A circularly permuted α-amylase-type α/β-barrel structure in glucan-synthesizing glucosyltransferases

E. Ann MacGregor[a], Hans M. Jespersen[b], Birte Svensson[c,*]

[a]*Department of Chemistry, University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada*
[b]*Department of Protein Chemistry, University of Copenhagen, DK-1353 Copenhagen, Denmark*
[c]*Department of Chemistry, Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-2500 Copenhagen Valby, Denmark*

Received 20 September 1995; revised version received 28 November 1995

Abstract A motif of amino acid residues, located at the active site and specific β-strands in α-amylases, is recognized in α-1,3- and α-1,6-glucan-synthesizing glucosyltransferases, leading to the conclusion that these enzymes contain an α/β-barrel closely related to the $(\beta/\alpha)_8$-fold of the α-amylase superfamily. The secondary structure elements of the transferase barrel, however, are circularly permuted to start with an α-helix equivalent to helix 3 in the α-amylases. Thus, the transferase counterpart of the long third β → α connection – constituting a domain in the α-amylases – is divided to precede and succeed the barrel. This architectural arrangement may be coupled to sucrose scission and glucosyl transfer. The involvement in the mechanism in glucosyltransferases of active site residues recurring in amylolytic enzymes is discussed.

Key words: α-Amylase protein superfamily; Circular permutation; Glucosyltransferase; Parallel α/β-barrel; Structure prediction

1. Introduction

The glucosyltransferases (EC 2.4.1.5) of several species of oral *Streptococci* are of interest because they synthesize extracellular glucans implicated in the development of dental plaque. These transferases are 1430–1470 amino acid residues long multidomain proteins and constitute a family of enzymes synthesizing, from sucrose, water-insoluble, mainly α-1,3-linked glucans (so-called mutans) or soluble dextrans containing principally α-1,6-linked glucose. Related glucansucrases from strains of *Leuconostoc mesenteroides* catalyse the synthesis of various glucans, containing also 1,2- and 1,4-glucosidic bonds. Understanding of the way in which the transferases catalyse the reaction:

$$\text{sucrose} + (\text{glucan})_n \rightarrow \text{fructose} + (\text{glucan})_{n+1}$$

is complicated by the fact that glucose from sucrose is added at the reducing end of a growing chain, but can also be transferred to the non-reducing end of acceptors, e.g. maltose or methyl α-D-glucoside [1–8].

*Corresponding author. Fax: (45) (3) 327-4708.

Abbreviations: CGTase, cyclodextrin glucanotransferase; E1, E2, ...E8, β-strands of the $(\beta/\alpha)_8$-barrel; G4ase, maltotetraose-producing amylase; GTF-I, glucosyltransferase that synthesizes insoluble glucan (so-called mutan); GTF-S, glucosyltransferase that synthesizes soluble dextran; H1, H2..., H8, helices of the $(\beta/\alpha)_8$-barrel; MBC/C, minimum base change per codon; TAA, Taka-amylase A.

Sequences of several glucosyltransferases from *Streptococci* are closely related [9]. Information is lacking on their three-dimensional structure, but short stretches of sequence similarities exist [10–12] between the glucosyltransferases and the α-amylases and related amylolytic enzymes that belong to a superfamily of multidomain proteins with a catalytic $(\beta/\alpha)_8$-barrel domain [13–22]. These barrel domains contain a few invariant residues involved in the enzyme mechanism and also β-strand-associated sequence motifs which provide a fingerprint of the enzyme origin and specificity [16–23]. Here we investigate further the relationship between glucosyltransferases and the α-amylase family, and conclude that the transferases also contain an α/β-barrel, but circularly permuted. Whereas the barrel domain in the α-amylases and related amylolytic enzymes begins with a β-strand, the α/β-barrel in the glucosyltransferases starts with a helix and ends with a β-strand equivalent to helix 3 and β-strand 3, respectively, in the α-amylases.

2. Materials and methods

Amino acid sequences of *Streptococcus downei* glucosyltransferases GTF-I and GTF-S, that synthesize insoluble and soluble dextrans, respectively [10,24] and *Leuconostoc mesenteroides* B512-F dextran sucrase [25], were aligned with those of Taka-amylase A, TAA – an α-amylase from *Aspergillus oryzae* [26] – *Bacillus circulans* cyclodextrin glucanotransferase (CGTase) [27], a maltotetraose-producing amylase from *Pseudomonas stutzeri* (G4ase) [16,28], *Pseudomonas amyloderamosa* isoamylase [29], *Bacillus cereus* oligo-1,6-glucosidase [30], *Streptococcus mutans* dextran glucosidase [31], and *E. coli* glycogen-branching enzyme [32]. The scoring methods of Dayhoff et al. [33], Levin et al. [34] and Johnson and Overington [35] were used to aid alignment where sequence similarities were low. Secondary structure of the $(\beta/\alpha)_8$-barrel fold was predicted as described earlier [17,18], using a combination of the Garnier procedure [36] (verified as suggested by Ellis and Milius [37]), hydrophobic cluster analysis [38], and direct comparison with three-dimensional structures of TAA, G4ase, CGTase, and oligo-1,6-glucosidase [13,16,19,39].

3. Results and discussion

3.1. Sequence alignment and $(\beta/\alpha)_8$-barrel elements in glucosyltransferases

The central segments of the sequences of two streptococcal glucosyltransferases and a *Leuconostoc* dextran sucrase are aligned, guided by invariant residues, with the known α/β-barrel domains of TAA, G4ase, CGTase, and oligo-1,6-glucosidase [13,16,19,39] and those earlier predicted by using the same procedure on isoamylase, dextran glucosidase, and glycogen branching enzyme [20]. The result of hydrophobic cluster analysis (not shown) supports the presence of barrel elements as given in Fig. 1.

SSDI 0014-5793(95)01428-4

Fig. 1. Amino acid sequence alignment of glucosyltransferases and enzymes of the α-amylase family. Amino acid residues are given in single-letter code and are numbered from the N-terminal end of the mature enzyme, except for GTFS and LEUC where numbering is for the protein with signal peptide attached. The enzymes are: GTFI = glucosyltransferase of *S. downei* synthesizing insoluble dextran; GTFS = glucosyltransferase of *S. downei* synthesizing soluble dextran; LEUC = dextran sucrase of *Leuconostoc mesenteroides* B512-F; TAKA = α-amylase of *A. oryzae*; CGT = CGTase of *B. circulans*; G4, G4ase of *P. stutzeri*; IA = isoamylase of *P. amyloderamosa*; OG = oligo-1,6-glucosidase of *B. cereus*; DG = dextran glucosidase of *S. mutans*; BE = glycogen-branching enzyme of *E. coli.* Double-underlining indicates the sequences used for calculation of MBC/C scores. E1–E8 and H1–H8 indicate approximate positions of $(\beta/\alpha)_8$-barrel elements, and letters in bold and italics show predicted locations of barrel elements in the glucosyltransferases and dextran sucrase, while letters in bold and und erlined show barrel elements identified by X-ray crystallography: OG [39], TAKA [13] CGT [19], and G4 [16]. ⋆ marks conserved amino acid residues.

This multiple alignment illustrates how the α/β-barrel of the glucosyltransferases and dextran sucrase is circularly permuted with respect to that of the α-amylase family. Since other streptococcal glucosyltransferases show high sequence homology with those included in the alignment [9] (Fig. 1) we conclude that these transferases contain a similar α/β-barrel domain. In α-amylases, the first structural element of the barrel domain, the β-strand E1, is located close to the N-terminus of the domain and followed by alternating helices and β-strands (H1, E2, H2,...H8). In contrast, in the amino acid sequence of the transferases, the first secondary structure element of the barrel domain appears to be a helix equivalent to H3 of the α-amylase family, followed by elements equivalent to E4, H4, E5,...H8, then a stretch of approximately 160 amino acid residues of unknown folding, and the barrel resumes with elements equivalent to E1, H1,...E3 of the α-amylases (Fig. 1). The locations of the principle α/β-barrel elements in the transferases are indicated in Fig. 1 as predicted by the procedure implemented for the α-amylase structural family [17,18,36,37], but although secondary structure elements are present in various $\beta \rightarrow \alpha$ connecting segments in the known crystal structures [14–16,19], the present prediction only concerns the regular barrel elements. Certain typical features of the α/β-barrel of the α-amylases are recognized in transferases, e.g. Gly and Pro flanking the second β-strand followed by a segment containing conserved residues such as Tyr^{82} of TAA, known to bind substrate in TAA [13]. Additional residues (Asp^{117} and Arg^{204} of TAA), invariant in the α-amylase enzymes superfamily [17,18,20] appear in the transferases.

Low overall sequence identity is typical of α-amylase family enzymes [17,18,20]. TAA and oligo-1,6-glucosidase, for example, for which structures have been determined by X-ray crystallography, share only 17% identical residues in the α/β-barrel domain, while here TAA and GTF-I share 15% sequence identity. It seems likely therefore that this sequence identity, although low, is significant. When overall sequence identity is low, minimum base change per codon scores can indicate sequence similarity [40,41], and hence structural relationships. MBC/C scores calculated for the glucosyltransferases and dextran sucrase relative to enzymes of known tertiary structure (TAA, G4ase, CGTase, and oligo-1,6-glucosidase) are 1.11–1.24 for the double-underlined segments in Fig. 1, i.e. those corresponding to secondary structure of the $(\beta/\alpha)_8$-barrel and to a conserved region in $\beta \rightarrow \alpha$ loop 2 [13,18,20]. Since scores under 1.29 are expected for related proteins [17,18,20,40,41], this indicates a close relationship between the transferases and the α-amylase family throughout the length of the α/β-barrel, and supports the predicted structures. These MBC/C scores are comparable to values found when comparing enzymes within the α-amylase family [18,20] and contrast with the score of 1.5 expected for unrelated proteins.

It has already been shown, by constructing and expressing circularly permuted genes for phosphoribosyl anthranilate isomerase, a single-domain $(\beta/\alpha)_8$-barrel enzyme, that protein variants with circularly permuted barrels can fold and exhibit activity [42]. Further, it has been speculated, from a study of several α/β-barrel enzymes, that circular permutations may already have occurred [43]. It is exciting, therefore, that natu-

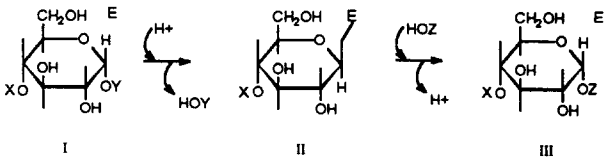

Fig. 2. General reaction catalysed by enzymes of the α-amylase family; X = H or α-linked glucosyl or maltodextrinyl group; Y or Z = 4- or 6-linked glucose or maltodextrin or isomaltodextrin, respectively; E = carboxyl group on enzyme.

rally occurring enzymes with permuted barrels have now been identified. The glucosyltransferases appear to contain domains preceding as well as succeeding the α/β-barrel. Certain members of the α-amylase family, e.g. α-1,6-endo-hydrolases and branching enzymes, similarly possess N-terminal extensions presumably folding as one or more separate domains [44]. The α/β-barrel in all family members is succeeded by one or more domains, of which one may be an identified starch granule binding domain [44].

3.2. Relation between active site and GTF-reactions

In α/β-barrel enzymes, the active site is constructed from loops joining C-termini of β-strands to N-termini of adjacent helices [45]. The active site in α-amylases is a wide cleft accommodating several glucose residues of the substrate and a large part of one side of the cleft is formed by a long $\beta \rightarrow \alpha$ segment [46], referred to as domain B, linking E3 to H3 [13–17]. Dextran sucrase and glucosyltransferases, with the permuted barrel, contain two segments of polypeptide chain with the potential to fulfill the role of domain B, i.e. the sequences preceding H3, and following E3, respectively (see Fig. 1).

Glucosyltransferases catalyse two different, but concerted reactions – scission of sucrose and subsequent glucosyl transfer. The transfer is complex, for it may take place to the reducing end of a growing dextran chain or to the nonreducing end of a so-called acceptor [1–8]. Experimental evidence supports the idea that sites for sucrose cleavage and glucosyl transfer are distinct [2,47]. Thus in an *L. mesenteroides* dextran sucrase the sequence immediately preceding the predicted H3 has been implicated in sucrose binding and cleavage [12], while in glucosyltransferases the C-terminal region is recognised as a dextran-binding domain and a segment of ca. 300 residues following E3 is necessary for full transferase activity [10,48–50]. It is, therefore, likely that the transferases utilise both regions equivalent to domain B in α-amylase, the N-terminal one being involved in sucrose scission and the C-terminal in glucosyl transfer to dextran. In fact sucrase activity is less affected by C-terminal deletion than glucosyltransferase, i.e. dextran synthesising, activity [48,50].

3.3. Acceptor reaction

Enzymes of the α-amylase family catalyse reactions of the type shown in Fig. 2 and share the ability to catalyse the scission and formation of bonds of α-configuration from C1 of a glucose residue to oxygen linked to C4 or C6 of another glucose residue. Acceptor reactions of the glucosyltransferases are similar i.e. an α-glycosidic bond is formed between glucose from sucrose and a carbon of the nonreducing end of an acceptor such as methyl α-D-glucoside. There is evidence that for α-amylases a covalent intermediate (II) may form between the glycon part of the substrate and a carboxyl group on the enzyme [51]. For glucosyl transferases strong evidence supports the occurrence of a glucosyl-glucosyltransferase intermediate like II [1,4,11,52,53]. The esterified carboxyl group of II has been identified as that of Asp^{415} of GTF-I or Asp^{437} of GTF-S [11]. By analogy, Asp^{551} of dextran sucrase and residues equivalent to Asp^{206} of TAA may be involved in covalent intermediates in different, but related enzymes. Here Y of Fig. 2 would represent the fructosyl moiety of sucrose, while Z could be an acceptor such as methyl-α-glucoside. Thus the mechanism may resemble that of the α-amylase family.

3.4. Comparison with residues at the catalytic site in the α-amylase family

The enzymes of the α-amylase family have three carboxyl groups, important for catalysis, at or near the C-termini of β-strands 4, 5 and 7, i.e. of the residues equivalent to Asp^{206}, Glu^{230} and Asp^{297} of TAA [3,11,16,19,22]. Equivalent residues occur invariantly in glucosyltransferases – Asp^{415}, Glu^{453} and Asp^{526} in GTF-I; Asp^{437}, Glu^{475} and Asp^{547} in GTF-S; and Asp^{551}, Glu^{589}, and Asp^{662} in dextran sucrase (Fig. 1). The first of these, implicated in formation of a glucosyl-enzyme intermediate, has been shown to be necessary for transferase activity [54] and we suggest the others are also important. Further, most α-amylase family enzymes contain two histidine residues (equivalent to His^{122} and His^{296} of TAA) critical for activity, that are found by site-directed mutagenesis to be important in transition state stabilisation [55,56]. The first is replaced in the transferases by a glutamine (Gln^{899} of GTF-I, Gln^{920} of GTF-S, and Gln^{1029} of dextran sucrase), while the second is conserved as His^{525} of GTF-I, His^{546} of GTF-S, and His^{661} of dextran sucrase.

In enzymes of the α-amylase family, amino acid residues in $\beta \rightarrow \alpha$ loops that join E4 to H4 and E5 to H5 are implicated in enzyme specificity and interact with the substrate aglycon (Y or Z of Fig. 2) [13,20,46,57]. Residues equivalent to 209–210 and 231–233 of TAA (see Fig. 1) are particularly important. Thus, in the glucosyltransferases the equivalent Asp^{418}–Asn^{419} and Ala^{454}–Trp^{455}–Ser^{456} of GTF-I, 440–441, 476–478 of GTF-S, and 554–555, 590–592 of dextran sucrase, are expected to be significant for binding of Z, i.e. acceptor binding and glucosyl transfer to the nonreducing end of an acceptor.

4. Conclusion

Although a complete explanation of the mechanism of action of the glucosyltransferases cannot yet be provided, the results and sequence alignment presented should help to pinpoint functionally important amino acid residues that are located at the active site. Their alteration by site-directed mutagenesis should provide information about the role of particular residues in acceptor reactions and/or dextran synthesis that may be useful in future protein engineering of glucosyltransferases. Furthermore, elucidation of the function of domains flanking the barrel may improve our insight into the mechanism of GTFases.

Acknowledgements: Prof. Roy R.B. Russell is thanked for assistance in acquiring sequences.

References

[1] Robyt, J.F. and Martin, P.J. (1983) Carbohydr. Res. 113, 301–315.
[2] Robyt, J.F. (1994) in: Enzymes for Carbohydrate Engineering, Proceedings of the '94 Agricultural Biotechnology Symposium on Enzymes for Carbohydrate Engineering, August 26 – The Research Center for New Bio-Materials in Agriculture, Suwon, Korea, pp. 1–26.
[3] Ditson, S.L. and Mayer, R.M. (1984) Carbohydr. Res. 126, 170–175.
[4] Luzio, G.A., Parnaik, V.K. and Mayer, R.M. (1983) Carbohydr. Res. 121, 269–278.
[5] Nisizawa, T., Takeuchi, K. and Imai, S. (1986) Carbohydr. Res. 147, 135–144.
[6] Binder, T.P. and Robyt, J.F. (1986) Carbohydr. Res. 154, 229–238.
[7] Fu, D. and Robyt, J.F. (1991) Carbohydr. Res. 217, 201–211.
[8] Bhattacharjee, M.K. and Mayer, R.M. (1993) Carbohydr. Res. 242, 191–201.
[8] Shiroza, T., Ueda, S. and Kuramitsu, H.K. (1987) J. Bacteriol. 169, 4263–4270.
[9] Giffard, P.M., Allen, D.M., Milward, C.P., Simpson, C.L. and Jacques, N.A. (1993) J. Gen. Microbiol. 139, 1511–1522.
[10] Ferretti, J.J., Gilpin, M.L. and Russell, R.R.B. (1987) J. Bacteriol. 169, 4271–4278.
[11] Mooser, G., Hefta, S.A., Paxton, R.J., Shively, J.E. and Lee, T.D. (1991) J. Biol. Chem. 266, 8916–8922.
[12] Funane, K., Shiraiwa, M., Hashimoto, K., Ichishima, E. and Kobayashi, M. (1993) Biochemistry 32, 13696–13702.
[13] Matsuura, Y., Kusunoki, M., Harada, W. and Kakudo, M. (1984) J. Biochem. (Tokyo) 95, 697–702.
[14] Qian, M., Haser, R. and Payan, F. (1993) J. Mol. Biol. 231, 785–799.
[15] Kadziola, A., Abe, J., Svensson, B. and Haser, R. (1994) J. Mol. Biol. 239, 104–121.
[16] Matsuura, Y. (1994) in: Enzyme Chemistry and Molecular Biology of Amylases and Related Enzymes (Yamamoto, T. ed.) pp. 137–145, CRC Press, Ann Arbor/London.
[17] MacGregor, E.A. (1988) J. Protein Chem. 7, 399–415.
[18] MacGregor, E.A. and Svensson, B. (1989) Biochem. J. 259, 145–152.
[19] Klein, C. and Schulz, G.E. (1991) J. Mol. Biol. 217, 737–750.
[20] Jespersen, H.M., MacGregor, E.A., Henrissat, B., Sierks, M.R. and Svensson, B. (1993) J. Protein Chem. 12, 791–805.
[21] Svensson, B. (1994) Plant Mol. Biol. 25, 141–157.
[22] Kizaki, H., Hata, Y., Watanabe, K., Katsube, Y. and Suzuki, Y. (1993) J. Biochem. (Tokyo) 113, 646–649.
[23] Janecek, S., MacGregor, E.A. and Svensson, B. (1995) Biochem. J. 305, 685–686.
[24] Gilmore, K.S., Russell, R.R.B. and Ferretti, J.J. (1990) Infect. Immun. 58, 2452–2458.
[25] Wilke-Douglas, M., Perchorowicz, J.T., Houck, C.M. and Thomas, B.R. (1989) PCT Int. Appl. WO 89 12,386; Chem. Abstr. (1990) 113, P72531y.
[26] Toda, H., Kondo, K. and Narita, N. (1982) Proc. Jpn. Acad. Ser B 58, 208–212.
[27] Nitschke, L., Heeger, K., Bender, H. and Schulz, G.E. (1990) Appl. Microbiol. Biotechnol. 33, 542–546.
[28] Fujita, M., Torigoe, K., Nakada, T., Tsusaki, K., Kubota, M., Saka, S. and Tsujisasa, Y. (1989) J. Bacteriol. 171, 1333–1339.
[29] Amemura, A., Chakraborty, R., Fujita, M., Noumi, T. and Futai, M. (1988) J. Biol. Chem. 263, 9271–9275.
[30] Watanabe, K., Kitamura, K., Iha, H. and Suzuki, Y. (1990) Eur. J. Biochem. 192, 609–620.
[31] Russell, R.R.B. and Ferretti, J.J. (1990) J. Gen. Microbiol. 136, 803–810.
[32] Baecker, P.A., Greenberg, E. and Preiss, J. (1986) J. Biol. Chem. 261, 8738–8743.
[33] Dayhoff, M.O., Baker, W.C. and Hunt, L.T. (1983) Methods Enzymol. 91, 524–545.
[34] Levin, J.M., Robson, B. and Garnier, J. (1986) FEBS Lett. 205, 303–308.
[35] Johnson, M.S. and Overington, J.P. (1993) J. Mol. Biol. 233, 716–738.
[36] Garnier, J., Osguthorpe, D.J. and Robson, B. (1978) J. Mol. Biol. 120, 97–120.
[37] Ellis, L.B.M. and Milius, R.P. (1994) CABIOS 10, 341–348.
[38] Gaboriaud, C., Bissery, V., Benchetrit, T. and Mornon, J.P. (1987) FEBS Lett. 224, 149–155.
[39] Watanabe, K., Masuda, T., Okashi, H., Mihara, H. and Suzuki, Y. (1994) Eur. J. Biochem. 226, 277–283.
[40] Rossmann, M.G., Moras, D. and Olsen, K.W. (1974) Nature 250, 194–199.
[41] Argos, P., Mahoney, W.C., Hermodson, M.A. and Hanei, M. (1981) J. Biol. Chem. 256, 4357–4361.
[42] Luger, K., Hommel, U., Herold, M., Hofsteenge, J. and Kirschner, K. (1989) Science 243, 206–210.
[43] Sergeev, Y. and Lee, B. (1994). J. Mol. Biol. 244, 168–182.
[44] Jespersen, H.M., MacGregor, E.A., Sierks, M.R. and Svensson, B. (1991) Biochem. J. 280, 51–55.
[45] Farber, G.K. (1993) Curr. Opin. Struct. Biol. 3, 409–412.
[46] Qian, M., Haser, R., Buisson, G., Duée, E. and Payan, F. (1994) Biochemistry 33, 6284–6294.
[47] Yamashita, Y., Hanada, N., Itoh-Andoh, M. and Takehara, T. (1989) FEBS Lett. 243, 343–346.
[48] Kato, C. and Kuramitsu, H.K. (1990) FEMS Microbiol. Lett. 72, 299–302.
[49] Wong, C., Hefta, S.A., Paxton, R.J., Shively, J.E. and Mooser, G. (1990) Infect. Immun. 58, 2165–2170.
[50] Abo, H., Matsumura, T., Kodama, T., Ohta, H., Fukui, K., Kato, K. and Kagawa, H. (1991) J. Bacteriol. 173, 989–996.
[51] Tao, B.Y., Reilly, P.J. and Robyt, J.F. (1989) Biochim. Biophys. Acta 995, 214–220.
[52] Parnaik, V.K., Luzio, G.A., Grahame, D.A., Ditson, S.L. and Mayer, R.M. (1983) Carbohydr. Res. 121, 257–268.
[53] Mooser, G. and Iwaoka, K.R. (1989) Biochemistry 28, 443–449.
[54] Kato, C., Nakano, Y., Lis, M. and Kuramitsu, H.K. (1992) Biochem. Biophys. Res. Commun. 189, 1184–1188.
[55] Søgaard, M., Kadziola, A., Haser, R. and Svensson, B. (1993) J. Biol. Chem. 268, 22480–22484.
[56] Nakamura, A., Haga, K. and Yamane, K. (1993) Biochemistry 32, 6624–6631.
[57] Strokopytov, B., Penninga, D., Rozeboom, H.J., Kalk, K.H., Dijkhuizen, L. and Dijkstra, B.W. (1995) Biochemistry 34, 2234–2240.