



Structural and Functional Insights Into CmGH1, a Novel GH39 Family β -Glucosidase From Deep-Sea Bacterium

Yanfang Shen^{1†}, Zhengyang Li^{1†}, Ying-Yi Huo², Luyao Bao¹, Baocai Gao¹, Peng Xiao¹, Xiaojian Hu¹, Xue-Wei Xu^{2*} and Jixi Li^{1*}

¹ State Key Laboratory of Genetic Engineering, Department of Neurology, School of Life Sciences, Huashan Hospital, Shanghai Engineering Research Center of Industrial Microorganisms, Fudan University, Shanghai, China, ² Key Laboratory of Marine Ecosystem and Biogeochemistry, Second Institute of Oceanography, State Oceanic Administration, Hangzhou, China

OPEN ACCESS

Edited by:

Baolei Jia,
Chung-Ang University, South Korea

Reviewed by:

Rameshwar Tiwari,
Ulsan National Institute of Science
and Technology, South Korea
Jing Xiao,
Qilu University of Technology, China

*Correspondence:

Xue-Wei Xu
xuxw@sio.org.cn
Jixi Li
lijixi@fudan.edu.cn

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Extreme Microbiology,
a section of the journal
Frontiers in Microbiology

Received: 16 August 2019

Accepted: 04 December 2019

Published: 20 December 2019

Citation:

Shen Y, Li Z, Huo Y-Y, Bao L,
Gao B, Xiao P, Hu X, Xu X-W and Li J
(2019) Structural and Functional
Insights Into CmGH1, a Novel GH39
Family β -Glucosidase From Deep-Sea
Bacterium. *Front. Microbiol.* 10:2922.
doi: 10.3389/fmicb.2019.02922

Glucosidases play key roles in many diseases and are limiting enzymes during cellulose degradation, which is an important part of global carbon cycle. Here, we identified a novel β -glucosidase, CmGH1, isolated from marine bacterium *Croceicoccus marinus* E4A9^T. In spite of its high sequence and structural similarity with β -xylosidase family members, CmGH1 had enzymatic activity toward *p*-nitrophenyl- β -D-glucopyranoside (*p*-NPG) and cellobiose. The K_m and K_{cat} values of CmGH1 toward *p*-NPG were 0.332 ± 0.038 mM and 2.15 ± 0.081 min⁻¹, respectively. CmGH1 was tolerant to high concentration salts, detergents, as well as many kinds of organic solvents. The crystal structure of CmGH1 was resolved with a 1.8 Å resolution, which showed that CmGH1 was composed of a canonical (α/β)₈-barrel catalytic domain and an auxiliary β -sandwich domain. Although no canonical catalytic triad residues were found in CmGH1, structural comparison and mutagenesis analysis suggested that residues Gln157 and Tyr264 of CmGH1 were the active sites. Mutant Q157E significantly increased its hydrolase activity up to 15-fold, whereas Y264E totally abolished its enzymatic activity. These results might provide new insights into understanding the different catalytic mechanism during evolution for β -glucosidases and β -xylosidases.

Keywords: CmGH1, β -glucosidase, crystal structure, characterization, *Croceicoccus marinus*, marine bacterium

INTRODUCTION

β -xylosidases (or xylan 1,4- β -xylosidases, EC 3.2.1.37) catalyze the hydrolysis of short chain xylooligosaccharide into xylose units, which play an important role in plant cell-wall hemicelluloses degradation and carbon cycle maintenance (Beg et al., 2001; Galbe and Zacchi, 2002). The depolymerization xylan to free xylose has a significant relevance for lignocellulose-based biofuels production by pentose-fermenting microorganisms and other industrial processes such as the paper

and food industries (Bosetto et al., 2016). According to the amino acid sequence and structure, β -xylosidases are generally assigned to glycoside hydrolase (GH) families 1, 3, 30, 39, 43, 51, 52, 54, 116, and 120 (Henrissat and Davies, 1997). With the development of industry, it is of great interest to explore different β -xylosidases with compatible stability in harsh conditions from bacteria and fungi (Joseleau et al., 1992; Jordan and Wagschal, 2010; Lagaert et al., 2011).

Croceicoccus marinus E4A9^T, the type species of the genus *Croceicoccus*, was isolated from deep-sea sediment at the East Pacific polymetallic nodule region (5,280 m depth, 2°C, 34 ‰ salinity) (Xu et al., 2009). The enzymes from marine bacteria might have good stress tolerance (Jiang et al., 2012; De Santi et al., 2016; Huang et al., 2016). A novel glycosidase gene *cmgh1* was identified based on the *in silico* analysis (Kim et al., 2004; de Pascale et al., 2008). Sequence analysis of the *cmgh1* gene showed that it shared 51% identity with the β -xylosidase from *Microbacterium testaceum* StLB037^T, indicating that CmGH1 might be a new β -xylosidase belonging to the GH39 family (Morohoshi et al., 2011).

Different with β -xylosidase, the β -glucosidases (or β -D-glucoside glucohydrolase, EC 3.2.1.21) have specific substrate-binding pattern and catalytic activities, and facilitate release of aromatic compounds, phytohormone activation, as well as cell wall oligosaccharide recycling that hydrolyze exogenous glucosides, glycolipids and oligosaccharides into β -D-glucose in microbials, animals and plants (Bhatia et al., 2002; Ketudat Cairns and Esen, 2010; Ketudat Cairns et al., 2015). β -glucosidases play key roles in the synthesis or hydrolysis of glycosidic bonds, which is the rate-limiting step in sugar metabolism and biofuels production (Bhatia et al., 2002; Singhania et al., 2013; Kuusk and Valjamae, 2017). β -glucosidases are well-characterized and generally classified into many GH families 1, 2, 3, 5, 9, 30, 39, and 116 (Henrissat and Davies, 1997). However, there is no β -glucosidase reported in GH39 family and the structural basis for substrate specificity is barely investigated. Here, we present the characterizations and crystal structure of CmGH1, which showed β -glucosidase activity, instead of β -xylosidase activity, to provide a structural basis for the development, and utilization of the marine sourced β -glucosidase.

RESULTS

Biochemical Properties of CmGH1

A putative open reading frame of 1458 bp (*cmgh1*), encoding a protein of 485 aa (CmGH1) with a theoretical molecular weight of 53.48 kDa and pI of 4.66, was identified from the whole-genome of strain *C. marinus* E4A9^T. According to the phylogenetic tree, the protein CmGH1 belongs to the GH39 family (Supplementary Figure S1). To investigate the catalytic characterization of CmGH1, the recombinant protein CmGH1 with a N-terminal His₆-SUMO tag was expressed in *Escherichia coli* BL21 (DE3) cells. After removal of the His₆-SUMO tag with Ulp1 enzyme as described in previous study (Kuang et al., 2017), the target protein CmGH1 was purified into 95% homogeneity by gel filtration chromatography

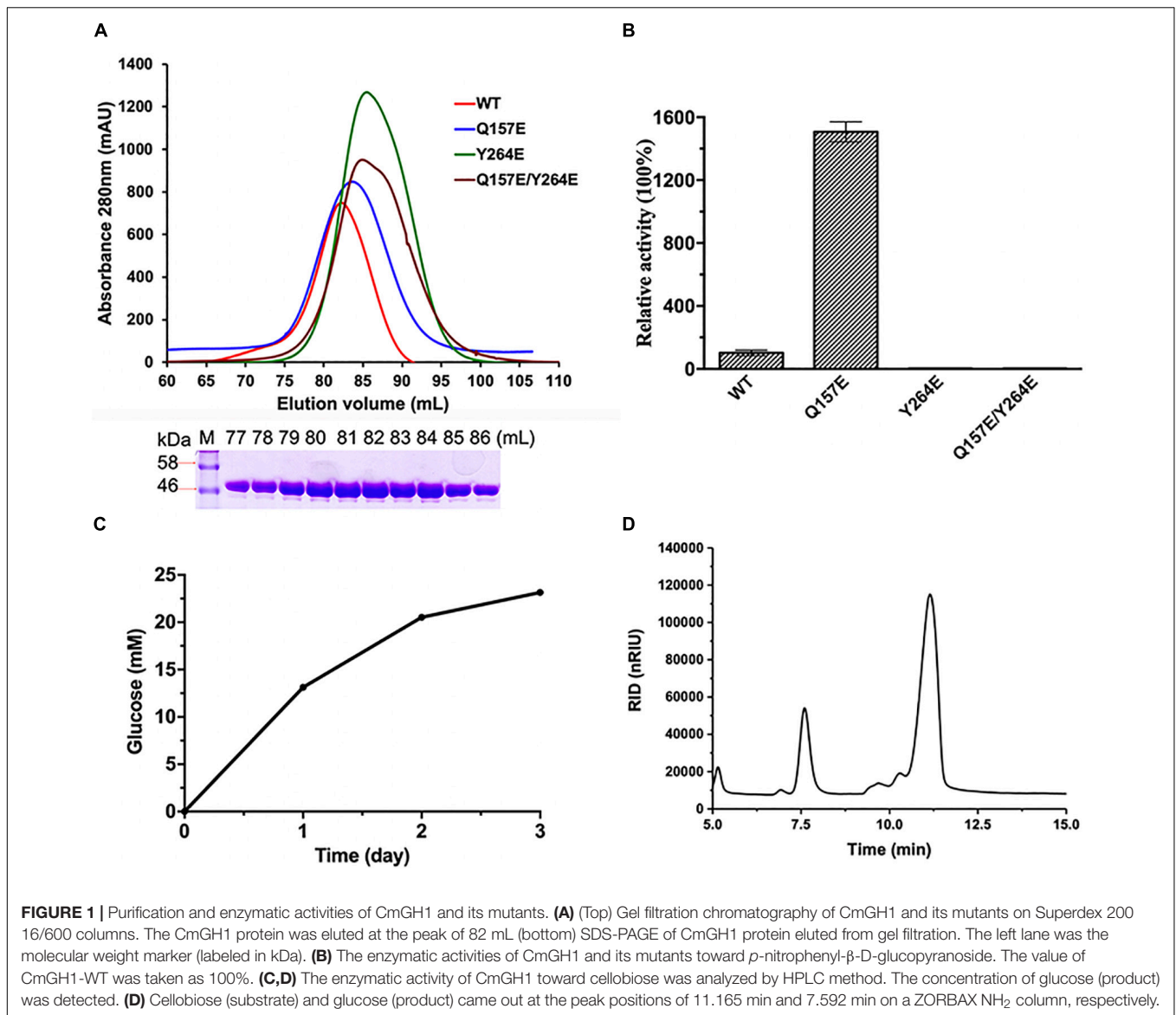
(Figure 1A). CmGH1 was firstly predicted to be a β -xylosidase based on sequence analysis, however, CmGH1 did not show β -xylosidase activity when using *p*-nitrophenyl- β -D-xylopyranoside, xylobiose and xylotriose as substrates, with the high sensitive HPLC method and 3,5-dinitrosalicylic acid (DNS) method (McCleary and McGeough, 2015). Therefore, CmGH1 was evaluated for saccharification potential with different substrates, which turned out that CmGH1 showed β -glucosidase activity when using *p*-nitrophenyl- β -D-glucopyranoside (*p*-NPG) and cellobiose as substrates (Figures 1B–D and Supplementary Figure S2).

The optimum reaction conditions of CmGH1 toward *p*-NPG were determined over a pH range of 6.0–11.0 and a temperature range of 25–65°C. CmGH1 showed the highest catalytic activity at pH 9.0 and 55°C (Figures 2A,B). The enzyme activity of CmGH1 was 1.46×10^{-2} U/mg, and the K_m and K_{cat} values were 0.332 ± 0.038 mM and 2.15 ± 0.081 min⁻¹, respectively.

Furthermore, various kinds of divalent cations, organic solvents and detergents were added into the reaction buffer to investigate the tolerance of CmGH1 (Figures 2C–F). The β -glucosidase activity of CmGH1 was completely abolished with the additions of Zn²⁺, Ni²⁺, Cd²⁺, Cu²⁺, and Ba²⁺, and only about 10–15% activity was remained with Ca²⁺, Sr²⁺, and Co²⁺. CmGH1 could tolerate Mn²⁺ and Mg²⁺ with the relative activity values at 64% and 88%, respectively. In addition, 10 mM EDTA severely suppressed the activity of CmGH1 which remained 10% relative activity (Figure 2C). Most detergents and organic solvents had little impacts on the catalytic reaction of CmGH1. The β -glucosidase activities were comparative with the blank under the addition of 1% detergents (Triton X-T14, Triton X-100, Tween20, Tween80, and SDS) and 5% organic solvents (DMSO, DMF, glycerin, methanol, acetone, acetonitrile, alcohol, and isopropanol). The activities attenuated significantly under conditions with SDS or Tween20 at 5% concentration, as well as glycerin, alcohol, isopropanol, DMF, methanol, and acetonitrile at 15% concentrations. However, 15% DMSO and acetone displayed little effects on the activity of CmGH1 (Figures 2D,E). Moreover, CmGH1 had enzymatic activity (about 30%) with addition of 2 M NaCl, but abolished its activity at higher salt concentration (Figure 2F). CmGH1 remained over 30% enzymatic activity with the presence of 200 mM glucose, however, completely abolished its activity at 500 mM or higher concentration (Figure 2G). In addition, CmGH1 reached its maximum activity with 1.0 mM *p*-NPG or higher concentration (Figure 2H).

Overall Structure of CmGH1

The crystal structure of CmGH1 was solved with a 1.8 Å resolution (Table 1). Diffraction dataset was integrated into the monoclinic space group P2₁ with two molecules per asymmetric unit. The two chains were identical and the root-mean-square deviation (RMSD) value for its backbone was 0.211 Å. The structure of CmGH1 was refined to the satisfied R_{work} and R_{free} values of 14.71% and 18.94%, respectively. The first four residues at the N-terminal region



and the residues from Val396 to Thr403 were invisible in chain A but were refined unambiguously in the electron-density map of chain B. Therefore, we will discuss the Chain B structure only.

CmGH1 was composed of the N-terminal domain (NTD), the C-terminal domain (CTD), and the middle catalytic domain (Figure 3A). The CmGH1 structure had eight α -helices and twenty β -stands, which combined into a classical $(\alpha/\beta)_8$ -barrel catalytic core domain and a β -sandwich accessory domain (Figure 3B). The core domain included eight parallel β -stands, namely β 2 (Leu53-Arg55), β 3 (Asn104-Ile109), β 4 (His151-Phe154), β 5 (Lys191-Val197), β 6 (Phe223-Tyr229), β 7 (Glu259-Leu262), β 7 (Tyr264-Ser266), and β 9 (Lys298-Phe301), surrounded by eight α -helices, namely α A (Leu43-Asn48), α B (Tyr86-Ser100), α C (Leu123-Val138), α D (Pro169-Val186), α E (Glu208-Gln216), α F (Asp238-Lys252), α G (Asp277-Met291), and α H (Ala329-Leu341). A β -hairpin motif consisting of β 13

(Val391-Ile398), β 14 (Gln401-Asn408) and the loop between these two β -stands protruded out from the barrel. The accessory domain included β 1 (Ile6-Asn12), β 10 (Thr344-Arg346), β 11 (Phe357-Arg363), β 12 (Thr369-Asn376), β 15 (Gly423-Gly429), β 16 (His438-Asp446), β 17 (His449-Glu452), β 18 (Leu454-Gly459), β 19 (Val464-Lys469), and β 20 (Gly474-Arg482) (Figure 3B).

Structural Comparison of CmGH1 With Other GH39 Family Members

Several GH39 family β -xylosidase crystal structures have been reported, including XynB from *Thermoanaerobacterium saccharolyticum* [PDB_ID: 1PX8 (Yang et al., 2004)], XynB1 from *Geobacillus stearothermophilus* (PDB_ID: 1W91), CcXynB2 from *Caulobacter crescentus* [PDB_ID: 4EKJ (Santos et al., 2012)], PslG from *Pseudomonas aeruginosa*

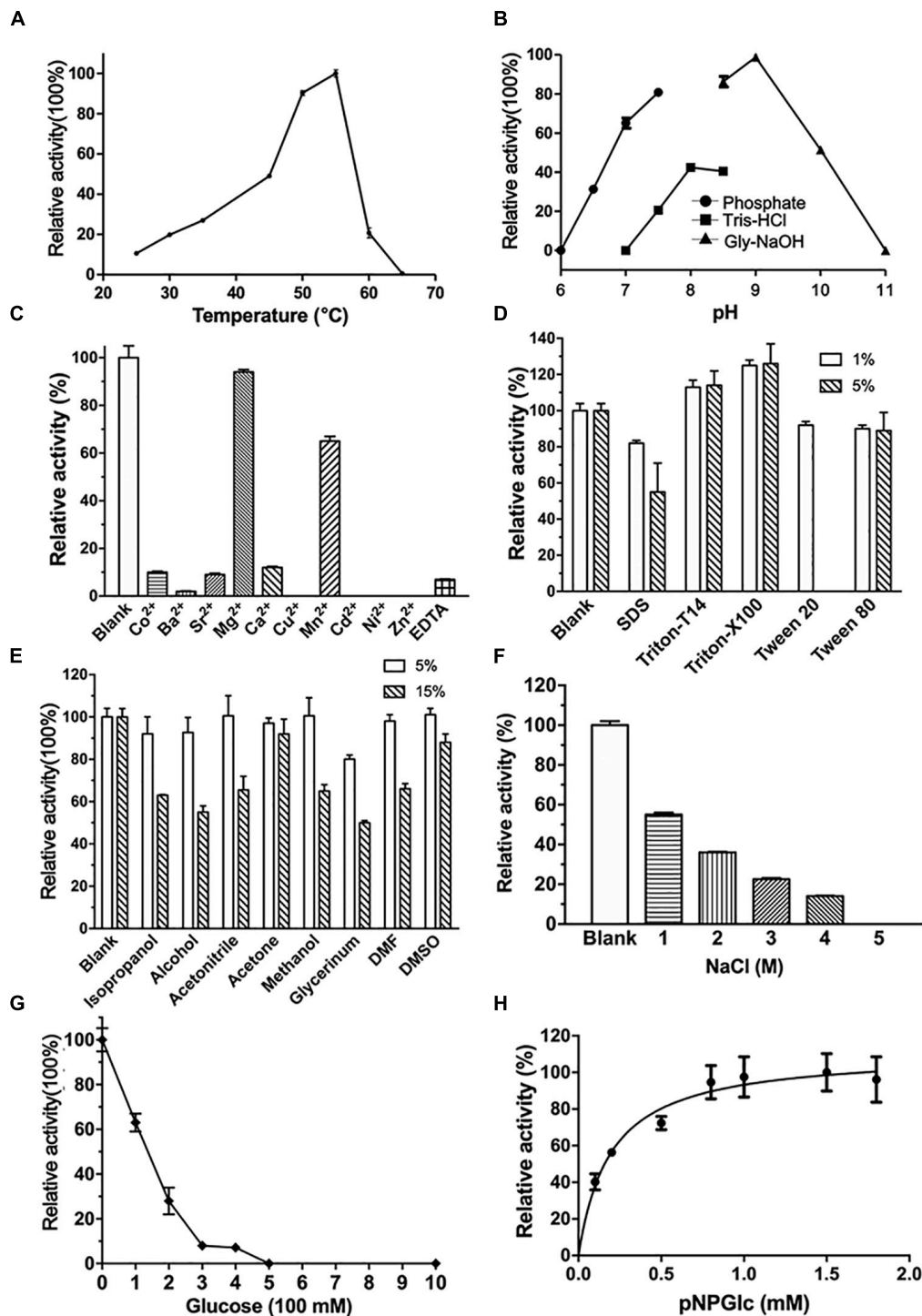


FIGURE 2 | Enzymatic characterization of CmGH1. **(A)** Effects of temperature on enzyme activity. The value obtained at 55°C was taken as 100%. **(B)** Effects of pH on enzyme activity. The value obtained at pH 9.0 was taken as 100%. **(C)** Effects of different metal ions on the enzymatic activity. The values obtained without ions in the reaction mixture were taken as 100%. **(D)** Effects of different detergents on the enzymatic activity. All kind of detergents were added at the final concentration of 1% and 5%, respectively. The values obtained without ions in the reaction mixture were taken as 100%. **(E)** Effects of different organic solvents on the enzymatic activity. All kind of organic solvents were added at the final concentration of 5% and 15%, respectively. The values obtained without ions in the reaction mixture were taken as 100%. **(F)** Effects of NaCl concentration on the activities. The value obtained without NaCl in the reaction mixture was taken as 100%. **(G)** Effects of products (glucose) concentration on the activities. The value obtained without glucose in the reaction mixture was taken as 100%. **(H)** Effects of substrate *p*-nitrophenyl- β -D-glucopyranoside on the activities. The value obtained with 1.5 mM *p*-nitrophenyl- β -D-glucopyranoside was taken as 100%.

TABLE 1 | Data collection and refinement statistics of CmGH1.

Items	CmGH1-Se
Data collection	
Wavelength	0.9793
Resolutions (Å)	48.35–1.80 (1.85–1.80) ^a
Space group	P 2 ₁
Unit cell (Å, °)	$a = 48.911, b = 95.911, c = 100.908$ $\alpha = \gamma = 90, \beta = 98.712$
Unique reflections	82639 (8048)
Completeness (%)	97.31 (95.41)
R_{merge} (%) ^b	13.7 (58.5)
I/σ (I)	15.03 (3.67)
Number of Se atoms	18
Refinement statistics	
Resolutions (Å)	48.35–1.80 (1.85–1.80)
Reflection used in refinement	82639 (8048)
R_{work} (%) ^c	14.71 (26.14)
R_{free} (%) ^d	18.94 (31.82)
Number of glycerol	2
Number of water	560
Number of protein residues	961
RMSD	
Bond lengths (Å)	0.007
Bond angles (°)	0.84
Average B-factor (Å ²)	19.75
Ramachandran favored (%)	96.43
Ramachandran allowed (%)	3.57
Ramachandran outliers (%)	0
Rotamer outliers (%)	0
PDB code	5Z3K

^aStatistics for the highest-resolution shell are shown in parentheses. ^b $R_{\text{merge}} = \sum |I_i - \bar{I}| / \sum I_i$, where I_i is the intensity of an individual reflection and \bar{I} is the average intensity of that reflection. ^c $R_{\text{work}} = \sum ||F_o| - |F_c|| / \sum |F_o|$, where F_o and F_c are the observed and calculated structure factors for reflections, respectively. ^d R_{free} was calculated as R_{work} using the 5% of reflections that were selected randomly and omitted from refinement.

[PDB_ID: 4ZN2 (Yu et al., 2015)] and GH39wh2 from *Bacteroides cellulosilyticus* (PDB_ID: 5JVK [Ali-Ahmad et al., 2017]). These structures share similar architectures, including a conservative (α/β)₈-barrel catalytic domain and a β -sandwich domain at the C-terminal region (Figures 4, 5A,B). Moreover, XynB and XynB1 have an additional long α -helix-containing auxiliary domain at the C-terminal region (Yang et al., 2004; Czjzek et al., 2005; Santos et al., 2012).

Structural superimposition showed that CmGH1 shared a similar architecture with other GH39 family β -xylosidases, despite that PslG and GH39wh2 only shared 17% and 13% identities with CmGH1 (Figures 4, 5A,B). PslG has one conservative (α/β)₈-barrel catalytic domain and one β -sandwich domain at C-terminal region, whereas GH39wh2 has one (α/β)₈-TIM barrel and two β -sandwich domains at N-terminal and C-terminal region. The N-terminal β -sandwich domain could afford minimal flexibility requiring to provide an additional substrate-binding interface and likely to strengthen substrate

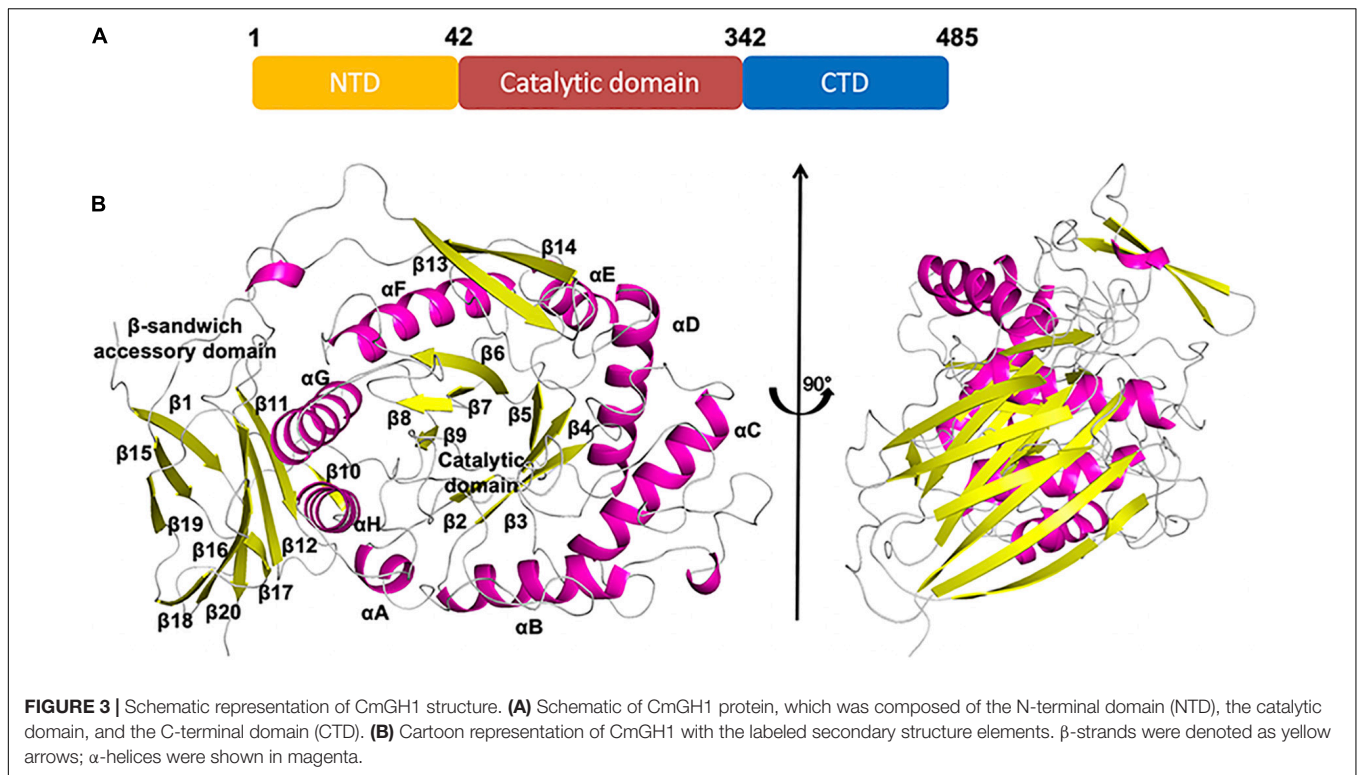
recognition (Ali-Ahmad et al., 2017; Figure 5B). In addition, structural comparisons among CmGH1, RlGH from *Rhizobium leguminosarum* (PDB_ID: 5NDX), and CvGH from *Caulobacter vibrioides* (PDB_ID: 4M29) showed that CvGH shared similar architecture with CmGH1, whereas RlGH had one more unknown function β -sandwich domain at C-terminal region (Supplementary Figure S3).

Catalytic Center and Active Sites of CmGH1

The catalytic pocket of GH39 family enzymes, such as PslG and GH39wh2, are surrounded by negative charged residues (Ali-Ahmad et al., 2017). The corresponding area in CmGH1 showed the same predominant negative surface potential cluster (Figure 6). Among GH39 family members, PslG and GH39wh2 acted as endoglycosidase with a large groove, whereas XynB, XynB1, CcXynB2, and α -L-iduronidase might act as exoglycosidase because of the small groove, which can only accept oligosaccharides as substrates (Yang et al., 2004; Santos et al., 2012). The groove distances among residues Phe114 and Tyr312, Pro271 and Leu163 of CmGH1 were 6.6 Å and 15.0 Å, respectively, whereas the corresponding distances in the groove of GH39wh2 were 19.3 Å and 22.4 Å (Tyr210 and Glu415, Asn251, and Asn370), and those in XynB were 5.4 Å and 14.6 Å (Tyr116 and Arg324, Tyr282 and Phe166) (Figure 6). Therefore, CmGH1 might be a new exoglycosidase when comparing with GH39wh2 and XynB.

The active sites of CmGH1 might locate within the central (α/β)₈ TIM barrel by structure comparison with other members. The active sites of XynB include the conserved Glu160 and nucleophile Glu277, whereas the corresponding residues in CmGH1 were Gln157 and Tyr264 (Figures 4, 7). Although no glutamic acid was found in the catalytic pocket of CmGH1, it might use a similar catalytic mechanism with the other members of GH39 family. Besides the active sites, CmGH1 did not have any conserved substrate binding residues with those reported in XynB (Ile151-Leu164) (Figure 4). When docking and computing CmGH1 with different substrates (cellobiose and xylobiopyranose) and products (D-glucose and D-xylopyranose) using AutoDock Tools4 program (Morris et al., 2009), the binding energy and inhibition constants of the CmGH1 complexes with cellobiose and glucose were lower than that with xylobiopyranose and xylopyranose, which meant CmGH1 might prefer to use cellobiose and glucose as substrates (Supplementary Figure S4 and Supplementary Table S1).

To verify the possible active sites of CmGH1, site-directed mutagenesis was performed to identify their roles in enzymatic activity. The mutant Q157E had 15 folds higher catalytic activity toward *p*-NPG than wild-type CmGH1, whereas the mutant Q157A had a similar value with wild-type CmGH1. The enzyme activity of CmGH1-Q157E was 2.04×10^{-1} U/mg, while the K_m , V_{max} , and K_{cat} values were 0.3542 ± 0.042 mM, 109.5 ± 3.29 $\mu\text{M}/\text{min}$, and 29.28 ± 0.88 min^{-1} , respectively. However, Y264E and Q157E/Y264E mutants abolished the enzymatic activities of CmGH1, which suggested that Tyr264 could contribute a vital role in the enzyme activity (Figure 1B).



DISCUSSION

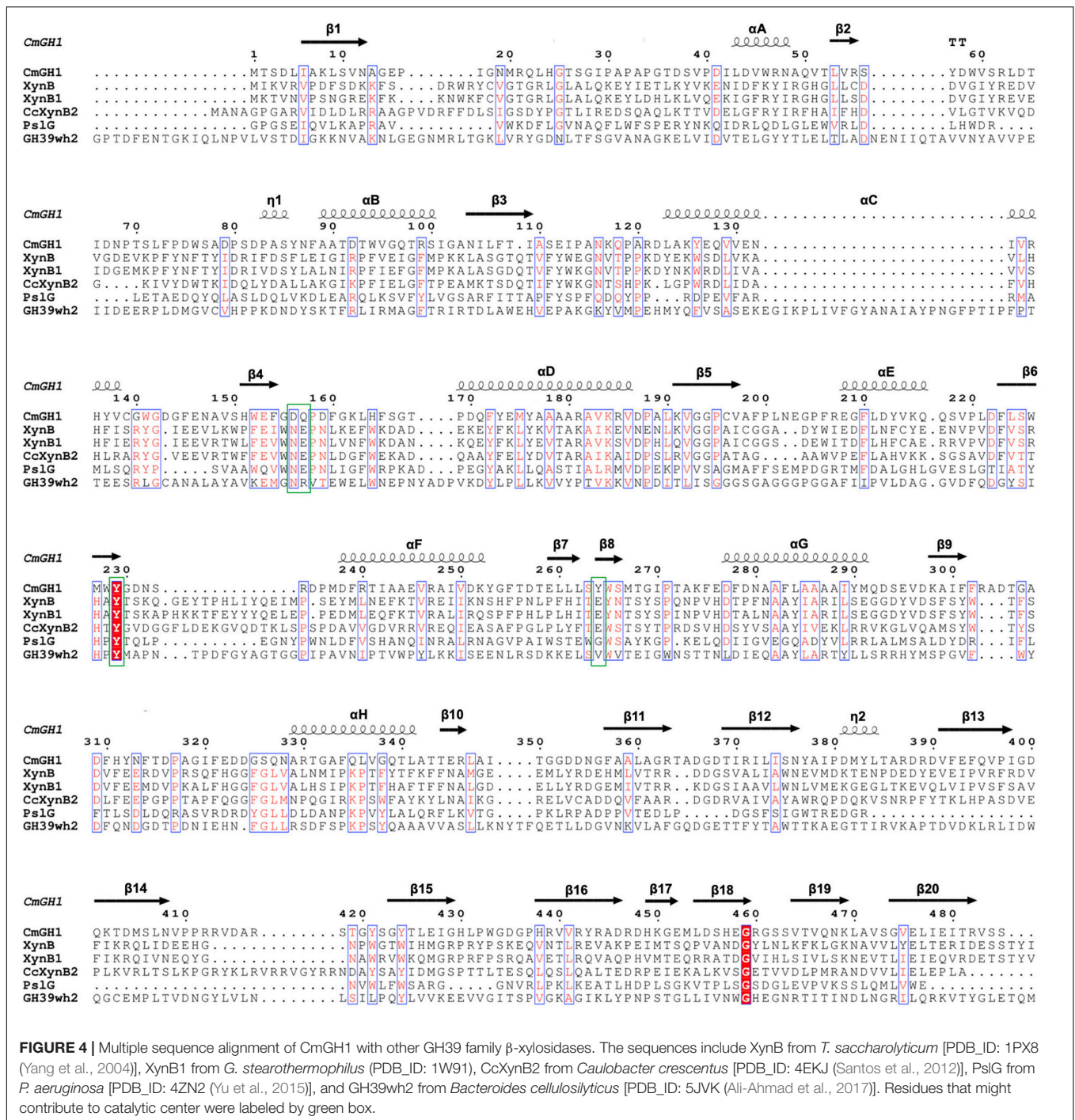
Glycosidases are a group of evolutional conserved enzymes that cleave the glycosidic bonds and have been classified into over 63 families according to their sequences (Sinnott, 1990; Henrissat and Bairoch, 1996). Based on substrate specificity, glycosidases are consisted of β -xylosidase, β -D-glucosidase, β -mannosidase, myrosinase, and so on (Mian, 1998). Although CmGH1 was predicted as a β -xylosidase, the biochemical results suggested CmGH1 might be a β -D-glucosidase belonging to the GH39 family. The auto-docking result also showed that CmGH1 preferred to combine cellobiose and glucose, which had lower binding energy and inhibition constant compared with xylobiopyranose and xylopyranose. Most characterized β -D-glucosidase showed highest activity in the acidic pH range and remained low activity under alkaline environment (Bhatia et al., 2002; Sorensen et al., 2013). CmGH1 had maximal activity at pH 9.0 and remained over 50% activity at pH 10.0 (Figure 2B). This uncommon feature might be due to the host line *C. marinus*, which lives in the marine environment and prefers a mildly alkaline condition. This property might have a potential application in many industrial processes of which the high pH catalytic condition required. The catalytic activity of Q157E was 15 times higher than that of wild type CmGH1, most probably because that Gln157 acted as an auxiliary catalytic proton donor and glutamine had a weaker proton donating ability than glutamic acid (Brooks et al., 2001; Vocadlo et al., 2002). The bacteria *C. marinus* E4A9^T might prefer a weaker acid/base residue to keep an insufficient catalytic activity, which was related to the evolution divergence and was an

adaptive selection for the extreme environment in deep sea (Xu et al., 2009).

As Y264E and Q157E/Y264E mutants abolished the enzymatic activities (Figures 1A,B), this suggested that Tyr264 might play a vital role in the catalytic site. In the GH83 and GH143 families, one Tyr residue was reported as nucleophile in the catalytic center. In corresponding with Glu277 of XynB, the Tyr264 of CmGH1 could act as a nucleophile during catalysis (Crennell et al., 2000; Ndeh et al., 2017).

As CmGH1 had the optimum pH at 9.0 (Gly-NaOH buffer) and abolished the enzymatic activities at acidic conditions (pH < 6.0), we suggest that Tyr264 might act as a nucleophile residue only under alkaline conditions. Although it is not clear how many amino acids are involved in the catalytic process of CmGH1, we found that Asp156, Glu157, Tyr229, Arg55, and Tyr264 might play a major role in catalytic reaction by superimposition with the substrate D-xylose in XynB (Figures 4, 7).

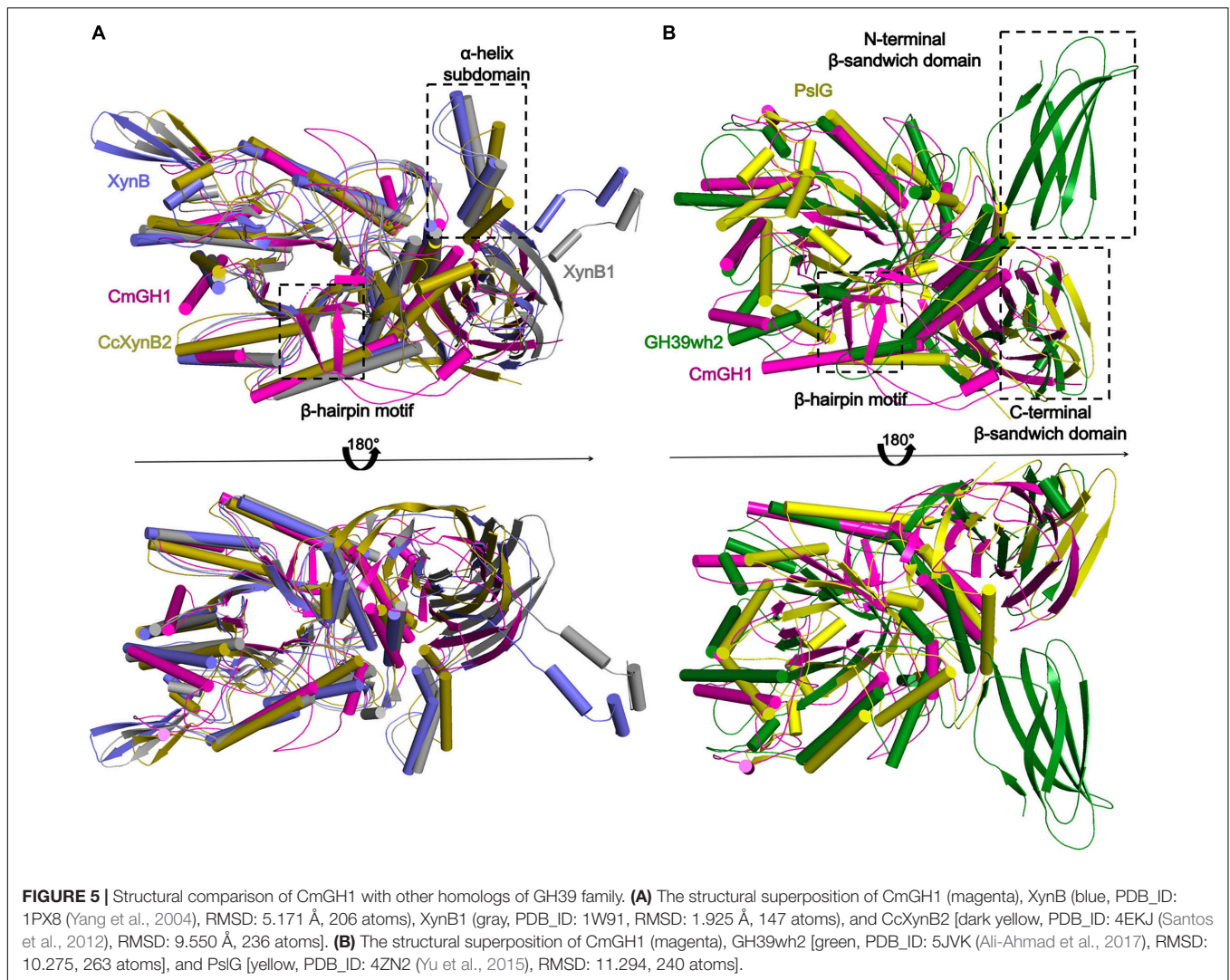
The superimposed CmGH1 with other members of GH39 families showed that there were some differences in architectures, apart from the conserved $(\alpha/\beta)_8$ -TIM barrel domain and β -sandwich domain. For example, the β -hairpin motif of CmGH1 extended from the catalytically active pocket and held onto one end, which was not reported in GH39 family. In CcXynB2, there was a β -hairpin motif connecting the sixth loop β 7 to α F and interacted with a long α -helix-containing loop that was the only one of the reported xylosidases of GH39 through polar and hydrophobic effects, which induced the β -hairpin motif adopting an open conformation (Santos et al., 2012). However, the β -hairpin motif of CmGH1 was far away from the nearby



small α -helix, indicating that there was no interaction between them and the β -hairpin motif adopted a closed conformation, which might be related to its substrate recognition (Figure 5).

The substrate-binding groove in CmGH1 was slightly larger than that in XynB and smaller than that in GH39wh2, which could be inferred by different states, as XynB and XynB1 appeared to be a tetramer in solution, whereas CcXynB2, Ps1G, and GH39wh2 formed monomers in solution (Yang et al., 2004;

Czjzek et al., 2005; Santos et al., 2012; Baker et al., 2015; Ali-Ahmad et al., 2017). In addition, the GH1 family members share (α/β)₈ barrel catalytic domain and use Glu as the catalytic residue, whereas the GH3 family members have no specific catalytic domain and use Glu or His or Asp as catalytic acid-base and nucleophile residue, respectively (Rye and Withers, 2000; Vasella et al., 2002). By compared with these glucosidases, CmGH1 did not belong to any known family of β -D-glucosidases.



In conclusion, CmGH1 showed glucosidase activity toward different substrates *p*-NPG and cellobiose, despite of that it shared sequence similarity with xylosidase. These findings provide new insights into understanding the different catalytic mechanism during evolution for β -glucosidases and β -xylosidases, as well as offer a structural and theoretical basis for modification of industrial enzymes.

MATERIALS AND METHODS

Bacterial Strains, Plasmids, and Media

Croceicoccus marinus E4A9^T, a member of *Erythrobacteraceae* family, was isolated from a deep-sea sediment sample that was collected from the East pacific polymetallic nodule region, and was cultivated in 2216 marine broth (BD, United States) at 30°C (Xu et al., 2009). Plasmid pSMT3 was stored in our lab and used as the vector for gene cloning and expression. *E. coli* strain DH5 α and BL21 (DE3) plysS (Transgen, China) were used as the host for cloning and protein expressing. *E. coli* cells were

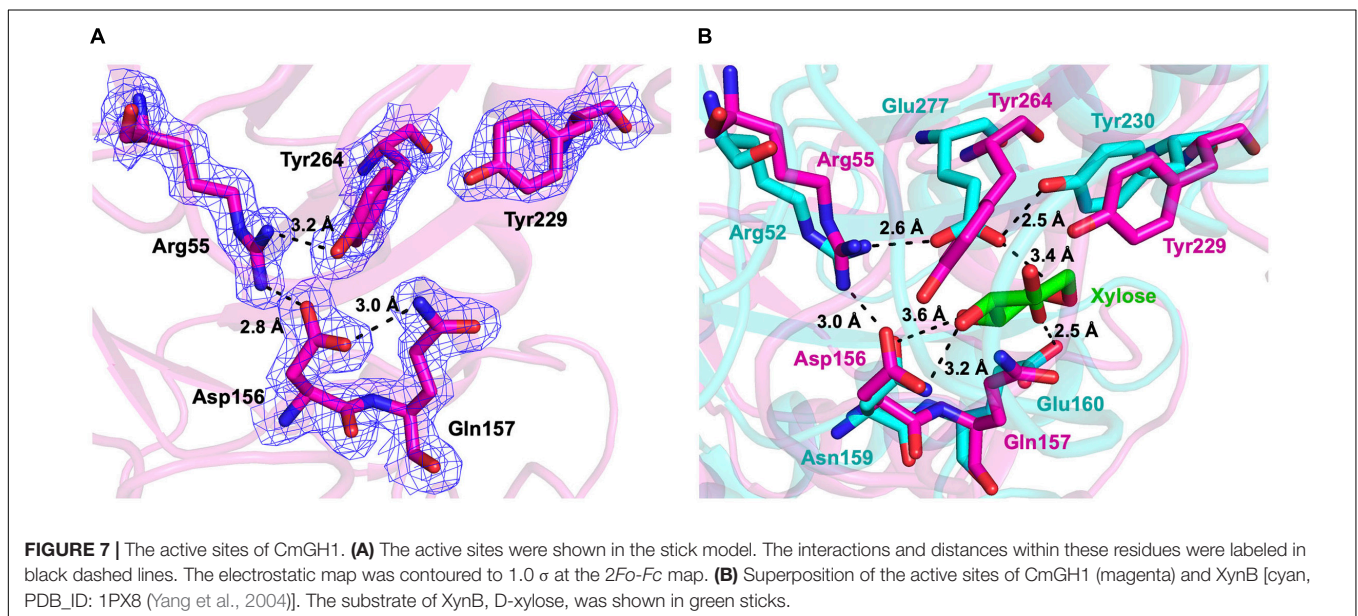
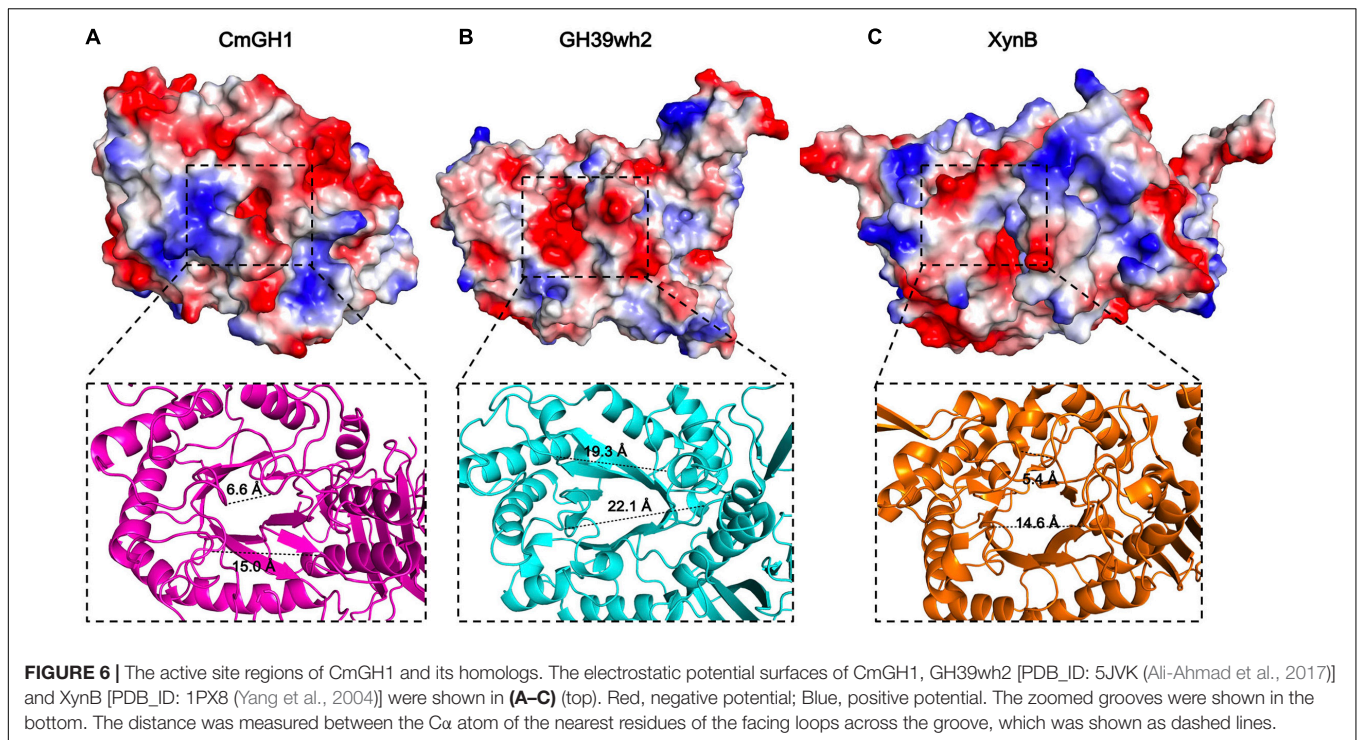
grown at 37°C in LB medium containing 10 g NaCl, 10 g tryptone and 5 g yeast extract (Sigma-Aldrich, United States) per liter at pH 7.0, and LB agar medium was added with 1.5–2.0% (w/v) agar.

Sequence Analysis

Multiple alignment of amino acid sequences of homologs was performed using ClustalX v.2 program (Larkin et al., 2007). Second structure alignment was generated by the ESript v.3.0 server (Robert and Gouet, 2014). The phylogenetic tree was constructed by neighbor-joining method using MEGA (Molecular Evolutionary Genetics Analysis) v.10.1 software (Felsenstein, 1985; Saitou and Nei, 1987; Kumar et al., 2018).

Cloning, Protein Expression, and Purification

The gene *cmgh1* (accession number: WP_066849948) from *C. marinus* was inserted into the pSMT3 vector. The wild type and mutants of CmGH1 plasmids were transformed into *E. coli* BL21



(DE3) *plysS* cells for protein expression. 0.5 mM isopropyl β -D-thiogalactoside was added into cells to induce protein expression at 16°C for 20 h when the OD₆₀₀ reached 0.8. Then the cells were harvested by centrifugation at 6000 rpm for 10 min at 4°C. The cell pellets were resuspended in a lysis buffer (50 mM Tris-HCl, 500 mM NaCl, 10 mM imidazole, 5% glycerol, 2 mM β -ME, pH 8.0) and were disrupted using a high-pressure homogenizer. The supernatant was purified by NTA affinity chromatography. After cleavage of the His₆-sumo tag with Ulp1 enzyme, the target proteins were obtained in the flow-through fractions.

Subsequently, CmGH1 protein was purified by gel-filtration chromatography (Superdex 200 16/600, GE, United States). The fractions were determined by SDS-PAGE and the concentration was measured by the method of Bradford with bovine serum albumin (BSA) as a standard (Bradford, 1976).

Selenomethionine (Se-Met) substituted CmGH1 protein was expressed as mentioned above. When OD₆₀₀ reached 0.4, the cells were harvested by centrifugation and resuspended in 100 mL M9 medium (47.7 mM Na₂HPO₄, 22 mM KH₂PO₄, 8.6 mM NaCl, and 28.2 mM NH₄Cl). The resuspended cells were centrifuged

and transferred into 500 mL fresh M9 medium supplemented with 50 µg/mL kanamycin and 30 mg/L Se-Met. After growing at 37°C for 1 h, the temperature was lowered to 16°C and the protein expression was induced by adding 0.5 mM IPTG for an additional 18 h. Then, the cells were harvested by centrifugation and purified as mentioned above.

Biochemical Characterization of CmGH1

The enzymatic activity of wild type CmGH1 and mutants were tested by *p*-nitrophenol method and DNS method (McCleary and McGeough, 2015). The enzymatic activities toward xylopyranoside and glucopyranoside were measured using *p*-nitrophenyl-β-D-xylopyranoside (*p*-NPX) and *p*-nitrophenyl-β-D-glucopyranoside (*p*-NPG) as substrates, respectively. The standard reaction buffer consisted of 100 µM purified CmGH1 and 1mM *p*-NPX (or *p*-NPG) in 200 µl buffer with different pH varying from 3.0 to 11.0, including 100 mM citrate buffer (pH 3.0-pH 6.5), 100 mM potassium phosphate buffer (pH 6.5-pH 7.5), 100 mM Tris-HCl buffer (pH 7.5-pH 9.0), and 100 mM Gly-NaOH buffer (pH 9.0-pH 11.0). The enzyme activity was determined by measuring the amount of released *p*-nitrophenol from 293K to 333K at 405 nm using SpectraMax M5 (Molecular Devices, United States). The absorbance of 405 nm was measured every 5 min and the total reaction time is 5 h.

The influences of cations on enzyme activity were examined in the presence of 10 mM Ba²⁺, Ca²⁺, Co²⁺, Cu²⁺, Mg²⁺, Mn²⁺, Ni²⁺, Sr²⁺, Zn²⁺, and the chelating agent EDTA. The effects of organic solvents were tested by using 5% and 15% isopropanol, alcohol, acetonitrile, acetone, methanol, glycerinum, dimethylformamide (DMF), or dimethyl sulfoxide (DMSO). The effects of detergents were determined by using 1% or 5% Tween20, Tween80, Triton X-100, or SDS. All measurements were performed in 100 mM Tris-HCl buffer (pH 7.5), and the enzyme activity in the blank group was defined as 100% without additives. The salt tolerance of CmGH1 was determined by adding 1 M, 2 M, 3 M, 4 M or 5 M NaCl to 100 mM Gly-NaOH buffer (pH 9.0). The enzyme activity in the blank group was defined as 100% without NaCl. The tolerance of products on enzymatic activities was measured in the presence of 0.1, 0.2, 0.3, 0.4, 0.5, and 1 M glucose. The influences of substrate were determined by adding 0.1, 0.2, 0.5, 0.8, 1.0, 1.5, and 1.8 mM *p*-NPG.

In the reaction system of DNS method, the amount of xylobiose or xylotriose was 1 mM, and the amount of CmGH1 was 0.1, 1, 10, and 50 mM, respectively. The reaction buffer pH varied from 3.0 to 11.0, including 100 mM citrate buffer (pH 3.0 -pH 6.5), 100 mM potassium phosphate buffer (pH 6.5 -pH 7.5), 100 mM Tris-HCl buffer (pH 7.5-pH 9.0), and 100 mM Gly-NaOH buffer (pH 9.0 -pH 11.0). The reaction temperature was 20–60°C, and the reaction time was 10 min, 30 min, 1 h, 2 h, 12 h, 24 h, 48 h, and 1 week. DNS reagents, which contained 10 g 3,5-dinitrosalicylic acid, 2 g phenol, 0.5 g sodium sulfide, 10 g sodium hydroxide and 80 g potassium sodium tartrate, were added to the completed reaction system and incubated in boiling water bath for 5 min. Then the solution was quickly cooled to room temperature. The reaction mixture

was measured with the absorbance at 540 nm wavelength using SpectraMax M5. The standard curve was measured with xylose at various concentration according to the above experimental method and the linear equation was $Y = 0.6567X + 0.0643$, $R^2 = 0.9984$.

The HPLC method was used to detect enzymatic activity of CmGH1 toward cellobiose and xylobiose. The reaction mixture contained 250 mM cellobiose (or 170 mM xylobiose), 1 mg/mL CmGH1 and 100 mM Gly-NaOH buffer (pH 9.0). The catalytic reaction was performed at 55°C for various times (1, 2 and 3 days for cellobiose, and 12 h for xylobiose). BSA (1mg/mL), instead of CmGH1, was added into the reaction mixture as a negative control. Ethanol was added into the mixture to quench this reaction. The samples were centrifuged for 10 min at 17,000 g and the supernatants were detected by HPLC on a ZORBAX Original 70 Å NH₂ column (Agilent, United States) with 75% acetonitrile as a mobile phase. The concentration of glucose was calculated by using the standard curve, $Y = 3 \times 10^{-5}X + 1.1617$ (Y: glucose concentration, X, peak area, $R^2 = 0.9978$).

Crystallization and Data Collection

The Se-Met substituted CmGH1 protein was crystallized using the hanging drop vapor diffusion method at 291K by mixing 1 µL of 5 mg/mL protein with 1 µL of reservoir solution, including 0.2 M ammonium acetate, 0.1 M Bis-Tris pH 5.5 and 25% PEG3350. Diffraction data were collected with crystals flashing-frozen in crystallization buffer supplemented with 25% (v/v) glycerol. The Se-Met derivative data sets were collected at BL17U1, 18U1, and 19U1 beamlines of the Shanghai Synchrotron Radiation Facility (Wang et al., 2018). Diffraction data were integrated and scaled using software HKL2000 or HKL3000 (Otwinowski and Minor, 1997; Minor et al., 2006).

Structure Determination and Refinement

The structure of CmGH1 was solved by single wavelength anomalous diffraction (SAD) method with a resolution 1.8 Å. In brief, the diffraction data were scaled with XDS and merged using AIMLESS from CCP4 program suite. A total of 18 Se atoms were found with the autobuild software from Phenix package (Adams et al., 2010). The automatic and manual refinement were performed using REFMAC5 and Coot softwares, respectively (Murshudov et al., 1997; Emsley and Cowtan, 2004). Finally, the R_{free} and R_{work} were refined to 0.1894 and 0.1471, respectively. The final model was checked by Procheck software and was deposited in the Protein Data Bank with ID number: 5Z3K. The refine statistics were summarized in **Table 1**. The structural models of CmGH1 complex with substrates and products were built using AutoDock Tools4 program¹ (Morris et al., 2009).

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the structure of CmGH1 was deposited in Protein data bank (PDB, <http://www.rcsb.org/>). The PDB id is 5Z3K.

¹<http://autodock.scripps.edu/>

AUTHOR CONTRIBUTIONS

YS, ZL, Y-YH, LB, PX, BG, XH, and X-WX performed the experiments and analyzed the data. JL designed the study and wrote the manuscript.

FUNDING

This work was supported by grants from the National Key Research and Development Project of China (2016YFA0500600 and 2015CB943300), the National Natural Science Foundation of China (31470724 and 41506183), and the Shanghai Committee of Science and Technology (18430711400).

REFERENCES

- Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 66(Pt 2), 213–221. doi: 10.1107/S0907444909052925
- Ali-Ahmad, A., Garron, M. L., Zamboni, V., Lenfant, N., Nurizzo, D., Henrissat, B., et al. (2017). Structural insights into a family 39 glycoside hydrolase from the gut symbiont *Bacteroides cellulosilyticus* WH2. *J. Struct. Biol.* 197, 227–235. doi: 10.1016/j.jsb.2016.11.004
- Baker, P., Whitfield, G. B., Hill, P. J., Little, D. J., Pestrak, M. J., Robinson, H., et al. (2015). Characterization of the *Pseudomonas aeruginosa* glycoside hydrolase PslG reveals that its levels are critical for Psl polysaccharide biosynthesis and biofilm formation. *J. Biol. Chem.* 290, 28374–28387. doi: 10.1074/jbc.M115.674929
- Beg, Q. K., Kapoor, M., Mahajan, L., and Hoondal, G. S. (2001). Microbial xylanases and their industrial applications: a review. *Appl. Microbiol. Biotechnol.* 56, 326–338. doi: 10.1007/s002530100704
- Bhatia, Y., Mishra, S., and Bisaria, V. S. (2002). Microbial beta-glucosidases: cloning, properties, and applications. *Crit. Rev. Biotechnol.* 22, 375–407. doi: 10.1080/07388550290789568
- Bosetto, A., Justo, P. I., Zanardi, B., Venzon, S. S., Graciano, L., dos Santos, E. L., et al. (2016). Research progress concerning fungal and bacterial beta-xylosidases. *Appl. Biochem. Biotechnol.* 178, 766–795. doi: 10.1007/s12010-015-1908-4
- Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254. doi: 10.1006/abio.1976.9999
- Brooks, D. A., Fabrega, S., Hein, L. K., Parkinson, E. J., Durand, P., Yogalingam, G., et al. (2001). Glycosidase active site mutations in human alpha-L-iduronidase. *Glycobiology* 11, 741–750. doi: 10.1093/glycob/11.9.741
- Crennell, S., Takimoto, T., Portner, A., and Taylor, G. (2000). Crystal structure of the multifunctional paramyxovirus hemagglutinin-neuraminidase. *Nat. Struct. Biol.* 7, 1068–1074. doi: 10.1038/81002
- Czjzek, M., Ben David, A., Bravman, T., Shoham, G., Henrissat, B., and Shoham, Y. (2005). Enzyme-substrate complex structures of a GH39 beta-xylosidase from *Geobacillus stearothermophilus*. *J. Mol. Biol.* 353, 838–846. doi: 10.1016/j.jmb.2005.09.003
- de Pascale, D., Cusano, A. M., Autore, F., Parrilli, E., di Prisco, G., Marino, G., et al. (2008). The cold-active Lip1 lipase from the Antarctic bacterium *Pseudoalteromonas haloplanktis* TAC125 is a member of a new bacterial lipolytic enzyme family. *Extremophiles* 12, 311–323. doi: 10.1007/s00792-008-0163-9
- De Santi, C., Leiros, H. K., Di Scala, A., de Pascale, D., Altermark, B., and Willassen, N. P. (2016). Biochemical characterization and structural analysis of a new cold-active and salt-tolerant esterase from the marine bacterium *Thalassospira* sp. *Extremophiles* 20, 323–336. doi: 10.1007/s00792-016-0824-z
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 60(Pt 12 Pt 1), 2126–2132. doi: 10.1107/S0907444904019158

ACKNOWLEDGMENTS

We thank Dr. H. Lv and Dr. J. Zhou for help with HPLC method and the staff from the BL17U1/18U1/19U1 beamlines at Shanghai Synchrotron Radiation Facility for assistance during data collection.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2019.02922/full#supplementary-material>

- Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39, 783–791. doi: 10.1111/j.1558-5646.1985.tb00420.x
- Galbe, M., and Zacchi, G. (2002). A review of the production of ethanol from softwood. *Appl. Microbiol. Biotechnol.* 59, 618–628. doi: 10.1007/s00253-002-1058-9
- Henrissat, B., and Bairoch, A. (1996). Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.* 316(Pt 2), 695–696. doi: 10.1042/bj3160695
- Henrissat, B., and Davies, G. (1997). Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol.* 7, 637–644. doi: 10.1016/s0959-440x(97)80072-3
- Huang, J., Huo, Y. Y., Ji, R., Kuang, S., Ji, C., Xu, X. W., et al. (2016). Structural insights of a hormone sensitive lipase homologue Est22. *Sci. Rep.* 6:28550. doi: 10.1038/srep28550
- Jiang, X., Huo, Y., Cheng, H., Zhang, X., Zhu, X., and Wu, M. (2012). Cloning, expression and characterization of a halotolerant esterase from a marine bacterium *Pelagibacterium halotolerans* B2T. *Extremophiles* 16, 427–435. doi: 10.1007/s00792-012-0442-3
- Jordan, D. B., and Wagschal, K. (2010). Properties and applications of microbial beta-D-xylosidases featuring the catalytically efficient enzyme from *Selenomonas ruminantium*. *Appl. Microbiol. Biotechnol.* 86, 1647–1658. doi: 10.1007/s00253-010-2538-y
- Joseleau, J. P., Cartier, N., Chambat, G., Faik, A., and Ruel, K. (1992). Structural features and biological activity of xyloglucans from suspension-cultured plant cells. *Biochimie* 74, 81–88. doi: 10.1016/0300-9084(92)90187-j
- Ketudat Cairns, J. R., and Esen, A. (2010). Beta-Glucosidases. *Cell Mol. Life Sci.* 67, 3389–3405. doi: 10.1007/s00018-010-0399-2
- Ketudat Cairns, J. R., Mahong, B., Baiya, S., and Jeon, J. S. (2015). Beta-Glucosidases: multitasking, moonlighting or simply misunderstood? *Plant Sci.* 241, 246–259. doi: 10.1016/j.plantsci.2015.10.014
- Kim, H. K., Jung, Y. J., Choi, W. C., Ryu, H. S., Oh, T. K., and Lee, J. K. (2004). Sequence-based approach to finding functional lipases from microbial genome databases. *FEMS Microbiol. Lett.* 235, 349–355. doi: 10.1016/j.femsle.2004.05.004
- Kuang, S., Zheng, J., Yang, H., Li, S., Duan, S., Shen, Y., et al. (2017). Structure insight of GSDMD reveals the basis of GSDMD autoinhibition in cell pyroptosis. *Proc. Natl. Acad. Sci. U.S.A.* 114, 10642–10647. doi: 10.1073/pnas.1708194114
- Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096
- Kuusk, S., and Valjamae, P. (2017). When substrate inhibits and inhibitor activates: implications of beta-glucosidases. *Biotechnol. Biofuels* 10:7. doi: 10.1186/s13068-016-0690-z
- Lagaert, S., Pollet, A., Delcour, J. A., Lavigne, R., Courtin, C. M., and Volckaert, G. (2011). Characterization of two beta-xylosidases from *Bifidobacterium adolescentis* and their contribution to the hydrolysis of prebiotic xylooligosaccharides. *Appl. Microbiol. Biotechnol.* 92, 1179–1185. doi: 10.1007/s00253-011-3396-y

- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404
- McCleary, B. V., and McGeough, P. (2015). A comparison of polysaccharide substrates and reducing sugar methods for the measurement of endo-1,4-beta-Xylanase. *Appl. Biochem. Biotechnol.* 177, 1152–1163. doi: 10.1007/s12010-015-1803-z
- Mian, I. S. (1998). Sequence, structural, functional, and phylogenetic analyses of three glycosidase families. *Blood Cells Mol. Dis.* 24, 83–100. doi: 10.1006/bcmd.1998.9998
- Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, M. (2006). HKL-3000: the integration of data reduction and structure solution - from diffraction images to an initial model in minutes. *Acta Crystallogr. D Biol. Crystallogr.* 62, 859–866. doi: 10.1107/S0907444906019949
- Morohoshi, T., Wang, W. Z., Someya, N., and Ikeda, T. (2011). Genome sequence of microbacterium testaceum StLB037, an N-Acylhomoserine lactone-degrading bacterium isolated from potato leaves. *J. Bacteriol.* 193, 2072–2073. doi: 10.1128/Jb.00180-11
- Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., et al. (2009). AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.* 30, 2785–2791. doi: 10.1002/jcc.21256
- Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* 53(Pt 3), 240–255. doi: 10.1107/S0907444996012255
- Ndeh, D., Rogowski, A., Cartmell, A., Luis, A. S., Basle, A., Gray, J., et al. (2017). Complex pectin metabolism by gut bacteria reveals novel catalytic functions. *Nature* 544, 65–70. doi: 10.1038/nature21725
- Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276, 307–326. doi: 10.1016/s0076-6879(97)76066-x
- Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* 42, W320–W324. doi: 10.1093/nar/gku316
- Rye, C. S., and Withers, S. G. (2000). Glycosidase mechanisms. *Curr. Opin. Chem. Biol.* 4, 573–580. doi: 10.1016/s1367-5931(00)00135-6
- Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425. doi: 10.1093/oxfordjournals.molbev.a040454
- Santos, C. R., Polo, C. C., Correa, J. M., Simao Rde, C., Seixas, F. A., and Murakami, M. T. (2012). The accessory domain changes the accessibility and molecular topography of the catalytic interface in monomeric GH39 beta-xylosidases. *Acta Crystallogr. D Biol. Crystallogr.* 68(Pt 10), 1339–1345. doi: 10.1107/S0907444912028491
- Singhania, R. R., Patel, A. K., Sukumaran, R. K., Larroche, C., and Pandey, A. (2013). Role and significance of beta-glucosidases in the hydrolysis of cellulose for bioethanol production. *Bioresour. Technol.* 127, 500–507. doi: 10.1016/j.biortech.2012.09.012
- Sinnott, M. L. (1990). Catalytic mechanisms of enzymatic glycosyl transfer. *Chem. Rev.* 90, 1171–1202. doi: 10.1021/cr00105a006
- Sorensen, A., Lubeck, M., Lubeck, P. S., and Ahring, B. K. (2013). Fungal Beta-glucosidases: a bottleneck in industrial use of lignocellulosic materials. *Biomolecules* 3, 612–631. doi: 10.3390/biom3030612
- Vasella, A., Davies, G. J., and Bohm, M. (2002). Glycosidase mechanisms. *Curr. Opin. Chem. Biol.* 6, 619–629. doi: 10.1016/s1367-5931(02)00380-0
- Vocadlo, D. J., Wicki, J., Rupitz, K., and Withers, S. G. (2002). A case for reverse protonation: identification of Glu160 as an acid/base catalyst in *Thermoanaerobacterium saccharolyticum* beta-xylosidase and detailed kinetic analysis of a site-directed mutant. *Biochemistry* 41, 9736–9746. doi: 10.1021/bi020078n
- Wang, Q. S., Zhang, K. H., Cui, Y., Wang, Z. J., Pan, Q. Y., Liu, K., et al. (2018). Upgrade of macromolecular crystallography beamline BL17U1 at SSRF. *Nucl. Sci. Technol.* 29:68. doi: 10.1007/s41365-018-0398-9
- Xu, X. W., Wu, Y. H., Wang, C. S., Wang, X. G., Oren, A., and Wu, M. (2009). *Croceicoccus marinus* gen. nov., sp. nov., a yellow-pigmented bacterium from deep-sea sediment, and emended description of the family Erythrobacteraceae. *Int. J. Syst. Evol. Microbiol.* 59(Pt 9), 2247–2253. doi: 10.1099/ij.s.0.004267-0
- Yang, J. K., Yoon, H. J., Ahn, H. J., Lee, B. I., Pedelacq, J. D., Liang, E. C., et al. (2004). Crystal structure of beta-D-xylosidase from *Thermoanaerobacterium saccharolyticum*, a family 39 glycoside hydrolase. *J. Mol. Biol.* 335, 155–165. doi: 10.1016/j.jmb.2003.10.026
- Yu, S., Su, T., Wu, H., Liu, S., Wang, D., Zhao, T., et al. (2015). PslG, a self-produced glycosyl hydrolase, triggers biofilm disassembly by disrupting exopolysaccharide matrix. *Cell Res.* 25, 1352–1367. doi: 10.1038/cr.2015.129

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Shen, Li, Huo, Bao, Gao, Xiao, Hu, Xu and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.