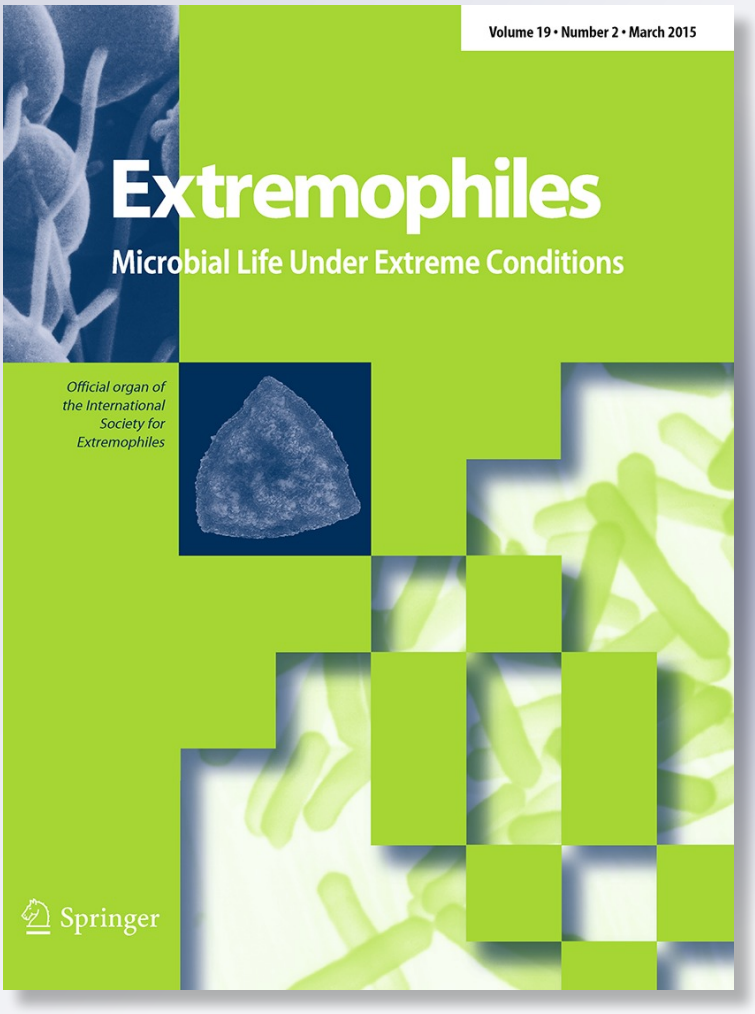
Novel characteristics of a carbohydrate-binding module 20 from hyperthermophilic bacterium

Il-Nam Oh, Jay-lin Jane, Kan Wang, Jong-Tae Park & Kwan-Hwa Park

Extremophiles
Microbial Life Under Extreme Conditions

ISSN 1431-0651
Volume 19
Number 2

Extremophiles (2015) 19:363-371
DOI 10.1007/s00792-014-0722-1



Springer

Extremophiles (2015) 19:363–371
DOI 10.1007/s00792-014-0722-1

ORIGINAL PAPER

Novel characteristics of a carbohydrate-binding module 20 from hyperthermophilic bacterium

Il-Nam Oh · Jay-lin Jane · Kan Wang · Jong-Tae Park · Kwan-Hwa Park

Received: 19 September 2014 / Accepted: 1 December 2014 / Published online: 10 January 2015

Abstract In this study, a gene fragment coding carbohydrate-binding module 20 (CBM20) in the amylopullulanase (APU) gene was cloned from the hyperthermophilic bacteria *Thermoanaerobacter pseudoethanolicus* 39E and expressed in *Escherichia coli*. The protein, hereafter Tp39E, possesses very low sequence similarity with the CBM20s previously reported and has no starch binding site 2. Tp39E did not demonstrate thermal denaturation at 50 °C; however, thermal unfolding of the protein was observed at 59.5 °C. A binding assay with Tp39E was conducted using various soluble and insoluble substrates, and starch was the best binding polysaccharide. Intriguingly, Tp39E bound, to a lesser extent, to soluble and insoluble xylan as well. The dissociation constant (K_d) and the maximum specific binding (B_{max}) of Tp39E to corn starch granules were 0.537 μM and 5.79 μM/g, respectively, at pH 5.5 and 20 °C. $_{99}APU_{1357}$ with a Tp39E domain exhibited 2.2-fold greater activity than a CBM20-truncation mutant when starch granules were the substrate. Tp39E was an independently thermostable CBM and had a considerable effect on APU activity in the hydrolysis of insoluble substrates.

Keywords Carbohydrate binding module · Thermostability · Starch · Xylan · Dissociation constant · Insoluble substrate

Communicated by A. Driessen.

I.-N. Oh · J.-T. Park (✉)
Department of Food Science and Technology, Chungnam National University, 99 Daehak-ro, Yuseong-gu, Daejeon 305-764, Korea
e-mail: jtpark@cnu.ac.kr

J. Jane
Department of Food Science and Human Nutrition, Iowa State University, Ames, IA, USA

K. Wang
Department of Agronomy, Iowa State University, Ames, IA 50011, USA

K.-H. Park
Department of Agricultural Biotechnology, Seoul National University, 599 Gwanak-ro, Gwanak-gu, Seoul 151-921, Korea

K.-H. Park
Department of Foodservice Management and Nutrition, Sangmyung University, 7 Hongji-dong, Jongno-gu, Seoul 110-743, Korea

Introduction

Many carbohydrate-active enzymes have non-catalytic modules that interact and bind to polysaccharides, such as starch, cellulose, and xylan. The modules are so-called carbohydrate-binding modules (CBMs) (Christiansen et al. 2009a; Chu et al. 2014; Guillén et al. 2010; Janeček et al. 2011; Machovič and Janeček 2006). The CBMs are defined as contiguous amino acid sequences that have carbohydrate-binding activity within carbohydrate-active enzymes and non-enzymatic proteins. Currently, there are 71 CBM families according to the CAZY site (http://www.cazy.org/Carbohydrate-Binding-Modules.html). The CBMs increase the catalytic efficiency of enzymes by binding to substrates (Guillén et al. 2010; Southall et al. 1999) or modulating the properties of regulatory proteins, such as AMPKb (Janeček et al. 2011; Koay et al. 2007, 2010). There are several CBMs that recognize starch-like carbohydrates. Among them, CBM20 was established earlier and intensively investigated. CBM20 is also known as a starch-binding domain (SBD) (Christiansen et al. 2009a). This type of CBM binds

to raw starch (starch granules) as well as gelatinized starch (Christiansen et al. 2009a; Paldi et al. 2003; Williamson et al. 1992; Ye et al. 2004). In CBM20, there are two starch-binding sites (SBS) that are thought to be functionally different. Binding site 1 contains two conserved tryptophans, and binding site 2 contains one conserved tryptophan in the canonical SBD of *Aspergillus niger* and other CBMs identified earlier (Abbott and van Bueren 2014; Christiansen et al. 2009a; Southall et al. 1999). Some newly identified CBM20s have tyrosine instead of tryptophan at SBS2; these domains have no or weak binding affinity to carbohydrates. CBM20s are mainly found in glycoside hydrolase families 13, 14, 15, and 77 (Christiansen et al. 2009a; Machovič and Janeček 2006, 2008). The CBM20 domain provides enhanced hydrolysis activity to these amylolytic enzymes as they bind to insoluble and soluble starches (Southall et al. 1999). Many putative CBM20s were found by amino acid sequence alignment in previous studies (Christiansen et al. 2009a; Janeček et al. 2011; Machovič and Janeček 2006, 2008), but few were characterized experimentally (Araki et al. 2009; Guillen et al. 2007; Lin et al. 2003; Southall et al. 1999).

Starch is the most abundant carbohydrate energy storage in plants. It consists of 0–30 % amylose and 70–100 % amylopectin, which are glucose polymers with α-1,4 and α-1,6-glycosidic linkages. Amylopectin has approximately 5 % α-1,6-linkage, but amylose has no more than 0.1 % of this type of bond. Fiber-type carbohydrates, i.e., cellulose, xylan, and β-glucan, have β-glycosidic linkages, and cellulose and xylan backbones are linear, relatively flat, and stretched. In contrast, starch exhibits a helical structure in solution (Christiansen et al. 2009a; Chung et al. 2011).

Various industries that use starch require high-temperature processes. Therefore, many types of thermostable enzymes have been studied for decades. One of them is amylopullulanase (APU), which possesses hydrolytic activity towards α-1,4 and α-1,6-glucosidic linkages. Many APUs have been studied for industrial applications (Ganghofner et al. 1998; Lin et al. 2008; Vieille and Zeikus 2001). Some of these APUs also have CBM; however, there are not enough experimental results on the roles of CBMs in APUs (Machovič and Janeček 2008). APU from *Thermoanaerobacter pseudoethanolicus* 39E was first reported by Saha et al. (1988). The APU consists of 1481 amino acids, and the deduced molecular weight based on the primary sequence is 166.23 kDa (Mathupala et al. 1993). The enzymatic properties of the APU were revealed, and the effects of the C-terminal end truncation on the enzyme were evaluated by several research groups (Lin et al. 2008, 2012; Lin and Leu 2002; Lin et al.). As the CBM database was built, a putative CBM20 was found in the C-terminal region of APU from *T. pseudoethanolicus* 39E. Currently, no studies have focused on the characterization of CBM20 in the enzyme. Therefore, in this study, the biochemical characteristics of putative CBM20 from *T. pseudoethanolicus* 39E were investigated to evaluate the possibility of an amylase fusion partner that can improve the enzymatic properties of amylases for industrial applications.

Materials and methods

Cloning, expression, and purification of recombinant proteins

The putative CBM20 (Tp39E) in the APU gene (gi:189047124) was amplified using primers Tp39EF (*Nde*I, 5′-CATATGGCATCAAATATAGTGAAAGC-3′; restriction site is underlined) and Tp39ER (*Hin*dIII, 5′-AAGCTTTTAGTCGCGCCAGCGTTGC-3′) from the APU of *T. pseudoethanolicus* 39E using PCR. For $_{99}$APU$_{1357}$ and $_{106}$APU$_{1062}$ (Fig. 5a), F primer (*Nde*I, 5′-CATGGGAACATATGGGAGTTCCTTCACAAGGTAAT-3′) and R primer (*Xba*I, 5′-GCAAGCGATCTAGATTAGTCTCTCCATCTTTGCACTG-3′) were designed for $_{99}$APU$_{1357}$, and F primer (*Nde*I, 5′-GAATTAATTCCATATGTTAAGCTTGCATCTTG-3′) and R primer (*Xba*I, 5′-CTCCACCGCGGTTCTAGACGCTCTAGTCGA-3′) were used for $_{106}$APU$_{1062}$. The other following procedures for the two proteins were the same as that of Tp39E.

The amplified PCR product was digested with restriction enzymes and ligated into a p6xHis119 vector (Ryu et al. 2005), and the recombinant plasmid harboring the Tp39E gene was designated as p6xHTp39E. *Escherichia coli* BL21(DE3) (Jeong et al. 2009) carrying p6xHTp39E was cultured in 1 L LB broth at 37 °C, and the cells were collected after 20 h incubation. The cell pellet was resuspended in 100 ml of lysis buffer (300 mM NaCl, 50 mM Tris buffer (pH 7.5), and 10 mM imidazole) and sonicated in an ice bath. The crude cell extract was centrifuged (13,000×*g*, 4 °C, 15 min), and the recombinant protein (Tp39E) in the supernatant was purified using nickel ion-nitrilotriacetic acid (Ni–NTA) affinity chromatography (Qiagen GmbH, Hilden, North Rhine-Westphalia, Germany). Then, the protein was further purified by Q-Sepharose anion exchange chromatography (GE Healthcare, Uppsala, Uppsala, Sweden) using fast protein liquid chromatography (FPLC, GE Healthcare).

Protein sequence analysis

Amino acid sequence of Tp39E was compared with CBMs in *A.* glucoamylase 1 (GA1, gi: 732556), *Aspergillus kawachii* GA1 (gi: 217815), *Bacillus cereus* β-amylase (gi: 3925826), *Bacillus circulans* cyclodextrin glucanotransferase (CGT, gi: 510492), *Bacilllus* TS-23 α-amylase (gi:

722279), *Deinococcus geothermalis* branching enzyme (BE, gi: 118572355), *Geobacillus stearothermophilus* APU (gi: 12006232) and CGT (gi: 399224), *Klebsiella pneumoniae* CGT (gi: 149179) and pullulanase (Pul, gi: 43913), *Streptococcus suis* APU (gi: 251820781), *Sulfolobus solfataricus* TreX (gi: 15898878), *Thermoanaerobacterium thermosulfurigenes* APU (gi: 1235795), *Thermobifida fusca* YX glycogen debranching enzyme (GDE, gi: 71916022), *Thermococcus* sp. B1001 CGT (gi: 6552352), *Thermococcus aggregans* Pul (gi: 8439491), *Thermotoga maritima* Pul (gi: 6225896), and *Thermus thermophilus* Pul (gi: 14861067) (Christiansen et al. 2009a; Janeček et al. 2011; Machovič and Janeček 2008). Sequence alignment was conducted using AlignX program in VectorNTI ver. 11.0 (Invitrogen, Carlsbad, CA, USA), and manual modification was added. Evolutionary tree was drawn by neighbor-joining method (Saitou and Nei 1987) using MEGA6 (Tamura et al. 2013). Domain architectures of proteins were predicted using SMART (Letunic et al. 2012; Schultz et al. 1998).

Protein assay

Protein concentration was determined according to the Bio-Rad Protein Assay (Cat# 500-0006, Bio-rad, Hercules, California, USA) using bovine serum albumin (BSA, BobOStar, Bobogen, Australia) as a standard. The reaction mixture consisted of 40 µl of diluted enzyme solution and 2 ml of the Bio-Rad Protein Assay dye reagent that was diluted tenfold. After 1 h reaction, absorbance of the samples was measured at 595 nm.

Determination of thermostability

The thermostability of Tp39E was determined at 50, 55, and 60 °C in sodium acetate buffer (pH 5.5, 50 mM). After incubation an hour, the protein was centrifuged to remove aggregated protein, and the residual protein in the supernatant was assayed using the Bradford protein assay. To determine the thermal unfolding temperature of Tp39E with or without β-cyclodextrin, differential scanning calorimetry (DSC) was conducted. The protein solution [2 mg/mL, pH 5.5 sodium acetate (50 mM)] was analyzed using the DSC1 system (Mettler-Tolledo, Schwerzenbach, Switzerland). Before analysis, Tp39E was incubated at pH 5.5 and 20 °C, then heated to 80 °C at 5 °C for 1 min. The endothermic peaks were analyzed and compared.

Binding features with various polysaccharides and pH conditions

The binding ability of Tp39E was assayed using the following insoluble polysaccharides (final concentration 10 mg/ml): granular corn starch, cellulose, and xylan. Each polysaccharide and Tp39E [final concentration 80 µg/ml in pH 5.5 sodium acetate (50 mM)] was mixed and incubated at 20 °C with gentle shaking (130 rpm) for 40 min, then centrifuged (13000×*g*) for 10 min. The supernatant was precipitated with trichloroacetic acid and analyzed using sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE). The protein concentration in the supernatant was measured using a Bradford protein assay. The binding properties of Tp39E were determined at various pH conditions from pH 4.0 to pH 7.0 using corn starch granule. A gel mobility shift assay was conducted to examine binding between Tp39E and soluble polysaccharides, amylopectin and xylan. The protein was incubated with or without each carbohydrate at pH 5.5 sodium acetate (50 mM) and 20 °C for 30 min, then analyzed using native-PAGE conditions.

Binding kinetics

The binding kinetics were determined using a one-site specific binding method (Mehmood et al. 2011). Various amounts (from 10 to 110 µg/mL) of Tp39E were added to a corn starch granule suspension (10 mg/mL) in sodium acetate buffer (50 mM, pH 5.5). The mixture was incubated at 20 °C for 40 min with shaking (130 rpm) and centrifuged at 13,500×*g* for 10 min. The protein concentration in the supernatant was determined by the protein assay, and the amount of absorbed protein was calculated by subtraction. Using the double-reciprocal plot of the binding and free protein, the dissociation constant (K_d) and maximal specific binding (B_{max}) were calculated.

Starch hydrolysis activity of amylopullulanase

Amylopectin (soluble) and starch granules (insoluble) were used for evaluating the effects of Tp39E on APU activity. Two APU truncated enzymes, one with a Tp39E domain and the other without, were incubated with amylopectin and starch granules at pH 5.5 and 50 °C. Reducing sugars produced by APU were determined using the copper-bicinchoninate method (Fox and Robyt 1991).

Results

Cloning and analysis of amino acid sequence

The putative CBM20 module, Tp39E in APU from *T. pseudoethanolicus* 39E, was cloned, and the protein was successfully expressed and purified. Tp39E showed an expected molecular size of 14 kDa by SDS-PAGE (data not shown). The primary structure of Tp39E was analyzed and compared with other CBM20s and CBM48s that have a close

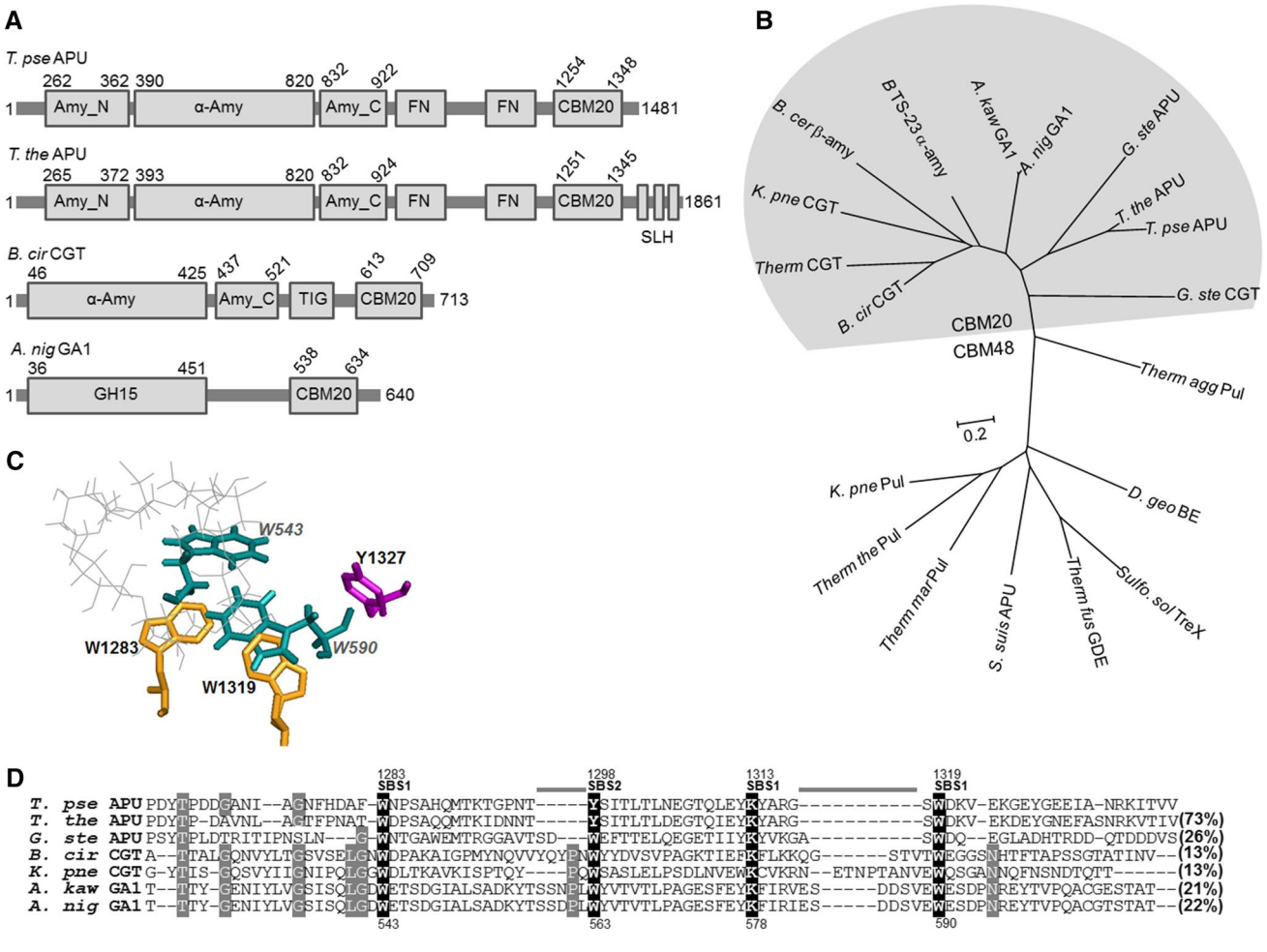


Fig. 1 Domain structure of the APU from *Thermoanaerobacter pseudoethanolicus* 39E is compared with several selected enzymes (**a**). The *numbers* in **a** indicate positions of amino acids. Domain names are abbreviations of amylase N domain (Amy_N), α-amylase catalytic domain (α-Amy), amylase C-domain (Amy_C), fibronectin-like domain, S-layer homologous domain (SLH), domains having immunoglobulin-like fold, and glycoside hydrolase family 15 catalytic domain (GH15). Evolutionary tree of Tp39E with 10 CBM20s and 8 CBM48s is shown in **b**. Superimposed positions of the two tryptophans in the SBS1 comparing Tp39E and the CBM20 of *Aspergillus niger* GA1 (PDB ID: 1AC0) are shown in **c**. The modeling of Tp39E structure was conducted based on 1AC0 (Sorimachi et al. 1997) as a template using Swiss-model Workspace (Arnold et al. 2006). Sequence alignment of Tp39E with 6 CBM20s is given in **d**. For the abbreviations, refer to the "Materials and methods". Numbers in **d** indicate positions of the conserved four amino acids in *T. pseudoethanolicus* 39E APU (top) and *A. niger* GA1 (bottom), and gray bars indicate missing regions in loops of CBM20s from APU

evolutionary relationship with CBM20. APU from *T. pseudoethanolicus* 39E has a multi-domain architecture, and other APUs have similar structures. In the APU, Tp39E is located at the C-terminal region. CBM20s in CGT and β-amylase are also observed at the C-terminus (Fig. 1a). Tp39E has relatively low sequence identity to other CBM20s, but it is separated from CBM48s and located close to the CBM20s of other APUs in the evolutionary tree (Fig. 1b). A contiguous 50 amino acid portion of Tp39E shares considerable sequence homology (20–30 %) with CBM48s (data not shown). Of the four well-conserved amino acid residues, Tp39E has 3 amino acids, W1283, K1313 and 1319W, conserved in SBS1. W563 (SBS2) in CBM20 from *A. niger* GA1 was replaced with Y1298 in Tp39E. Y1327 in Tp39E is expected to be located close to SBS1 (Fig. 1c, d). Tp39E has relatively short loops before Y1298 (SBS2) and W1319 (SBS1) (Fig. 1d). The former loop is a counterpart of a loop in CBM20 of the GA1 that exhibits a large conformational change when binding starch (Sorimachi et al. 1997).

Thermostability

Purified Tp39E, the CBM20 from APU of *T. pseudoethanolicus,* was examined to evaluate independent thermostability of the single domain. Tp39E was very stable at 50 °C, and no denatured protein was detected after 1 h of incubation. At 55 °C, the *D* value of the domain was estimated as 7.5 h. However, the stability of Tp39E decreased drastically

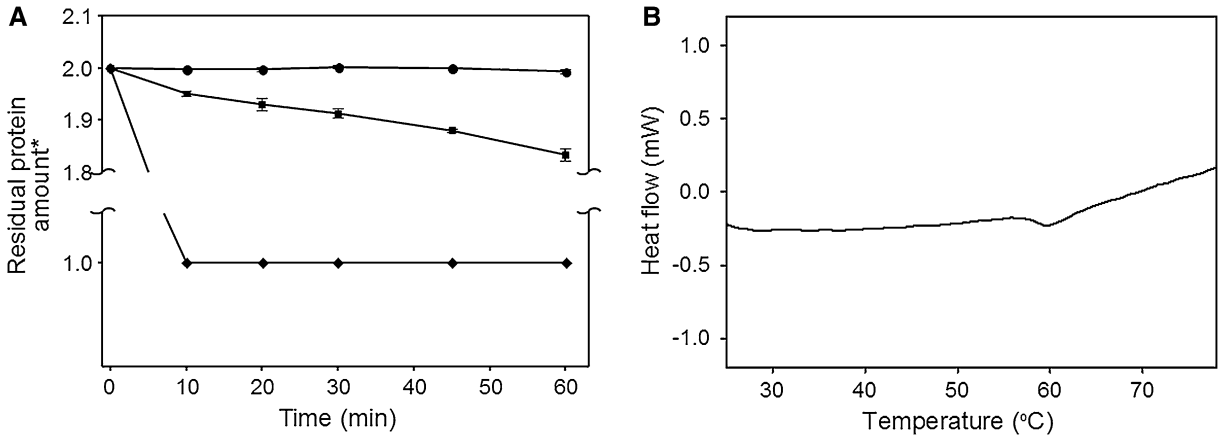


Fig. 2 Thermostability of Tp39E was determined using heat treatment at different temperatures (**a**) and DSC thermography (**b**). *Error bar* in **a** shows standard deviations

at 60 °C (Fig. 2a). The thermal unfolding characteristics of Tp39E were analyzed using DSC. On the DSC thermogram, the heat flow for melting of Tp39E started at 57.8 °C, but the closing temperature was unclear. The thermal unfolding temperature, i.e., peak temperature on the curve, was 59.5 °C (Fig. 2b).

Binding assay using various polysaccharides

At first, binding property of Tp39E was roughly screened using insoluble polysaccharides such as starch granules, cellulose, and xylan. An optimal binding condition assay was conducted using starch granules because Tp39E bound more tightly to starch granules than the other polysaccharides (data not shown). At pH 5.5, Tp39E showed the highest binding ratio. The CBM20 retained its binding power to the starch granule over a broad pH range (Fig. 3). The optimal pH of Tp39E was the same as that for hydrolysis activity of the APU from which the CMB20 originated (Lin and Leu 2002). The binding efficiency of Tp39E to insoluble polysaccharides was further evaluated. Approximately 40 % of Tp39E was removed by incubating with the starch granule, but the protein amount was not changed when cellulose was used. Intriguingly, more than a quarter of Tp39E was removed after incubation with insoluble xylan (Fig. 4a, b). The interaction between Tp39E and xylan was also observed when soluble xylan was used. In the gel mobility shift assay, Tp39E with amylopectin and xylan showed retarded migration. Mobility shift was greater in amylopectin than in xylan, but Tp39E clearly recognized and bound to xylan (Fig. 4c).

Binding kinetics

Kinetic binding parameters between Tp39E and starch granules were determined. The dissociation constant (K_d)

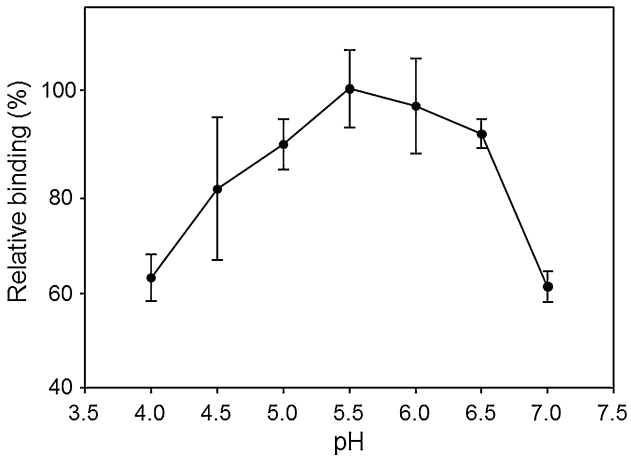


Fig. 3 Binding assay of Tp39E using the starch granule as the ligand at different pH conditions. *Error bar* indicates standard deviations

was 0.536 μM, and the maximal specific binding (B_{max}) was 5.79 μM/g starch. Tp39E exhibited a much smaller K_d value than CBM20 from GA in previous reports (Table 1). This means that Tp39E recognizes and binds to starch granules in a more efficient manner than GA CBM20s, shown in Table 1. The B_{max} of Tp39E was greater than GA CBM20s as well. However, CBM20 from β-amylase has greater binding efficiency and higher maximal specific binding (Table 1).

Effects of Tp39E on the catalytic properties of APU

To evaluate the role of Tp39E, two types of APU mutants were constructed and their catalytic properties on soluble and insoluble substrates were examined. $_{106}$APU$_{1062}$ with no Tp39E domain was compared with the other mutant, $_{99}$APU$_{1357}$ with the CBM20 domain (Fig. 5a).

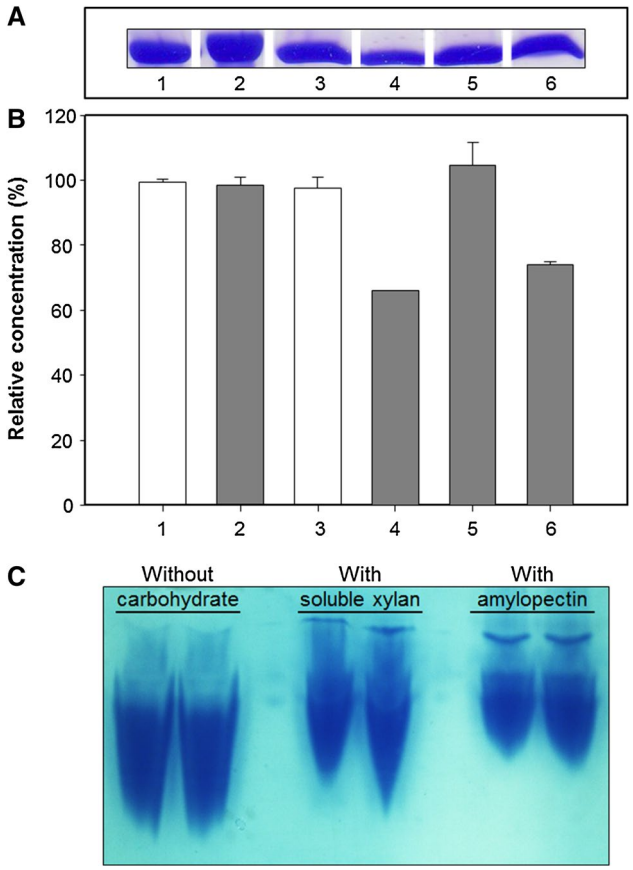


Fig. 4 Binding of Tp39E to different insoluble ligands was examined using SDS-PAGE (**a**) and the residual protein amount was determined (**b**). *Lane 1* BSA only, *lane 2* BSA with the starch granule, *lane 3* Tp39E only, *lanes 4, 5,* and *6* Tp39E with starch granule, cellulose, or xylan, respectively. The Tp39E gel-mobility assay was conducted using soluble ligands, amylopectin, and xylan (**c**)

$_{99}$APU$_{1357}$ showed greater activity than $_{106}$APU$_{1062}$ to both amylopectin and the starch granules. The difference in the molar specific activity between the two mutants was greater when the insoluble starch granules were used (Fig. 5b). The time-lapsed reactions of the two mutants and starch granules corresponded with the activity assays (Fig. 5c). $_{99}$APU$_{1357}$ hydrolyzed starch granules more efficiently and continuously than the other mutant.

Discussion

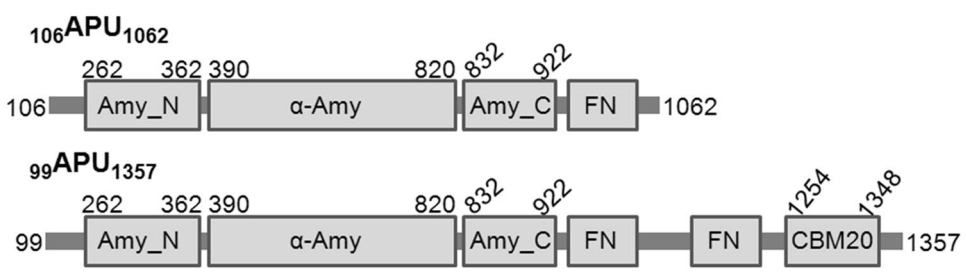
CBM20s have close relationship to CBM48s in its evolutionary history as well as the structure and the function (Janeček et al. 2011). Among CBM20s, Tp39E is located nearby the border of the two CBM families (Fig. 1b). CBM20s that recognize starch granules are usually found in extracellular amylolytic enzymes such as α-amylase, CGT, and β-amylase. Most of the APUs are extracellular enzymes as well. Some APU are found to be cell surface-bound amylases that have specific repeated amino acid regions. Tp39E is located at the C-terminal end of the extracellular APU. However, the APU from *T. pseudoethanolicus* 39E does not have a sequence identified as a domain or a motif for cell surface-binding (Fig. 1a). Among the 11 conserved amino acids selected based on earlier studies (Janeček et al. 2011; Machovič and Janeček 2006), some are conserved but others are replaced in Tp39E, including five missing amino acids in the loop at SBS2 that can cause loss of binding site 2. This type of deletion, not the exact sequence but the deletion position, are found in CBM20s in glucan and water dikinase (Christiansen et al. 2009b). In the *A. niger* GA1, the loop exhibits a large conformational change when binding to the β-CD. The movement of the loop causes the approach of Tyr556 that has critical a role in SBS2 to Asp554 and Lys555 (Sorimachi et al. 1997; Williamson et al. 1997). There are a few CBM20s that have no carbohydrate-binding affinity at SBS2. For example, CBM20s from *G. stearothermophilus* maltogenic α-amylase and *B. cereus* β-amylase exhibit no binding carbohydrates at SBS2 in X-ray crystallography (Dauter et al. 1999; Mikami et al. 1999; Oyama et al. 1999).

The thermostabilities of CBM20s isolated as an independent domain were not easily found in previous studies. Of the CBM20s, the thermal properties of CBM20 from *A. niger* GA1 is well characterized using mutagenesis studies. The thermal unfolding temperature of *A. niger* GA1 CBM20 is 52.7 °C, but decreased to 42.7 °C when two cysteines were replaced by Gly or Ser to disrupt a disulfide bond (Sugimoto et al. 2007). Intriguingly, Tp39E has no disulfide bond but showed much greater thermostability than the GA1 CBM20. The domain size of Tp39E is similar to that of the GA1 CBM20 and has fewer bulk-hydrophobic

Table 1 Comparison of binding kinetics parameters between selected CBM20s

Enzyme	Family	Origin	Substrate	K_d (μM)	B_{max} (μM/g)	Reference
APU	CBM20	*Thermoanaerobacter pseudoethanolicus* 39E	Corn starch	0.536 ± 0.060	5.79 ± 0.23	This study
GA	CBM20	*Aspergillus niger*	Corn starch	19.6	2.14	(Williamson et al. 1992)
GA	CBM20	*Aspergillus niger* B1	Corn starch	3.2	0.56	(Paldi et al. 2003)
GA	CBM20	*Aspergillus niger* B1	Potato starch	3.3	0.08	(Paldi et al. 2003)
β-Amylase	CBM20	*Bacillus cereus* var.	Corn starch	0.09	16.1	(Ye et al. 2004)

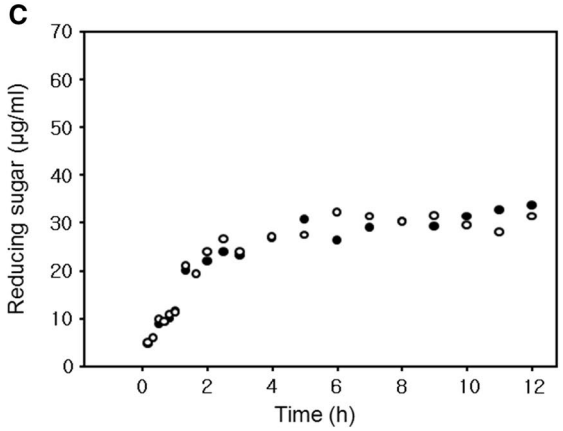
A



B

Substrate	Molar specific activity (mU/nmole protein)		Ratio ($_{99}$APU$_{1357}$/ $_{106}$APU$_{1062}$)
	$_{106}$APU$_{1062}$	$_{99}$APU$_{1357}$	
Amylopectin	242 ± 19	404 ± 10	1.67
Starch granule	2.72 ± 0.03	5.99 ± 0. 31	2.20

C



D

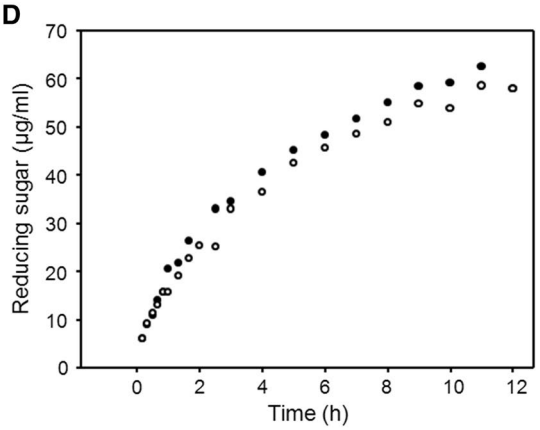


Fig. 5 Domain structures of two mutants to evaluate the roles of Tp39E in enzymatic catalysis (**a**). The numbers in **a** indicate positions of amino acids. For the full name of domain, refer to Fig. 1. The activities of mutants on soluble and insoluble substrates were determined (**b**). Time-lapsed reactions of $_{106}$APU$_{1062}$ (**c**) and $_{99}$APU$_{1357}$ (**d**) using starch granules as the substrate were compared. *Closed and open circles* indicate repeated analyses

amino acids, i.e., Phe, Tyr, and Trp (Fig. 1d). It is feasible that Tp39E has a unique strategy to increase the thermal stability of a small domain.

CBM20s are well-known B-type structures that recognize and bind at least two subsites in single polysaccharides with helical structures or at least no planar structure. These types of CBMs do not recognize planar ligands such as cellulose and xylan (Pell et al. 2003; Simpson et al. 2002; Szabó et al. 2001). A-type CBMs, which are found in cellulases and hemicellulases, recognize these carbohydrates using hydrophobic binding at the planar surface (Boraston et al. 2004). A-type CBMs often bind several types of ligands that have β-1,4-glycosidic linkage as their backbone structure (Duan et al. 2010). However, to date, no CBM20s have been identified that bind two different types of ligands, such as starch and xylan. The xylan-binding domains belong to CBM families, such as 4, 6, 13, and 22, which are different from the SBD families (Pell et al. 2003). The binding affinity of Tp39E to xylan was weaker than to starch, but the experimental results proved that there must be a specific binding between Tp39E and xylan (Fig. 4). The mechanism of binding to xylan needs to be investigated by further experiments. Due to the very low sequence identities of Tp39E to CBM20s in the Protein Structure Data Bank (http://www.pdb.org), modeling of a 3-dimensional structure was not reliable for this purpose. To determine the 3-dimensional structure of Tp39E, X-ray crystallography is required.

The effects of the CBM20 domain Tp39E on enzymatic catalysis have been unclear until this study was conducted. In previous studies, APU from *T. pseudoethanolicus* 39E was investigated to determine the catalytic properties and the

thermostability of the enzyme, but CBM20 at the C-terminal end was not included (Lin and Leu 2002; Mathupala and Zeikus 1993). Most studies were conducted using C-terminal truncation mutants that had no CBM20 domains. Lin et al. predicted that the APU does not need the CBM20 domain for its enzyme catalysis (Lin et al. 2012). However, they did not use any APU mutants having the CBM20 domain for comparison. This study clearly shows that the CBM20 of APU has some role in the hydrolysis of both soluble and insoluble substrates. Removing Tp39E from the APU decreased the hydrolysis rates of amylopectin (40 %) and decreased more when starch granule was used (55 %) (Fig. 5b). The binding affinity of Tp39E to soluble and insoluble substrates could provide more opportunities for the APU to interact with its substrates. This proximity effect was suggested and accepted in many studies (Bolam et al. 1998; Hervé et al. 2010). Substrate structure disruption by CBM20 was also reported in the GA1 CBM20 (Southall et al. 1999). The pre-incubation of the CBM20 domain with corn starch granules enhanced the hydrolysis of the insoluble substrate by GA2 with no CBM20. They presumed that two SBS are needed for the disruption of the starch structure. Based on the sequence analysis, Tp39E is likely to have only one SBS and is expected to have no role in the disruption of starch structure. Tp39E, however, exhibited stronger binding affinity to starch granules than some CBM20s (Table 1).

The results presented in this study clearly demonstrate the roles of Tp39E in the catalytic properties of APU. This thermostable CBM20 could be used as a fusion partner to improve the catalytic properties of thermostable amylases in starch conversion.

Acknowledgments This study was carried out with the support of Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2011-0014801) and the Next Generation BioGreen21 Program (SAAC, No. PJ009086) of Rural Development Administration, Republic of Korea.

References

Abbott DW, van Bueren AL (2014) Using structure to inform carbohydrate binding module function. Curr Opin Struct Biol 28:32–40

Araki Y, Karita S, Tanaka A, Kondo M, Goto M (2009) Characterization of family 17 and family 28 carbohydrate-binding modules from *Clostridium josui* Cel5A. Biosci Biotechnol Biochem 73:1028–1032

Arnold K, Bordoli L, Kopp J, Schwede T (2006) The swiss-model workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22:195–201

Bolam D et al (1998) *Pseudomonas* cellulose-binding domains mediate their effects by increasing enzyme substrate proximity. Biochem J 331:775–781

Boraston A, Bolam D, Gilbert H, Davies G (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem J 382:769–781

Christiansen C, Abou Hachem M, Janeček Š, Viksø-Nielsen A, Blennow A, Svensson B (2009a) The carbohydrate-binding module family 20–diversity, structure, and function. FEBS J 276:5006–5029

Christiansen C, Hachem MA, Glaring MA, Viksø-Nielsen A, Sigurskjold BW, Svensson B, Blennow A (2009b) A CBM20 low-affinity starch-binding domain from glucan, water dikinase. FEBS Lett 583:1159–1163

Chu CH, Li KM, Lin SW, Chang MDT, Jiang TY, Sun YJ (2014) Crystal structures of starch binding domain from *Rhizopus oryzae* glucoamylase in complex with isomaltooligosaccharide: insights into polysaccharide binding mechanism of CBM21 family. Proteins 82:1079–1085

Chung H-J, Liu Q, Lee L, Wei D (2011) Relationship between the structure, physicochemical properties and in vitro digestibility of rice starches with different amylose contents. Food Hydrocoll 25:968–975

Dauter Z et al (1999) X-ray structure of Novamyl, the five-domain "maltogenic" α-amylase from *Bacillus stearothermophilus*: maltose and acarbose complexes at 1.7 Å resolution. Biochemistry 38:8385–8392

Duan C-J, Liu J-L, Wu X, Tang J-L, Feng J-X (2010) Novel carbohydrate-binding module identified in a ruminal metagenomic endoglucanase. Appl Environ Microbiol 76:4867–4870

Fox JD, Robyt JF (1991) Miniaturization of three carbohydrate analyses using a microsample plate reader. Anal Biochem 195:93–96

Ganghofner D, Kellermann J, Staudenbauer WL, Bronnenmeier K (1998) Purification and properties of an amylopullulanase, a glucoamylase, and an α-glucosidase in the amylolytic enzyme system of *Thermoanaerobacterium thermosaccharolyticum*. Biosci Biotechnol Biochem 62:302–308

Guillen D et al (2007) Alpha-amylase starch binding domains: cooperative effects of binding to starch granules of multiple tandemly arranged domains. Appl Environ Microbiol 73:3833–3837

Guillén D, Sánchez S, Rodríguez-Sanoja R (2010) Carbohydrate-binding domains: multiplicity of biological roles. Appl Microbiol Biotechnol 85:1241–1249

Hervé C, Rogowski A, Blake AW, Marcus SE, Gilbert HJ, Knox JP (2010) Carbohydrate-binding modules promote the enzymatic deconstruction of intact plant cell walls by targeting and proximity effects. Proc Natl Acad Sci U S A 107:15293–15298

Janeček Š, Svensson B, MacGregor E (2011) Structural and evolutionary aspects of two families of non-catalytic domains present in starch and glycogen binding proteins from microbes, plants and animals. Enzyme Microb Technol 49:429–440

Jeong H et al (2009) Genome Sequences of *Escherichia coli* B strains REL606 and BL21 (DE3). J Mol Biol 394:644–652

Koay A, Rimmer KA, Mertens HD, Gooley PR, Stapleton D (2007) Oligosaccharide recognition and binding to the carbohydrate binding module of AMP-activated protein kinase. FEBS Lett 581:5055–5059

Koay A et al (2010) AMPK β subunits display isoform specific affinities for carbohydrates. FEBS Lett 584:3499–3503

Letunic I, Doerks T, Bork P (2012) SMART 7: recent updates to the protein domain annotation resource. Nucleic Acids Res 40:D302–D305

Lin F-P, Leu K-L (2002) Cloning, expression, and characterization of thermostable region of amylopullulanase gene from *Thermoanaerobacter ethanolicus* 39E. Appl Biochem Biotechnol 97:33–44

Lin LL, Lo HF, Chi MC, Ku KL (2003) Functional expression of the raw starch-binding domain of *Bacillus* sp. strain TS-23 α-amylase in recombinant Escherichia coli. Starch-Stärke 55:197–202

Lin H-Y, Chuang H-H, Lin F-P (2008) Biochemical characterization of engineered amylopullulanase from *Thermoanaerobacter ethanolicus* 39E-implicating the non-necessity of its 100 C-terminal amino acid residues. Extremophiles 12:641–650

Lin F-P, Ho Y-H, Lin H-Y, Lin H-J (2012) Effect of C-terminal truncation on enzyme properties of recombinant amylopullulanase from *Thermoanaerobacter pseudoethanolicus*. Extremophiles 16:395–403

Machovič M, Janeček Š (2006) The evolution of putative starch-binding domains. FEBS Lett 580:6349–6356

Machovič M, Janeček Š (2008) Domain evolution in the GH13 pullulanase subfamily with focus on the carbohydrate-binding module family 48. Biologia 63:1057–1068

Mathupala SP, Zeikus JG (1993) Improved purification and biochemical characterization of extracellular amylopullulanase from *Thermoanaerobacter ethanolicus* 39E. Appl Microbiol Biotechnol 39:487–493

Mathupala SP, Lowe SE, Podkovyrov SM, Zeikus JG (1993) Sequencing of the amylopullulanase (apu) gene of *Thermoanaerobacter ethanolicus* 39E, and identification of the active site by site-directed mutagenesis. J Biol Chem 268:16332–16344

Mehmood MA, Xiao X, Hafeez FY, Gai Y, Wang F (2011) Molecular characterization of the modular chitin binding protein Cbp50 from *Bacillus thuringiensis* serovar konkukian. Antonie Van Leeuwenhoek 100:445–453

Mikami B, Adachi M, Kage T, Sarikaya E, Nanmori T, Shinke R, Utsumi S (1999) Structure of raw starch-digesting *Bacillus cereus* β-amylase complexed with maltose. Biochemistry 38:7050–7061

Oyama T, Kusunoki M, Kishimoto Y, Takasaki Y, Nitta Y (1999) Crystal structure of β-amylase from *Bacillus cereus* var. mycoides at 2.2 Å resolution. J Biochem 125:1120–1130

Paldi T, Levy I, Shoseyov O (2003) Glucoamylase starch-binding domain of *Aspergillus niger* B1: molecular cloning and functional characterization. Biochem J 372:905–910

Pell G, Williamson MP, Walters C, Du H, Gilbert HJ, Bolam DN (2003) Importance of hydrophobic and polar residues in ligand binding in the family 15 carbohydrate-binding module from *Cellvibrio japonicus* Xyn10C. Biochemistry 42:9316–9323

Ryu S-I, Park C-S, Cha J, Woo E-J, Lee S-B (2005) A novel trehalose-synthesizing glycosyltransferase from *Pyrococcus horikoshii*: molecular cloning and characterization. Biochem Biophys Res Commun 329:429–436

Saha B, Mathupala S, Zeikus J (1988) Purification and characterization of a highly thermostable novel pullulanase from *Clostridium thermohydrosulfuricum*. Biochem J 252:343–348

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Schultz J, Milpetz F, Bork P, Ponting CP (1998) SMART, a simple modular architecture research tool: identification of signaling domains. Proc Natl Acad Sci U S A 95:5857–5864

Simpson PJ, Jamieson SJ, Abou-Hachem M, Karlsson EN, Gilbert HJ, Holst O, Williamson MP (2002) The solution structure of the CBM4-2 carbohydrate binding module from a thermostable *Rhodothermus marinus* xylanase. Biochemistry 41:5712–5719

Sorimachi K, Gal-Coëffet M-FL, Williamson G, Archer DB, Williamson MP (1997) Solution structure of the granular starch binding domain of *Aspergillus niger* glucoamylase bound to β-cyclodextrin. Structure 5:647–661

Southall SM, Simpson PJ, Gilbert HJ, Williamson G, Williamson MP (1999) The starch-binding domain from glucoamylase disrupts the structure of starch. FEBS Lett 447:58–60

Sugimoto H, Nakaura M, Kosuge Y, Imai K, Miyake H, Karita S, Tanaka A (2007) Thermodynamic effects of disulfide bond on thermal unfolding of the starch-binding domain of *Aspergillus niger* glucoamylase. Biosci Biotechnol Biochem 71:1535–1541

Szabó L, Jamal S, Xie H, Charnock SJ, Bolam DN, Gilbert HJ, Davies GJ (2001) Structure of a family 15 carbohydrate-binding module in complex with xylopentaose evidence that xylan binds in an approximate 3-fold helical conformation. J Biol Chem 276:49061–49065

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729

Vieille C, Zeikus GJ (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. Microbiol Mol Biol Rev 65:1–43

Williamson G, Belshaw NJ, Williamson MP (1992) O-glycosylation in *Aspergillus glucoamylase*. Conformation and role in binding. Biochem J 282:423–428

Williamson MP, Le Gal-Coëffet M-F, Sorimachi K, Furniss CS, Archer DB, Williamson G (1997) Function of conserved tryptophans in the *Aspergillus niger* glucoamylase 1 starch binding domain. Biochemistry 36:7535–7539

Ye Z, Miyake H, Tatsumi M, Nishimura S, Nitta Y (2004) Two additional carbohydrate-binding sites of β-amylase from *Bacillus cereus* var. mycoides are involved in hydrolysis and raw starch-binding. J Biochem 135:355–363