CrossMark

Structural analysis of the α-glucosidase HaG provides new insights into substrate specificity and catalytic mechanism

Xing Shen,[a,b] Wataru Saburi,[c] Zuoqi Gai,[d] Koji Kato,[d] Teruyo Ojima-Kato,[e] Jian Yu,[d] Keisuke Komoda,[d] Yusuke Kido,[c] Hirokazu Matsui,[c] Haruhide Mori[c] and Min Yao[b,d,f]*

[a]School of Light Industry and Food Science, South China University of Technology, 381 Wushan Road, Tianhe District, Guangzhou, Guangdong 510640, People's Republic of China, [b]Graduate School of Life Science, Hokkaido University, Kita 10, Nishi 8, Kita-ku, Sapporo, Hokkaido 060-0810, Japan, [c]Research Faculty of Agriculture, Hokkaido University, Kita 9, Nishi 9, Kita-ku, Sapporo, Hokkaido 060-8689, Japan, [d]Faculty of Advanced Life Science, Hokkaido University, Kita 10, Nishi 8, Kita-ku, Sapporo, Hokkaido 060-0810, Japan, [e]Graduate School of Bioagricultural Sciences, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan, and [f]Department of Pharmacology, Basic Medical College of Zhengzhou University, 100 Science Avenue, Zhengzhou, Henan 450001, People's Republic of China.
*Correspondence e-mail: yao@castor.sci.hokudai.ac.jp

α-Glucosidases, which catalyze the hydrolysis of the α-glucosidic linkage at the nonreducing end of the substrate, are important for the metabolism of α-glucosides. *Halomonas* sp. H11 α-glucosidase (HaG), belonging to glycoside hydrolase family 13 (GH13), only has high hydrolytic activity towards the α-(1→4)-linked disaccharide maltose among naturally occurring substrates. Although several three-dimensional structures of GH13 members have been solved, the disaccharide specificity and α-(1→4) recognition mechanism of α-glucosidase are unclear owing to a lack of corresponding substrate-bound structures. In this study, four crystal structures of HaG were solved: the apo form, the glucosyl-enzyme intermediate complex, the E271Q mutant in complex with its natural substrate maltose and a complex of the D202N mutant with D-glucose and glycerol. These structures explicitly provide insights into the substrate specificity and catalytic mechanism of HaG. A peculiar long β→α loop 4 which exists in α-glucosidase is responsible for the strict recognition of disaccharides owing to steric hindrance. Two residues, Thr203 and Phe297, assisted with Gly228, were found to determine the glycosidic linkage specificity of the substrate at subsite +1. Furthermore, an explanation of the α-glucosidase reaction mechanism is proposed based on the glucosyl-enzyme intermediate structure.

1. Introduction

α-Glucosidase (EC 3.2.1.20) hydrolyzes the nonreducing-end α-glucosidic linkages of its substrate to release α-D-glucose. This enzyme is important for the generation of D-glucose from carbon sources, including maltodextrin and sucrose, in various living organisms. The substrate specificity of α-glucosidases differs greatly depending on the source of the enzymes. According to substrate specificity, α-glucosidases are classified into three groups: group 1 enzymes prefer heterogeneous substrates such as sucrose and p-nitrophenyl α-D-glucopyranoside to homogeneous substrates such as maltooligosaccharides, while group 2 enzymes prefer homogeneous substrates to heterogeneous substrates; group 3 enzymes prefer homogenerous substrates, similar to group 2 enzymes,

but have a high activity towards long-chain substrates (Chiba, 1988).

α-Glucosidase from *Halomonas* sp. strain H11 (HaG) belongs to group 2 and is an enzyme that was found by screening with high transglucosylation activity towards low-molecular-weight compounds such as glycerol and ethanol (Ojima, Saburi *et al.*, 2012). α-Glucosides, α-glucosyl glycerol and 5-α-D-glucosylgingerol have been efficiently synthesized by transglucosylation of HaG with maltose as the glucosyl donor substrate (Ojima, Aizawa *et al.*, 2012; Ojima, Saburi *et al.*, 2012). HaG has high regioselectivity for the α-(1→4)-glucosidic linkage in the hydrolytic reaction. Interestingly, this enzyme is highly specific for disaccharides, unlike other bacterial α-glucosidases, which are generally more active towards oligosaccharides longer than disaccharides (Hung *et al.*, 2005; Kelly & Fogarty, 1983; Nakao *et al.*, 1994). Furthermore, HaG is significantly activated by monovalent cations such as K^+, Rb^+, Cs^+ and NH_4^+. According to the sequence-based classification of glycoside hydrolases, HaG belongs to GH family 13 (GH13) containing various retaining glycosidases and glycosyltransferases such as α-amylase (EC 3.2.1.1) and cyclodextrin glucanotransferase (EC 2.4.1.19) (Henrissat & Davies, 1997). In further classification of the GH13 family, HaG falls into subfamily 23 (Stam *et al.*, 2006), which contains α-transglucosidase from *Xanthomonas campestris* WU-9701 (XgtA), which efficiently synthesizes various α-glucosides by transglucosylation (Nakagawa *et al.*, 2000; Sato *et al.*, 2000; Kurosu *et al.*, 2002).

The members of the GH13 family share three common domains: domain A, a catalytic domain formed by a $(\beta/\alpha)_8$-barrel fold; domain B, a loop-rich domain connected to β3 and α3 of domain A; and domain C, a domain made of β-strands following domain A (MacGregor *et al.*, 2001). The catalytic nucleophile (Asp) and the acid/base catalyst (Glu) are situated at the C-termini of the fourth β-strand and the fifth β-strand of domain A, respectively. The catalytic reaction of the GH13 family enzymes is believed to proceed through a double-displacement mechanism (MacGregor *et al.*, 2001; Uitdehaag *et al.*, 1999). The general acid/base catalyst donates a proton to the glucosidic O atom, and the catalytic nucleophile attacks the anomeric centre of the glucosyl residue to form a β-glycosyl enzyme intermediate. A deprotonated general acid/base catalyst assists in the nucleophilic attack of a water molecule on the anomeric centre of the intermediate and the release of the product as the α-anomer. In transglucosylation, the OH group of acceptor molecules such as sugars, alcohol and glycerol nucleophilically attacks the intermediate in place of water.

The GH13 family contains two types of exo-glucosidases. The first are α-(1→4)-specific enzymes, called α-glucosidases, and the other are α-(1→6)-specific enzymes such as oligo-1,6-glucosidases (EC 3.2.1.10) and dextran glucosidases (EC 3.2.1.70). We have previously reported the three-dimensional structure of *Streptococcus mutans* dextran glucosidase (SmDG) complexed with isomaltotriose and discussed the substrate-recognition mechanism of the α-(1→6)-glucosidic linkage (Hondoh *et al.*, 2008). In contrast, no substrate-bound

structure of an α-(1→4)-specific glucosidase is available, and thus the substrate-recognition mechanism of α-(1→4)-specific glucosidases remains unclear.

In this study, we investigated the substrate recognition and catalytic mechanism of HaG. Four structures of HaG are reported here: the apo form (HaG), the glucosyl-enzyme intermediate structure (glucosyl-HaG), the E271Q mutant in complex with its natural substrate maltose (E271Q-Mal) and the complex of the D202N mutant with both D-glucose and glycerol (D202N-Glc-Gol). This is the first report of both the native reaction intermediate of an α-glucosidase and of an α-(1→4)-linkage substrate complex in the GH13 family, and also the first report of a crystal structure of a subfamily 23 enzyme. These results reveal new insights into the substrate specificity and catalytic mechanism of α-glucosidase.

2. Methods

2.1. Plasmid construction, expression and purification of recombinant proteins

Wild-type HaG was expressed in *Escherichia coli* BL21(DE3) cells harbouring the recombinant plasmid pFLAG-CTS and purified using a DEAE-650M anion-exchange column (Tosoh, Tokyo, Japan) as described elsewhere (Shen *et al.*, 2014). Expression plasmids for two active-site mutants, D202N and E271Q, were created by site-directed mutagenesis with a Primestar Mutagenesis Basal Kit (Takara Bio, Otsu, Japan). The mutant enzymes were prepared using the same protocol as for the wild-type HaG (Shen *et al.*, 2014).

2.2. Crystallization and data collection

Crystals of wild-type HaG were grown in 0.1 *M* HEPES–NaOH buffer pH 7.0 containing 0.02 *M* magnesium chloride and 20% polyacrylic acid (PAA) 5100 sodium salt as described previously (Shen *et al.*, 2014).

The crystallizations of the complexes were performed by co-crystallization or soaking. The crystal of the glucosyl-HaG intermediate was obtained by co-crystallization with sucrose in 0.1 *M* HEPES pH 7.5, 22% PEG 6000, 0.5 *M* LiCl. Drops consisted of 0.5 µl 10 mg ml^{-1} wild-type HaG, 0.5 µl 2.0 m*M* sucrose and 1.0 µl reservoir solution. The crystal of the E271Q-Mal complex was obtained by soaking the crystals of the E271Q mutant in the same crystallization solution as HaG containing 1 m*M* maltose for 15 min. The D202N-Glc-Gol complex was co-crystallized with glucose and glycerol, both at final concentrations of 2 m*M*, in the same crystallization condition.

All crystals used for diffraction experiments were cryoprotected in advance by rapid soaking in the mother liquor containing 15%(*v/v*) glycerol and were then flash-cooled under a stream of liquid nitrogen at 100 K. For E271Q-Mal, 1 m*M* maltose was also included in the cryoprotectant buffer.

While diffraction data for the D202N-Glc-Gol crystals were collected on beamline 44XU at SPring-8, Hyogo, Japan, other data were collected on beamlines 5A, 17A or NE3A of Photon Factory, Ibaraki, Japan. All crystals, including those of wild-

Table 1
Crystallization, data-collection and refinement statistics.

	HaG	Glucosyl-HaG	E271Q-Mal	D202N-Glc-Gol
Ligand added	—	Sucrose	Maltose	Glucose and glycerol
Method	—	Co-crystallization	Soaking	Co-crystallization
Data collection				
Wavelength (Å)	1.0	1.0	1.0	1.0
Resolution range (Å)	50–2.15 (2.28–2.15)	50–1.47 (1.56–1.47)	50–2.50 (2.65–2.50)	50–3.00 (3.18–3.00)
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12$	$P2_12_12_1$
Unit-cell parameters (Å)	$a = 60.5$, $b = 119.6$, $c = 178.0$	$a = 63.2$, $b = 103.8$, $c = 176.9$	$a = 111.5$, $b = 181.1$, $c = 51.9$	$a = 60.6$, $b = 119.6$, $c = 178.4$
No. of reflections	518379	1436347	180927	196381
No. of unique reflections	71192	197034	37261	26750
Multiplicity	7.3 (7.4)	7.3 (7.2)	4.9 (4.9)	7.3 (7.4)
Completeness (%)	99.7 (98.7)	99.8 (98.9)	99.4 (99.0)	99.8 (99.0)
R_{meas} (%)	8.9 (52.6)	5.6 (56.7)	11.1 (68.7)	16.7 (61.9)
$\langle I/\sigma(I) \rangle$	20.8 (5.8)	25.2 (3.9)	14.6 (2.5)	15.8 (5.5)
Wilson B factor (Å^2)	35.3	21.2	42.5	38.3
No. of protein molecules in asymmetric unit	2	2	2	2
Refinement				
Resolution range (Å)	50–2.15	50–1.47	50–2.50	50–3.00
No. of reflections used	67524	197027	35293	25311
No. of atoms				
Protein	8584	8596	7690	8584
Water	649	1503	452	0
Others	48	49	48	50
Average B, all atoms (Å^2)	35.4	19.9	42.3	42.1
R/R_{free} (%)	18.3/21.9	15.6/17.9	20.2/24.1	19.8/24.4
R.m.s.d., bond lengths (Å)	0.008	0.010	0.003	0.002
R.m.s.d., angles (°)	1.131	1.060	0.789	0.665
Ramachandran plot (%)				
Favoured	97.7	97.9	95.1	96.8
Allowed	2.4	2.2	3.5	3.2
Outliers	0	0	1.4	0

type HaG, belonged to the same space group, $P2_12_12_1$, with different unit-cell parameters, apart from the E271Q-Mal complex, which belonged to space group $P2_12_12$. All diffraction data were indexed, integrated, scaled and merged using the *XDS* (Kabsch, 2010) and *CCP*4 (Winn *et al.*, 2011) packages. The crystals of HaG, glucosyl-HaG, E271Q-Mal and D202N-Glc-Gol diffracted to 2.15, 1.47, 2.50 and 3.00 Å resolution, respectively, and the Wilson B factors for these data sets were 35.3, 21.2, 42.5 and 38.3 Å^2, respectively. The statistics of data processing are summarized in Table 1.

2.3. Structure determination and refinement

The structure of HaG was determined by the molecular-replacement (MR) method with *Phaser* (McCoy *et al.*, 2007), using the monomer structure of the isomaltulose synthase PalI from *Klebsiella* sp. LX3 (PDB entry 1m53; 36% identity to HaG; Zhang *et al.*, 2003) as a search model (Shen *et al.*, 2014).

The monomer structure of HaG was then used as a search model in the MR method to solve the structures of glucosyl-HaG and E271Q-Mal. To confirm that the E271Q-Mal crystal belonged to a different space group from the others, the MR calculation was performed using different space groups such as $P2_12_12_1$, $P2_12_12$ or $P2_1$. An initial rigid-body refinement followed by a jelly-body refinement was then performed with *REFMAC* (Murshudov *et al.*, 2011). The structure of D202N-

Glc-Gol was obtained by rigid-body refinement using the HaG model.

All of the models were manually modified using *Coot* (Emsley *et al.*, 2010) and refined with *PHENIX* (Adams *et al.*, 2010). After several cycles of refinement, electron-density blobs for ligands appeared in both $2F_o - F_c$ and $F_o - F_c$ maps, and all ligand models were manually built based on $2F_o - F_c$ and $F_o - F_c$ maps. The refinement of glucosyl-HaG was tested with several conditions including multiple conformations with and without restraints. Finally, refinement was performed with an additional bond restraint between the C1 atom of the glucosyl residue and the O$^{\delta1}$ atom of the nucleophile Asp202; a bond length set to 1.4 Å with a large σ (10.0 Å) provided the best result and the distance between the C1 and O$^{\delta1}$ atoms converged at 2.15 Å. Active sites in the enzyme-complex structures were inspected and all ligands were checked by OMIT refinement. R_{work} and R_{free} were monitored during the refinement processing and the latter was calculated from 5% of reflections that were randomly selected and not used for refinement. The statistics of the refinement are summarized in Table 1.

2.4. Alignment and figure preparation

The sequences of HaG and other enzymes were aligned using *ClustalW* (Thompson *et al.*, 1994), and the amino-acid

sequence alignment around $\beta \rightarrow \alpha$ loop 4 (Leu201–Arg246) between HaG and other enzymes was rendered using *ESPript* (Gouet *et al.*, 2003). All of the three-dimensional structural drawings were prepared with *PyMOL* (v.1.5.0.4; Schrödinger).

2.5. Coordinates

The atomic coordinates and structure factors were deposited in the Protein Data Bank; the accession codes for HaG, glucosyl-HaG, E271Q-Mal and D202N-Glc-Gol are 3wy1, 3wy2, 3wy4 and 3wy3, respectively.

3. Results and discussion

3.1. Overall structure

All four crystal structures contain dimers in the asymmetric unit and all HaG molecules were nearly fully built except for the E271Q-Mal complex (see below). Similar to the members of the GH13 family, the structure of HaG consists of the main catalytic domain A (residues 1–106 and 175–465), a loop-rich domain B (residues 107–174) and a highly conserved domain C (residues 466–538) formed by two antiparallel β-sheets (Fig. 1*a*). An electron-density blob which appears to be an ion and is coordinated by the $O^{\delta 1}$ atoms of Asp23, Asp27 and Asp31, the O atom of Val29 and a water molecule was found in all structures. Moreover, this electron-density blob is weaker in the glucosyl-HaG structure than in the other structures. Considering the crystallization condition (0.02 *M* magnesium chloride was added, apart from in the case of glucosyl-HaG) together with the coordination bond and refinement results, this blob was assigned as Mg^{2+}. Like Ca^{2+} in GH-family enzymes, such an Mg^{2+} ion could also stabilize the structure (Kobayashi *et al.*, 2011).

Domain A is the catalytic domain, which is formed of a classical TIM barrel (Banner *et al.*, 1975) and is sandwiched between domains B and C. Domain B contains one α-helix and three β-strands; it stretches from domain A and forms the wall of the catalytic pocket together with domain A. Domain C appears to stabilize the conformation of the whole structure, but its function remains unclear (Zhang *et al.*, 2003). In HaG, there are four phenylalanines (Phe479, Phe492, Phe516 and Phe534) in domain C, all of which are located on the surface in

contact with domain A, forming a hydrophobic interface that stabilizes the entire structure.

3.2. Active site

The overall structure of HaG displays the highest similarity to a member of the GH13 family, the sucrose isomerase MutB (PDB entry 2pwh; Ravaud *et al.*, 2007) from *Pseudomonas mesoacidophila* MX-45, with a root-mean-square deviation (r.m.s.d.) of 1.9 Å for 556 C^{α} atoms as analyzed using the *DALI* server (Holm & Rosenström, 2010). Interestingly, although the architecture of subsite −1 at the bottom of the active-site pockets of both enzymes is similar, the pockets of subsite +1 of the active site are different (Fig. 1*b*). Moreover, while Phe256 and Phe280 in MutB (MutB_Phe256 and MutB_Phe280) are conserved in sucrose isomerases and are essential in controlling intramolecular transglucosylation
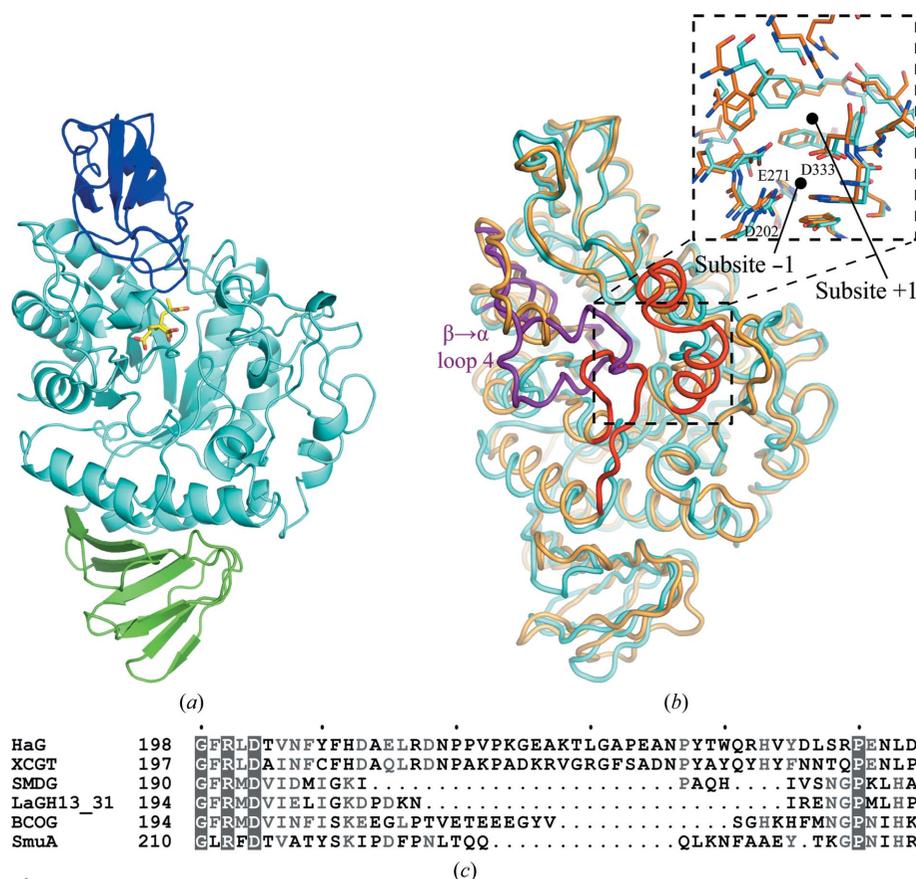
Figure 1
The structure and sequence alignment of HaG. (*a*) Overall structure of HaG. Domains A, B and C are shown in cyan, blue and green, respectively. One polyacrylic acid molecule is shown in yellow. (*b*) Superposition of HaG (cyan) with MutB (bright orange). The long $\beta \rightarrow \alpha$ loop 4 of HaG is shown in purple and the other main differences between the two structures are indicated in red. An enlargement of the active sites is shown in the dotted box. (*c*) Amino-acid sequence alignment of $\beta \rightarrow \alpha$ loop 4 between HaG and other enzymes. The sequences from the top to the bottom belong to the α-glucosidase HaG from *Halomonas* sp. H11 (GenBank accession No. BAL49684), the α-glucosidase XcGT from *Xanthomonas campestris* WU-9701 (GenBank accession No. BAC87873), the dextran glucosidase SmDG from *Streptococcus mutans* (GenBank accession No. BAE79634), the glucan-(1→6)-α-glucosidase LaGH13_31 from *Lactobacillus acidophilus* NCFM (GenBank accession No. WP_029779818), the oligo-1,6-glucosidase BcOG from *Bacillus cereus* (GenBank accession No. WP_002044712) and the isomaltulose synthase PrSmuA from *Protaminobacter rubrum* (GenBank accession No. EKF64560).

(Ravaud *et al.*, 2007, 2009), the corresponding residue to MutB_Phe256, Gly273, in HaG does not interact with the substrate. Phe297 of HaG corresponds to MutB_Phe280 and is positioned further from the substrate in HaG–substrate complex structures. Such structural characteristics may be reflected in their transglucosylation specificities.

The active-site pocket of HaG consists of 15 residues. Among these residues, eight residues, including the catalytic nucleophile Asp202, the general acid/base catalyst Glu271, Asp333, Arg200, His105, His332, Phe297 and Tyr65, are conserved in the GH13 family. Six residues, Asp62, Arg400, Phe166, Thr203, Phe206 and Phe147, are conserved in the homologues of HaG and another residue, Gly228, is only found in HaG. The distance between the catalytic nucleophile ($O^{\delta 1}$ of Asp202) and the general acid/base catalyst ($O^{\epsilon 2}$ of Glu271) is 5.6 Å, as observed in other retaining glycosidases (Davies & Henrissat, 1995). The loop (Leu201–Arg246) between the fourth β-sheet and α-helix ($\beta \rightarrow \alpha$ loop 4) in the catalytic A domain in HaG, which is longer than that in other GH13 members (Fig. 1*c*), is located just above the subsite +1 glucosyl moiety and covers a major part of the active-site entrance (Fig. 1*b*).

The conformations of the active sites described above are very similar in all current structures (r.m.s.d. within 0.21 Å for 15 C$^\alpha$ atoms), which means that the reaction does not require a conformational change in either the binding residues or the reaction residues of HaG. This suggests that less energy is required to start the catalytic reaction for HaG than for other enzymes, for example, *Geobacillus* sp. HTA-462 α-glucosidase (GSJ; PDB entry 2ze0), which is assumed to form a more compact active site by conformational change for a catalytic reaction (Shirai *et al.*, 2008). This may explain why HaG exhibits more than 50% of the maximum activity even at 277 K, whereas GSJ shows a thermophilic pattern.

3.3. E271Q-Mal substrate complex

To obtain the substrate-bound complex, the general acid/base Glu271 was mutated to the inactive Gln271 (E271Q). This residue was located at subsite +1 and thus does not influence the substrate binding at subsite −1 (Ravaud *et al.*, 2007). The structure of the inactive acid/base mutant in complex with its natural substrate (E271Q-Mal) allowed us to precisely define the manner of substrate recognition and binding.

The structure of E271Q-Mal was determined in space group $P2_12_12$ with two complexes in the asymmetric unit, one of which (chain *B*) has a large disordered part (114 residues in total) for which electron density is invisible (residues 23–29, 64–78, 222–230, 331–345 and 373–440) in domain A. This disordered part includes half of the catalytic pocket wall, and no maltose molecule was found here. Interestingly, similar disorder was also found in another E271Q-Mal crystal structure obtained by soaking with maltose, the space group of which is also $P2_12_12$, but under a different crystallization condition consisting of 0.1 *M* HEPES pH 7.5, 20% PEG 6000, 0.5 *M* LiCl (data not shown). In another complete monomer

(chain *A*), the $F_o − F_c$ map clearly showed the existence of a maltose molecule (Fig. 2*a*). Below, we explain the manner of maltose binding based on this monomer.

The glucose residue at subsite −1 is fixed in place by six conserved residues: His105, Asp62, Arg200, Arg400, Asp333 and His332 (Fig. 2*b*). A salt bridge is formed between Asp62 and Arg400, which are important for the recognition of the nonreducing-end glucosyl residue. Such salt bridges not only exist in glucosidase but also in other exo-enzymes such as amylosucrase and sucrose phosphorylase (Hondoh *et al.*, 2008). Two residues, Thr203 and Phe297, were found to be correlated with glycosidic linkage recognition. Thr203 was hypothesized to be one of the important residues for linkage specificity (Yamamoto *et al.*, 2004, 2011). α-(1→4)-Specific glucosidases generally have Thr or Ala at this position,
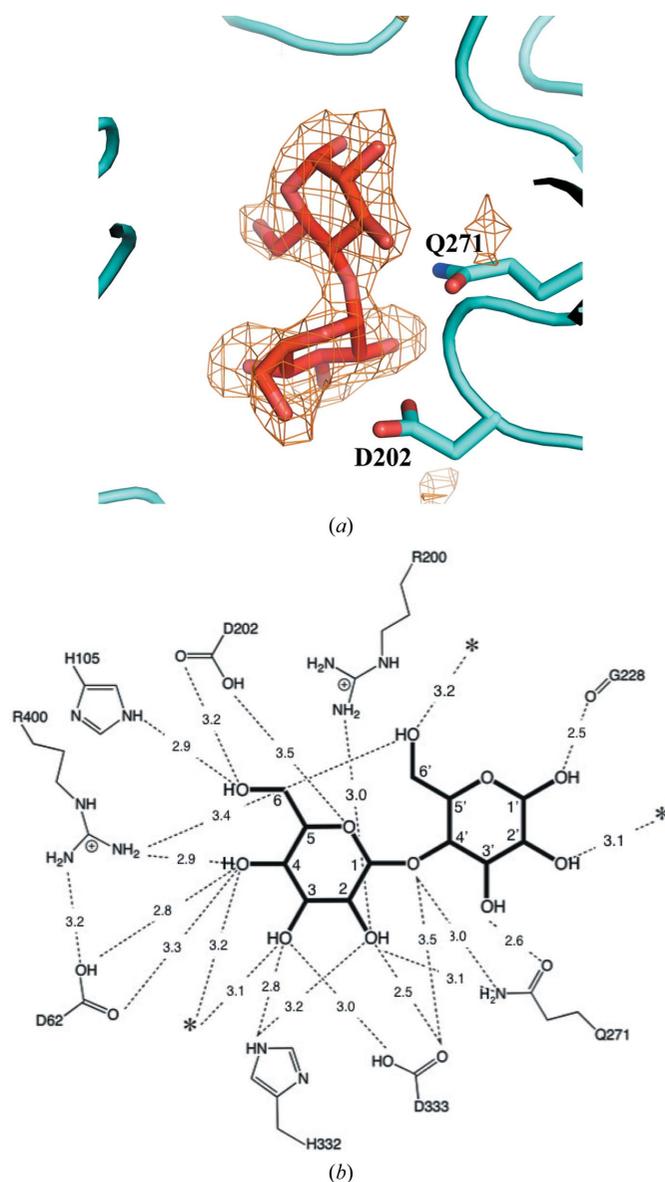
Figure 2
Close-up of the active site of the E271Q-Mal complex. (*a*) The $F_o − F_c$ OMIT map of maltose (contoured at 3.0 σ). (*b*) Contact between maltose and E271Q. Hydrogen bonds between specific atoms are indicated by dotted lines. Water molecules are indicated by asterisks.

whereas α-(1→6)-specific glucosidases have Val. The super-position of E271Q-Mal with isomaltase from *Saccharomyces cerevisiae* in complex with isomaltose (PDB entry 3axh; Yamamoto *et al.*, 2011) revealed that the reducing end of maltose cannot bind to the subsite +1 of isomaltase because of steric hindrance from Val216, whereas Thr203 of HaG leaves sufficient space for both maltose and isomaltose (Fig. 3*a*). Conversely, Phe297 of HaG can obstruct isomaltose from binding to HaG but is suitable for maltose binding. In isomaltase, the phenylalanine at the same position (Phe303) adopts a different conformation (Fig. 3*a*) from Phe297 in HaG, which forms the inner wall of subsite +1 through hydrophobic interactions with Phe206, Ile146 and Phe166. The O atom of another nonconserved residue, Gly228, which is located in the middle of the long β→α loop 4, interacts with O1′ of maltose at a distance of 2.5 Å (Fig. 2*b*). In dextran glucosidase, Lys275 and Glu371 form hydrogen bonds to the O2 and O3 atoms of

the glucose residue at subsite +1 (Hondoh *et al.*, 2008); in HaG, however, only Gly228 is positioned within hydrogen-bond interaction distance between substrate and enzyme at subsite +1, where it has an auxiliary function in substrate binding and recognizes only α-maltose. The superposition of E271Q-Mal and MutB–sucrose (PDB entry 2pwe; Ravaud *et al.*, 2007) showed that an apparent torsion of the glucose residues at subsite +1 between maltose and sucrose causes Gly228 to be closer to maltose than to sucrose (Fig. 3*b*), which may relate to selective activity towards maltose, because sucrose is a poor substrate for HaG. Phe297 may also lead to a low activity of HaG towards sucrose because of space limits (2.6 Å).

The β→α loop 4 of the catalytic domain A of α-glucosidase has been proposed to be responsible for the substrate specificity by reaction kinetics and molecular modelling (Noguchi *et al.*, 2003). In SmDG, a short β→α loop 4 was suggested to contribute to the high activity for long-chain substrates
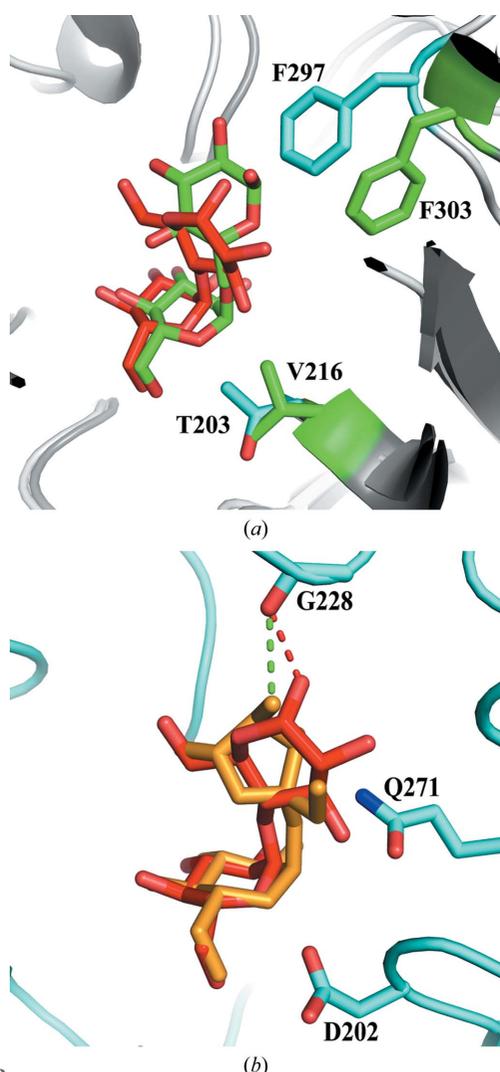
Figure 3
Superposition of the active sites of HaG structures. (*a*) E271Q-Mal superposed on isomaltase–isomaltose. (*b*) E271Q-Mal superposed on MutB–sucrose. The E271Q mutant is shown in light blue in both (*a*) and (*b*). Isomaltase is shown in green, whereas MutB is not shown. Maltose in (*a*) and (*b*), isomaltose in (*a*) and sucrose in (*b*) are shown in red, green and bright orange, respectively.
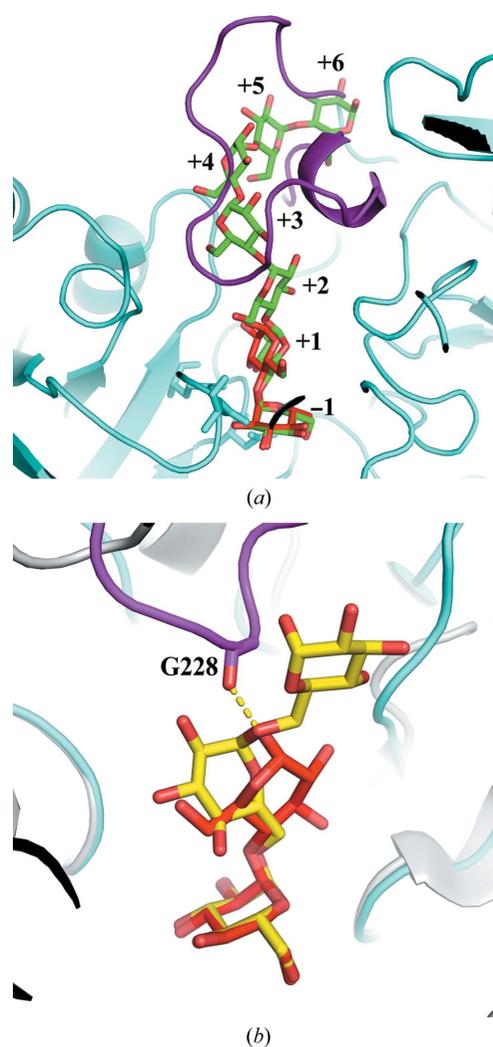
Figure 4
Long β→α loop 4 at the active site. (*a*) The model of the active site was built by superposing E271Q-Mal on amylosucrase from *N. polysaccharea* in complex with maltoheptaose. (*b*) The model of the active site was built by superposing E271Q-Mal on SmDG in complex with isomaltotriose. The long loops of the E271Q mutant, maltose, maltoheptaose and isomaltotriose are shown in purple, red, green and yellow, respectively.

(Saburi *et al.*, 2006). In contrast, HaG has a longer $\beta\rightarrow\alpha$ loop 4 than the other GH13 family exo-glucosidases. α-Glucosidase from *Xanthomonas campestris* WU-9701 (XcGT; GenBank accession No. BAC87873), which hydrolyzes almost solely disaccharides, as does HaG (Itou *et al.*, 2001), also has a similar length of putative $\beta\rightarrow\alpha$ loop 4 to that in HaG (Fig. 1*c*). The long $\beta\rightarrow\alpha$ loop 4 of HaG covers the majority of the entrance, obstructing the long-chain substrate (Figs. 4*a* and 5*a*). Steric hindrance can be observed directly on superposing E271Q-Mal with amylosucrase from *Neisseria polysaccharea* (Skov *et al.*, 2002; NpAS) in complex with maltoheptaose (PDB entry 1mw0). The long loop is very close to the glucose residue at subsite +2 and completely occupies subsites +3 and +4

(Fig. 4*a*). Similarly, the O6 atom of isomaltotriose is also very close (2.0 Å) to the O atom of Gly228 on the long loop when E271Q-Mal is superposed on SmDG in complex with isomaltotriose (PDB entry 2zid; Hondoh *et al.*, 2008; Fig. 4*b*). The disaccharide specificity of HaG is caused by the steric hindrance provided by the long $\beta\rightarrow\alpha$ loop 4. Furthermore, this long $\beta\rightarrow\alpha$ loop 4 may also be involved in acceptor specificity during the transglucosylation of HaG. HaG can add a glucosyl group to 6-gingerol, but only to the OH of the β-hydroxy keto group (the OH group on the alkyl side chain) rather than to the phenolic OH group (Ojima, Saburi *et al.*, 2012). In comparison to α-glucosidase GSJ from *Geobacillus* sp., which has a catalytic pocket that is more widely open, HaG only had a capacity for the entrance of the slender alkyl moiety of 6-gingerol into the narrow channel at the entrance of the catalytic pocket (Fig. 5*b*). In contrast, GSJ can utilize various acceptors, even rather bulky molecules such as curcumin (Fig. 5*c*; Shirai *et al.*, 2008).

3.4. Glucosyl-HaG intermediate structure

Crystals of HaG that diffracted to 1.47 Å resolution were grown in the presence of sucrose. The $F_o - F_c$ electron-density map clearly showed a glucosyl moiety trapped in subsite −1 by interacting with Asp202 (Fig. 6), suggesting that a reaction-intermediate state was captured in this complex structure. The reaction mechanism of α-glucosidase has been studied for more than 30 years (Chiba, 1997; Davies & Henrissat, 1995). The glucosyl-enzyme intermediate of GH13 α-glucosidase has been trapped by using its inactivator (McCarter & Withers, 1996), and structures of intermediates have been obtained by using general acid/base mutant enzymes with substrate analogues such as cyclodextrin glycosyltransferase from

(*a*)

(*b*)

(*c*)
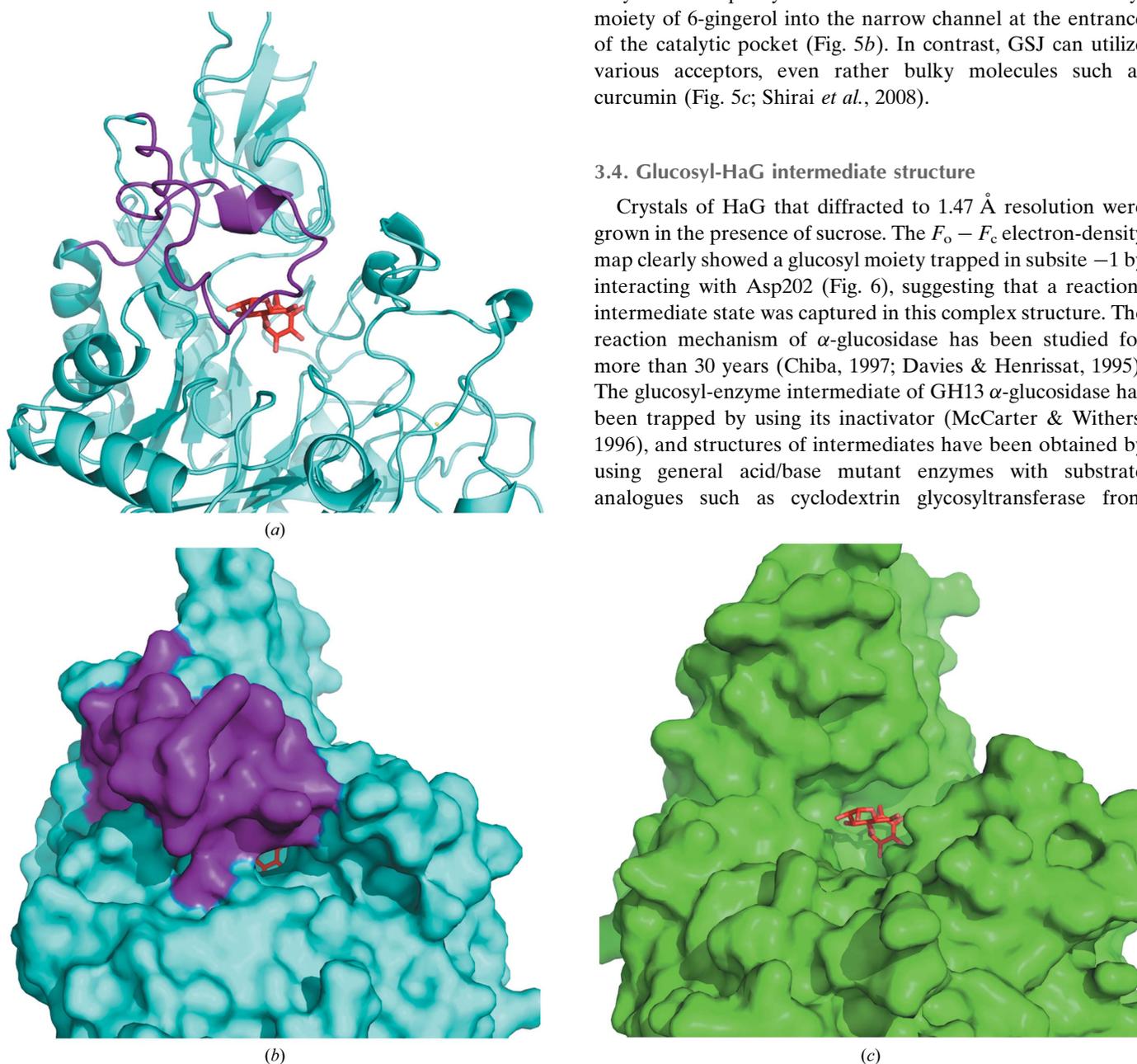
Figure 5
Long $\beta\rightarrow\alpha$ loop 4 of HaG in comparison with GSJ. (*a*) The $\beta\rightarrow\alpha$ loop 4 of HaG covers the active-site entrance. (*b*) Active-site pocket surface of HaG. (*c*) Active-site pocket surface of GSJ. The long loop is shown in purple and the substrate maltose is shown in red.

Bacillus circulans (*Ba*CGTase; PDB entry 1cxl; Uitdehaag *et al.*, 1999), NpAs (Jensen *et al.*, 2004; PDB entry 1s46) and SmDG (Kobayashi *et al.*, 2015; PDB entry 4wlc). In contrast to these structures obtained with inactive mutant enzymes and substrate analogues, the glucosyl-HaG structure strongly suggested the formation of a glucosyl-enzyme intermediate in its native form. We can speculate on the reasons why this intermediate was trapped in the crystal as the following: (i) the rate of deglycosylation is very slow without a monovalent cation (Ojima, Saburi *et al.*, 2012), (ii) sucrose is a poor substrate for HaG and (iii) the crystallization buffer does not contain monovalent cation at pH 7.5, which is higher than the optimum pH (6.5) for activity.

The glucosyl ring was tightly fixed in the same position with the nonreducing end at subsite −1 in E271Q-Mal and adopts a $^4C_1$ chair conformation (Fig. 6). No obvious conformational change was observed in comparison with HaG. However, this glucosyl-HaG structure showed some different properties compared with the other reported reaction models. As mentioned in §2.3, the distance between the C1 of the glucosyl residue and $O^{\delta 1}$ of the nucleophile Asp202 converged to 2.15 Å on refinement without restraints. Considering an ambiguous electron-density blob which is located close to the glucosyl residue at the position of fructosyl part of the sucrose (Fig. 6; it is impossible to fit any molecule in this blob), the glucosyl-HaG structure reflects a situation in which some of the HaG molecules are in a covalent bonded state and others are in a dissociated state (enzyme–substrate/product complex) in the crystal. Unlike the inactive acid/base mutants without activity, the glucosyl-HaG intermediate should be a representation of the real reaction process. The distance between the C1 of the glucosyl ring and $O^{\delta 1}$ of Asp202 is 3.1 Å in the E271Q-Mal structure, which is much longer than that in the glucosyl-HaG intermediate. This suggests that the C2—C1—O5—C5 torsion angle of the glucosyl ring is altered to enable the catalytic reaction, as was also encountered in the intermediate structure of *Ba*CGTase (Uitdehaag *et al.*, 1999). This torsion angle is −49.3° for the glucosyl ring at subsite −1 of maltose in E271Q-Mal and −64.7° for the glucosyl ring in the glucosyl-HaG intermediate. The torsion angle is flattened to allow C1 to approach closer to the catalytic Asp202.

3.5. D202N-Glc-Gol complex

Another structure of the inactive D202N mutant in complex with one glucose molecule and one glycerol molecule (D202N-Glc-Gol) allows us to propose the transglucosylation process for producing α-D-glucosylglycerol (αGG; Fig. 7a). The conformation of the protein in the D202N-Glc-Gol complex is nearly identical to that of HaG, with an r.m.s.d. of 0.14 Å for C^α atoms. The glucose occupies subsite −1, whereas the glycerol appears in subsite +1 (Fig. 7b). The distance between C1 of the glucose and $O^{\delta 1}$ of Asp202 is 3.2 Å, which is

Figure 6
Close-up of the active site of the glucosyl-HaG structure (stereoview). The $F_o − F_c$ OMIT map is contoured at 3.0σ.

(*a*) (*b*) (*c*)
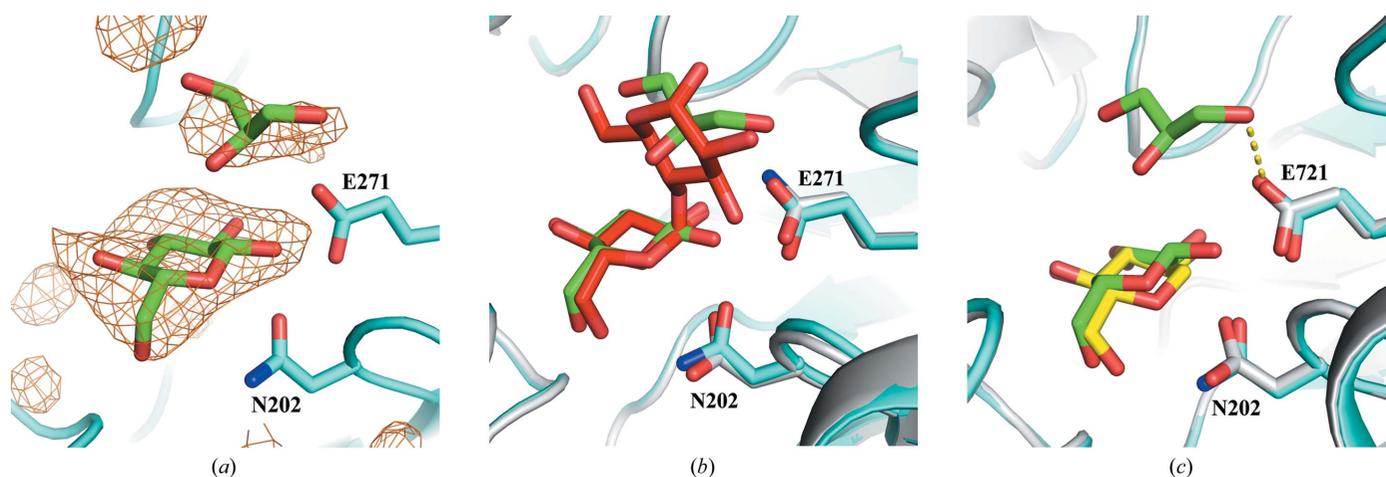
Figure 7
Structure of the active site in the D202N-Glc-Gol complex. (*a*) The OMIT map of glucose and glycerol (3.0σ). (*b*) Superposition of D202N-Glc-Gol and E271Q-Mal. (*c*) Superposition of D202N-Glc-Gol and glucosyl-HaG. D202N is shown in light blue, whereas E271Q and HaG are shown in grey, in both (*a*) and (*b*). Maltose and the glucosyl residue are shown in red and yellow, respectively. Glucose and glycerol are shown in green.
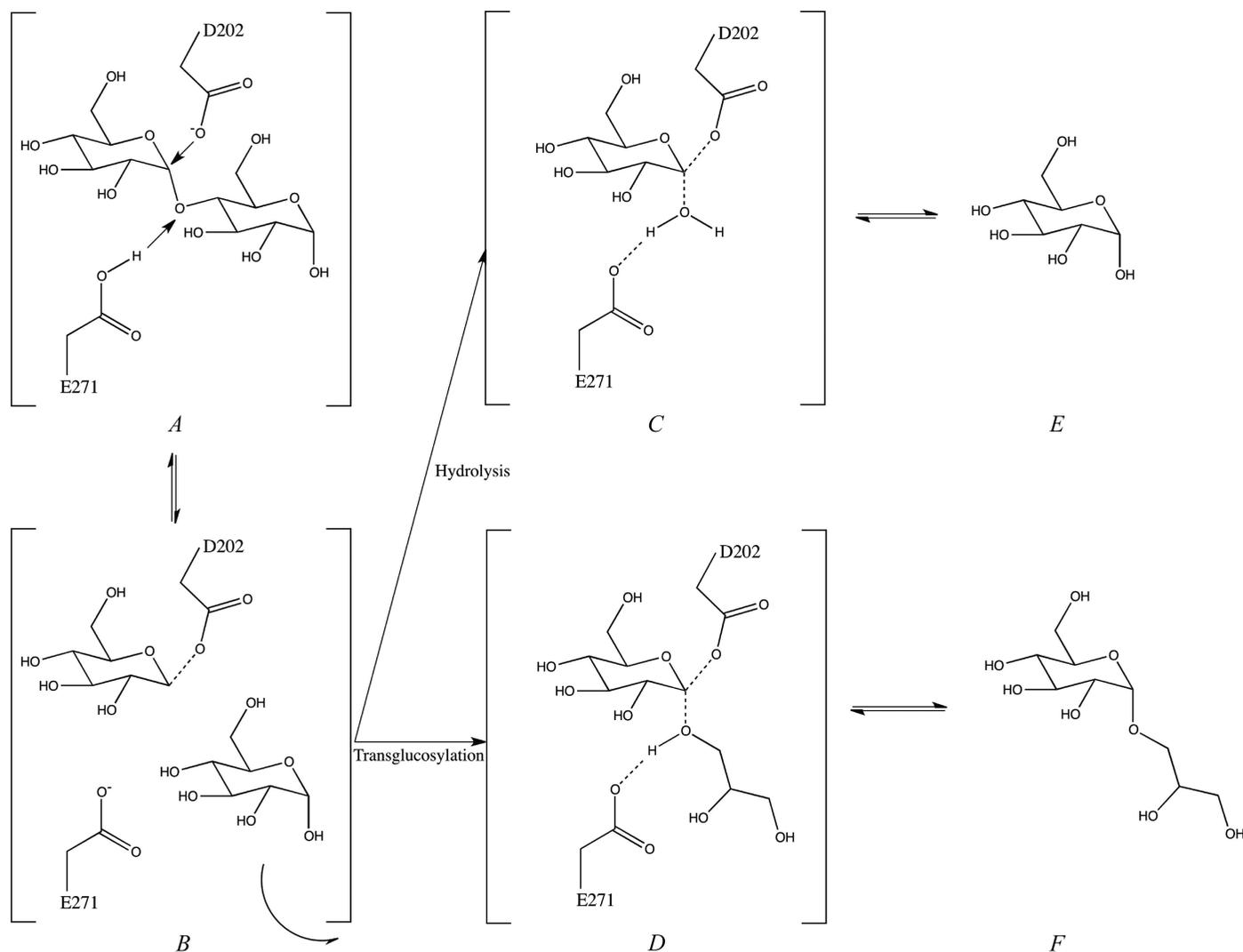
Figure 8
The reaction mechanism based on complex structures of HaG.

nearly equal to that in the E271Q-Mal structure. The C2—C1—O5—C5 torsion angle of the glucose in D202N-Glc-Gol is $-65.7°$, which is similar to that of the glucosyl ring in the glucosyl-HaG intermediate. However, a small rotation between the glucose in D202N-Glc-Gol and the glucosyl residue in the glucosyl-HaG intermediate can be observed, as the distance between the two C1 atoms is 0.9 Å (Fig. 7c). It is considered that the glucosyl ring would move away from the catalytic site after the hydrolytic reaction.

The glycerol molecules in the two chains of the dimer are in different positions, and only one glycerol molecule of the dimer interacts with the acid/base Glu271 at 3.2 Å (another is at 4.5 Å; Fig. 7c), suggesting that the glycerol moves freely within a certain range in the active-site pocket before transglucosylation. For the glycerol molecule that is close to Glu271, the 1-OH is 3.2 Å from $O^{\delta 2}$ of Glu271, closer than the 2-OH (4.3 Å). Thus, the proton of the 1-OH may be more easily accepted by Glu271 as a general base catalyst during transglucosylation. Thus, in the transglucosylation process

producing αGG, the 1-OH of glycerol is more glucosylated than the 2-OH (Ojima et al., 2012).

4. Conclusion

Structural analysis provides new insights for substrate specificity and the catalytic mechanism of HaG. Combining the findings of our previous studies, we propose an explanation of the reaction mechanism of HaG (steps A–F), as shown in Fig. 8. Based on structural comparisons of the apo form and E271Q-Mal with homologues, it is considered that in step A a long $\beta\rightarrow\alpha$ loop 4 is responsible for the strict recognition of disaccharides owing to steric hindrance at subsite +2 (Fig. 4a). In addition, the specificity of the glycosidic linkage at subsite +1 is determined by two residues, Thr203 and Phe297, and is also assisted by another residue, Gly228.

Furthermore, the glucosyl-HaG intermediate complex represents a mixed situation of the covalent-bonded state (step B) and dissociated state (steps C and E). In the process

of transglucosylation (steps D and F), glycerol can move freely within a certain range in the active-site pocket. The 1-OH of glycerol is more glucosylated than the 2-OH in the trans-glucosylation process to produce αGG because the proton of the 1-OH may be more easily accepted by Glu271.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.

Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., Pogson, C. I., Wilson, I. A., Corran, P. H., Furth, A. J., Milman, J. D., Offord, R. E., Priddle, J. D. & Waley, S. G. (1975). *Nature (London)*, **255**, 609–614.

Chiba, S. (1988). *Handbook of Amylases and Related Enzymes*, edited by The Amylase Research Society of Japan, pp. 104–105. Oxford: Pergamon Press.

Chiba, S. (1997). *Biosci. Biotechnol. Biochem.* **61**, 1233–1239.

Davies, G. & Henrissat, B. (1995). *Structure*, **3**, 853–859.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Gouet, P., Robert, X. & Courcelle, E. (2003). *Nucleic Acids Res.* **31**, 3320–3323.

Henrissat, B. & Davies, G. (1997). *Curr. Opin. Struct. Biol.* **7**, 637–644.

Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.

Hondoh, H., Saburi, W., Mori, H., Okuyama, M., Nakada, T., Matsuura, Y. & Kimura, A. (2008). *J. Mol. Biol.* **378**, 913–922.

Hung, V. S., Hatada, Y., Goda, S., Lu, J., Hidaka, Y., Li, Z., Akita, M., Ohta, Y., Watanabe, K., Matsui, H., Ito, S. & Horikoshi, K. (2005). *Appl. Microbiol. Biotechnol.* **68**, 757–765.

Itou, Y., Shimura, S., Tsugane, T., Yoshida, K., Nakagawa, H., Kirimura, K., Kino, K. & Usami, S. (2001). Japanese Patent 2001-46096.

Jensen, M. H., Mirza, O., Albenne, C., Remaud-Simeon, M., Monsan, P., Gajhede, M. & Skov, L. K. (2004). *Biochemistry*, **43**, 3104–3110.

Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.

Kelly, C. T. & Fogarty, W. (1983). *Process Biochem.* **18**, 6–12.

Kobayashi, M., Hondoh, H., Mori, H., Saburi, W., Okuyama, M. & Kimura, A. (2011). *Biosci. Biotechnol. Biochem.* **75**, 1557–1563.

Kobayashi, M., Saburi, W., Nakatsuka, D., Hondoh, H., Kato, K., Okuyama, M., Mori, H., Kimura, A. & Yao, M. (2015). *FEBS Lett.* **589**, 484–489.

Kurosu, J., Sato, T., Yoshida, K., Tsugane, T., Shimura, S., Kirimura, K., Kino, K. & Usami, S. (2002). *J. Biosci. Bioeng.* **93**, 328–330.

MacGregor, E. A., Janeček, Š. & Svensson, B. (2001). *Biochim. Biophys. Acta*, **1546**, 1–20.

McCarter, J. D. & Withers, S. G. (1996). *J. Biol. Chem.* **271**, 6889–6894.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Nakagawa, H., Dobashi, Y., Sato, T., Yoshida, K., Tsugane, T., Shimura, S., Kirimura, K., Kino, K. & Usami, S. (2000). *J. Biosci. Bioeng.* **89**, 138–144.

Nakao, M., Nakayama, T., Kakudo, A., Inohara, M., Harada, M., Omura, F. & Shibano, Y. (1994). *Eur. J. Biochem.* **220**, 293–300.

Noguchi, A., Nakayama, T., Hemmi, H. & Nishino, T. (2003). *Biochem. Biophys. Res. Commun.* **304**, 684–690.

Ojima, T., Aizawa, K., Saburi, W. & Yamamoto, T. (2012). *Carbohydr. Res.* **354**, 59–64.

Ojima, T., Saburi, W., Yamamoto, T. & Kudo, T. (2012). *Appl. Environ. Microbiol.* **78**, 1836–1845.

Ravaud, S., Robert, X., Watzlawick, H., Haser, R., Mattes, R. & Aghajari, N. (2007). *J. Biol. Chem.* **282**, 28126–28136.

Ravaud, S., Robert, X., Watzlawick, H., Haser, R., Mattes, R. & Aghajari, N. (2009). *FEBS Lett.* **583**, 1964–1968.

Saburi, W., Mori, H., Saito, S., Okuyama, M. & Kimura, A. (2006). *Biochim. Biophys. Acta*, **1764**, 688–698.

Sato, T., Nakagawa, H., Kurosu, J., Yoshida, K., Tsugane, T., Shimura, S., Kirimura, K., Kino, K. & Usami, S. (2000). *J. Biosci. Bioeng.* **90**, 625–630.

Shen, X., Saburi, W., Gai, Z.-Q., Komoda, K., Yu, J., Ojima-Kato, T., Kido, Y., Matsui, H., Mori, H. & Yao, M. (2014). *Acta Cryst.* F**70**, 464–466.

Shirai, T., Hung, V. S., Morinaka, K., Kobayashi, T. & Ito, S. (2008). *Proteins*, **73**, 126–133.

Skov, L. K., Mirza, O., Sprogøe, D., Dar, I., Remaud-Simeon, M., Albenne, C., Monsan, P. & Gajhede, M. (2002). *J. Biol. Chem.* **277**, 47741–47747.

Stam, M. R., Danchin, E. G., Rancurel, C., Coutinho, P. M. & Henrissat, B. (2006). *Protein Eng. Des. Sel.* **19**, 555–562.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.

Uitdehaag, J. C., Mosi, R., Kalk, K. H., van der Veen, B. A., Dijkhuizen, L., Withers, S. G. & Dijkstra, B. W. (1999). *Nature Struct. Biol.* **6**, 432–436.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

Yamamoto, K., Miyake, H., Kusunoki, M. & Osaki, S. (2011). *J. Biosci. Bioeng.* **112**, 545–550.

Yamamoto, K., Nakayama, A., Yamamoto, Y. & Tabata, S. (2004). *Eur. J. Biochem.* **271**, 3414–3420.

Zhang, D., Li, N., Lok, S.-M., Zhang, L.-H. & Swaminathan, K. (2003). *J. Biol. Chem.* **278**, 35428–35434.