doi:10.1016/j.jmb.2008.03.016 J. Mol. Biol. (2008) 378, 913–922

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Substrate Recognition Mechanism of α-1,6-Glucosidic Linkage Hydrolyzing Enzyme, Dextran Glucosidase from *Streptococcus mutans*

Hironori Hondoh[1]*, Wataru Saburi[1], Haruhide Mori[1], Masayuki Okuyama[1], Toshitaka Nakada[2], Yoshiki Matsuura[3] and Atsuo Kimura[1]

[1]*Research Faculty of Agriculture, Hokkaido University, Sapporo 060-8589, Japan*

[2]*Faculty of Science and Engineering, Ritsumeikan University, Kusatsu, Shiga 525-8577, Japan*

[3]*Institute for Protein Research Osaka University, Suita, Osaka 565-0871, Japan*

Received 24 November 2007; received in revised form 7 March 2008; accepted 10 March 2008
Available online 18 March 2008

We have determined the crystal structure of *Streptococcus mutans* dextran glucosidase, which hydrolyzes the α-1,6-glucosidic linkage of isomaltooligosaccharides from their non-reducing ends to produce α-glucose. By using the mutant of catalytic acid Glu236→Gln, its complex structure with the isomaltotriose, a natural substrate of this enzyme, has been determined. The enzyme has 536 amino acid residues and a molecular mass of 62,001 Da. The native and the complex structures were determined by the molecular replacement method and refined to 2.2 Å resolution, resulting in a final *R*-factor of 18.3% for significant reflections in the native structure and 18.4% in the complex structure. The enzyme is composed of three domains, A, B and C, and has a $(\beta/\alpha)_8$-barrel in domain A, which is common to the α-amylase family enzymes. Three catalytic residues are located at the bottom of the active site pocket and the bound isomaltotriose occupies subsites −1 to +2. The environment of the glucose residue at subsite −1 is similar to the environment of this residue in the α-amylase family. Hydrogen bonds between Asp60 and Arg398 and O4 atom of the glucose unit at subsite −1 accomplish recognition of the non-reducing end of the bound substrate. The side-chain atoms of Glu371 and Lys275 form hydrogen bonds with the O2 and O3 atoms of the glucose residue at subsite +1. The positions of atoms that compose the scissile α-1,6-glucosidic linkage (C1, O6 and C6 atoms) are identical with the positions of the atoms in the scissile α-1,4 linkage (C1, O4 and C4 atoms) of maltopentaose in the α-amylase structure from *Bacillus subtilis*. The comparison with the α-amylase suggests that Val195 of the dextran glucosidase and the corresponding residues of α-1,6-hydrolyzing enzymes participate in the determination of the substrate specificity of these enzymes.

Keywords: dextran glucosidase; *Streptococcus mutans*; α-amylase family; crystal structure; substrate recognition

Edited by G. Schultz

Introduction

Starch hydrolases and related enzymes are classified into the α-amylase family, which includes α-amylases, cyclodextrin glucanotransferases, oligo-1,6-glucosidases (O16G), pullulanases, isoamylases, branching enzymes and neopullulanases.[1,2] The enzymes of this family possess common three-dimensional structural features. The degree of amino acid sequence similarity between members of the α-amylase family is not very high. However, these enzymes have four short conserved regions I–IV in common, which include the catalytic residues and other amino acids essential for stabilization of the transition state.[2,3] All members of this family of

*Corresponding author. E-mail address: hondoh@abs.agr.hokudai.ac.jp.

Abbreviations used: DGase, dextran glucosidase; IG3, isomaltotriose; O16G, oligo-1,6-glucosidase; AS, amylosucrase; SP, sucrose phosphorylase.

enzymes have a catalytic domain (domain A), containing a $(\beta/\alpha)_8$-barrel, with a small domain (domain B) inserted between the third β-strand and the third α-helix of domain A. The specificities of these enzymes for the cleavage of glucosidic linkages can be classified into one of two types: α-1,4-glucosidic linkage-specific or α-1,6-glucosidic linkage-specific.

One of the best-known α-1,4 linkage hydrolyzing enzymes is α-amylase. The crystal structure of Taka-amylase was determined in 1984,[4] and numerous structural studies of α-amylases complexed with products, substrate analogues and natural substrates have been carried out to clarify the structural basis for substrate specificity.[5–15] The structure of the natural substrate complex revealed the details of the interaction between α-amylase and its substrate.[13] The conformation of the glucosidic linkage of the bound substrate (or analogue) at the cleavage site of these enzymes is kinked, allowing the substrate to fit into the active site cleft. The significance of this kinked conformation of the scissile bond has been described for the structure of amylosucrase (AS).[16] The kinked conformation allows the lone pair of the glucosidic oxygen (O4 atom) of the scissile linkage to point toward the general acid catalytic residue. This configuration allows the minimization of the proton-transfer energy barrier during catalysis.

In comparison to the extensive research in the field of α-amylases, there are only a few structural studies that describe enzymes hydrolyzing α-1,6-linkages. The crystal structure of O16G from *Bacillus cereus* was determined at 3.0 Å resolution,[17] and refined to 2.0 Å resolution.[18] The structure of isoamylase, which hydrolyzes the α-1,6-glucosidic linkage at the branching point of amylopectin and glycogen, has been determined.[19] To date, however, only substrate-free structures are available, and the mechanism by which these enzymes recognize their substrates remains obscure. Recently, an α-1,6 bond cleavage model of a neopullulanase, which hydrolyzes both α-1,4 and α-1,6 linkages, has been advanced by using structures of the enzyme in complex with natural substrates.[20] Because neopullulanase cleaves primarily α-1,4 linked substrates, this model cannot be extended readily to understand substrate recognition by enzymes hydrolyzing α-1,6 linkages. The crystal structures of pullulanase, which is a well-known starch-debranching enzyme, in complex with several oligosaccharides have been reported very recently.[21] Although these studies clearly demonstrate the parallel binding of maltooligosaccharides, which are used as the model for binding of branched substrates at the active site, the conformation of the α-1,6-glucosidic linkage in the cleavage site is unclear.

Dextran glucosidase (DGase) from *Streptococcus mutans*, a member of the α-amylase family of enzymes, hydrolyzes α-1,6-glucosidic linkages at the non-reducing end of panose, isomaltooligosaccharides and dextran to produce α-glucose.[22,23] The enzyme shows the highest k_{cat}/K_m value toward panose and the order of substrate preference is panose>isomaltotriose>isomaltotetraose>isomaltose. The kinetic parameters for a series of isomaltooligosaccharides show that the catalytic efficiency decreases with increasing chain length.[24] The native DGase contains 536 amino acid residues (molecular mass 62,001 Da) and its amino acid sequence has a high particularly similarity to O16Gs in the α-amylase family. DGase is capable of degrading relatively long isomaltooligosaccharides and even dextran, while O16Gs act only on short-chain substrates despite sharing a high level of sequence identity.[25–28] Our recent studies have shown that Trp238 and short β→α loop 4 of DGase are important for the high level of activity of this enzyme toward long-chain substrates.[24] Another notable feature of the DGase is a relatively high level of transglucosylation ability, which may be useful for oligosaccharide synthesis.

Here, we report the crystal structures of DGase from *S. mutans* in uncomplexed form and a mutant (E236Q) in which catalysis is compromised in complex with a natural substrate (isomaltotriose: IG3) at 2.2 Å resolution. These structures reveal the residues that mediate specific recognition of α-1,6-glucosidic linkages and dictate the substrate specificity of α-amylase family enzymes.

Results

Overall structure

The native and substrate complex structures of DGase have been determined by the molecular replacement method and refined to a crystallographic *R*-factor (R_{free}) of 18.3% (22.2%) and 18.4% (22.3%) at 2.2 Å resolution, respectively. DGase consists of domains A, B and C, as shown in Fig. 1. Domain A contains three catalytic residues Asp194, Glu236 and Asp313, and a $(\beta/\alpha)_8$-barrel structure, which are conserved among α-amylase family enzymes. Domain B contains two helices and three strands, and forms part of the wall of the active site pocket. Surrounded by the two loops of domain B (residue numbers 130–142, and 156–168) and β→α loops 4, 5, 6 and 7 of domain A (residue numbers 195–212, 237–240, 269–283, and 308–315), the active site cleft of DGase forms a pocket similar to those seen in the structures of O16G,[18] AS[29] and sucrose phosphorylase (SP).[30] The catalytic residues Asp194, Glu236 and Asp313 are located at the bottom of this pocket. The extra α-helixes protruding near the C-terminal end of the $(\beta/\alpha)_8$-barrel provide the amino acid residues Glu371 and Arg398 that are important for substrate recognition. Domain C consists of an antiparallel, β-sheet motif with one helix. Bound Tris and IG3 molecules are identified as electron density at the bottom of the active site pocket of the native structure and the substrate complex, respectively. A molecule of glycerol, which was used as a cryo-protectant, is found in the active site pocket of the native enzyme. The histidine tag at the C terminus could not be refined, suggesting a non-definite conformation. The r.m.s deviation between corresponding main-chain

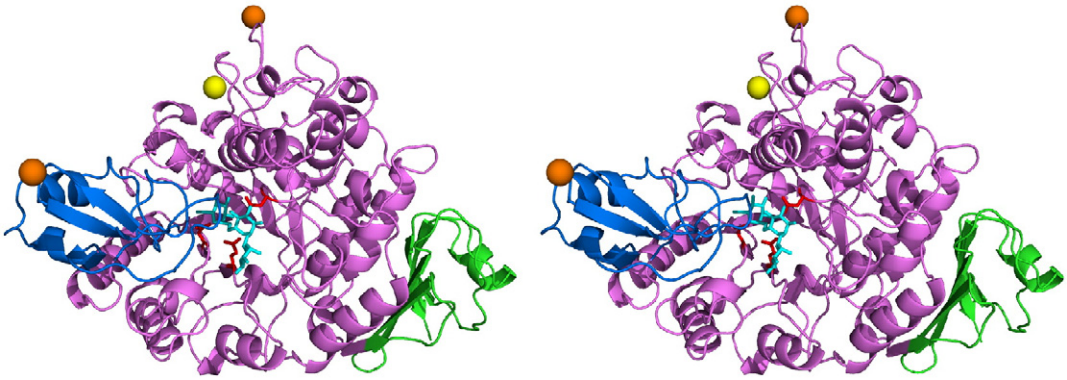

Fig. 1. Fold of DGase from *Streptococcus mutans*. A ribbon representation (stereo view) of DGase. Domain A is shown in magenta, domain B is shown in blue and domain C is shown in green. Bound IG3 molecules at the active site pocket and three of the catalytic residues are shown using stick models in cyan and red, respectively. A calcium ion in the $(\beta/\alpha)_8$ barrel is shown as a yellow sphere, and the other calcium ions located between the symmetry-related enzyme molecules are shown as orange spheres. The figure was produced with Pymol [http://www.pymol.org].

atoms of the native and complex structures is 0.22 Å. The conformation of the active site residues is almost identical between the native structure and that of the complex, but obvious differences are found in the side-chain of Trp238. The contact with the bound IG3 molecule results in a movement of the side-chain atoms of Trp238, which is therefore shifted approximately 1 Å away from the cleavage site.

In the crystallographic asymmetric unit, three calcium ions, which are bound to enzymes, are found (Fig. 1). Two of them are located between symmetry-related enzyme molecules. These two calcium ions would be released from these sites of DGase in the solution because each ion is fixed by two symmetry-related enzyme molecules. The other calcium ion is bound at the periphery of the $(\beta/\alpha)_8$-barrel of an enzyme. The calcium ion is bound to Asp21 OD1, Asn23 OD1, Asp25 OD1 and OD2, Ile27 O, Asp29 OD2, and two water molecules. The distances between the calcium ion and these atoms in the native DGase are 2.30 Å, 2.27 Å, 2.45 Å, 2.84 Å, 2.55 Å, 2.52 Å, 2.50 Å, and 2.35 Å, respectively. Therefore, this calcium could be held tightly by this site in solution. The conformation of this site in the substrate complex is similar to that of native enzyme, but the electron density of the bound water molecule was poor. Thus, the water molecule is not included in the substrate-complex model. O16G shows the identical conformation at this site, but a bound calcium ion is not included in the model of O16G. Since the crystallization solution of O16G contains EDTA,[18] the calcium ion in O16G may have been chelated and therefore unable to bind. All the calcium-binding sites of DGase are far from the active site pocket and not identical with the conserved calcium ion-binding sites among α-amylases.[31]

Structure in the active site pocket of the DGase/IG3 complex

The inactive E236Q mutant was used to crystallize and obtain the enzyme–substrate complex. F_{obs}–F_{calc} omit map of the active site pocket of the E236Q–IG3 complex is shown in Fig. 2. The bound IG3 molecule was identified as electron density in the region of the active site pockets of the catalytic acid mutant. The glucose residues in the bound IG3 occupy three subsites, −1 through +2. The unbroken electron density reveals that IG3 was not hydrolyzed, and the α-1,6-glucosidic linkage at the active site crosses over the active site of DGase. This is the first model of an α-1,6-glucosidic linkage-hydrolyzing member of the α-amylase family in complex with a natural substrate. The electron density of the reducing end glucose residue of IG3 is not clearly defined, most likely because of the flexibility of this terminal residue. The model of amino acid residues surrounding IG3 is shown in Fig. 2, and a schematic drawing of E236Q–IG3 interactions is shown in Fig. 3. The dihedral angles of bound IG3 are summarized in

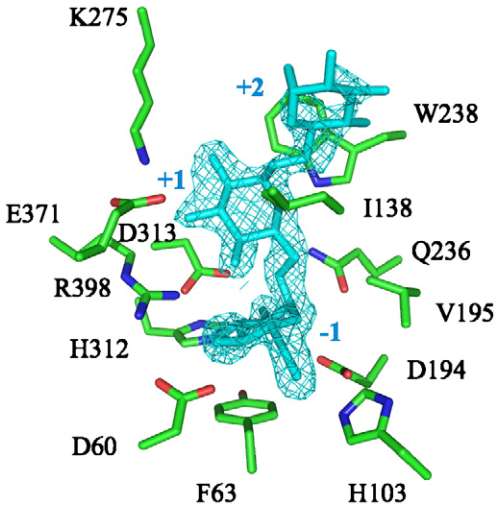

Fig. 2. The F_{obs}–F_{calc} omit map of the bound IG3. The IG3 molecule was found in the active site pocket of the catalytic acid mutant E236Q enzyme. IG3 is shown in cyan. The numbers in cyan denote subsite numbering. The contour level of the map is 3.0σ. The figure was generated using CNS and Pymol [http://www.pymol.org].

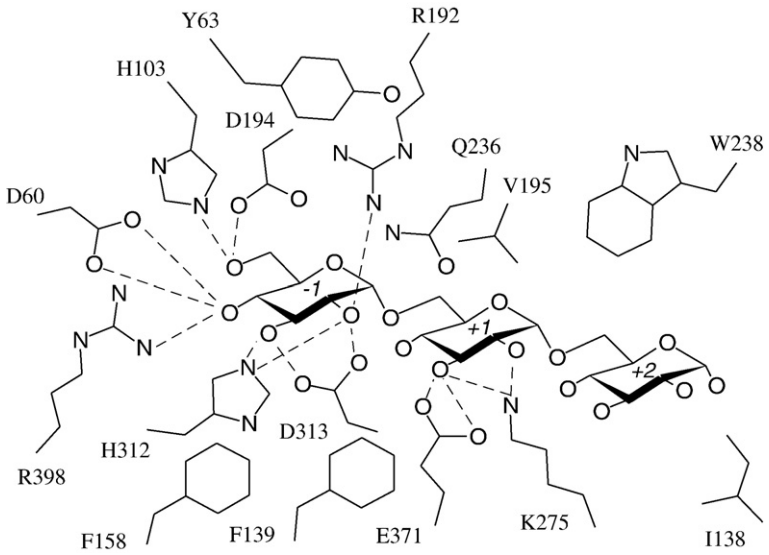

Fig. 3. Schematic drawing of the IG3 binding. Hydrogen bonds are indicated by dotted lines. The numbers in the glucose units denote subsite numbering.

Table 1. The dihedral angles in this work are defined as $\phi = O5_n\text{-}C1_n\text{-}O6_{n-1}\text{-}C6_{n-1}$, $\psi = C1_n\text{-}O6_{n-1}\text{-}C6_{n-1}\text{-}C5_{n-1}$ and $\omega = O6_{n-1}\text{-}C6_{n-1}\text{-}C5_{n-1}\text{-}C4_{n-1}$. In subsite −1, the glucose residue stacks against Tyr63, and is recognized by an extensive hydrogen bonding network involving residues Arg192, His312 and His103. The catalytically important residue Asp313 is included in this hydrogen bonding network. The side-chain atoms of residue Asp194, which acts as the catalytic nucleophile, also forms a hydrogen bond with the O6 atom of the glucose residue bound in subsite −1. The distance between C1 of the glucose unit and Asp192OD1 is close to 3.1 Å. The side-chain atoms of the mutated catalytic acid residue Gln236 are in close contact with the O4 atom of the scissile bond, at a distance of 3.2 Å. This configuration of the glucose residue bound to subsite −1 is conserved within the entire α-amylase family enzymes.[19] In addition, three hydrogen bonds were formed in the side chains of Asp60 and Arg398 and the glucose in subsite −1, in which the non-reducing end of the substrate is identified. The side chains of Asp60 and Arg398 make a salt-bridge with each other and engage in hydrogen bonds with the O4 atom of the glucose unit bound in subsite −1. At subsite +1, the side chains of Lys275 and Glu371 form hydrogen bonds with the O2 and O3 atoms of the glucose residue. Moreover, the side chains of Phe139, Phe158 and Trp238 make hydrophobic contacts with the same glucose unit. Notably, the reducing end glucose has no hydrogen bond with the enzyme, and only a few hydrophobic interactions between the side chains of Trp238 and Ile138.

Table 1. Torsion angles around the glucosidic bonds of IG3 bound to E236Q DGase

Glucosidic linkage	ϕ (°) ($O5_n\text{-}C1_n\text{-}O6_{n-1}\text{-}C6_{n-1}$)	ψ (°) ($C1_n\text{-}O6_{n-1}\text{-}C6_{n-1}\text{-}C5_{n-1}$)	ω (°) ($O6_{n-1}\text{-}C6_{n-1}\text{-}C5_{n-1}\text{-}C4_{n-1}$)
−1/+1	42	158	36
+1/+2	59	141	158

The binding mode of Tris to the active site of the native *S. mutans* DGase is different from that in a psychrophilic α-amylase from *Alteromonas haloplanctis*,[32] and in a chimeric α-amylase from *Bacillus licheniformis* and *Bacillus amyloliquefaciens*.[33] The amino group of Tris has only one hydrogen bond with catalytic residue Asp194, while many hydrogen bonds were found in the other structures mentioned above. All of the hydroxy groups of the Tris molecule are involved in hydrogen bonding interactions with surrounding residues. The O2 and O3 atoms of Tris are situated at the same place as the O4 and O6 atoms of the non-reducing end glucose residue of IG3 bound to E236Q mutant. Also, a glycerol molecule is found at subsite +1 in the native structure and the O3 atom bonds with Lys275 and Glu371 in a way similar to that seen for the E236Q–IG3 complex.

One striking observation is that a row of water molecules was found at the bottom of the active site pocket. This line of water molecules connects the active site pocket and the surface of the enzyme. Identical paths were found in both of the structures of the native enzyme and the E236Q–IG3 complex. A close-up view of this path for the E236Q–IG3 complex is shown in Fig. 4. This path is located near Asp60 and Arg398, which form a salt-bridge and bind the non-reducing terminus of bound substrate. These three water molecules are designated from the inner entrance of this path toward the outer entrance as Wat1 through Wat3. The distances between these water molecules are 3.11 Å and 3.16 Å, respectively. Both Wat1 and Wat2 are bound to Asp60 OD1. Wat1 also forms a hydrogen bond with Arg398 NH2 and Arg402 NH2, and contacts the O3 and O4 atoms of the glucose residue in subsite −1. Wat2 is bound by Asn61 ND2, Asp369 OD1, Arg398 NH1 and Arg402 NH1. Wat3 forms hydrogen bonds with the main-chain nitrogen atom of Ile370 and the main-chain carbonyl oxygen atoms of Asp59 and Asp368. A

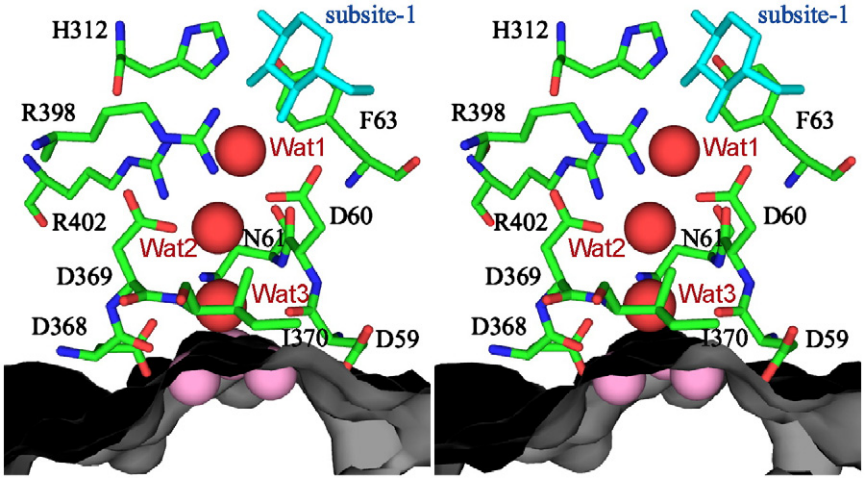

Fig. 4. The row of water molecules at the bottom of the active site pocket. Water molecules in the path are shown as red spheres, and those not in the path are shown in pink. The glucose residue of IG3 bound to subsite −1 is shown in cyan. The molecular surface is drawn in gray. The figure was drawn with Pymol [http://www.pymol.org].

possible role for this orientation of water molecules will be discussed later.

Discussion

Substrate specificity of α-1,6 linkage-hydrolyzing enzymes

The IG3 substrate was bound and recognized by the enzyme through an extensive network of hydrogen bonds, as mentioned above. The arrangement of amino acid residues at subsite −1 is common within the α-amylase family of enzymes. Therefore, the differences in specificity for α-1,4 or α-1,6 linkages among α-amylase family members are thought to have originated from the structure at subsite +1 or subsite +2. To compare the differences between recognition of substrates having α-1,4 or α-1,6 linkages, the models at the active site and bound substrate in DGase, and in α-amylase from *Bacillus subtilis*[13] (PDB ID 1bag) are shown in Fig. 5. Four hydrogen bonds from Lys275 and Glu371 to the O2 and O3 atoms of the glucose residue bound to subsite +1 were observed in the substrate–enzyme complex structure of DGase. In the structures of many α-amylases, there is a hydrogen bond to the O2 atom of this glucose residue from a His residue in subsite +1.[5,10,13] However, the orientations of these bonds are not the same as those shown in Fig. 5. The side chains of Lys275 and Glu371 of DGase make hydrogen bonds with the glucose residue from the left in the figure, while His180 of α-amylase from *B. subtilis* bind the residue from the right-hand side.[13] Consequently, the location of subsite +1 for these enzymes is different, and the orientations of glucosidic linkages between glucose units in subsites +1 and +2 are opposite. The hydrogen bonds at subsite +1 in DGase appear to stretch the scissile bond across the cleavage site. With these hydrogen bonds, the conformation of the α-1,6-glucosidic linkage seen here for *S. mutans* DGase bound across the cleavage site is similar to that seen for α-1,4 linkages bound to α-amylases. The torsion angle ϕ in the scissile bond of bound IG3 in DGase is 43°, which is not much different from the equivalent angles O5-O1-O4-C4 of maltoheptaose bound to AS (33°),[34] and of maltopentaose bound to α-amylase (26°),[13] and of maltotetraose bound to neopullulanase (34° and 43°).[20] The positions of the C1, O6 and C6 atoms of the scissile α-1,6-linkage seen here are identical with those of the C1, O4 and C4 atoms of α-1,4 linkage bound to α-amylase from *B. subtilis*. The consequence is that the lone pair of the O6 atom in the glucosidic linkage faces the acidic hydrogen atom of Glu236 of wild-type DGase and thereby reduces the proton-transfer energy barrier during catalysis in a manner similar to that seen for AS.[16]

The role of Val195 may be discussed by comparison with the structure of the substrate–enzyme complex of α-amylase from *B. subtilis*.[13] Bulky hydrophobic residues are conserved in O16G, isoamylase and pullulanase, which all hydrolyze α-1,6-glucosidic linkages (Table 2). In contrast, there is a relatively small residue at this position in most α-amylases, which specifically hydrolyze α-1,4 linkages. The bound α-1,6 linked substrate has a hydrophobic contact with the Val residue. This interaction would provide an attractive force that is favorable for the α-1,6 linked substrate binding. Therefore, Val195 or bulky hydrophobic residues found in other enzymes acting on α-1,6 linkages would make a significant contact for substrate specificity with the scissile α-1,6 linkage of the bound substrate. Besides, these residues may cause steric hindrance to the binding of α-1,4 linked substrates. If the α-1,4 linked substrate observed in the structure of the α-amylase from *B. subtilis*[13] is superimposed over the active site of DGase while maintaining the

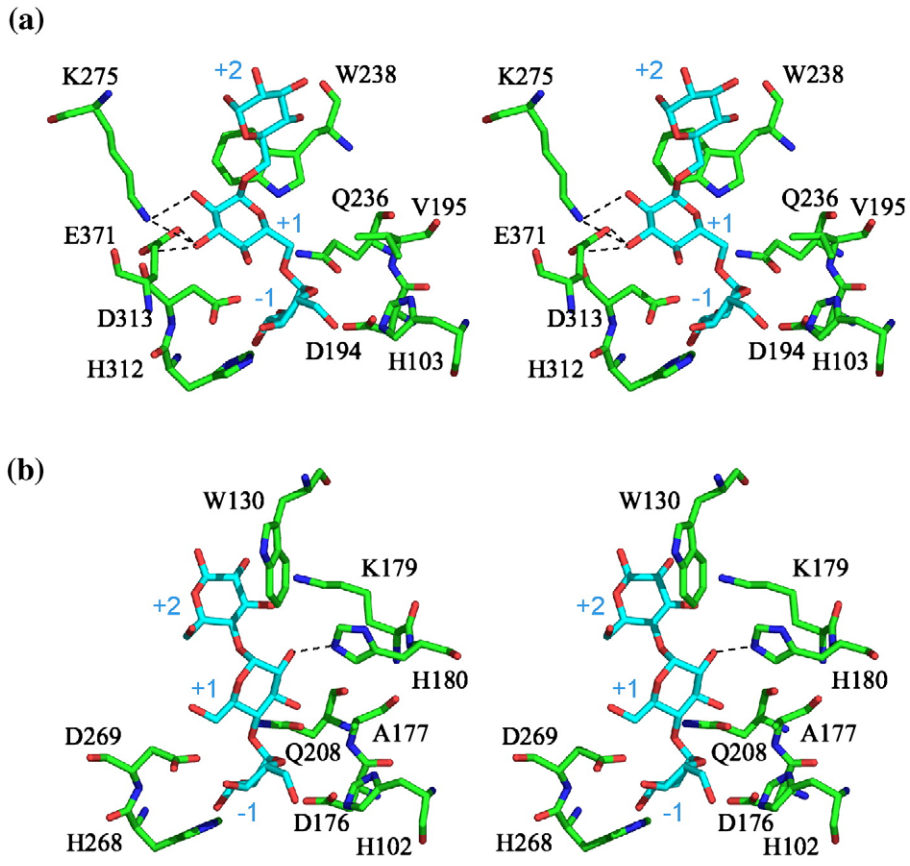

Fig. 5. Comparison of substrate binding by DGase (a) to substrate binding by α-amylase from *Bacillus subtilis* (PDB ID 1bag) (b). Bound substrates are drawn in cyan. The subsite numbers are denoted in cyan. Hydrogen bonds to the glucose unit in subsite +1 are indicated by dotted lines.

bound conformation, Val195 appears to clash with the O3 atom of the glucose moiety at subsite +1 (the distance between those atoms being about 1 Å).

Recognition of the non-reducing end of substrate

In the substrate–enzyme complex, the O4 atom of the non-reducing end glucose residue of bound IG3 forms hydrogen bonds to the side-chain atoms of Asp60 and Arg398. These two residues also form a salt-bridge with each other. An equivalent salt-bridge is found in the structures of O16G, AS and SP, which all have pocket-shaped active sites like DGase. Moreover, identical hydrogen bonds between enzymes and their substrates have been observed in the substrate–enzyme complex structures of AS[16] and SP.[35] In the *exo*-type enzymes, such as DGase, O16G, AS and SP, recognition of the non-reducing end seems to be accomplished predominantly by these interactions. Although hydrogen atoms are not visible in the present X-ray analysis, the hydrogen bonding network including hydrogen atoms can be predicted as follows. The OD2 atom of Asp60 will attract the hydrogen atom of O4 (denoted as H4) in the non-reducing end glucose, and the NH2 group of Arg398 will bind with the O4 atom of the glucose unit from the opposite direction. Because Asp60 orients the H4 atom of the substrate, Arg398 can approach the O4 atom without hindrance. Besides, if all of these atoms are in the same plane, attractive forces between the enzyme and substrate will be enhanced, since the bond distances are minimal. A similar arrangement was observed in the structure of maltotetraose-forming exo-amylase.[9] In this structure, the OD1 atom of Asp160 and the main-chain nitrogen of Gly158 at subsite −4 bind to the H4 and O4 atoms of the non-reducing end glucose residue of bound substrate, respectively. For strict recognition of the substrate, the H4 and O4 atoms of the non-reducing end glucose residue of the substrate are bound by appropriate atoms. These specific hydrogen bonds confer selectivity and may enable high-affinity binding of the non-reducing end glucose residue to the binding site of *exo*-type glycosidase.

A water drain for the active site pocket

A row of water molecules that connects the pocket with the outside of the enzyme was observed at the bottom of the active site pocket (Fig. 4). An identical row of water molecules is seen in the structures of O16G, AS and SP, which also all have a pocket-shaped active site. As far as we are aware, a similar path of water molecules has not been reported for α-amylases. The level of amino acid sequence identity of DGase with these other enzymes, O16G, AS and

Table 2. Multiple sequence alignment of the amino-acid sequences around Val195 of *Streptococcus mutans* dextran glucosidase

Enzyme	Origin	Sequence	Position	Specificity
Dextran glucosidase	*Streptococcus mutans*	GFRMD**V**IDM	190-198	α-1,6
Oligo-1,6-glucosidase	*Bacillus cereus*	GFRMD**V**INF	195-203	α-1,6
Oligo-1,6-glucosidase	*Bacillus coagulans*	GWRMD**V**IGS	195-203	α-1,6
Oligo-1,6-glucosidase	*Bacillus thermoglucosidius*	GFRMD**V**INM	195-203	α-1,6
Oligo-1,6-glucosidase	*Bacillus* sp. F5	GWRMD**V**IGS	193-201	α-1,6
Isomaltase	*Saccharomyces cerevisiae*	GFRID**V**GSL	211-219	α-1,6
Isoamylase	*Pseudomonas amyloderamosa*	GFRFD**L**ASV	397-405	α-1,6
Isoamylase	*Flavobacterium* sp.	GFRFD**L**ASV	406-414	α-1,6
Pullulanase	*Bacillus subtilis* 168S	GFRFD**L**LGI	402-410	α-1,6
Pullulanase	*Klebsiella pneumoniae*	GFRFD**L**MGY	690-698	α-1,6
α-Amylase	*Aspergillus oryzae*	GLRID**T**VKH	223-231	α-1,4
α-Amylase	*Sus scrofa* (pancreas)	GFRID**A**SKH	193-201	α-1,4
α-Amylase	*Homo sapiens* (pancreas)	GFRLD**A**SKH	208-216	α-1,4
α-Amylase	*Bacillus licheniformis*	GFRLD**A**VKH	256-264	α-1,4
α-Amylase	*Hordeum vulgare* isozyme 1	AWRLD**F**ARG	200-208	α-1,4
α-Amylase	*Pseudomonas stutzeri*	GFRFD**F**VRG	210-218	α-1,4
α-Glucosidase	*Geobacillus stearothermophilus*	GFRID**A**ISH	195-203	α-1,4
α-Glucosidase	*Saccharomyces cereviciae*	GFRID**T**AGL	210-218	α-1,4
CGTase	*Bacillus circulans* 251	GIRMD**A**VKH	252-260	α-1,4
CGTase	*Bacillus clarkii*	GIRVD**A**VAH	244-252	α-1,4
iGT	*Bacillus circulans*	GARVD**A**AKL	261-269	α-1,4
Branching enzyme	*Escherichia coli*	ALRVD**A**VAS	401-409	α-1,4
CmmA	*Arthrobacter globiformis* M6	GIRVD**A**ARS	254-262	α-1,4
MTSase	*Arthrobacter ramosus* S34	ALRVD**G**DEL	143-151	α-1,4
MTHase	*Arthrobacter ramosus* S34	GLRLD**A**VHA	246-254	α-1,4
Amylosucrase	*Neisseria polysaccharea*	ILRMD**A**VAF	290-298	α-1,4
Neopullulanase	*Geobacillus stearothermophilus*	GWRLD**V**ANE	324-332	α-1,4 α-1,6
TVAII	*Thermoactinomyces vulgaris* R47	GWRLD**V**ANE	321-329	α-1,4 α-1,6

The catalytic nucleophile and the amino acid corresponding to Val195 of dextran glucosidase are indicated by underlined and bold letters, respectively. CGTase, cyclodextrin glucanotransferase; iGT, isocyclomaltooligosaccharide glucanotransferase; CmmA, 6-maltosyltransferase; MTSase, maltooligosyltrehalose synthase; MTHase, maltooligosyltrehalose hydrolase.

SP, is not high in all cases. The amino acid identities (similarities) are 51.7% (68.2%), 15.6% (27.2%) and 15.4% (27.3%), respectively. Despite such varying degrees of identity (similarity), these enzymes all share an identical water path. So, it is possible that the water path has an important role in the enzyme reaction, and we believe that the path may affect substrate binding. When a substrate enters into an active site pocket, water molecules within the pocket must be displaced from their binding sites. If an active site pocket has only one entrance, water molecules in the pocket may be forced out through the entrance when a substrate enters. From the structure of the DGase substrate–enzyme complex, the active site pocket appears so tight that there appears to be no space to accommodate both the incoming substrate and outgoing water molecules. Therefore, the active site pocket could have another path for draining the water molecules from the active site. These water molecules could be pushed out from the pocket through the drain by the incoming substrate. We conclude that the water path at the bottom of the active site pocket may work as the water drain. The role of the water drain should be as follows. First, water molecules will be bound to their stable position in the active site pocket. When the substrate comes into the active site pocket, the bound water molecules in the pocket could be forced out through the water drain located at the bottom of the active site pocket by the incoming substrate. Finally, the water molecules are drained out from the active site pocket when the incoming substrate is bound tightly to the enzyme.

Comparison with oligo-1,6-glucosidase

Although DGase shares 52% amino acid sequence identity with O16G, their preference for substrates of different lengths is known. Of several isomaltooligosaccharides, isomaltotriose is the best substrate for both enzymes. If substrates are longer than three residues, the k_{cat}/K_m values for processing these longer substrates by O16G are reduced considerably.[26] Nevertheless, DGase does hydrolyze relatively longer isomaltooligosaccharides and dextran.[24] In our previous study, the significant role of Trp238 and the short β→α loop 4 of DGase in chain-length preference was described.[24] The mutants in which Trp238 was replaced by Ala, Pro and Asn, and a chimeric enzyme in which the β→α loop 4 (Lys201–Asn210) was exchanged for the corresponding parts of *B. subtilis* O16G (Lys206-Asn229), had lower preferences for long-chain substrates than the native enzyme. To clarify, from a structural point of view, the different preferences of DGase and O16G for substrates of different lengths, we have compared the crystal structures of these two enzymes using the program Chimera.[36] The residues that are important for substrate binding at subsite −1 are almost identical, as shown in Fig. 6. Since the positions of the Lys and

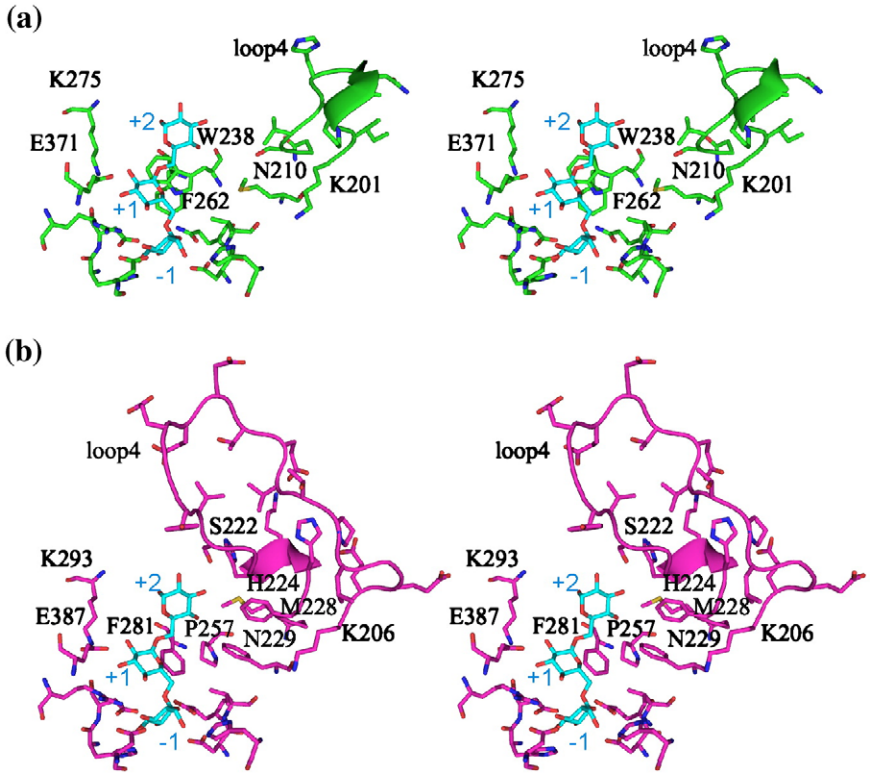

Fig. 6. Comparison of substrate binding by DGase (a) and O16G (b). The IG3 molecule bound to DGase is superimposed on the structure of O16G. Stick models of DGase in green and O16G in magenta. Bound IG3 molecules are shown in cyan. The subsite numbers are denoted in cyan.

Glu residues in subsite +1 are identical in DGase (K275 and E371) and O16G (K293 and E387), similar hydrogen bonding networks would be formed in subsite +1 of these two enzymes. Trp238 of DGase and Phe281 of O16G are found in similar positions. These residues would form the subsite +1 of DGase and O16G, respectively. The side chain of Trp238 makes a hydrophobic contact with the glucosidic linkage spanning subsites +1 and +2, while Phe281 may not interact with the identical glucosidic linkage because the location of the side chain is not appropriate for contact with the linkage. Moreover, the conformation of the long-chain substrate bound to O16G should be different from the conformation when bound to DGase because amino acid residues Ser222, His224 and Met228 on the β→α loop 4 of O16G will clash with the glucose unit in subsite +2. Therefore, Trp238 of DGase would confer favorable recognition of substrate in subsites +1 and +2, which are favorable for binding longer substrates, while the interactions of longer-chain substrates with the residues of the β→α loop 4 of O16G may prevent efficient substrate binding.

Materials and Methods

Preparation and purification

Recombinant DGase with a His tag was produced in *E. coli* BL21 (DE3) CodonPlus RIL (Stratagene, USA) containing the expression plasmid derived from pET23d (Novagen, Germany) as described.[37] Isopropyl β-D-thiogalactopyranoside was added (0.1 mM) to the cultures, and cells were cultured for another 24 h at 18 °C. Recombinant DGase was purified from cells harvested from 1 l of culture broth by Ni-chelating Sepharose column chromatography as described.[35] The enzyme was eluted with buffer A (20 mM sodium acetate buffer, pH 6.2) containing 300 mM imidazole, and dialyzed against buffer A. The enzyme was purified to homogeneity using DEAE Toyopearl column chromatography. The enzyme was eluted using a linear gradient (0–500 mM) of NaCl in buffer A. Purified DGase was dialyzed against 0.5×buffer A. The mutant DGase (E236Q) was generated by megaprimer PCR,[38] using the wild-type DGase expression plasmid as template and an antisense primer: 5′-TGCTCCCCAAGTTTGCCCCA-CAGTCAG-3′ to introduce the desired mutation. The gene was inserted into pET23d. The desired mutation and correct sequence of the inserted DNA was verified. Production of the E236Q in *E. coli* and purification were done as for the wild type.

Crystallization and data collection

Crystallization and data collection of uncomplexed native DGase were performed by the hanging-drop, vapor-diffusion method at 20 °C as described.[37] The reservoir solution was 100 mM Tris–HCl buffer (pH 7.5) containing 120 mg/ml of polyethylene glycol 6000 and 200 mM calcium chloride. DGase was dissolved in 0.5× buffer A at a concentration of 10 mg/ml. Crystallization droplets were prepared by mixing protein solution and reservoir solution to a final volume of 8 μl at a ratio of 5:3 (v/v). The E236Q mutant and IG3 were co-crystallized to

Table 3. Data collection and refinement statistics

	Native	E236Q mutant (IG3-complex)
A. *Crystal data*		
Space group	Orthorhombic $P2_12_12_1$	
a (Å)	72.72	72.49
b (Å)	86.47	82.53
c (Å)	104.30	104.30
Resolution (Å)[a]	2.2 (2.32–2.20)	2.2 (2.32–2.20)
No. measured reflections	189,588	221,042
No. unique reflections	32,892	32,499
I/σ (I)	10.1 (4.4)	9.6 (3.9)
$R_{r.i.m.}$(%)[a,b]	6.8 (18.4)	8.1 (21.4)
Completeness (%)[a]	96.8 (88.7)	100 (100)
B. *Refinement*		
Resolution limits (Å)	10–2.2	10–2.2
R_{cryst} (%)/R_{free} (%)	18.3/22.2	18.4/22.4
No. protein atoms	4374	4374
No. oligosaccharide atoms	–	34
No. water molecules	405	340
No. calcium ions	3	3
No. other atoms	14	–
r.m.s. deviations from ideal		
Bond lengths (Å)	0.005	0.007
Bond angles (°)	1.2	1.3
B-values		
Average (Å^2)	15.9	13.9
Water molecules (Å^2)	22.6	5.0
Oligosaccharides (Å^2)	–	5.7

[a] Values for highest resolution shells are shown in parentheses.
[b] $R_{r.i.m} = \Sigma_{hkl}\sqrt{N/N-1}\Sigma_i|Ii(hkl) - \overline{I(hkl)}|/\Sigma_{hkl}\Sigma_i Ii(hkl)$, where N is the redundancy and Ii is the intensity of the Ith observation.

prepare crystals of the DGase–IG3 complex. The reservoir solution in this co-crystallization was the same as that used for native DGase, with a concentration of IG3 in the crystallization droplet of 25 mM. For data collection under cryogenic conditions, crystals were soaked in a cryoprotectant solution of 100 mM Tris–HCl (pH 7.5), 200 mg/ml of polyethylene glycol 6000, 200 mM calcium chloride and 370 mg/ml of glycerol. For crystals of the DGase–IG3 complex, 25 mM IG3 was added.

Crystals were flash-frozen in a stream of cold nitrogen gas (100 K) just before data collection. The intensity data of native and substrate-complexed crystals were collected using an R-AXIS IV^{++} image-plate detector with CuKα radiation from a Rigaku ultra X 18 rotating-anode generator equipped with an Osmic mirror system. All data collection was performed at 100 K. Diffraction data sets were processed using the programs Mosflm[39] and SCALA in the CCP4 suite.[40] The data collection statistics are summarized in Table 3.

Structure determination and crystallographic refinement

The native structure was determined at a resolution of 2.2 Å by the molecular replacement method using the structure of O16G as a search model with the program AMoRe[41] in the CCP4 suite. The substrate–enzyme complex structure was also solved at 2.2 Å resolution but using the native structure as a search model. All model building and manual mode corrections were carried out using the programs CNS[42] and COOT.[43] The stereochemistry of the models was verified using the program PROCHECK.[44] The refinement statistics are given in Table 3.

Protein Data Bank accession codes

The atomic coordinates and structure factors of the free DGase and the E236Q–IG3 complex have been deposited in the Protein Data Bank under accession codes 2zic and 2zid, respectively.

Acknowledgement

This work was supported by a Grant-in-Aid from the Japan Society for the Promotion of Science Fellows.

References

1. Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **280**, 309–316.
2. Kuriki, T. & Imanaka, T. (1999). The concept of the α-amylase family: Structural similarity and common catalytic mechanism. *J. Biosci. Bioeng.* **87**, 557–565.
3. Svensson, B. (1994). Protein engineering in the α-amylase family: catalytic mechanism, substrate specificity, and stability. *Plant Mol. Biol.* **25**, 141–157.
4. Matsuura, Y., Kusunoki, M., Harada, W. & Kakudo, M. (1984). Structure and possible catalytic residues of Taka-amylase A. *J. Biochem. (Tokyo)*, **95**, 697–702.
5. Qian, M., Haser, R., Buisson, G., Duee, E. & Payan, F. (1994). The active center of a mammalian α-amylase. Structure of the complex of a pancreatic α-amylase with a carbohydrate inhibitor refined to 2.2 Å resolution. *Biochemistry*, **33**, 6284–6294.
6. Larson, S. B., Greenwood, A., Cascio, D., Day, J. & McPherson, A. (1994). Refined molecular structure of pig pancreatic α-amylase at 2.1 Å resolution. *J. Mol. Biol.* **235**, 1560–1584.
7. Kadziola, A., Abe, J., Svensson, B. & Haser, R. (1994). Crystal and molecular structure of barley α-amylase. *J. Mol.Biol.* **239**, 104–121.
8. Qian, M., Haser, R. & Payan, F. (1995). Carbohydrate binding sites in a pancreatic α-amylase-substrate complex, derived from X-ray structure analysis at 2.1 Å resolution. *Protein Sci.* **4**, 747–755.
9. Yoshioka, Y., Hasegawa, K., Matsuura, Y., Katsube, Y. & Kubota, M. (1997). Crystal structure of a mutant maltotetraose-forming exo-amylase cocrystallized with maltopentaose. *J. Mol. Biol.* **271**, 619–628.
10. Brzozowski, A. M. & Davies, G. J. (1997). Structure of the *Aspergillus oryzae* α-Amylase complexed with the inhibitor acarbose at 2.0 Å resolution. *Biochemistry*, **36**, 10837–10845.
11. Kadziola, A., Søgaard, M., Svensson, B. & Haser, R. (1998). Molecular structure of a barley α-amylase-inhibitor complex: implications for starch binding and catalysis. *J. Mol. Biol.* **278**, 205–217.
12. Vallée, F., Kadziola, A., Bourne, Y., Juy, M., Rodenburg, K. W., Svensson, B. & Haser, R. (1998). Barley α-amylase bound to its endogenous protein inhibitor BASI: crystal structure of the complex at 1.9 Å resolution. *Structure*, **6**, 649–659.
13. Fujimoto, Z., Takase, K., Doui, N., Momma, M., Matsumoto, T. & Mizuno, H. (1998). Crystal structure of a catalytic-site mutant α-amylase from *Bacillus*

subtilis complexed with maltopentaose. *J. Mol. Biol.* **277**, 393–407.
14. Hasegawa, K., Kubota, M. & Matsuura, Y. (1999). Roles of catalytic residues in α-amylases as evidenced by the structures of the product-complexed mutants of a maltotetraose-forming amylase. *Protein Eng.* **12**, 819–824.
15. MacGregor, E. A., Janecek, S. & Svensson, B. (2001). Relationship of sequence and structure to specificity in the α-amylase family of enzymes. *Biochim. Biophys. Acta*, **1546**, 1–20.
16. Mirza, O., Skov, L. K., Remaud-Simeon, M., Potocki de Montalk, G., Albenne, C., Monsan, P. & Gajhede, M. (2001). Crystal structures of amylosucrase from *Neisseria polysaccharea* in complex with D-glucose and the active site mutant Glu328Gln in complex with the natural substrate sucrose. *Biochemistry*, **40**, 9032–9039.
17. Kizaki, H., Hata, Y., Watanabe, Y., Katsube, Y. & Suzuki, Y. (1993). Polypeptide folding of *Bacillus cereus* ATCC7064 oligo-1,6-glucosidase revealed by 3.0 Å resolution X-ray analysis. *J. Biochem. (Tokyo)*, **113**, 646–649.
18. Watanabe, K., Hata, Y., Kizaki, H., Katsube, Y. & Suzuki, Y. (1997). The refined crystal structure of *Bacillus cereus* oligo-1,6-glucosidase at 2.0 Å resolution: structural characterization of proline-substitution sites for protein thermostabilization. *J. Mol. Biol.* **269**, 142–153.
19. Katsuya, Y., Mezaki, Y., Kubota, M. & Matsuura, Y. (1998). Three-dimensional structure of *Pseudomonas* isoamylase at 2.2 Å resolution. *J. Mol. Biol.* **281**, 885–897.
20. Hondoh, H., Kuriki, T. & Matsuura, Y. (2003). Three-dimensional structure and substrate binding of *Bacillus stearothermophilus* neopullulanase. *J. Mol. Biol.* **326**, 177–188.
21. Mikami, B., Iwamoto, H., Malle, D., Yoon, H. Y., Demirkan-Sarikaya, E., Mezaki, Y. & Katsuya, Y. (2006). Crystal structure of pullulanase: Evidence for parallel binding of oligosaccharides in the active site. *J. Mol. Biol.* **359**, 690–707.
22. Linder, L. & Sund, M. L. (1981). Characterization of dextran glucosidase (1,6-α-D-glucan glucohydrolase) of *Streptococcus mitis*. *Caries Res.* **15**, 436–444.
23. Russell, R. R. & Ferretti, J. J. (1990). Nucleotide sequence of the dextran glucosidase (*dexB*) gene of *Streptococcus mutans*. *J. Gen. Microbiol.* **136**, 803–810.
24. Saburi, W., Mori, H., Saito, S., Okuyama, M. & Kimura, A. (2006). Structural elements in dextran glucosidase responsible for high specificity to long chain substrate. *Biochim. Biophys. Acta*, **1764**, 688–698.
25. Suzuki, Y., Aoki, R. & Hayashi, H. (1982). Assignment of a *p*-nitrophenyl-α-D-glucopyranoside-hydrolyzing-α-glucosidase of *Bacillus cereus* ATCC7064 to an exo-oligo-1,6-glucosidase. *Biochim. Biopys. Acta*, **704**, 476–483.
26. Suzuki, Y. & Tomura, Y. (1986). Purification and characterization of *Bacillus coagulans* oligo-1,6-glucosidase. *Eur. J. Biochem.* **158**, 77–83.
27. Suzuki, Y., Fujii, H., Uemura, H. & Suzuki, M. (1987). Purification and characterization of extremely thermostable *exo*-oligo-1,6-glucosidase from a caldoactive *Bacillus* sp. KP 1228. *Starch*, **39**, 17–23.
28. Suzuki, Y., Yuki, T., Kishigami, T. & Abe, S. (1976). Purification and properties of extracellular α-glucosidase of a thermophile, *Bacillus thermoglucosidius* KP 1006. *Biochim. Biophys. Acta*, **445**, 386–397.
29. Skov, L. K., Mirza, O., Henriksen, A., Potocki de Montalk, G., Remaud-Simeon, M., Sarçabal, P. *et al.* (2001). Amylosucrase, a glucan-synthesizing enzyme from the α-amylase family. *J. Biol. Chem.* **276**, 25273–25278.
30. Sproge, D., van den Broek, L. A. M., Mirza, O., Kastrup, J. S., Voragen, A. G. J., Gajhede, M. & Skov, L. K. (2004). Crystal structure of sucrose phosphorylase from *Bifidobacterium adolescentis*. *Biochemistry*, **43**, 1156–1162.
31. Nielsen, J. E. & Borchert, T. V. (2000). Protein engineering of bacterial α-amylases. *Biochim. Biophys. Acta*, **1543**, 253–274.
32. Aghajari, N., Feller, G., Gerday, C. & Haser, R. (1998). Crystal structures of the psychrophilic α-amylase from *Alteromonas haloplanctis* in its native form and complexed with an inhibitor. *Protein Sci.* **7**, 564–572.
33. Brzozowski, A. M., Lawson, D. M., Turkenburg, J. P., Bisgaard-Frantzen, H., Svendsen, A., Borchert, T. V. *et al.* (2000). Structural analysis of a chimeric bacterial α-Amylase. High-resolution analysis of native and ligand complexes. *Biochemistry*, **39**, 9099–9107.
34. Skov, L. K., Mirza, O., Sprogøe, D., Dar, I., Remaud-Simeon, R., Albene, C. *et al.* (2002). Oligosaccharide and sucrose complexes of amylosucrase. *J. Biol. Chem.* **277**, 47741–47747.
35. Mirza, O., Skov, L. K., Sprogøe, D., van den Broek, L. A. M., Beldman, G., Kastrup, J. S. & Gajhede, M. (2006). Structural rearrangements of sucrose phosphorylase from *Bifidobacterium adolescentis* during sucrose conversion. *J. Biol. Chem.* **281**, 35576–35584.
36. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). UCSF Chimera – a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.
37. Saburi, W., Hondoh, H., Unno, H., Okuyama, M., Mori, H., Nakada, T. *et al.* (2007). Crystallization and preliminary X-ray analysis of *Streptococcus mutans* dextran glucosidase. *Acta Crystallogr. F*, **63**, 774–776.
38. Sarkar, G. & Sommer, S. S. (1990). The "megaprimer" method of site-directed mutagenesis. *BioTechniques*, **8**, 404–407.
39. Leslie, A. G. W. (1992). *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography*, vol. 26, pp. 30, Daresbury Laboratory, Warrington, UK.
40. Collaborative Computational Project, Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D*, **50**, 760–763.
41. Navaza, J. (1994). *AMoRe*: an automated package for molecular replacement. *Acta Crystallogr. A*, **50**, 157–163.
42. Brünger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). *Crystallography & NMR system*: a new software suite for macromolecular structure determination. *Acta Crystallogr. D*, **54**, 905–921.
43. Emsley, P. & Cowtan, K. (2004). *Coot*: model-building tools for molecular graphics. *Acta Crystallogr. D*, **60**, 2126–2132.
44. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *PROCHECK*: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.