novel pattern of repeat units at its N-terminus. In the present study the nature and location of repeat units in GTFs and DSRs were closely examined and compared with ASR.

* Corresponding author. Tel.: +421 (7) 5249 2990; Fax: +421 (7) 5249 3824; E-mail: umikstef@savba.sk

0378-1097/00/$20.00 
PII: S0378-1097(00)00408-0

## 2. Materials and methods

Sequences were retrieved from GenBank or the European Bioinformatics Institute (EBI) databases. Sequences examined, and their accession numbers, were *Streptococcus mutans* GTF-B (AAA88588), GTF-C (AAA88589) and GTF-D (AAA26895); *Streptococcus gordonii* GTF-G (AAC43483); *Streptococcus downei* GTF-I (AAC41412) and GTF-S (AAA63063); *Streptococcus sobrinus* GTF-T (D13928); *Streptococcus salivarius* GTF-J (CAA77900), GTF-K (CAA77901), GTF-L (AAC41412) and GTF-M (AAC41413); *Streptococcus oralis* GTF-R (BAA95201); *Leuconostoc mesenteroides* ASR (CAB76565), DSR-A (JC5473), DSR-B (AAB95433) and DSR-S (I09598). Dot matrix comparisons were performed with the DNAMAN package (Lynnon BioSoft) and multiple alignments with CLUSTAL-W on the EBI server. Pattern recognition databases were searched through the Interpro (Integrated Resource of Protein Domains and Functional Sites) resource at http://www.ebi.ac.uk/interpro.

## 3. Results and discussion

### 3.1. N-terminal repeat elements in alternansucrase

The catalytic core domains of glucansucrases are well conserved and in ASR this domain starts at residue 350, aligning closely from this point on with GTF and DSR sequences. Immediately preceding the catalytic core, however, we have identified three repeat elements (Fig. 1). Two repeats are of 64 aa while the third is 52 aa long. Fig. 2 also shows that all three repeats align well with the most conserved residues of the 'A' repeats first identified in the C-terminal glucan binding domain of GTF-I and present in the C-terminus of all other glucansucrases sequenced [2,7,10]. The upstream variable region of ASR that extends from the signal peptide to position 162 shows no apparent match in sequence or structure to variable regions of other glucansucrases.

### 3.2. Relationship to other repeat elements

In the streptococcal GTFs, a variety of C-terminal repeat elements have been described (A, B, C, D), the only one of which found in all glucansucrase is the A repeat. The alignment in Fig. 1 shows that this conserved element is part of the N-terminal repeats in ASR. A similar 20-aa motif has been recognised in the cell wall amidase of *Streptococcus pneumoniae* and its bacteriophages and repeat-containing modules have been shown to be responsible for binding to choline in the lipoteichoic acid of this organism [11]. More recently, further choline binding proteins have been recognised in *S. pneumoniae* and represent a new means of anchorage to the cell surface [12]. Besides their occurrence in *S. pneumoniae*, repeats with the same conserved residues are found in the toxins of *Clostridium difficile* that recognise the trisaccharide isoglobotriose in the gut lining [8,13,14] and in surface-associated proteins from *Erysipelothrix rhusiopathiae*, *Clostridium acetobutylicum* and *Peptostreptococcus micros* [15–17]. The widespread occurrence of this distinctive pattern of conserved residues in a range of proteins has led to the incorporation of this 20-aa motif as the CW 'putative cell wall binding domain' in the Pfam motif-seeking program [18]. Analysis of the ASR sequence by Pfam through the Interpro interface located CW repeats in ASR and Fig. 1 shows how these align, with a triplet of CW motifs corresponding to the location of the N-terminal repeats in ASR. The most highly conserved region comes in the first part of the triplets and this corresponds to the A repeat first identified in GTF-I.

### 3.3. N-terminal repeats in dextransucrases

Examination of the N-terminal regions of two dextransucrases from *L. mesenteroides*, DSR-B and DSR-S, shows that they both contain repeats N-terminal to the start of the catalytic core (Fig. 2) aligning to the second and third repeats on ASR. In marked contrast, none of the streptococcal GTFs have these long repeats at their N-terminus. However, the existence of a small region of similarity to the C-terminal repeats was previously noted [7,19] and is illustrated in Fig. 2. All GTFs, DSRs and ASR can be aligned from this point, which represents the start of the highly conserved catalytic core domain found in all the glucansucrases [2].

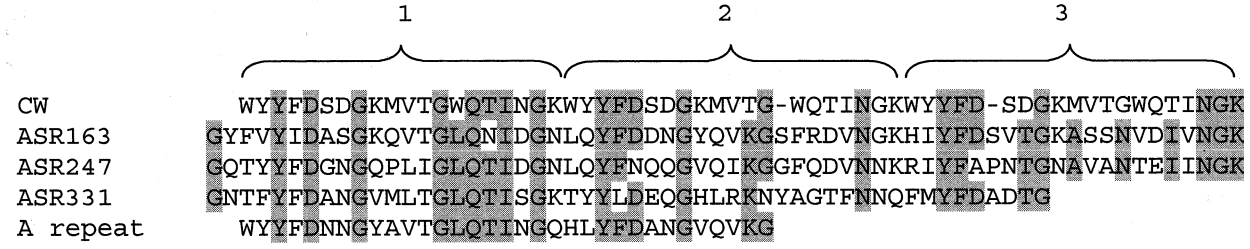


Fig. 1. Alignment of the three repeat elements from the N-terminal region of ASR of *L. mesenteroides*, the A repeat present in streptococcal GTF and the CW motif identified in the Pfam database. The most highly conserved residues are indicated by shading.

```
ASR          ------------------------------------------GYFVYIDASGKQVTGLQ  179

ASR          NIDGNLQYFDDNGYQVKGSFRDVNGKHIYFDSVTGKASSNVDIVNGKAQGYDAQGNQLKK  239

ASR          SYVADSSGQTYYFDGNGQPLIGLQTIDGNLQYFNQQGVQIKGGFQDVNNKRIYFAPNTGN  299
DSR-B        ---------WYYVTSDNTLAKGLTTVDNHKQYFDNNGVQAKGQFVTDNSKTYYLDPNSGN  229
DSR-S        ---------WYYYFDDGKNAKGLSTIDNNIQYFYESGKQAKGQYVTIDNQTYYFDKGSGD  247

ASR          AVANTEIINGKLQGRDANGNQVKNAFSKDVAGNTFYFDANGVMLTGLQTISGKTYYLDEQG  360
DSR-B        AVTGIQQIGSQTLAFNDNGEQVFADFYTAPDGKTYYFDDKGQATIGLKAINGHNYYFDSLG  259
DSR-S        ELTGLQSIDGNIVAFNDEGQQIFNQYYQSENGTTYYFDDKGHAATGIKNIEGKNYYFDNLG  277
DSR-A        ------------------------------------------------VDGKVYFYGDDG   17
GTF-B        ------------------------------------------------IDGKYYYYDNNG  176
GTF-C        ------------------------------------------------VNGKYYYYKEDG  201
GTF-D        ------------------------------------------------IDGKYYYIGSDG  188
GTF-G        ------------------------------------------------IDGKKYYVQDDG  242
GTF-I        ------------------------------------------------VDGKYYYYDQDG  174
GTF-J        ------------------------------------------------IDGKYYYVNEDG  190
GTF-K        ------------------------------------------------KDGKYYYLLEDG  185
GTF-L        ------------------------------------------------INGKQYYVEDDG  220
GTF-M        ------------------------------------------------IDGKQYYVEN-G  284
GTF-R        ------------------------------------------------IDGKNYYVQDDG  240
GTF-S        ------------------------------------------------VDGKTYYVDANG  164
GTF-T        ------------------------------------------------IDGKFYYVMADG  178
```

Fig. 2. Multiple alignment of the start of the highly conserved catalytic core domain and immediately upstream sequences of glucansucrases. The N-terminal repeats identified in alternansucrase (Fig. 1) are underlined and conserved identical residues shaded. Within the variable domain, a dash indicates that the sequence shows no similarity to any of the other sequences.

### 3.4. C-terminal repeats in alternansucrase

All known GTF and DSR sequences contain between three and six A repeats in the C-terminal one-third of the molecule. In contrast, ASR has only a single well defined A repeat starting at position 1413. ASR does, nevertheless, have its own novel and distinctive set of repeat elements at its C-terminus, shown in Fig. 3. ASR thus presents a number of interesting and distinctive features that set it apart from the other glucansucrases – multiple long N-terminal repeats, only a single C-terminal A repeat, and novel short repeats.

```
1679   DGLFLNAPY
1757   DGLFLNAPY
1600   DGLFLNAPY
1836   DGLWLNAPY
1508   DGLFANAPY
1916   DGMFKTAPY
1995   DGVFSGAPY
```

Fig. 3. Alignment the seven unique repeat sequences from the C-terminal domain of ASR.

### 3.5. Function of repeat domains

The C-terminal repeat domain of streptococcal GTFs binds to glucans and this is proposed to contribute to sucrose-dependent adherence [1,7,20,21]. In some GTFs, but not all, it is also essential for the retention of glucan synthesis activity and the stimulatory effect of dextran [22,23]. It is clear, however, that the N-terminal repeats identified in this paper are not essential for glucan synthesis; DSR-A is a naturally truncated enzyme that has no N-terminal repeats or variable region [24] and in GTF it is possible by genetic manipulation to remove all the variable region while retaining function [20,22]. Indeed, even highly conserved residues at the start of the conserved catalytic core are not essential for activity [22,25]. *Leuconostoc* DSRs also bind to dextran but the regions of the molecule involved in binding have not yet been identified. Identification of regions of repeats will help the design of N- and C-terminal truncation experiments to address this question and throw further light on their functional significance.

It is not yet clear whether the similarity between repeats in glucansucrases and CW motifs is a consequence of convergent or divergent evolution though their widespread occurrence in different contexts complies with the definition of an independent, mobile domain. In surface proteins, wall lytic enzymes and bacteriophages of *S. pneumo-*

*niae*, the binding module is generally found at the C-terminus of enzyme domains but has recently been identified at the N-terminus of an autolytic lysozyme [26] and is also functional as an independent module [27]. This report shows that repeats in glucansucrases similarly can be found in different relationships to catalytic domains; they also occur in proteins without enzymatic activity such as the glucan binding protein of *S. mutans* and the dextranase inhibitor of *S. sobrinus* [19,28].

In *S. pneumoniae* and *C. acetobutylicum*, blocks of CW motifs make up a choline binding module that serves to anchor or target proteins to the cell surface [17]. This anchoring function is also proposed to be important in surface proteins of *E. rhusiopathiae* and *P. micros* though in these cases the ligand has not been identified [15,16]. The glucansucrases of oral streptococci and *Leuconostoc* are generally regarded as extracellular enzymes but nevertheless they can also be detected in a cell-associated form. C-terminal deletions result in an increase in the secretion of GTF [29,30]. The observations described here suggest that further exploration is required into the contribution of N- and C-terminal repeats to surface localisation as well as to the complex process of glucan synthesis and binding. Identification of repeat sequences will thus be of value in guiding experiments involving truncation and mutagenesis of these complex enzymes.

## Acknowledgements

Š.J. is grateful to FEBS for a Short-term Fellowship to visit the Carlsberg Laboratory and VEGA (Slovakia) for Grant 2/6045/1999-2001. B.S. and R.R.B.R. received support from the European Union (BIO4-CT98-0022).

## References

[1] Colby, S.M. and Russell, R.R.B. (1997) Sugar metabolism by mutans streptococci. J. Appl. Microbiol. 83, S80–S88.
[2] Monchois, V., Willemot, R.M. and Monsan, P. (1999) Glucansucrases: mechanism of action and structure-function relationships. FEMS Microbiol. Rev. 23, 131–151.
[3] MacGregor, E.A., Jespersen, H.M. and Svensson, B. (1996) A circularly permuted α-amylase-type α/β-barrel structure in glucan-synthesizing glucosyltransferases. FEBS Lett. 378, 263–266.
[4] Janeček, Š. (1997) α-Amylase family: molecular biology and evolution. Progr. Biophys. Mol. Biol. 67, 67–97.
[5] Russell, R.R.B. (1990) Molecular genetics of glucan metabolism in oral streptococci. Arch. Oral Biol. 35, S53–S58.
[6] Monchois, V., Lakey, J.H. and Russell, R.R.B. (1999) Secondary structure of *Streptococcus downei* GTF-I glucansucrase. FEMS Microbiol. Lett. 177, 243–248.
[7] Ferretti, J.J., Gilpin, M.L. and Russell, R.R.B. (1987) Nucleotide-sequence of a glucosyltransferase gene from *Streptococcus sobrinus* MFE28. J. Bacteriol. 169, 4271–4278.
[8] Wren, B.W. (1991) A family of clostridial and streptococcal ligand-binding proteins with conserved C-terminal repeat sequences. Mol. Microbiol. 5, 797–803.
[9] Argüello-Morales, M.A., Remaud-Simeon, M., Pizzut, S., Sarçabal, P., Willemot, R.M. and Monsan, P. (2000) Sequence analysis of the gene encoding alternansucrase, a sucrose glucosyltransferase from *Leuconostoc mesenteroides* NRRL B-1355. FEMS Microbiol. Lett. 182, 81–85.
[10] Vickerman, M.M., Sulavik, M.C., Nowak, J.D., Gardner, N.M., Jones, G.W. and Clewell, D.B. (1997) Nucleotide sequence analysis of the *Streptococcus gordonii* glucosyltransferase gene, *gtfG*. DNA Seq. 7, 83–95.
[11] García, P., García, J.L., García, E., Sánchez-Puelles, J.M. and López, R. (1990) Modular organization of the lytic enzymes of *Streptococcus pneumoniae* and its bacteriophages. Gene 86, 81–88.
[12] Rosenow, C., Ryan, P., Weiser, J.N., Johnson, S., Fontan, P., Ortqvist, A. and Masure, H.R. (1997) Contribution of novel choline-binding proteins to adherence, colonization and immunogenicity of *Streptococcus pneumoniae*. Mol. Microbiol. 25, 819–829.
[13] von Eichel-Streiber, C. and Sauerborn, M. (1990) *Clostridium difficile* toxin A carries a C-terminal repetitive structure homologous to the carbohydrate binding region of streptococcal glycosyltransferases. Gene 96, 107–113.
[14] von Eichel-Streiber, C., Sauerborn, M. and Kuramitsu, H.K. (1992) Evidence for a modular structure of the homologous repetitive C-terminal carbohydrate-binding sites of *Clostridium difficile* toxins and *Streptococcus mutans* glucosyltransferases. J. Bacteriol. 174, 6707–6710.
[15] Kremer, B.H.A., Bijlsma, J.J.E., Kusters, J.G., de Graaff, J. and van Steenbergen, T.J.M. (1999) Cloning of *fibA*, encoding an immunogenic subunit of the fibril-like surface structure of *Peptostreptococcus micros*. J. Bacteriol. 181, 2485–2491.
[16] Makino, S., Yamamoto, K., Murakami, S., Shirahata, T., Uemura, K., Sawada, T., Wakamoto, H. and Morita, H. (1998) Properties of repeat domain found in a novel protective antigen, SpaA of *Erysipelothrix rhusiopathiae*. Microb. Pathogen. 25, 353.
[17] Sanchez-Beato, A.R., Ronda, C. and García, J.L. (1995) Tracking the evolution of the bacterial choline-binding domain: molecular characterization of the *Clostridium actobutylicum* NCIB 8052 *cspA* gene. J. Bacteriol. 177, 1098–1103.
[18] Bateman, A., Birney, E., Durbin, R., Eddy, S.R., Howe, K.L. and Sonnhammer, E.L. (2000) The Pfam protein families database. Nucleic Acids Res. 28, 263–266.
[19] Banas, J.A., Russell, R.R.B. and Ferretti, J.J. (1990) Sequence-analysis of the gene for the glucan-binding protein of *Streptococcus mutans* Ingbritt. Infect. Immun. 58, 667–673.
[20] Abo, H., Matsumura, T., Kodama, T., Ohta, H., Fukui, K., Kato, K. and Kagawa, H. (1991) Peptide sequences for sucrose splitting and glucan binding within *Streptococcus sobrinus* glucosyltransferase (water-insoluble glucan synthetase). J. Bacteriol. 173, 989–996.
[21] Mooser, G. and Wong, C. (1988) Isolation of a glucan-binding domain of glucosyltransferase (1,6-α-glucan synthase) from *Streptococcus sobrinus*. Infect. Immun. 56, 880–884.
[22] Monchois, V., Vignon, M. and Russell, R.R.B. (1999) Isolation of key amino acid residues at the N-terminal end of the core region *Streptococcus downei* glucansucrase, GTF-I. Appl. Microbiol. Biotechnol. 52, 660–665.
[23] Monchois, V., Argüello-Morales, M.A. and Russell, R.R.B. (1999) Isolation of an active catalytic core of *Streptococcus downei* MFe28 GTF-I glucosyltransferase. J. Bacteriol. 181, 2290–2292.
[24] Monchois, V., Willemot, R.M., Remaud-Simeon, M., Croux, C. and Monsan, P. (1996) Cloning and sequencing of a gene coding for a novel dextransucrase from *Leuconostoc mesenteroides* NRRL B-1299 synthesizing only α(1–6) and α(1–3) linkages. Gene 182, 23–32.
[25] Tsumori, H., Minami, T. and Kuramitsu, H.K. (1997) Identification of essential amino acids in the *Streptococcus mutans* glucosyltransferases. J. Bacteriol. 179, 3391–3396.
[26] García, P., González, M.P., García, E., García, J.L. and López, R. (1999) The molecular characterization of the first autolytic lysozyme of *Streptococcus pneumoniae* reveals evolutionary mobile domains. Mol. Microbiol. 33, 128–138.

[27] Sánchez-Puelles, J.M., Sanz, J.M., García, J.L. and García, E. (1990) Cloning and expression of gene fragments encoding the choline-binding domain of pneumococcal murein hydrolases. Gene 89, 69–75.

[28] Sun, J.W., Wanda, S.Y., Camilli, A. and Curtiss, R. (1994) Cloning and DNA-sequencing of the dextranase inhibitor gene (*dei*) from *Streptococcus sobrinus*. J. Bacteriol. 176, 7213–7222.

[29] Kato, C. and Kuramitsu, H.K. (1991) Molecular-basis for the association of glucosyltransferases with the cell-surface of oral streptococci. FEMS Microbiol. Lett. 79, 153–158.

[30] Vickerman, M.M. and Clewell, D.B. (1997) Deletions in the carboxy-terminal region of *Streptococcus gordonii* glucosyltransferase affect cell-associated enzyme activity and sucrose-associated accumulation of growing cells. Appl. Environ. Microbiol. 63, 1667–1673.