Acta Crystallographica Section D
Biological Crystallography
ISSN 1399-0047

Structural features underlying the selective cleavage of a novel exo-type maltose-forming amylase from *Pyrococcus* sp. ST04

Kwang-Hyun Park,[a,b]‡ Jong-Hyun Jung,[c,d]‡ Sung-Goo Park,[a] Myeong-Eun Lee,[c] James F. Holden,[e] Cheon-Seok Park[c]* and Eui-Jeon Woo[a,b]*

[a]Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon 305-806, Republic of Korea, [b]Department of Bio-analytical Science, University of Science and Technology, Daejeon 305-333, Republic of Korea, [c]Graduate School of Biotechnology and Institute of Life Science and Resources, Kyung Hee University, Yongin 446-701, Republic of Korea, [d]Research Division for Biotechnology, Korea Atomic Energy Research Institute, Jeongeup 580-185, Republic of Korea, and [e]Department of Microbiology, University of Massachusetts, Amherst, MA 01003, USA

‡ These authors contributed equally to this work.

Correspondence e-mail: cspark@khu.ac.kr, ejwoo@kribb.re.kr

Received 27 January 2014
Accepted 25 March 2014

PDB reference: maltose-forming α-amylase, 4cmr

A novel maltose-forming α-amylase (PSMA) was recently found in the hyperthermophilic archaeon *Pyrococcus* sp. ST04. This enzyme shows <13% amino-acid sequence identity to other known α-amylases and displays a unique enzymatic property in that it hydrolyzes both α-1,4-glucosidic and α-1,6-glucosidic linkages of substrates, recognizing only maltose units, in an exo-type manner. Here, the crystal structure of PSMA at a resolution of 1.8 Å is reported, showing a tight ring-shaped tetramer with monomers composed of two domains: an N-domain (amino acids 1–341) with a typical GH57 family $(\beta/\alpha)_7$-barrel fold and a C-domain (amino acids 342–597) composed of α-helical bundles. A small closed cavity observed in proximity to the catalytic residues Glu153 and Asp253 at the domain interface has the appropriate volume and geometry to bind a maltose unit, accounting for the selective exo-type maltose hydrolysis of the enzyme. A narrow gate at the putative subsite +1 formed by residue Phe218 and Phe452 is essential for specific cleavage of glucosidic bonds. The closed cavity at the active site is connected to a short substrate-binding channel that extends to the central hole of the tetramer, exhibiting a geometry that is significantly different from classical maltogenic amylases or β-amylases. The structural features of this novel exo-type maltose-forming α-amylase provide a molecular basis for its unique enzymatic characteristics and for its potential use in industrial applications and protein engineering.

1. Introduction

Starch, glycogen and related carbohydrates play an essential role in the metabolism of animals, plants and microorganisms. Amylolytic enzymes produced by microorganisms are structurally diverse and are extensively distributed among strains and species. They form a large group of enzymes, including α-amylases (EC 3.2.1.1), β-amylases (EC 3.2.1.2), glucoamylases (EC 3.2.1.3), pullulanases (EC 3.2.1.41) and amylopullulanases (EC 3.2.1.–). The anomer-retaining glycoside hydrolases that act on the α-1,4-glucosidic or α-1,6-glucosidic linkage of starch or amylose are mainly categorized into the following three families: glycoside hydrolase (GH) families 13, 57 and 77 (Imamura *et al.*, 2001; Cantarel *et al.*, 2009). Most α-amylases are assigned to the GH13 family, which are multidomain $(\alpha/\beta)_8$-barrel enzymes that exhibit large variation in the structural motif with an array of substrate specificities (MacGregor *et al.*, 2001). GH77 is closely related to GH13 in both amino-acid sequence and three-dimensional structure (Przylas *et al.*, 2000). The GH57 family is known to be essential for carbohydrate utilization (Xavier *et al.*, 1999).

© 2014 International Union of Crystallography

This family includes α-amylases (EC 3.2.1.1), amylopullulanases (EC 3.2.1.1/41), branching enzymes (EC 2.4.1.18), 4-α-glucanotransferases (EC 2.4.1.25) and α-galactosidases (EC 3.2.1.22), and is mainly composed of thermostable enzymes from extremophiles (Janeček & Blesák, 2011). Most industrial starch processes require high temperatures for liquefaction and saccharification, prompting a search for new sources of thermostable amylolytic enzymes (Niehaus *et al.*, 1999).

Recently, we reported a novel exo-type maltose-forming α-amylase from *Pyrococcus* sp. ST04 (Py04_0872; PSMA; Jung *et al.*, 2014). *Pyrococcus* sp. ST04 is a hyperthermophilic, anaerobic and heterotrophic archaeon that was isolated from a deep-sea hydrothermal sulfide chimney (Ver Eecke *et al.*, 2009). Based on a genomic analysis, the amylolytic enzymes in *Pyrococcus* are composed of GH13 and GH57 family members. Although *Pyrococcus* sp. ST04 possesses only GH57 amylases, it can efficiently utilize starch and its degradation products, such as maltose, as primary carbon sources (Oslowski *et al.*, 2011). PSMA belongs to the GH57 family and contains five conserved sequence motifs (Zona *et al.*, 2004) (Supplementary Fig. S1[1]). While more than 1000 genes from bacteria and archaea have been identified as encoding GH57 enzymes, only 18 of these proteins have been characterized and the structural information available for the GH57 family is still limited. To date, the following four structures have been determined: AmyC from *Thermotoga maritima* (PDB entry 2b5d; Dickmanns *et al.*, 2006), branching enzymes from *Thermus thermophilus* (PDB entry 1ufa; Palomo *et al.*, 2011) and *Thermococcus kodakarensis* KOD1 (PDB entry 3n8t; Santos *et al.*, 2011) and 4-α-glucanotransferase from *Thermococcus litoralis* (PDB entry 1k1y; Imamura *et al.*, 2003). These structures showed that the GH57 amylase commonly consists of two domains: the N-domain and the C-domain.

PSMA shares less than 13% amino-acid sequence identity with other known GH enzymes. PSMA hydrolyzes maltose units from short oligosaccharides as well as from starch and glycogen, exhibiting maltose-producing amylase-like activity. To date, there are two representative maltose-producing amylases: β-amylase and maltogenic amylase. PSMA is distinct from these two amylases in many aspects. β-Amylase is a GH14 family enzyme that only catalyzes the cleavage of α-1,4-glucosidic bonds of polysaccharides and liberates maltose from the nonreducing ends of starch, amylose and maltodextrins *via* an inverting mechanism (Mikami *et al.*, 1999; Cantarel *et al.*, 2009). Many GH13 family maltogenic amylases produce maltose from the reducing end of the substrates and cleave cyclodextrin substrates (Kim *et al.*, 1999), whereas PSMA removes maltose from the nonreducing end of substrates and cannot cleave cyclodextrin. PSMA primarily hydrolyzes short substrates, such as the hydrolysis of maltotriose to maltose, as well as long substrates such as amylose and amylopectin. PSMA is an exo-type glucan hydrolase with α-1,4-glucan and α-1,6-glucan hydrolytic activities. It is inhibited competitively by the reaction product maltose. Owing to its unique features such as its substrate specificity for short oligosaccharides, its hydrolysis pattern of dual glycosyl linkages, its hyperthermostability and its optimal activity at low pH, PSMA has considerable potential for industrial application (Jung *et al.*, 2014). To date, maltose production by maltogenic amylase and β-amylase has been limited in the starch-processing industry because of the α-1,6-linkage of the substrate, which results in the formation of β-limited dextrin. *In vivo*, PSMA appears to bridge the gap between the amylases that hydrolyze larger maltodextrins and α-glucosidase, which feeds maltose into glycolysis by hydrolyzing small glucans into maltose units. PSMA is the only enzyme with a debranching function in the cytoplasm of *Pyrococcus* sp. ST04 (Jung *et al.*, 2012). Despite its novel properties, its mechanism of recognition of only maltose units and hydrolysis of both glucosidic bonds has not been investigated. Here, we describe the first crystal structure of this unique α-amylase and show the structural basis of its enzymatic function in terms of substrate selection and binding.

2. Materials and methods

2.1. Cloning and site-directed mutagenesis

The cloning of the ORF (Py04_0872) for the maltose-forming α-amylase (PSMA) from *Pyrococcus* sp. ST04 has been described previously (Jung *et al.*, 2014). The E153Q, D253N, F218A, F452A, W453A and E580Q mutants were generated with the QuikChange site-directed mutagenesis kit (Agilent Technologies, Santa Clara, California, USA). The mutations were verified by sequencing the respective genes (Macrogen).

2.2. Expression of PSMA

2.2.1. Wild-type and mutant PSMA. Recombinant PSMA and its mutants were expressed in the *Escherichia coli* BL21-CodonPlus(DE3)-RP strain. Expression was induced by adding 0.5 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) when the optical density (600 nm) of the culture reached 0.55. To obtain sufficient expression, induction was performed at 310 K for 24 h. After induction, the cells were harvested by centrifugation (10 000*g*, 20 min, 277 K) and resuspended in lysis buffer (20 m*M* sodium phosphate, 500 m*M* NaCl, 20 m*M* imidazole pH 7.0).

2.2.2. Selenomethionine labelling of PSMA. To express selenomethionine-labelled wild-type PSMA, pET-PSMA was transformed into the methionine-auxotrophic *E. coli* strain B834(DE3) with the pRARE plasmid for efficient expression of archaeal protein. The transformant was grown on M9 mineral medium supplemented with 34 μg ml^{-1} chloramphenicol, 100 μg ml^{-1} ampicillin and 0.4% glucose as the carbon source.

2.3. Purification and activity assay

The purification of PSMA and its activity assay were performed as described previously (Jung *et al.*, 2014). For the activity assay, maltotriose and G2-β-cyclodextrin were used as

[1] Supporting information has been deposited in the IUCr electronic archive (Reference: MH5131).

substrates to determine the α-1,4-bond and α-1,6-bond hydrolysis activities of PSMA, respectively. The reaction was performed at 358 K for 5 min using 10 m*M* maltotriose as a substrate with or without 1 m*M* of each inhibitor. To analyze the hydrolysis products, thin-layer chromatography (TLC) was used. 10 m*M* of each substrate was reacted with 2.8 µg enzyme at 353 K for 10 h and the reaction samples were spotted onto the plate and developed with a solvent system consisting of 2-propanol/ethyl acetate/water [3:1:1(*v*:*v*:*v*)]. The TLC plate was dried and then visualized by soaking in 0.3%(*w*/*v*) *N*-(1-naphthyl)-ethylenediamine and 5%(*v*/*v*) H_2SO_4 in methanol.

2.4. Crystallization and data collection

PSMA crystallization trials were conducted using the sitting-drop method at 291 K. We mixed 1 µl 15 mg ml^{-1} PSMA solution with an equal volume of reservoir solution consisting of 1.0 *M* ammonium citrate dibasic, 0.8 *M* sodium acetate trihydrate pH 4.6. Before the data were collected, the crystals were cryocooled to 95 K using a cryoprotectant consisting of the mother liquor supplemented with 25% glycerol. The PSMA data set was collected on the MX7A synchrotron beamline at the Pohang Accelerator Laboratory, Pohang, Republic of Korea. The crystals diffracted to a resolution of 1.8 Å and data were collected over 365° of crystal rotation at 1° intervals.

2.5. Structure determination and refinement

Diffraction data were processed and scaled using *HKL-2000*. The structure was determined using the single-wavelength anomalous dispersion method with *phenix.autosol* (Adams *et al.*, 2010). The resulting model was refined through model rebuilding using *PHENIX*. *Coot* was used for stereographic manual refinement and model building (Emsley *et al.*, 2010). The structure was validated with *PROCHECK*. Molecular images, including cartoon and stick figures, were produced using *PyMOL* (http://www.pymol.org). The detailed statistics of data collection and refinement are listed in Table 1.

2.6. Molecular modelling

Isopanose ligands were modelled using 4-α-glucanotransferase complexed with acarbose (PDB entry 1k1y) from *T. litoralis* as the template. The ligand-complex models were initially obtained by superposition of PSMA onto 4-α-glucanotransferase. The initial complex models were subjected to energy minimization followed by 1 ps of molecular dynamics at 3008 K after equilibration. They were finally minimized to a maximum derivative of 1.0 kcal per step using the *DISCOVER* module in the *Insight II* program (Accelrys, San Diego, California, USA) with Amber force fields (Supplementary Fig. S2).

Table 1
Data-collection and structure-solution parameters.

Values in parentheses are for the highest resolution shell.

Data statistics	
Crystal type	Selenomethionine-labelled
Source	MX7A, Pohang Accelerator Laboratory
Wavelength	1.0000
Space group	$C222_1$
Unit-cell parameters (Å, °)	a = 117.639, b = 370.794, c = 82.135, α = 90, β = 90, γ = 90
Unique reflections	13716
$R_{\rm merge}$† (%)	6.9 (49.8)
Completeness (%)	99.1 (97.4)
Average $I/\sigma(I)$	52.6 (6.7)
Refinement statistics	
Resolution range (Å)	33.18–1.80
R factor‡ (%)	19.81
$R_{\rm free}$§ (%)	22.45
R.m.s.d., bond lengths (Å)	0.008
R.m.s.d., angles (°)	1.188
Average B value (Å^2)	32.2
No. of atoms	10845
No. of water atoms	946
Ramachandran plot, residues in (%)	
Most favoured region	98.15
Additionally allowed region	1.85
Outlier region	0
PDB code	4cmr

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the ith observation of reflection hkl and $\langle I(hkl)\rangle$ is the average over symmetry equivalents. ‡ R factor = $\sum_{hkl} ||F_{\rm obs}| - |F_{\rm calc}|| / \sum_{hkl} |F_{\rm obs}|$, where $F_{\rm obs}$ and $F_{\rm calc}$ are the observed and calculated structure-factor amplitudes, respectively. § $R_{\rm free}$ was calculated using 5% of the data.

3. Results and discussion

3.1. Structure determination

The PSMA crystals belonged to the orthorhombic space group $C222_1$. The Matthews coefficient ($V_{\rm M}$ = 3.18 Å^3 Da^{-1}) and the corresponding solvent content of 61.28% suggested the presence of two molecules in the asymmetric unit. The crystal structure of PSMA was determined by the single-wavelength anomalous dispersion method using selenomethionine-containing PSMA. The initial electron-density map obtained from the experimental phases after solvent flattening was of high quality and allowed the auto-tracing of 334 of the 597 amino acids present in the polypeptide chain. After several rounds of refinement, the electron-density map for the remaining region became traceable. The final model of PSMA at a 1.8 Å resolution encompassed all N- and C-terminal residues. The asymmetric unit in the crystal contained two molecules that were almost identical (r.m.s.d. of 0.207 Å between 591 C$^\alpha$ atoms). The structure was refined to a free R value of 22.45% with good stereochemistry. The detailed statistics for the structure are given in Table 1.

3.2. Overall structure of PSMA

The overall structure of PSMA can be separated into two major parts: an N-domain (residues 1–341) and a C-domain (residues 342–597) composed of a total of 26 α-helices and 11 β-strands that account for 48 and 11% of the total residues, respectively (Fig. 1*a*). The PSMA monomer has approximate dimensions of 50 × 50 × 75 Å. Examination of the N-domain

showed that it contains the conserved domains found in typical GH57 α-amylases. The core of the N-domain is a $(\beta/\alpha)_7$-barrel in which seven central parallel β-strands ($\beta1$, $\beta4$, $\beta5$, $\beta6$, $\beta7$, $\beta8$ and $\beta9$) are flanked by helices. An elongated loop (residues 16–52) of the N-domain interacts with and stabilizes the C-domain primarily by hydrophobic interactions (Fig. 1*b*). The C-domain is composed of ten α-helices and two short β-strands, showing significant variation in structure from other GH57 α-amylases. Resembling an elaborated four-helix bundle, the C-domain is stabilized by a hydrophobic core and is kept close to the $(\beta/\alpha)_7$-barrel by a number of polar interactions. The two domains assemble to form a groove in the active site. No metal ions could be detected in the electron-density map. Although the PSMA gene encodes a 70 191 Da polypeptide, purified PSMA was eluted as an approximately 260 kDa protein in gel-filtration chromatography, indicating that PSMA is a homotetramer, and the determined structure confirms that the tetrameric structure is in a ring formation with a dimer-of-dimers arrangement. In the dimeric structure, extensive interactions including Asn185–Gly327, Glu224–Lys241 and Gly327–Asn185 hydrogen bonds and a Glu224–Lys241 salt bridge between two $(\beta/\alpha)_7$-barrels of the N-domains comprise the dimer associations. Each monomer in a dimer is oriented by bilateral symmetry, with the active site of each monomer facing the other side. Further hydrophobic interactions between the two dimers at the C-domain contribute to the formation of the tetrameric ring. The PSMA tetramer measured 104 × 75 Å in width and 59 Å in height, with an inner diameter of the ring of 26 Å. In the tetrameric arrangement there is a buried interface of 3169 Å^2, corresponding to 13% of the solvent-accessible surface per monomer (Fig. 1*c*). The effect of oligomerization on the PSMA activity is still unknown. A homologue from *Thermococcus* sp. CL1 that shares 55% sequence similarity with

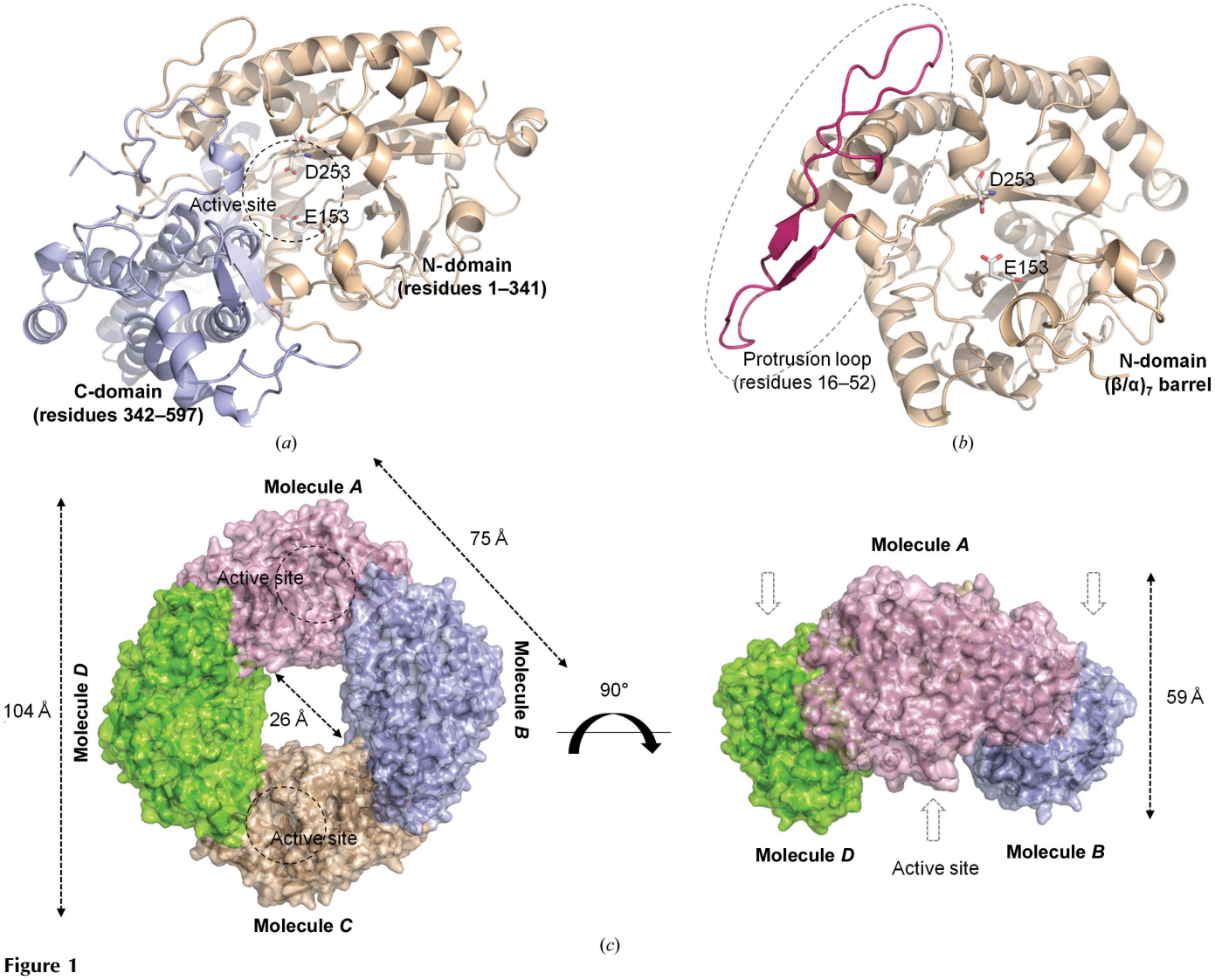


Figure 1
Overall structure of PSMA. (*a*) Overall structure of the PSMA monomer (N-domain, wheat; C-domain, blue). The two catalytic residues (Asp253 and Glu153) that are strictly conserved in the GH57 family α-amylases are shown in stick representation. (*b*) The $(\beta/\alpha)_7$-barrel of the N-domain with a long protruding loop (residue 16–52) interacts with and stabilizes the C-domain. (*c*) Top-down (left) and side views (right) of the PSMA tetrameric ring structure. The ring dimensions are marked. Monomers *A*–*D* are coloured differently. Arrows indicate the position of the active site of each molecule.

PSMA and exists in a monomeric form exhibited a similar cleavage activity towards both α-1,4-glucosidic and α-1,6-glucosidic bonds, suggesting possible activity in a monomeric form (Jeon *et al.*, 2014).

3.3. Structural comparison with GH57 and maltose-producing amylases

Four GH57-family protein structures were available, including those of an 4-α-glucanotransferase, branching enzymes and AmyC, an α-amylase from *T. maritima*. All of these proteins are endo-acting enzymes that randomly catalyze the cleavage of α-1,4-glucosidic bonds. *DALI* database analysis identified the 4-α-glucanotransferase from *T. litoralis* (PDB entry 1k1y) as the closest homologue of PSMA, with a *Z*-score of 26.6 (3.4 Å r.m.s.d. over 370 C^{α} atoms). Although the PSMA N-domain exhibits a similar structure to the N-domain of 4-α-glucanotransferase after superposition (6.9 Å r.m.s.d. over 217 C^{α} atoms), the C-domain of PSMA is composed entirely of α-helices with the exception of two short β-strands, whereas the C-domain of 4-α-glucanotransferase is entirely composed of long β-strands. Most glucanotransferase enzymes contain a structural lid architecture located close to the active site that stabilizes the

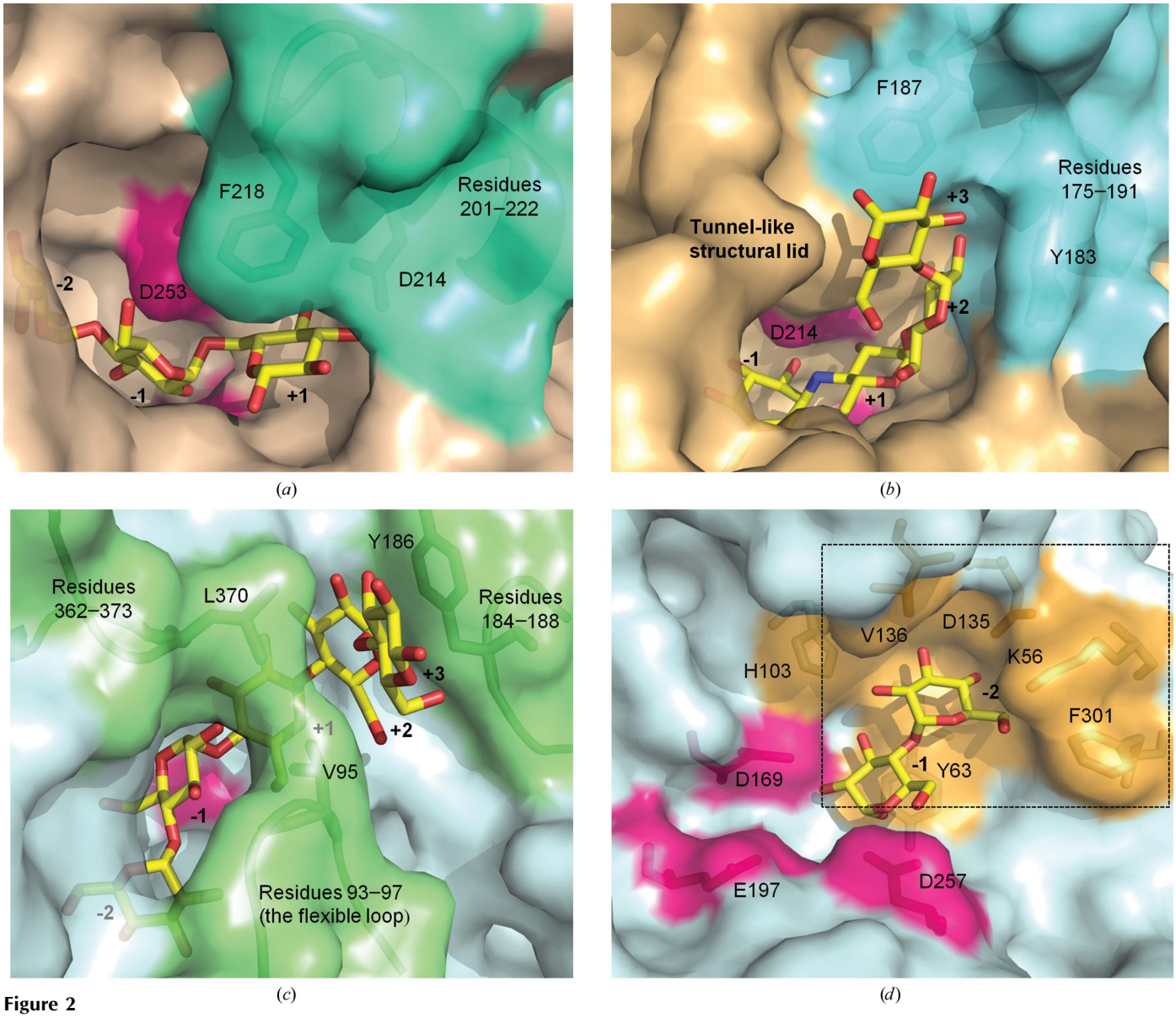


Figure 2
Comparison of the substrate-binding groove. (*a*) Surface diagram of the substrate-binding groove of PSMA bound with a hypothetical model of isopanose with an α-1,6-linkage between subsites −1 and +1. The region of the helix–loop (residues 201–222) forming the barrier at the entrance of the active site is coloured green. (*b*) Surface diagram of the GH57 family 4-α-glucanotransferase (PDB entry 1k1y) bound to acarbose. The corresponding barrier at the entrance to the active site in PSMA is coloured cyan. (*c*) Surface diagram of β-amylase in complex with the α-1,4-linked maltopentaose substrate (PDB entry 1itc; Miyake *et al.*, 2003). Flexible loops at the active site are coloured light green. (*d*) Surface representations of AmyB with maltose modelled at subsites −1 to −2 (PDB entry 4gkl; Jun *et al.*, 2013). The extra sugar-binding site that recognizes the nonreducing end of the substrate is highlighted in orange. Catalytic residues are shown in pink.

acceptor molecule. The PSMA active site lacks the tunnel-like structural lid found in 4-α-glucanotransferase, possibly explaining the absence of transferase activity in PSMA (Figs. 2*a* and 2*b*).

β-Amylase, maltogenic amylase and AmyB are well known enzymes that produce maltose as their main product. Maltogenic amylase and AmyB have been classified as GH13 family members, while β-amylase is assigned to the GH14 family. The crystal structures of these maltose-producing amylases showed that although the proteins have a common $(\beta/\alpha)_8$-barrel scaffold with an invariant active site located on top of the barrel, they have distinct and significantly different active-site cavity features. Maltogenic amylase is an endo-type amylase with an additional N-terminal domain with a narrow and deep groove at the active site that allows it to hydrolyze cyclodextrin and cleave a maltose unit from the reducing end of substrates (Kim *et al.*, 1999). The active sites of most maltogenic amylases are formed by dimerization and are located at the interface between the core $(\beta/\alpha)_8$ domain and the C-domain from one subunit and the N-domain from another subunit. However, the active site of PSMA is located at the domain interface within a monomer. β-Amylase has three long flexible loops that form a deep active-site cleft, and the conformation of the flexible loops determines the binding of different substrate such as cyclodextrin and maltose (Mikami *et al.*, 1993; Miyake *et al.*, 2003; Fig. 2*c*). This loop is not observed in PSMA, and *B*-factor analysis does not indicate a significant conformational change of the loops around the active site. The active site of AmyB shows a wide cavity that includes a putative extra sugar-binding site that recognizes the nonreducing end of the substrate (Jun *et al.*, 2013; Fig. 2*d*). The topology of the active site located deep in the pocket of the $(\beta/\alpha)_8$-barrel is in common, but the extra sugar-binding site is not observed in PSMA. From comparison with other classical maltose-forming α-amylases, PSMA belongs to the GH57 family yet appears to recognize the maltose unit in a significantly different mode from other family members.

3.4. Catalytic residues

Multiple sequence alignment and structural analysis indicated that the conserved residues Glu153 and Asp253 are catalytic residues located at the centre of the $(\beta/\alpha)_7$-barrel (Janeček & Kuchtová, 2012; Supplementary Fig. S1). The former residue may act as a nucleophile and the latter as an acid/base catalyst. To investigate the catalytic function of this enzyme, we mutated glutamate to glutamine (E153Q) and aspartate to asparagine (D253N). The D253N mutant and the E153Q mutant showed neither α-1,4- nor α-1,6-glucosidic bond hydrolysis activity, supporting an essential role for these residues. They are located in a deep pocket at the central axis of the $(\beta/\alpha)_7$-barrel domain. Residues Glu153 and Asp253, located in conserved motifs III and IV, respectively,

superposed well with the corresponding residues in other GH57 family proteins and display the same rotamer conformation. The geometry and distance between Glu153 C^{α} and Asp253 C^{α} is 10.35 Å, which is similar distance to that in 4-α-glucanotransferase (10.7 Å), branching enzyme (10.3 Å) and AmyC (10.6 Å) in the GH57 family, suggesting an α-retaining mechanism for the catalytic activity of PSMA (Imamura *et al.*, 2001; Palomo *et al.*, 2011; Fig. 3*a*). We predicted the pK_a value of the nucleophile Glu153 in the active site using the *PROPKA* program (Rostkowski *et al.*, 2011). The prediction showed the pK_a value of the catalytic residue to be 5.17, which is consistent with the biochemical pH profile and is comparable to the pK_a value of industrial β-amylase (5.4).

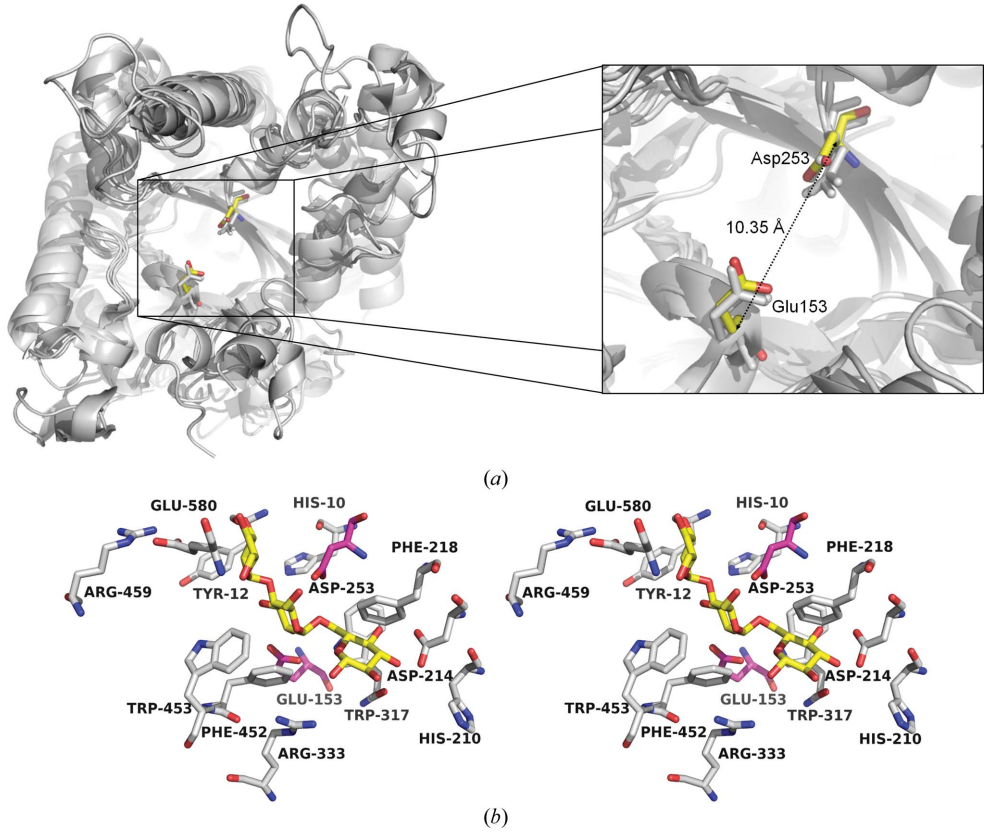


Figure 3
Active site of PSMA. (*a*) Superposition of the $(\beta/\alpha)_7$-barrel domain with three GH57 α-amylase structures (grey; PDB entries 2b5d, 1k1y and 1ufa). The active sites in the $(\beta/\alpha)_7$ domains are enlarged in the black box and the catalytic residues (yellow) are indicated. (*b*) Stereo representation of the active site with catalytic residues (pink) and other key residues (white) shown as sticks.

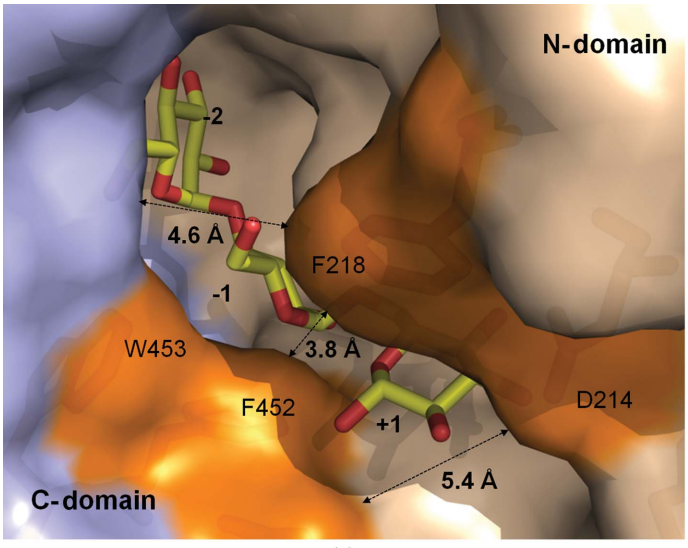
3.5. Small enclosed cavity in the active site

The crystal structure revealed a unique enclosed cavity around the active site (Fig. 4*a*). This enclosed cavity is composed of three major loops (residues 10–14, 114–116 and 254–256) from the $(\beta/\alpha)_7$-barrel of the N-domain and two α-helices (residue 449–453 and 577–583) from the C-domain. A long loop (residue 316–326) in the N-domain is juxtaposed with the C-domain, forming the entrance to the active site, and C-terminal residues 577–583 line the surface of the cavity around the putative subsite −2. The pocket is approximately 11 Å in length and 500 Å^3 in volume and accommodates two α-1,4-linked sugar units (Fig. 4*b*). Analysis of the electrostatic distribution reveals a negative surface charge inside the pocket, and the charge distribution is stretched along the domain interface to the catalytic residues. To investigate the substrate-binding mode, we modelled a substrate in the active site of PSMA. An isopanose molecule consisting of three glucose units with α-1,4- and α-1,6 linkages was modelled in the active site of PSMA based on the complex structure of 4-α-glucanotransferase and acarbose (PDB entry 1k1y; Supplementary Fig. S2). In the substrate-docking model, isopanose

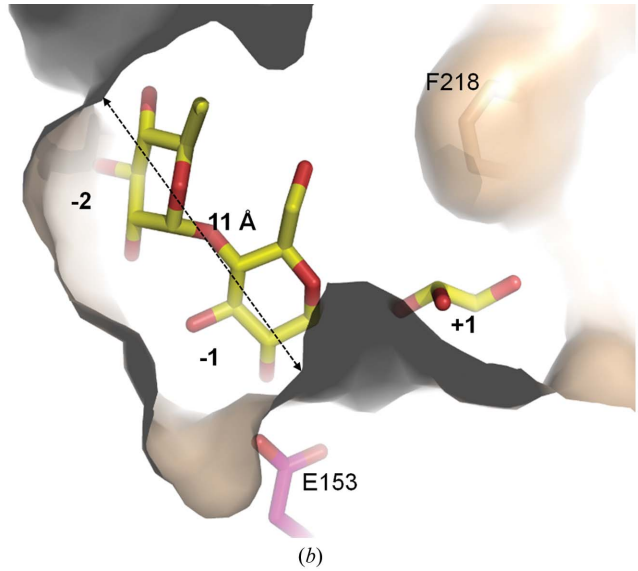
Table 2
The specific activity of wild-type PSMA and mutants.

N.D., not detected.

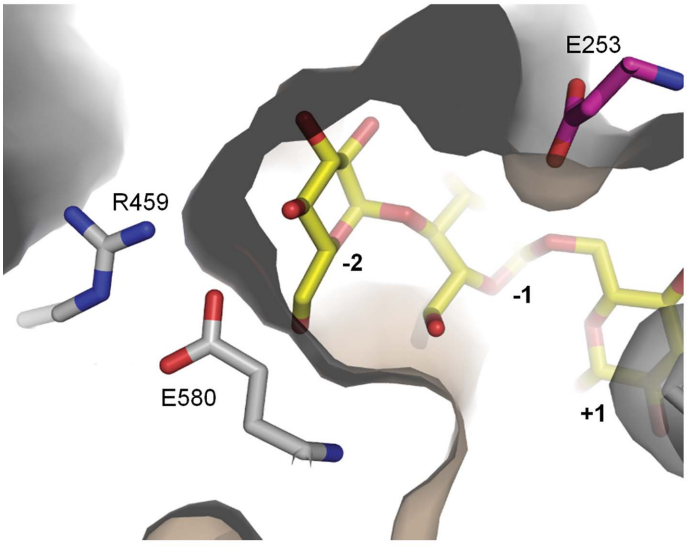
Enzyme	α-1,4-Glucosidic bond hydrolysis (μmol min^{-1} mg^{-1})	α-1,6-Glucosidic bond hydrolysis (μmol min^{-1} mg^{-1})
Wild type	3.93 ± 0.33	15.06 ± 1.27
F218A	14.00 ± 0.67	14.57 ± 1.53
F452A	N.D.	<0.02
W453A	0.04 ± 0.01	1.08 ± 0.07
E580Q	0.25 ± 0.01	2.32 ± 0.05



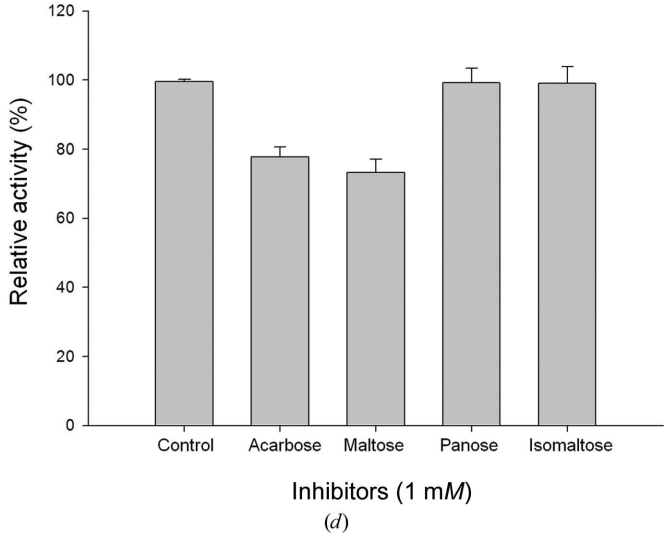
(*a*)



(*b*)



(*c*)



(*d*)

Figure 4
Small enclosed cavity in the active site. (*a*) Residues at the gate of the cavity and the width of the channel are indicated. (*b*) Surface representation of the small enclosed cavity at the active site with the hypothetical isopanose (yellow). Putative subsite numbers are indicated. (*c*) Glu580 and Asp459, located at the nonreducing end of the isopanose model, are shown. (*d*) Inhibition of α-1,4-hydrolysis activity by acarbose, maltose, panose and isomaltose.

fits in the active-site cleft with the nonreducing end of the isopanose placed inside the pocket and bound in the putative subsites −2, −1 and +1. Hydrophobic aromatic residues, such as Phe218, Trp317, Phe452 and Trp453, are located in the active site at the entrance of the channel and play a role in substrate recognition and binding. His10 and Tyr12, which are highly conserved in GH57 structures, are located at the bottom of the active site and are involved in polysaccharide binding at subsite −1. Phe452 and Trp453 are located in the C-domain and interact with the sugar through stacking interactions at the −2 and −1 subsites that place the substrate in position at the front of the catalytic residues (Fig. 3*b*). We mutated these two amino acids individually and assayed the activity of the mutants. As shown in Table 2, the mutations resulted in a significant loss of the hydrolysis activity of α-1,4-glucosidic and α-1,6-glucosidic bonds, supporting the structural analysis of substrate binding in the active site.

In the docking model, the nonreducing glucose end of isopanose was placed in close proximity to the side chains of Arg459 and Glu580, which form a salt bridge at subsite −2 (Fig. 4*c*). Some exo-type amylases, such as dextran glucosidase and oligo-1,6-glucosidase, also contain equivalent salt bridges composed of Arg and Asp in equivalent locations to form a hydrogen bond to the substrate (Yamamoto *et al.*, 2010). To investigate the functional role of the residues at subsite −2, we mutated Glu580 to glutamine (E580Q) and assayed the

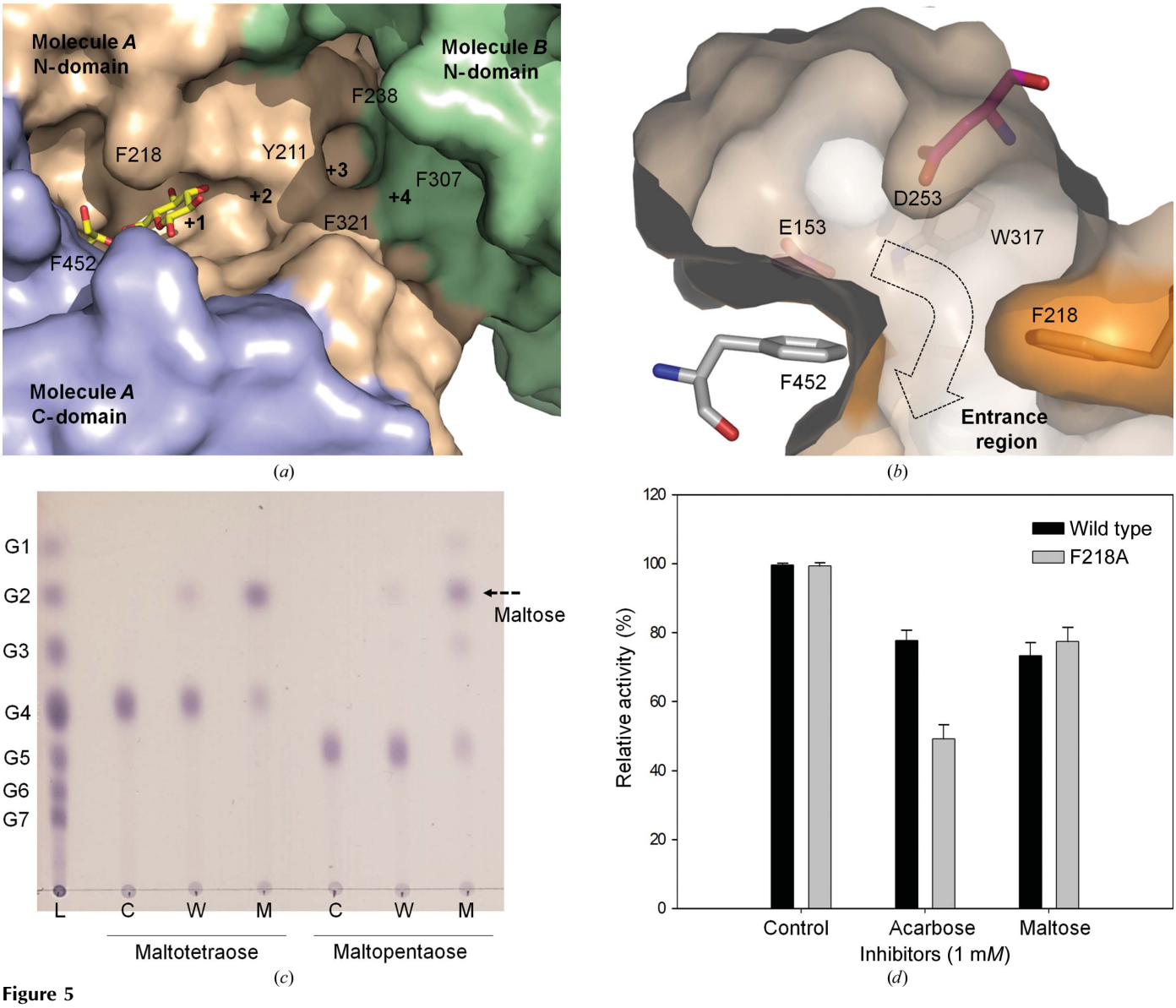


Figure 5
Substrate specificity. (*a*) Surface diagram of the substrate-binding channel that extends from the active site to the central hole of the tetrameric ring. Putative subsites from +1 to +4 are depicted. N-domains from molecule *B* are in green. (*b*) Dissection of the surface diagram of the substrate-binding groove. Phe218 and Phe452 form a physical barrier at the gate of the cavity and are shown as sticks with the active-site residues Asp253 and Glu153 (pink). The twisted geometry of subsite +1 resulting from Phe218 is highlighted by an arrow. (*c*) TLC analysis of the hydrolysis reaction products of maltotetraose and maltopentaose substrates for the wild type (W) and the F218A mutant (M) is shown. (*d*) Inhibition of α-1,4-hydrolysis activity by acarbose and maltose for the wild type (WT) and the F218A mutant.

activity of the mutant. The specific activity of the E580Q mutant was significantly decreased by 90% compared with the wild-type enzyme, suggesting the functional importance of the residue at the −2 subsite. The previous inhibition report showed that the PSMA active site has a high affinity towards maltose but a low affinity towards isomaltose or panose substrates (Fig. 4*d*). Based on the structural analysis and the mutational study, we propose that the −2 subsite is the main substrate-binding site and the interaction at the −2 subsite is essential for formation of the substrate–enzyme complex. Isomaltose and panose molecules were not hydrolyzed by PSMA, which can hydrolyze both α-1,4-glucosidic and α-1,6-glucosidic bonds. Both isomaltose and panose lack the nonreducing glucose moiety of the α-1,4-glucosidic bond for interaction with the −2 subsite, suggesting that substrate binding at the −2 subsite is critical for the enzymatic activity. This may also explain the low activity of PSMA towards G1-β-cyclodextrin because the substrate does not contain a sugar moiety capable of binding to the −2 subsite.

3.6. Substrate specificity

Analysis of the structure reveals a distinct groove-shaped channel that connects the active site to the central hole of the tetramer ring. The channel has a width of 8.6 Å and a length of 18 Å and has a slightly bent geometry at the interface between the two N-domains of the dimer. The surface contains several aromatic residues such as Tyr211 and Phe321 from one monomer and Phe238 and Phe307 from the other monomer (Fig. 5*a*). We attempted to determine the complex structure by soaking the crystal in a solution that contained maltose. Although it was difficult to model the ligand, the determined structure showed significant electron density inside the channel that was absent in the native structure (Supplementary Fig. S3), suggesting that the channel is a substrate-binding site. PSMA efficiently hydrolyzes short chains, which are generally inefficiently hydrolyzed by other amylases. The channel corresponding to putative subsites +1 to +4 may be able to bind short substrates such as G3–G5.

The entrance to the active site is covered with aromatic amino acids. Among them, Phe218 in the loop of the helix–loop–helix motif (residues 201–222) is located right at the gate of the active site with its aromatic side chain protruding perpendicular to the channel. Phe452 is positioned at the opposite side of the channel, shaping the cavity entrance and forming putative subsite +1, which is narrow and twisted relative to the catalytic site (Fig. 5*b*). Most α-1,6-glycosidases have a narrow and twisted geometry around the active-site entrance at subsite +1, a geometry suitable for the accommodation of a α-1,6-glycosidic bond, while that of α-1,4-glycosidase is wide open and straight (Yamamoto *et al.*, 2010; Abe *et al.*, 2005; Pijning *et al.*, 2012). Superimposition with the GH57 family 4-α-glucanotransferase revealed the obvious protruding architecture of Phe218 and the resultant twisted shape of PSMA. Modelling the structure with an isopanose ligand showed that Phe218 stacks on the glucose ring of isopanose at subsite +1, positioning the α-1,6-bond of the glucose moiety oriented towards the catalytic site and the 1-OH group towards the channel groove, possibly accounting for the higher activity towards the α-1,6-glucosidic linkage. The twisted architecture at subsite +1 causes a steric hindrance towards substrates, such as acarbose, in which the validamine and deoxyhexose moieties bind subsites −1 and +1, respectively, in other amylases (Roujeinikova *et al.*, 2002). We investigated whether acarbose binds and inhibits the activity of PSMA. Acarbose was not cleaved by PSMA and only marginally inhibited the activity, causing a 20% decrease in the cleavage activity in the presence of 1 m*M* acarbose, a concentration at which most of the hydrolysis activity of classical α-amylases is inhibited. To further investigate the relevance of the geometry at subsite +1 to substrate specificity, we mutated Phe218 to alanine (F218A) in order to widen and straighten the shape of the entrance. Surprisingly, the F218A mutant exhibited a significantly increased hydrolysis activity towards α-1,4-glucosidic bonds, increasing from 3.93 ± 0.33 to 14.00 ± 0.67 U mg^{-1} (Table 2). The mutation increased the α-1,4-glucosidase activity 3.65-fold in comparison to a marginal decrease in the α-1,6-glucosidase activity, suggesting that Phe218 is primarily involved in facilitating the cleavage of the α-1,6-glucosidic linkage. The hydrolysis activity significantly increased not only towards maltotriose but also towards maltotetraose and maltopentaose (Fig. 5*c*). In addition, the F281A mutant exhibited a sharp increase in acarbose inhibition, from 20 to 50%, in comparison to the wild type, indicating that the altered geometry at subsite +1 affected the binding of acarbose (Fig. 5*d*). These data demonstrate that Phe218 is a major modulator of substrate binding at subsite +1. Mutation of this residue to generate a wider gate resulted in a selective increase in the α-1,4-glycosidase activity, which may lead to a potential application in protein engineering for the modulation of substrate specificities.

4. Conclusion

We present the structure of a novel exo-type maltose-forming amylase from *Pyrococcus* sp. ST04 with a small enclosed cavity at the active site and a distinctive substrate-binding channel extending from the cavity. This structure can explain the novel cleavage pattern of maltose units from short carbohydrate substrates in an exo-type manner. Mutation of Phe218, which is located at the gate of the active site, to alanine to widen and straighten the channel geometry resulted in a selective increase in the α-1,4-bond cleavage activity, suggesting its involvement in and possible modulation of substrate specificity. The structural and biochemical analysis of PSMA in this study provides a molecular basis for the unique enzymatic properties observed and highlights the potential use of this novel exo-type maltose-forming amylase in the production of maltose.

This research was supported in part by the Converging Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (No. 2013K000249/2011-0017040) and

partially by a grant from the KRIBB Research Initiative Program.

References

Abe, A., Yoshida, H., Tonozuka, T., Sakano, Y. & Kamitori, S. (2005). *FEBS J.* **272**, 6145–6153.

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

Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B. (2009). *Nucleic Acids Res.* **37**, D233–D238.

Dickmanns, A., Ballschmiter, M., Liebl, W. & Ficner, R. (2006). *Acta Cryst.* D**62**, 262–270.

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

Imamura, H., Fushinobu, S., Jeon, B.-S., Wakagi, T. & Matsuzawa, H. (2001). *Biochemistry*, **40**, 12400–12406.

Imamura, H., Fushinobu, S., Yamamoto, M., Kumasaka, T., Jeon, B.-S., Wakagi, T. & Matsuzawa, H. (2003). *J. Biol. Chem.* **278**, 19378–19386.

Janeček, S. & Blesák, K. (2011). *Protein J.* **30**, 429–435.

Janeček, S. & Kuchtová, A. (2012). *FEBS Lett.* **586**, 3360–3366.

Jeon, E.-J., Jung, J.-H., Seo, D.-H., Jung, D.-H., Holden, J. F. & Park, C.-S. (2014). *Enzyme Microb. Technol.* **60**, 9–15.

Jun, S.-Y., Kim, J.-S., Choi, K.-H., Cha, J. & Ha, N.-C. (2013). *Acta Cryst.* D**69**, 442–450.

Jung, J.-H., Lee, J.-H., Holden, J. F., Seo, D.-H., Shin, H., Kim, H.-Y., Kim, W., Ryu, S. & Park, C.-S. (2012). *J. Bacteriol.* **194**, 4434–4435.

Jung, J.-H., Seo, D.-H., Holden, J. F. & Park, C.-S. (2014). *Appl. Microbiol. Biotechnol.* **98**, 2121–2131.

Kim, J.-S., Cha, S.-S., Kim, H.-J., Kim, T.-J., Ha, N.-C., Oh, S. -T., Cho, H.-S., Cho, M.-J., Kim, M.-J., Lee, H.-S., Kim, J.-W., Choi, K. Y., Park, K.-H. & Oh, B.-H. (1999). *J. Biol. Chem.* **274**, 26279–26286.

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

Mikami, B., Adachi, M., Kage, T., Sarikaya, E., Nanmori, T., Shinke, R. & Utsumi, S. (1999). *Biochemistry*, **38**, 7050–7061.

Mikami, B., Hehre, E. J., Sato, M., Katsube, Y., Hirose, M., Morita, Y. & Sacchettini, J. C. (1993). *Biochemistry*, **32**, 6836–6845.

Miyake, H., Kurisu, G., Kusunoki, M., Nishimura, S., Kitamura, S. & Nitta, Y. (2003). *Biochemistry*, **42**, 5574–5581.

Niehaus, F., Bertoldo, C., Kahler, M. & Antranikian, G. (1999). *Appl. Microbiol. Biotechnol.* **51**, 711–729.

Oslowski, D. M., Jung, J.-H., Seo, D.-H., Park, C.-S. & Holden, J. F. (2011). *Appl. Environ. Microbiol.* **77**, 3169–3173.

Palomo, M., Pijning, T., Booiman, T., Dobruchowska, J. M., van der Vlist, J., Kralj, S., Planas, A., Loos, K., Kamerling, J. P., Dijkstra, B. W., van der Maarel, M. J., Dijkhuizen, L. & Leemhuis, H. (2011). *J. Biol. Chem.* **286**, 3520–3530.

Pijning, T., Vujičić-Žagar, A., Kralj, S., Dijkhuizen, L. & Dijkstra, B. W. (2012). *Acta Cryst.* F**68**, 1448–1454.

Przylas, I., Tomoo, K., Terada, Y., Takaha, T., Fujii, K., Saenger, W. & Sträter, N. (2000). *J. Mol. Biol.* **296**, 873–886.

Rostkowski, M., Olsson, M. H., Søndergaard, C. R. & Jensen, J. H. (2011). *BMC Struct. Biol.* **11**, 6.

Roujeinikova, A., Raasch, C., Sedelnikova, S., Liebl, W. & Rice, D. W. (2002). *J. Mol. Biol.* **321**, 149–162.

Santos, C. R., Tonoli, C. C., Trindade, D. M., Betzel, C., Takata, H., Kuriki, T., Kanai, T., Imanaka, T., Arni, R. K. & Murakami, M. T. (2011). *Proteins*, **79**, 547–557.

Ver Eecke, H. C., Kelley, D. S. & Holden, J. F. (2009). *Appl. Environ. Microbiol.* **75**, 242–245.

Xavier, K. B., Peist, R., Kossmann, M., Boos, W. & Santos, H. (1999). *J. Bacteriol.* **181**, 3358–3367.

Yamamoto, K., Miyake, H., Kusunoki, M. & Osaki, S. (2010). *FEBS J.* **277**, 4205–4214.

Zona, R., Chang-Pi-Hin, F., O'Donohue, M. J. & Janeček, S. (2004). *Eur. J. Biochem.* **271**, 2863–2872.