ELSEVIER

FEMS Microbiology Letters 177 (1999) 243–248

FEMS MICROBIOLOGY LETTERS

Secondary structure of *Streptococcus downei* GTF-I glucansucrase

Vincent Monchois [a,*], Jeremy H. Lakey [b], Roy R.B. Russell [a]

[a] *Department of Oral Biology, The Dental School, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4BW, UK*

[b] *Department of Biochemistry and Genetics, The Medical School, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4HH, UK*

Received 16 June 1999; accepted 25 June 1999

Abstract

Multiple sequence alignment and structure prediction of glucansucrases produced by oral streptococci and *Leuconostoc mesenteroides* showed that all have common structural features, with three major domains. There is no conservation of primary sequence or structure in the N-terminal variable region. Sequence-based structure prediction combined with circular dichroism spectrum analysis of purified truncated forms of *Streptococcus downei* GTF-I revealed that the core catalytic region has a defined structure consistent with the proposed $(\alpha/\beta)_8$-barrel structure. The C-terminal domain is a mixed structure with significant amounts of β-sheet and random-coil. This information contributes to the development of our understanding of structure-function relationships in glucansucrases.

Keywords: Glucosyltransferase; Glucansucrase; Glycosylhydrolase; Glucan-binding; *Streptococcus*; *Leuconostoc*

1. Introduction

Glucansucrases (E.C. 2.4.1.5) are extracellular enzymes from oral streptococci (commonly named glucosyltransferases, GTFs) or *Leuconostoc mesenteroides* (commonly named dextransucrases, DSRs) that catalyze the synthesis from sucrose of high molecular mass D-glucose polymers called glucans. Properties of these glucans differ, depending on molecular size, nature of the glucosyl linkages, and degree of branching. In the presence of acceptor molecules in the reaction medium, glucansucrases catalyze the synthesis of low molecular mass oligosaccharides [1]. Glucansucrases from oral streptococci are of profound medical interest as they are key enzymes in the formation of dental plaque and induction of human dental caries [2,3] while glucansucrases of *Leuconostoc* are of industrial interest for the production of dextran and oligosaccharides that have potential as prebiotic food constituents [4]. Comparative studies of the enzymes should help explain their mechanism, throw light on their evolution and facilitate structural studies.

Multiple alignment of deduced amino acid sequences of glucansucrases has allowed recognition of a common pattern of organization [1,4,5]. They are of high molecular mass, around 160 kDa. With

* Corresponding author. Tel.: +44 (191) 222 8391; Fax: +44 (191) 222 6137; E-mail: vincent.monchois@ncl.ac.uk

0378-1097/99/$20.00
PII: S0378-1097(99)00317-1

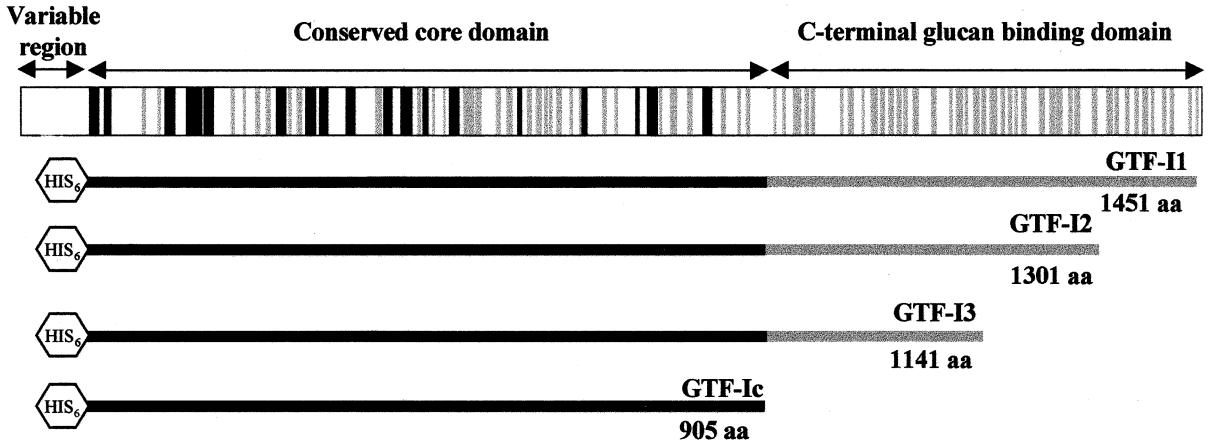

Fig. 1. Domain structure of *S. downei* GTF-I and location of predicted secondary structure. Black shades correspond to predicted α-helices and grey shades to predicted β-sheets. The lower part of the figure shows the extent of the truncated forms of the enzyme used in CD studies.

the exception of DSR-A of *L. mesenteroides* B1299 [6], they possess an N-terminal highly variable region containing 127–232 amino acids followed by a conserved central core region of about 900 amino acids including catalytic residues, and a C-terminal glucan-binding domain covering about 400 amino acids and containing a series of tandem repeats [7,8]. No 3-dimensional (3-D) structure of a glucansucrase has yet been determined but molecular modeling studies based on sequence similarities between glucansucrases and glycosylhydrolases of the α-amylase family have suggested that the core region contains a cyclically permuted form of the $(\alpha/\beta)_8$ TIM-barrel structure [9,10].

Streptococcus downei is one of the mutans group of oral streptococci [11]. It expresses four GTFs including GTF-I which produces α(1-3)-rich glucan [8,12]. Engineered derivatives of GTF-I possessing the core region alone or with various lengths of the glucan-binding domain (Fig. 1) have been expressed in *Escherichia coli* and purified by the use of a histidine affinity tag [13]. Analysis of the properties of these forms showed that neither the N-terminal variable region nor the glucan-binding domain played a significant part in determining the enzyme activity or stability of GTF-I [13].

In this study, the secondary structure pattern of glucansucrases was explored by combining for the first time circular dichroism spectroscopy of purified variants of GTF-I with multiple sequence alignments and molecular modelling.

2. Materials and methods

2.1. Production and purification of truncated forms of GTF-I

Production and purification of GTF-Ic, GTF-I3, GTF-I2 and GTF-I1 were performed as previously described [13], the enzymes being purified to homogeneity from sonicated *E. coli* extracts under native conditions using Ni-nitrilotriacetic acid agarose resin (Qiagen). Protein concentrations were determined by UV absorption at 280 nm using calculated extinction coefficients.

2.2. Computational analysis

The evolutionary tree of deduced amino acid sequences was constructed using CLUSTALW [14] and was drawn using the Treeview program [15]. Secondary structure analysis for each GTF sequence was achieved using the PREDATOR program [16].

2.3. Circular dichroism (CD)

Samples of GTF-I proteins were prepared in 50

mM Tris-HCl pH 7.5 at concentrations around 0.5 mg ml^{-1}. The CD spectrum of each sample solution was recorded at 20°C with a Jobin Yvon CD6 spectrophotometer using a 0.2 mm pathlength cell and protein concentrations $\approx$0.5 mg ml^{-1}. Data are presented in units of $\Delta\varepsilon$ (M^{-1} cm^{-1}). Twenty scans were collected at 0.5 nm intervals and smoothed using a five point FFT procedure in Origin™ graphics software. Secondary structure was predicted using the self-consistent algorithm of Sreerama and Woody [17], part of the DICROPROT package [18].

3. Results and discussion

3.1. Relationships between glucansucrases

To examine the relationship between the variable and core regions of glucansucrases from oral strep-

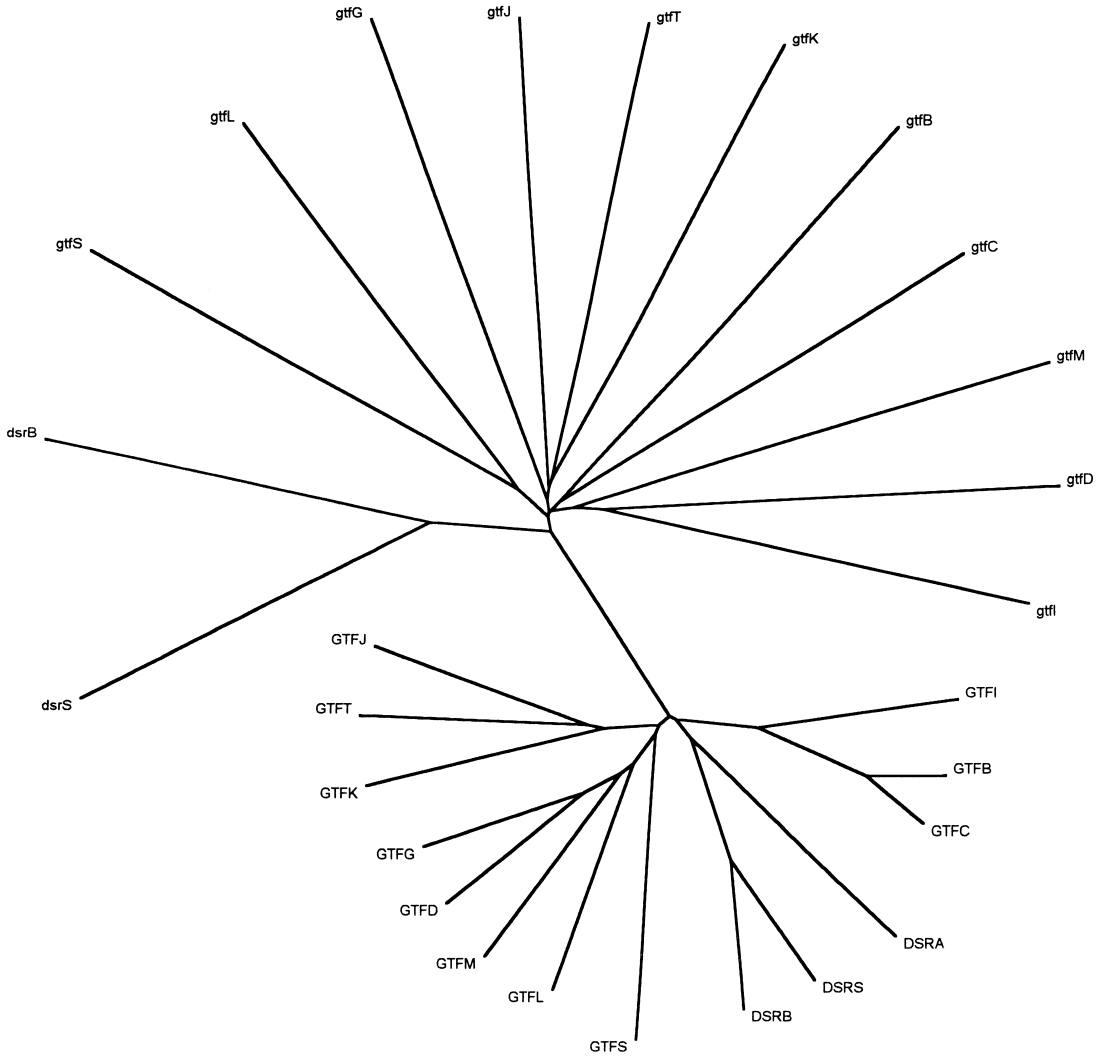

Fig. 2. Phylogenetic tree of the variable (top part of figure) and core (lower part of figure) domains of glucansucrases. GTF-B, C, D are from *S. mutans*; GTF-G from *S. gordonii*; GTF-S, I from *S. downei*; GTF-J, K, L, M from *S. salivarius*. DSR-A, B and S are produced by *L. mesenteroides*. Branch lengths are proportional to the sequence divergence.

tococci and *L. mesenteroides*, an evolutionary tree was calculated by aligning together the relevant sections of sequences available in the databases (Fig. 2). The C-terminal glucan-binding domains were not included because of the considerable variation in their lengths, which makes such analysis inappropriate. The N-terminal variable regions showed great heterogeneity, with no conservation of primary sequence or of predicted secondary structure. This suggests that there is no selective pressure for maintenance of an essential structure associated with function. Experimentally, it has been shown that enzymic function is retained even if the variable region is deleted [5,13,19] and argues against the proposal by Devulapalle et al. [10] that this region contains important functional or structural elements conserved amongst different glucansucrases.

Each GTF appears to have a distinctive variable region and in those cases where it seems probable that gene duplication has occurred (e.g. GTFB/C and J/K [5,20]) divergence of variable regions has taken place after duplication while the cores remain closely similar (Fig. 2). In *Streptococcus mutans*, where sequences are available from multiple isolates [21], the variable regions distinctive for each GTF are essentially the same in all isolates, indicating that each gene retains its distinctive features, regardless of the genetic background.

It is interesting to note that variability was observed in glucansucrases from both oral streptococci and *L. mesenteroides*. It has previously been proposed that the variation seen in the streptococcal GTFs might represent antigenic variation to avoid immune challenge in the mouth. However, the discovery that the glucansucrases of *L. mesenteroides*, a free-living soil organism, show similar variation is evidence of some other mechanism for maintaining variability. The role of the variable region therefore remains unknown, but the possibility that it is a form of proprotein associated with secretion or folding is worthy of investigation.

3.2. Secondary structure prediction

In contrast to the variable regions, the core and the glucan-binding regions possess distinct secondary structure patterns identified by multiple alignment of all the glucansucrase sequences combined with secondary structure prediction. The core region is predicted to contain alternating α-helices and β-sheets, which agrees with the presence of a TIM-barrel superfold structure proposed by MacGregor et al. [9]. Fig. 3 illustrates the results for GTF-I but, as would be expected from the close sequence alignment of core regions [1], the predicted helices and sheets align in all the glucansucrases.

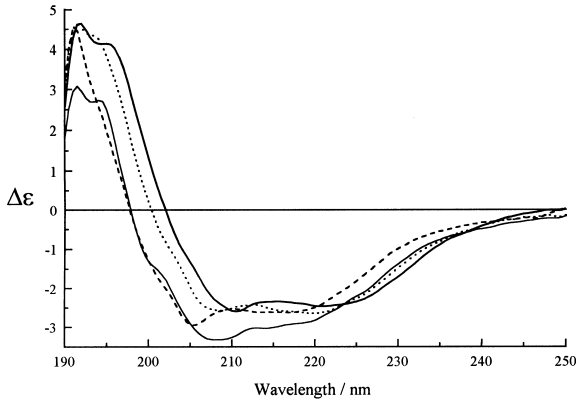

Fig. 3. Circular dichroism spectra of GTF-I. GTF-I1 (line), GTF-I2 (dashed line), TF-I3 (dotted line), and GTF-Ic (bold line).

Previous sequence analysis has shown that the C-terminal glucan-binding regions are composed of a series of related but non-identical tandem repeats [1,7,8]. While these result in short stretches of conserved structure, the number and arrangement of repeats varies with each enzyme. In addition, between these homologous repeat units are stretches of varying length of non-conserved sequence suggesting considerable structural flexibility within this domain, even though all these enzymes bind to glucan. This flexibility in the glucan-binding region is also revealed by the structure prediction algorithm, which indicates it to be composed almost entirely of β-sheets (Fig. 3).

3.3. Circular dichroism analysis of GTF-I derivatives

In order to confirm the prediction for secondary structure, far-UV circular dichroism spectra were measured for each of the purified derivatives of GTF-I produced by *S. downei*: GTF-I1, GTF-I2, GTF-I3 and GTF-Ic (Fig. 1).

The far-UV CD spectrum of GTF-Ic showed two

Table 1
Content of different types of secondary structure in GTF-I and truncated derivatives calculated from the data illustrated in Fig. 3

Protein	Length (aa)	α-helix		β-structure		Other	
		n	%	*n*	%	*n*	%
GTF-Ic	905	280	31	371	41	244	27
GTF-I3	1141	262	23	570	50	308	27
GTF-I2	1301	312	24	572	44	416	32
GTF-I1	1451	406	28	609	42	435	30

Numbers of amino acids (*n*) and percent of content (%).

negative bands at 210 and 222 nm, characteristic of the presence of α-helical secondary structure (Fig. 3). However, the low intensity of the signal (−2.5 Δε) indicated that the fraction of core region amino acids in defined secondary structure elements was not high. This observation is consistent with the computational analysis of the data [17,18] which indicated that the α-helix content was about 30% with a similar content of β-sheet (Table 1). This information thus supports the structure prediction proposed for the core region of glucansucrases.

The CD spectra of GTF-I1, GTF-I2 and GTF-I3, with increasing lengths of the C-terminal domain, all showed an increased intensity at 215 nm compared to GTF-Ic, indicating an increase in β-sheet content (Fig. 3). Of particular note is the shift in the wavelength at which crossover occurs on the zero axis. For both α- and β-structures this occurs in the range 202–205 nm and indeed the core fits this pattern. However, as the length of C-terminal domain increases in each of the variants, there is a shift of the crossover point to a lower wavelength. Since random-coil structure is unique in displaying a negative signal between 200 nm and 190 nm [22], this shift is most likely to be due to an increase in flexibility in the C-terminal region of the protein. This shift occurred particularly between GTF-I2 and GTF-I3 and may correspond to the large insertions between repeat units seen in GTF-I primary sequence [8]. The increase in the minimum at 215 nm together with the sequence-based prediction (Table 1) suggests that β-sheets form a higher proportion of the C-terminal glucan-binding domain of GTF-I than the core region.

Recently, Haas et al. [23] have described the CD spectrum of the glucan-binding protein (GBP-A) from *S. mutans*, which has strong sequence similarity to the C-terminal glucan-binding domains of GTFs [25]. GPB-A far-UV spectra contain a distinctive aromatic side chain contribution, which is clearly not present in our data. This is likely to be due to specific interactions between single residues [24] and its absence from our data is not indicative of a significantly different fold. This feature complicates analysis of GBP-A secondary structure but it also is predicted to contain a high content of β-sheets and negligible α-helix [23].

Thus, we have experimentally found that while the catalytic core region of GTF-I has a defined structure consistent with the proposed $(\alpha/\beta)_8$-barrel structure, the C-terminal domain is a mixed structure indicative of significant amounts of β-sheet and random-coil. According to the multiple alignment combined with structure prediction of glucansucrases of *L. mesenteroides* and oral streptococci, this pattern of organization appears to be a common feature of all the enzymes of this family. This information will be of value in determining structure/function relationships of these important enzymes.

Acknowledgements

This work was supported by Wellcome Trust grant 049554 and the European project BIOTECH CT98-0022.

References

[1] Monchois, V., Willemot, R.M. and Monsan, P. (1999) Glucansucrases: mechanism of action and structure-function relationships. FEMS Microbiol. Rev. 23, 131–151.
[2] Loesche, W.J. (1986) Role of *Streptococcus mutans* inhuman dental decay. Microbiol. Rev. 50, 353–380.
[3] Colby, S.M. and Russell, R.R.B. (1997) Sugar metabolism by

mutans streptococci. J. Appl. Microbiol. Symp. Suppl. 83, 80S–88S.
[4] Djouzi, Z., Andrieux, C., Pelenc, V., Somarriba, S., Popot, F., Paul, F., Monsan, P. and Szylit, O. (1995) Degradation and fermentation of α-glucooligosaccharides by bacterial strains from human colon: in vitro and in vivo studies in gnotobiotic rats. J. Appl. Bacteriol. 79, 117–127.
[5] Giffard, P.M., Allen, D.M., Milward, C.P., Simpson, C.L. and Jacques, N.A. (1993) Sequence of the *gtfK* gene of *Streptococcus salivarius* ATCC 25975 and evolution of the *gtf* genes of oral streptococci. J. Gen. Microbiol. 139, 1511–1522.
[6] 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.
[7] Giffard, P.M. and Jacques, N.A. (1994) Definition of a fundamental repeating unit in streptococcal glucosyltransferase glucan-binding regions and related sequences. J. Dent. Res. 73, 1133–1141.
[8] 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.
[9] 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.
[10] Devulapalle, K.S., Goodman, S.D., Gao, Q., Hemsley, A. and Mooser, G. (1997) Knowledge-based model of a glucosyltransferase from the oral bacterial group of mutans streptococci. Protein Sci. 6, 2489–2493.
[11] Whiley, R.A. and Beighton, D. (1998) Current classification of the oral streptococci. Oral Microbiol. Immun. 13, 195–216.
[12] Russell, R.R.B., Gilpin, M.L., Mukasa, H. and Dougan, G. (1987) Characterization of glucosyltransferase expressed from a *Streptococcus sobrinus* gene cloned in *Escherichia coli*. J. Gen. Microbiol. 133, 935–944.
[13] Monchois, V., Arguello-Morales, M.A. and Russell, R.R.B. (1999) Isolation of an active catalytic core of *Streptococcus downei* MFe28 glucosyltransferase, GTF-I. J. Bacteriol. 181, 2290–2292.
[14] Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTALW: Improving the sensitivity of progressive sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.
[15] Page, R.D.M. (1996) TREEVIEW: An application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. 12, 357–358.
[16] Frishman, D. and Argos, P. (1997) Seventy-five percent accuracy in protein secondary structure prediction. Proteins 27, 239–245.
[17] Sreerama, N. and Woody, R.W. (1994) Protein secondary structure from circular dichroism spectroscopy. Combining variable selection principle and cluster analysis with neural network, ridge regression and self-consistent methods. J. Mol. Biol. 242, 497–507.
[18] Deleage, G. and Geourjon, C. (1993) An interactive graphic program for calculating the secondary structures content of proteins from circular dichroism spectrum. Comp. Appl. Biosci. 9, 197–199.
[19] Abo, H., Matsumura, T., Kodama, T., Ohta, H., Fukui, K., Kato, H. and Kagawa, K. (1991) Peptide sequences for sucrose splitting and glucan binding within *Streptococcus sobrinus* glucosyltransferase (water-insoluble glucan synthetase). J. Bacteriol. 173, 989–996.
[20] Ueda, S. and Kuramitsu, H.K. (1988) Molecular basis for the spontaneous generation of colonization-defective mutants of *Streptococcus mutans*. Mol. Microbiol. 2, 135–140.
[21] Fujiwara, T., Terao, Y., Hoshino, T., Kawabata, S., Ooshima, T., Sobue, S., Kimura, S. and Hamada, S. (1998) Molecular analyses of glucosyltransferase genes among strains of Streptococcus mutans. FEMS Microbiol. Lett. 161, 331–336.
[22] Johnson, W.C. (1990) Protein secondary structure and circular dichroism: a practical guide. Proteins 7, 205–214.
[23] Haas, W., MacColl, R. and Banas, J.A. (1998) Circular dichroism analysis of the glucan binding domain of *Streptococcus mutans* glucan binding protein-A. Biochim. Biophys. Acta 1384, 112–120.
[24] Woody, R.W. and Dunker, A.K. (1996) Aromatic and cystine side-chain circular dichroism in proteins. In: Circular Dichroism and the Conformational Analysis of Biomolecules (Fasman, G.D., Ed.), pp. 109–157. Plenum Press, New York.
[25] 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.