


Structure of the Catalytic Domain of α -L-Arabinofuranosidase from *Coprinopsis cinerea*, CcAbf62A, Provides Insights into Structure–Function Relationships in Glycoside Hydrolase Family 62

Takashi Tonozuka¹  · Yutaro Tanaka¹ ·
Shunsaku Okuyama¹ · Takatsugu Miyazaki^{1,2} ·
Atsushi Nishikawa¹ · Makoto Yoshida³

Received: 12 June 2016 / Accepted: 26 August 2016 /
Published online: 2 September 2016
© Springer Science+Business Media New York 2016

Abstract α -L-Arabinofuranosidases, belonging to the glycoside hydrolase family (GH) 62, hydrolyze the α -1,2- or α -1,3-bond to liberate L-arabinofuranose from the xylan backbone. Here, we determined the structure of the C-terminal catalytic domain of CcAbf62A, a GH62 α -L-arabinofuranosidase from *Coprinopsis cinerea*. CcAbf62A is composed of a five-bladed β -propeller, as observed in other GH62 enzymes. The structure near the active site of CcAbf62A is also highly homologous to those of other GH62 enzymes. However, a calcium atom in the catalytic center interacts with an asparagine residue, Asn279, which is not found in other GH62 enzymes. In addition, some residues in subsites +3R, +2NR, +3NR, and +4NR of CcAbf62A are not conserved in other GH62 enzymes. In particular, a histidine residue, His221, is uniquely observed in subsite +2NR of CcAbf62A, which is likely to influence the substrate specificity. The results obtained here suggest that the amino acid residues that interact with the xylan backbone vary among the GH62 enzymes, despite the high similarity of their overall structures.

Keywords *Coprinopsis cinerea* · α -L-arabinofuranosidase · Arabinoxylan · GH62 · Hemicellulose

✉ Takashi Tonozuka
tonozuka@cc.tuat.ac.jp

¹ Department of Applied Biological Science, Tokyo University of Agriculture and Technology, Tokyo 183-8509, Japan

² Present address: Research Institute of Green Science and Technology, Shizuoka University, Shizuoka 422-8529, Japan

³ Department of Environmental and Natural Resource Science, Tokyo University of Agriculture and Technology, Tokyo 183-8509, Japan

Introduction

Lignocellulosic plant biomass contains polymers of cellulose, hemicellulose, and lignin bound together to form a complex. Plant biomass is considered a renewable energy resource, and carbohydrate-hydrolyzing enzymes are useful tools for biomass degradation [1]. Hemicellulose is a heteropolymer composed of a variety of sugars and typically consists of a β -1,4-xylopyranosyl backbone partially substituted with L-arabinofuranosyl, 4-O-methylglucuronosyl, and acetyl side chains [2]. In the case of L-arabinofuranosyl substituents, L-arabinofuranosyl residues are linked via either α -1,2- or α -1,3-bond [3] and ferulic acid is esterified to the O5 position of some L-arabinofuranosyl residues [4].

α -L-Arabinofuranosidases (EC 3.2.1.55) hydrolyze α -L-arabinofuranosyl residues from L-arabinosides and exhibit a variety of substrate specificities [5]. Most α -L-arabinofuranosidases are classified into glycoside hydrolase (GH) families GH43, GH51, GH54, and GH62 [6]. Among these, GH62 is a relatively small family, and all of the characterized GH62 enzymes are α -L-arabinofuranosidases that hydrolyze the α -1,2- or α -1,3-bond to liberate L-arabinofuranose from the xylan backbone. These enzymes have been proposed to hydrolyze sugars through an inverting mechanism [7, 8]. In recent years, three-dimensional structures of several GH62 α -L-arabinofuranosidases have been determined, and the enzymes contain the five-bladed β -propeller fold [7–10]. Similar catalytic five-bladed β -propeller domains have been found in GH43 enzymes [11, 12], and the two families, GH43 and GH62, are categorized into clan GH-F.

A basidiomycete, *Coprinopsis cinerea*, is known as a model mushroom-forming organism [13]. The entire genome has been sequenced [14] and shows that *C. cinerea* possesses three GH62 proteins, CC1G_01577, CC1G_01578, and CC1G_15259. GH62 enzymes have been phylogenetically clustered into two subfamilies, GH62_1 and GH62_2 [7]. All the three enzymes, CC1G_01577, CC1G_01578, and CC1G_15259, belong to subfamily GH62_1 and their amino acid sequences share ~60 % identity. CC1G_01577 and CC1G_15259 have a signal peptide sequence and a carbohydrate binding-module belonging to family 1 (CBM1) in their N-termini, while no signal sequence and CBM1 are found in CC1G_01578 [15]. In the previous paper, we designated CC1G_01577 as *CcAbf62A* and reported the cDNA cloning and characterization of this protein [15]. The results of the study indicated that the presence of a feruloyl esterase, *CcEst1*, enhances the activity of *CcAbf62A* for arabinoxylan, but the final amounts of reducing sugar in the course of arabinoxylan degradation with or without *CcEst1* are observed to be almost equal. Here, we determined the crystal structure of the catalytic domain (residues 82–397) of *CcAbf62A*, providing insight into the structure-function relationship.

Materials and Methods

Construction of the Expression Plasmid

The cDNA of *CcAbf62A* was obtained as described [15]. The DNA fragment encoding amino acid residues 82–397 was amplified by PCR using the cDNA of *CcAbf62A* as a template and the primers 5'-TT CAT ATG CTC CCA TCC AGC TTC AGG TGG A-3' and 5' -T TGC GGC CGC ACA AGC GGA GTT GGT TTG AG-3' (the restriction sites of NdeI and NotI are underlined). The amplified fragment was then ligated into the pET-

21a(+)-vector (Merck Millipore, Darmstadt, Germany) for heterologous expression in *Escherichia coli*. The resultant recombinant protein was designed to have a His-tag (AAALEHHHHHH) at the C-terminus.

Protein Expression and Purification

E. coli strain BL21 (DE3) was transformed with the obtained plasmid. The transformant was grown in 1-L Luria-Bertani (LB) medium containing 50 µg/mL ampicillin at 37 °C until the absorbance at 600 nm (A_{600}) reached 0.6. Protein expression was induced with isopropyl- β -D-thiogalactopyranoside at a final concentration of 0.2 mM for 18 h at 18 °C. The cells were harvested and resuspended in 30 mL of 20 mM Tris-HCl buffer (pH 7.5), followed by sonication for 2 min on ice. After centrifugation to remove insoluble material, the supernatant was applied onto a nickel (Ni^{2+}) nitrilotriacetic acid (Ni-NTA) agarose (QIAGEN, Hilden, Germany) column equilibrated with the same buffer. The column was washed with the same buffer, and the recombinant protein was eluted with the same buffer containing 50 mM imidazole. The protein fraction was dialyzed against 10 mM Tris-HCl buffer (pH 7.5), and the purified protein yielded a single band on sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The concentration of the purified protein was determined by measuring the absorbance at 280 nm, using the molar extinction coefficient (1 mg/mL = 2.50) calculated by the ExPasy ProtParam server (<http://web.expasy.org/protparam/>).

Crystallization, Data Collection, and Structure Determination

The purified protein was concentrated to 10 mg/mL in 10 mM Tris-HCl (pH 7.5) using an Amicon Ultra-15 centrifugal unit (Merck Millipore, Darmstadt, Germany). Needle-shaped crystals were obtained at 20 °C using the hanging-drop vapor diffusion method, in which 1.0 µL protein solution was mixed with an equal volume of the crystallization reservoir containing 0.1 M 2-(*N*-morpholino)ethanesulfonic acid (MES)-NaOH (pH 6.5), 0.1 M NaBr, 3 % (w/v) polyethylene glycol 20,000 (Sigma-Aldrich, St. Louis, MO, USA), and 20 % (w/v) polyethylene glycol 3350 (Hampton Research, Aliso Viejo, CA, USA). To obtain the complex structure of lead, the crystal was transferred for 20 min to a solution containing 10 mM lead(II) acetate in the reservoir solution. The harvested crystals were cryo-protected in the reservoir solution supplemented with 20 % (v/v) glycerol, and flash-frozen in liquid nitrogen. The diffraction data were collected at the AR-NE3A and AR-NW12A beamlines at the Photon Factory (Tsukuba, Japan). All data were processed and scaled using HKL2000 [16]. At the time that this study was being carried out, several crystal structures of GH62 enzymes had been reported, and thus, the structure was solved by molecular replacement using the program MOLREP [17] in the CCP4 suite [18], and a model of GH62 α -L-arabinofuranosidase from *Podospira anserina*, *Pod_ansAbf62A*, (PDB id, 4N4B) was employed as a probe model. Automated model building was performed with the program ARP/wARP [19]. The model was refined using REFMAC5 [20] in the CCP4 suite, and manual adjustment and rebuilding of the model were carried out using the program COOT [21]. Validation of the structures was performed using MolProbity [22]. Figures were prepared using PyMOL (<http://www.pymol.org/>). The data collection and refinement statistics are summarized in Table 1. The coordinates and structure factors have been deposited in the Protein Data Bank (PDB) under the accession codes 5B6S and 5B6T.

Results and Discussion

Structure Determination of *CcAbf62A*

The crystal structure of the catalytic domain of *CcAbf62A* (hereafter simply referred to as *CcAbf62A*) was determined at 1.70-Å resolution, and that of the catalytic domain of *CcAbf62A* soaked with $\text{Pb}(\text{CH}_3\text{COO})_2$ was also determined at 1.48-Å resolution (Table 1). The two structures are almost isomorphous, and thus, the following descriptions are based on *CcAbf62A* in the unsoaked form, unless otherwise stated.

The crystal belongs to the space group $P2_1$, which contains two molecules, Mol-A and Mol-B, in an asymmetric unit. In the structure of *CcAbf62A*-Pb, one lead atom is seen in the catalytic site of Mol-A (Fig. 1a). The Ramachandran plot was calculated with the MolProbity server (Table 1). Only one residue, His346, in both Mol-A and Mol-B was identified as an outlier, and the electron density for the residue was well-defined. His346 is a conserved residue in GH62, as described later, and the corresponding residue in *Scytalidium thermophilum* (*Scy_the*) Abf62C (abbreviations of GH62 enzymes are listed in Table 3) also reportedly adopts a disallowed Ramachandran conformation [10]. The $2F_o - F_c$ electron density maps contoured at 1σ show continuous density for residues 82–398 of both Mol-A and Mol-B, and there is no significant difference between the two molecules in the catalytic cleft. Structural analysis using the PISA server [23] indicated that there was no specific interaction to form an oligomeric structure, suggesting that *CcAbf62A* exists as a monomer in solution. The descriptions hereafter are based on Mol-A.

Overall Structure of *CcAbf62A*

The overall structure of *CcAbf62A* is indicated in Fig. 1b. *CcAbf62A* is composed of a five-bladed (blades I–V) β -propeller as observed in other GH62 enzymes. The $\text{C}\alpha$ backbone of *CcAbf62A* was superimposed onto those of *Pod_ansAbf62A* (PDB ID, 4N4B) [7], *Scy_theAbf62C* (PDB ID, 4PVI) [10], and *Streptomyces coelicolor* (*Str_coe*) Abf62A (PDB ID, 3WN2) [8], illustrating that the folds of these enzymes are essentially identical (Fig. 1c). The β -Strands that comprise the β -propeller blades are numbered as $\beta 1$ – $\beta 20$ (Fig. 2, top), based on the numbering scheme for *Pod_ansAbf62A* [7]. *CcAbf62A* had a disulfide bridge of Cys363–Cys397 in blade V (Fig. 1b), like those seen in *Pod_ansAbf62A* and *Str_coeAbf62A*.

A structural similarity search was performed by using the DALI server [24] (Table 2). Aside from the GH62 enzymes, *CcAbf62A* is similar to GH43, GH32, and GH68, as described previously, and these families are categorized as the GH43_62_32_68 superfamily [25]. *CcAbf62A* also shows homology to GH130 [26] and GH117 [27]. All of these GH families consist of five-bladed β -propeller structures. Among the characterized GH43 enzymes, *CcAbf62A* most resembles the *Streptomyces avermitilis* exo-1,5- α -L-arabinofuranosidase (*SaAraf43A*; PDB ID, 3AKH) [28] and the *Cellvibrio japonicus* 1,2- α -L-arabinofuranosidase (PDB ID, 3QEF) [29].

GH62 enzymes have been proposed to be inverting glycoside hydrolases, and two Asp and one Glu residues form a catalytic triad [7, 30]. Asp109, Asp224, and Glu276 in *CcAbf62A* were identified as a general base, a $\text{p}K_a$ modulator, and a general acid, respectively (Table 3). With one exception [10], a calcium atom is located in the center of the five-bladed β -propeller in GH62 enzymes, and a conserved His residue holds the calcium atom (Fig. 3a). The corresponding calcium atom is also seen in *CcAbf62A*, and His346 functions as the calcium

Table 1 Data collection and refinement statistics

	<i>CcAbf62A</i>	<i>CcAbf62A</i> -Pb
Data collection		
Beamline	PF-AR NW12A	PF-AR NE3A
Wavelength (Å)	1.0	0.95064
Space group	<i>P</i> 2 ₁	<i>P</i> 2 ₁
Cell dimensions		
<i>a</i> (Å)	47.7	47.9
<i>b</i> (Å)	77.7	77.8
<i>c</i> (Å)	74.2	74.5
β (°)	93.0	93.0
Resolution range (Å)	74.1–1.70 (1.76–1.70) ^a	27.6–1.48 (1.53–1.48) ^a
Measured reflections	216,952	321,802
Unique reflections	59,017	90,154
Redundancy	3.7 (3.5) ^a	3.6 (3.1) ^a
Completeness (%)	99.4 (97.2) ^a	99.4 (97.3) ^a
$\langle I/\sigma(I) \rangle$	15.1 (2.7) ^a	24.5 (3.8) ^a
R_{merge}	0.108 (0.545) ^a	0.073 (0.441) ^a
Refinement		
R_{work}	0.139	0.150
R_{free}	0.170	0.175
Root mean square deviation (rmsd)		
Bond lengths (Å)	0.009	0.011
Bond angles (°)	1.38	1.51
Ramachandran plot (Molprobit)		
Favored (%)	96.5	96.5
Allowed (%)	3.2	3.2
Outliers (%)	0.3	0.3
Number of atoms		
Protein	5022	5034
Metal atom	2 (Ca, 2)	3 (Ca, 2; Pb, 1)
Glycerol	42	36
Water	675	619
Average <i>B</i> (Å ²)		
Protein	13.6	11.1
Metal atom	13.9	17.0
Glycerol	21.7	18.5
Water	26.2	21.3
PDB id	5B6S	5B6T

^a The values for the highest resolution shells are listed in parentheses

holder. This His residue has been proposed to form the catalytic core together with the catalytic triad residues [10].

It has been reported that *Scy_theAbf62C* possesses no calcium atom in the catalytic cleft, and the presence of a cysteine residue, Cys233, could result in the weak binding of the metal

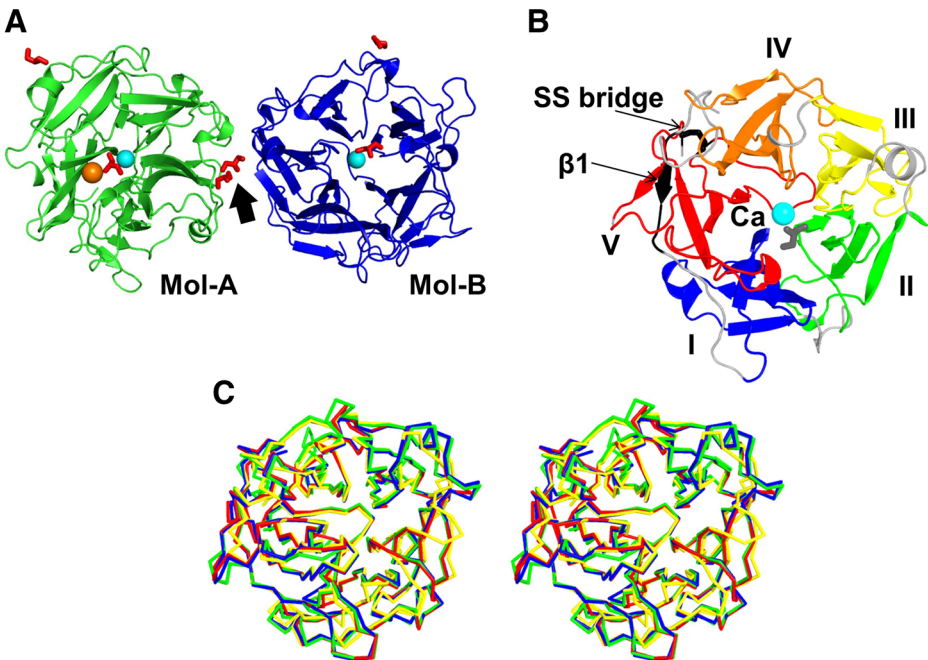


Fig. 1 Overall structure of *CcAbf62A*. **a** Ribbon model of *CcAbf62A*-Pb in an asymmetric unit. Mol-A (green), Mol-B (blue), glycerol molecules (red), lead atom (orange), and calcium atoms (cyan) are shown. Two glycerol molecules found near Tyr135, Tyr150, and Tyr151 are indicated with a black arrow. **b** Five blades I–V (blue, green, yellow, orange, and red, respectively) comprising the β -propeller fold. β 1 strand (black), SS bridge (black), and calcium atom (cyan) are indicated. **c** Stereo view of the superimposition of the $C\alpha$ backbones of *CcAbf62A* (red), *Scy_theAbf62C* (green; PDB id, 4PVI), *Pod_ansAbf62A* (blue; PDB id, 4N4B), and *Str_coeAbf62A* (yellow; PDB id, 3WN2)

ion [10]. In most of the GH62 enzymes, Gln or Cys is observed in this position, whereas Asn279 is identified as the equivalent residue in *CcAbf62A* (Table 3). The effect of divalent cations on the activities of two enzymes from *Scy. thermophilum*, *Scy_theAbf62A* and *Scy_theAbf62C*, has been investigated. The presence of Ca^{2+} or Mg^{2+} resulted in only small changes in the activities of the two enzymes, while the presence of 2 mM of Ni^{2+} , Co^{2+} , Zn^{2+} , Cu^{2+} , or Mn^{2+} inhibited both enzymes. In particular, a significant decrease was observed in the presence of Zn^{2+} or Cu^{2+} , both of which have a relatively large atomic radius [10]. The structure of *CcAbf62A*-Pb Mol-A shows that Pb^{2+} binds to the catalytic acid residue, Glu276, and does not interact with the calcium holder, His346 (Fig. 3a). It is likely that metal atoms having relatively large radii occupy positions different from that of the calcium atom.

Seven and six glycerol molecules from the cryoprotectant solution were identified in the *CcAbf62A* and *CcAbf62A*-Pb asymmetric units, respectively. A glycerol molecule each is located in the catalytic cleft of both Mol-A and Mol-B (Fig. 1a) and forms the same contacts with *CcAbf62A* (Fig. 3a). The other glycerol molecules were found on the protein surface (Fig. 1a). It is likely that some of these glycerol molecules are artifacts, as they are present near the crystal contacts. Recently, however, a secondary carbohydrate binding site, composed of Trp23 and Tyr44, has been identified in *Aspergillus nidulans* α -L-arabinofuranosidase, a member of the GH62_2 subfamily [31]. The two residues are not conserved in *CcAbf62A*; instead, Tyr135, Tyr150, and Tyr151 of *CcAbf62A* are located near the corresponding position

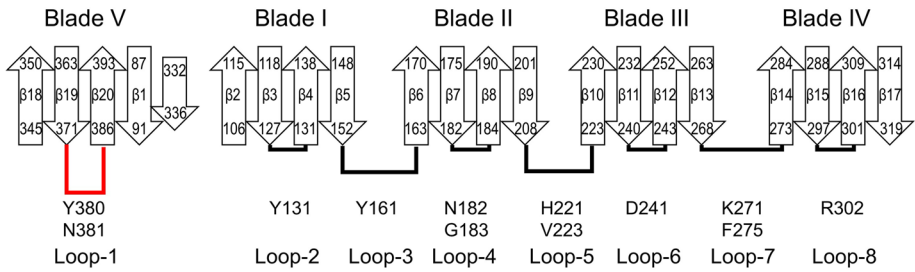
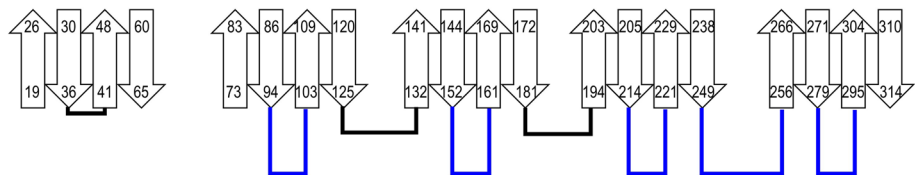
CcAbf62A**SaAraf43A**

Fig. 2 Topology of the β -propeller folds of *CcAbf62A* and *SaAraf43A*. Amino acid residue numbers are given at each end of the β -strands. Eight loops located at the catalytic cleft of the enzymes are indicated as Loop-1 to Loop-8. Loop-1 in *CcAbf62A* (shown in red) is longer than that in *SaAraf43A*, while Loop-2, Loop-4, Loop-6, Loop-7, and Loop-8 in *SaAraf43A* (shown in blue) are longer than those in *CcAbf62A*. In *CcAbf62A*, β -strands composed of the β -propeller fold are indicated as β 1– β 20, and an additional β -strand (residues 332–336) is shown. Amino acid residues, which are present in Loop-1 to Loop-8 and also listed in Table 3, are indicated

Table 2 Summary of structural similarity search using the DALI server

Enzyme	PDB id	CAZy	Z-score	r.m.s.d. (Å)	Aligned residues	Sequence identity (%)
<i>Podospora anserina</i> α -L-arabinofuranosidase (<i>Pod_ansAbf62A</i>)	4N4B	GH62	54.8	0.7	314	68
<i>Streptomyces avermitilis</i> exo-1,5- α -L-arabinofuranosidase (<i>SaAraf43a</i>)	3AKH	GH43	23.4	2.8	251	15
<i>Cellvibrio japonicus</i> α -1,2-arabinofuranosidase	3QEF	GH43	23.3	2.7	239	14
β -1,4-Mannopyranosyl-chitobiose phosphorylase from uncultured organism	4UDJ	GH130	22.3	2.9	250	11
<i>Zobellia galactanivorans</i> α -1,3-L-neoagarooligosaccharide hydrolase	4U6B	GH117	22.0	2.8	255	11
<i>Arthrobacter ureafaciens</i> levan fructotransferase	4FFF	GH32	21.8	3.1	254	9
<i>Microbacterium saccharophilum</i> β -fructofuranosidase	3WPY	GH68	20.5	3.3	264	9
Human cytosolic sialidase	1VCU	GH33	17.1	3.6	251	11

Table 3 Comparison of amino acid residues in the substrate binding site of GH62 enzymes

Organism	<i>Coprinopsis cinerea</i>	<i>Coprinopsis cinerea</i>	<i>Coprinopsis cinerea</i>	<i>Podospira anserina</i>	<i>Scytalidium thermophilum</i>	<i>Streptomyces coelicolor</i>	<i>Streptomyces thermoviolaceus</i>	<i>Ustilago maydis</i>
Abbreviation	CcAbf62A CC1G_01577	CC1G_01578	CC1G_15259	Pod_ans Abf62A	Scy_the Abf62C	Str_coe Abf62A	Sr_the Abf62A	Ust_may Abf62A
PDB ID	5B6S	–	–	4N4B	4PV1	3WN2	4O8O	4N2R
Subfamily	GH62_1	GH62_1	GH62_1	GH62_1	GH62_1	GH62_2	GH62_2	GH62_2
Sequence identity (%) ^a	–	68	62	63	62	41	40	39
N-terminal CBM	CBM1	No	CBM1	No	No	CBM13	No	No
Subsite –1	K108 Q181 E294 Q370 Y380	K65 Q138 E251 Q329 Y339	CBM1 K109 Q182 E295 Q372 Y385	No K48 Q121 E233 Q310 Y320	K54 Q127 E251 Q328 Y338	K201 Q269 E379 Q451 Y461	K87 Q156 E266 Q339 Y349	No K35 Q103 E213 Q285 Y296
Catalytic residue	D109 D224 E276 H346 N279	D66 D181 E233 H305 C236	D110 D225 E277 H348 C280	D49 D165 E216 H285 C219	D55 D171 E230 H303 C233	D202 D309 E361 H427 Q364	D88 D196 E248 H315 Q251	D36 D143 E195 H261 Q198
Calcium holder	K271 D241 F275 R302 Y131	R228 D198 F232 R259 Y88	R272 D242 F276 R303 W132	R211 D182 F215 R241 Y71	I225 D194 W229 R259 Y77	K356 D326 F360 R386 Y224	R243 D213 F247 R274 W111	Q190 D160 F194 R220 Y58
Subsite +3R	H221 V223 N381	Y178 V180 N340	Y222 L224 G386	Y162 V164 N321	Y168 V170 N339	T305 I308 N462	T192 I195 D350	T139 I142 D297
Subsites +2R and +1								
Subsite +2NR								

Table 3 (continued)

Organism	<i>Coprinopsis cinerea</i>	<i>Coprinopsis cinerea</i>	<i>Coprinopsis cinerea</i>	<i>Podospora anserina</i>	<i>Scytalidium thermophilum</i>	<i>Streptomyces coelicolor</i>	<i>Streptomyces thermoviolaceus</i>	<i>Ustilago maydis</i>
Subsites +3NR and +4NR	Y161 N182 G183	Y118 N139 D140	Y162 D183 N184	Y101 N122 G123	Y107 N128 G129	W270	W157	W104

^a The values were calculated using the Clustal Omega server (<http://www.ebi.ac.uk/Tools/msa/clustalo/>)

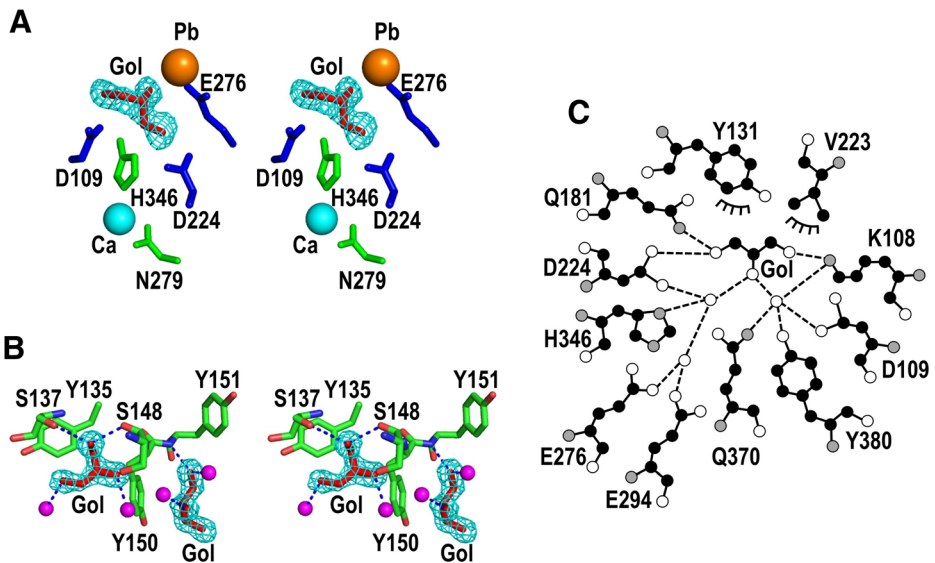


Fig. 3 Glycerol (labeled as Gol)-bound structure of *CcAbf62A*. **a** A glycerol molecule (red) found in the active site in *CcAbf62A*-Pb. A calcium atom (cyan), a lead atom (orange), catalytic residues (blue), and residues interacting with calcium (green) are indicated. The electron density ($F_o - F_c$, 3σ) is shown in cyan. **b** Two glycerol molecules (red) found near Tyr135, Tyr150, and Tyr151. Hydrogen bonds (blue dotted line) and water molecules (magenta) are indicated. The electron density ($F_o - F_c$, 3σ) is shown in cyan. **c** Schematic drawing of the amino acid residues interacting with glycerol in the active site. White circle, oxygen atom; black circle, carbon atom; gray circle, nitrogen atom; dashed line, hydrogen bond

of the secondary carbohydrate binding site in *A. nidulans* α -L-arabinofuranosidase, and two glycerol molecules in Mol-A are located close to these Tyr residues (Fig. 3b). The three Tyr residues are also found in the GH62_1 subfamily enzymes, *Pod_ansAbf62A* and *Scy_theAbf62C*, suggesting that the three Tyr residues might participate in substrate binding.

Ligand-Bound Model of *CcAbf62A*

Several ligand-bound structures of GH62 enzymes have been reported. We constructed a ligand-bound model of *CcAbf62A* using *Ustilago maydis* (*Ust_may*) Abf62A complexed with α -L-arabinofuranose (Ara) (PDB ID, 4N2R) [7], and *Str_coeAbf62A* complexed with xylopentose (X_5) [8] (PDB ID, 3WN2). The three structures were superimposed and the coordinates of Ara and X_5 were then placed in *CcAbf62A*. To probe the amino acid residues involved in each subsite, residues within 4 Å from Ara or X_5 were calculated using the program NCONT of CCP4, and the corresponding residues in other GH62 enzymes are listed (Table 3). The subsite numbers are based on those described [8]; Ara is accommodated by subsite -1, the enzymatic cleavage occurs between subsites -1 and +1, and X_5 interacts with subsites +4NR, +3NR, +2NR, +1, and +2R from the non-reducing end to the reducing end. Amino acid residues potentially located at subsite +3R and those potentially interacting with the calcium atom, namely the residues corresponding to Arg211 in *Pod_ansAbf62A* and those corresponding to Cys233 in *Scy_theAbf62C*, are also listed in Table 3.

A glycerol molecule binds at subsite -1 in both Mol-A and Mol-B of *CcAbf62A* (Fig. 3a), and the structure O1-C1-C2-(O2)-C3-O3 of glycerol is similar to that formed by atoms O3-C3-C4-(O4)-C5-O5 of Ara (Fig. 4a, b). The interaction between *CcAbf62A* and glycerol was

analyzed with the programs Ligplot [32] and Coot, indicating that as many as 11 residues participate in the binding of glycerol (Fig. 3c). Based on the ligand-bound model of *CcAbf62A*, the amino acid residues forming hydrogen bonds with glycerol (Lys108, Asp109, Gln181, Asp224, Glu276, Glu294, His346, Gln370, and Tyr380) appear to participate in subsite -1, and these residues are fully conserved among GH62 enzymes (Table 3).

In the previous research, a feruloyl esterase, *CcEst1*, was found to promote the activity of *CcAbf62A* against feruloyl arabinoxytan, whereas the amount of reducing sugar in the late stage of the reaction was the same regardless of the presence or absence of *CcEst1* [15]. There is a space, which could potentially accommodate a ferulate residue, around atoms C5–O5 of Ara in the ligand-bound model of *CcAbf62A*, as two amino acid residues surrounding the space are identified as tyrosine residues, Tyr131 and Tyr161 (Fig. 4b), whereas bulky Trp residues are found in the corresponding positions in some of the enzymes (Table 3). For example, the corresponding two residues in *Streptomyces thermoviolaceus* (*Str_the*) *Abf62A* are identified as Trp111 and Trp157 (Fig. 4c).

While the residues in subsites -1 and +1 are well conserved among GH62 enzymes, relatively large variations are found in subsites +3R, +2NR, +3NR, and +4NR. It is interesting to note that a His residue, His221, is not conserved in the other GH62 enzymes, and the corresponding residue is mostly Tyr or Thr. The imidazole ring of His221 stacks against the pyranose ring of xylose at subsite +2NR, and His221 appears to be the most critical residue for the binding of subsite +2NR (Fig. 4b). In contrast, the residue at the equivalent position in *Str_theAbf62A*, Thr192, does not directly interact with the xylan backbone (Fig. 4c). We have reported that *CcAbf62A* does not hydrolyze *p*-nitrophenyl α -L-arabinofuranoside [15], which is commonly used for measurement of α -L-arabinofuranosidase activity. Site-directed mutagenesis of *Scy_theAbf62C* indicated that alteration of Tyr168, the residue corresponding to His221 in *CcAbf62A*, leads to decrease of the activity for *p*-nitrophenyl α -L-arabinofuranoside [10]. Also, comparative properties of two GH62_1 α -L-arabinofuranosidases, ABFI and ABFII, from *Aspergillus fumigatus* have been reported [33]; ABFI has a very high K_m value,

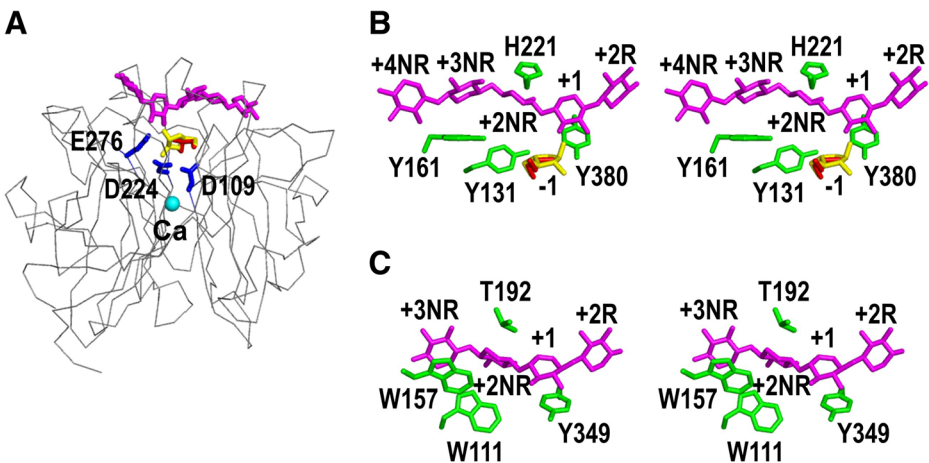


Fig. 4 Substrate bound model of *CcAbf62A*. **a** α backbone representation of *CcAbf62A*. Models of α -L-arabinofuranose (yellow) and xylopentaose (magenta) are placed on the structure. Red stick, glycerol; blue stick, catalytic residue; cyan ball, calcium atom. **b**, **c** Comparison of some key residues interacting with the xylan backbone in *CcAbf62A* (**b**) and those in *Str_theAbf62A* (PDB id, 4O8P) (**c**). Magenta stick, xylopentaose model (**b**) or xylotetraose (**c**); yellow stick, α -L-arabinofuranoside model; red stick, glycerol. Subsite numbers (-1, +1, +2R, +2NR, +3NR, and +4NR) are indicated

94 mM for *p*-nitrophenyl α -L-arabinofuranoside, whereas ABFII exhibits a low K_m value, 3.9 mM, for the same substrate. The residues equivalent to His221 in *CcAbf62A* are identified as Asn157 (ABFI) and Tyr222 (ABFII), suggesting that the presence of Tyr residue may be critical for the activity for *p*-nitrophenyl α -L-arabinofuranoside. It is likely that the presence of the His residue in subsite +2NR may result in the inactivity of *CcAbf62A* for the substrate *p*-nitrophenyl α -L-arabinofuranoside.

Sequence alignment of the three GH62 enzymes from *C. cinerea*: *CcAbf62A*, CC1G_01578, and CC1G_15259, indicates that CC1G_01578 and CC1G_15259 have additional typical amino acid residues found in other GH62 enzymes; the corresponding residues of His221 and Asn279 in *CcAbf62A* are identified as Tyr (CC1G_01578, Tyr178; CC1G_15259, Tyr222) and Cys (CC1G_01578, Cys236; CC1G_15259, Cys280), respectively. It is likely that the subsite affinities of *CcAbf62A* for the xylan backbone, and the effect of metal ions are different from those of other GH62 enzymes, including CC1G_01578 and CC1G_15259.

Comparison to GH43 Enzymes

The similarity of the β -propeller fold between GH62 and GH43 has been documented. As *CcAbf62A* showed high structural similarity with *SaAraf43A* [28] among the GH43 enzymes in the DALI search, the structures of *CcAbf62A* and *SaAraf43A* were compared. The β -strand backbones, which consist of the five-bladed β -propeller structures of *CcAbf62A* and *SaAraf43A*, are structurally identical, while the positions of the N- and C-termini were different between the two structures. The first β -strand, designated β_1 , of *CcAbf62A* is present at the identical position to the fourth β -strand of *SaAraf43A*, resulting in formation of a so-called “molecular Velcro” [7] at blade V in *CcAbf62A* (Fig. 1b). Also, an additional short β -strand, comprising residues 332–336, forms a parallel β -sheet with β_1 in blade 1 of *CcAbf62A* (Fig. 2, top).

Eight loops located at the entrance of the catalytic cleft of *CcAbf62A* adopt structures different from those of *SaAraf43A*, and these loops are designated Loop-1 to Loop-8 (Fig. 2). All the amino acid residues in subsites +3R, +2R, +1, +2NR, +3NR, and +4NR, which interact with the xylan backbone, are found in these eight loops. It is interesting to note that the lengths of Loop-2, Loop-4, Loop-6, Loop-7, and Loop-8 in *CcAbf62A* are shorter than the corresponding loops in *SaAraf43A*, which allows the molecular surface of *CcAbf62A* to form a

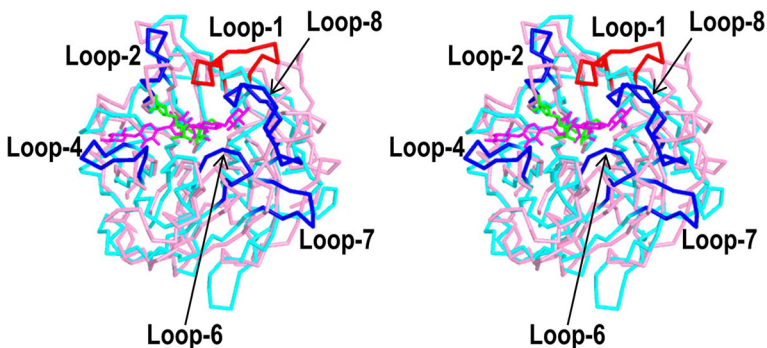


Fig. 5 Stereo view of the superimposition of the $C\alpha$ backbones of *CcAbf62A* (pink) and *SaAraf43A* (cyan). Loop-1 in *CcAbf62A* (shown in red) is longer than that in *SaAraf43A*, whereas Loop-2, Loop-4, Loop-6, Loop-7, and Loop-8 in *SaAraf43A* (shown in blue) are longer than those in *CcAbf62A*

xylan binding cleft. In contrast, Loop-1 in *CcAbf62A* appears to be longer than that in *SaAraf43A* (Fig. 5). Loop-1 appears to be critical for the enzymatic activity of *CcAbf62A*, since Tyr380 is located in Loop-1. Tyr380 forms hydrogen bonds with the catalytic residue, Asp109, via a water molecule (Fig. 3c), and this Tyr residue is strictly conserved among GH62 enzymes (Table 1). Mutation of the corresponding residue in *Str_coeAbf62A*, Tyr461, has been reported to cause a drastic decrease in activity [8].

Conclusions

The crystal structure of *CcAbf62A* was determined in this study. The structure reveals that residues in subsites +3R, +2NR, +3NR, and +4NR of *CcAbf62A* are relatively not conserved compared to those of other GH62 enzymes. In particular, a His residue, His221, is uniquely found in subsite +2NR of *CcAbf62A*, which may be responsible for the inactivity for *p*-nitrophenyl α -L-arabinofuranoside. In addition, although a calcium atom is observed in the structure of *CcAbf62A*, the residue interacting with calcium is identified as an Asn residue, Asn279, which is different from the Cys or Gln residues found in other GH62 enzymes. Because of the lack of the bulky Trp residue in the substrate binding site, there is a space near the catalytic center of *CcAbf62A*, which is likely to be capable of accommodating feruloyl L-arabinose.

There are three GH62 enzymes, *CcAbf62A*, CC1G_01578, and CC1G_15259 from *C. cinerea*, and two of them, *CcAbf62A* and CC1G_15259, possess a carbohydrate-binding module belonging to CBM1 in their N terminus. In the majority of the GH62 enzymes, however, no CBM1 module was found, as shown in Table 3, and CC1G_01578 is more like a typical α -L-arabinofuranosidase. *CcAbf62A* therefore appears to be structurally unique, despite its high structural similarity to other GH62 enzymes. It is not uncommon for a fungal genome to possess multiple GH62 genes [10, 33]. The results obtained here suggest that amino acid residues interacting with the xylan backbone are not conserved among the GH62 enzymes; this strategy is likely to be suitable for hydrolyzing a wide variety of arabinoxylan structures by multiple α -L-arabinofuranosidases.

Acknowledgments This work was supported in part by a Grant-in-Aid for Scientific Research (no. 16K07687 to T.T. and no. 15H04526 to M.Y.) from the Japan Society for the Promotion of Science. This work has been performed under the approval of the Photon Factory Program Advisory Committee (Nos. 2014G512 and 2016G013).

References

1. Jordan, D. B., Bowman, M. J., Braker, J. D., Dien, B. S., Hector, R. E., Lee, C. C., et al. (2012). Plant cell walls to ethanol. *Biochemical Journal*, *442*, 241–252.
2. Pauly, M., Gille, S., Liu, L., Mansoori, N., de Souza, A., Schultink, A., et al. (2013). Hemicellulose biosynthesis. *Planta*, *238*, 627–642.
3. Kormelink, F. J., Gruppen, H., Viëtor, R. J., & Voragen, A. G. (1993). Mode of action of the xylan-degrading enzymes from *Aspergillus awamori* on alkali-extractable cereal arabinoxylans. *Carbohydrate Research*, *249*, 355–367.
4. Gopalan, N., Rodríguez-Duran, L. V., Saucedo-Castaneda, G., & Nampoothiri, K. M. (2015). Review on technological and scientific aspects of feruloyl esterases: a versatile enzyme for biorefining of biomass. *Bioresource Technology*, *193*, 534–544.
5. Lagaert, S., Pollet, A., Courtin, C. M., & Volckaert, G. (2014). β -xylosidases and α -L-arabinofuranosidases: accessory enzymes for arabinoxylan degradation. *Biotechnology Advances*, *32*, 316–332.

6. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., & Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research*, *42*, D490–D495.
7. Siguier, B., Haon, M., Nahoum, V., Marcellin, M., Burlet-Schiltz, O., Coutinho, P. M., et al. (2014). First structural insights into α -L-arabinofuranosidases from the two GH62 glycoside hydrolase subfamilies. *Journal of Biological Chemistry*, *289*, 5261–5273.
8. Maehara, T., Fujimoto, Z., Ichinose, H., Michikawa, M., Harazono, K., & Kaneko, S. (2014). Crystal structure and characterization of the glycoside hydrolase family 62 α -L-arabinofuranosidase from *Streptomyces coelicolor*. *Journal of Biological Chemistry*, *289*, 7962–7972.
9. Wang, W., Mai-Gisondi, G., Stogios, P. J., Kaur, A., Xu, X., Cui, H., et al. (2014). Elucidation of the molecular basis for arabinoxylan-debranching activity of a thermostable family GH62 α -L-arabinofuranosidase from *Streptomyces thermoviolaceus*. *Applied and Environmental Microbiology*, *80*, 5317–5329.
10. Kaur, A. P., Nocek, B. P., Xu, X., Lowden, M. J., Leyva, J. F., Stogios, P. J., et al. (2015). Functional and structural diversity in GH62 α -L-arabinofuranosidases from the thermophilic fungus *Scytalidium thermophilum*. *Microbial Biotechnology*, *8*, 419–433.
11. Nurizzo, D., Turkenburg, J. P., Charnock, S. J., Roberts, S. M., Dodson, E. J., McKie, V. A., et al. (2002). *Cellvibrio japonicus* α -L-arabinanase 43 A has a novel five-blade β -propeller fold. *Nature Structural Biology*, *9*, 665–668.
12. Proctor, M. R., Taylor, E. J., Nurizzo, D., Turkenburg, J. P., Lloyd, R. M., Vardakou, M., et al. (2005). Tailored catalysts for plant cell-wall degradation: redesigning the exo/endo preference of *Cellvibrio japonicus* arabinanase 43 A. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 2697–2702.
13. Muraguchi, H., Umezawa, K., Niihara, M., Yoshida, M., Kozaki, T., Ishii, K., et al. (2015). Strand-specific RNA-seq analyses of fruiting body development in *Coprinopsis cinerea*. *PLoS One*, *10*, e0141586.
14. Stajich, J. E., Wilke, S. K., Ahrén, D., Au, C. H., Birren, B. W., Borodovsky, M., et al. (2010). Insights into evolution of multicellular fungi from the assembled chromosomes of the mushroom *Coprinopsis cinerea* (*Coprinus cinereus*). *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 11889–11894.
15. Hashimoto, K., Yoshida, M., & Hasumi, K. (2011). Isolation and characterization of CcAbf62A, a GH62 α -L-arabinofuranosidase, from the basidiomycete *Coprinopsis cinerea*. *Bioscience, Biotechnology, and Biochemistry*, *75*, 342–345.
16. Otwinowski, Z., & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology*, *276*, 307–326.
17. Vagin, A., & Teplyakov, A. (2010). Molecular replacement with MOLREP. *Acta Crystallographica*, *D66*, 22–25.
18. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., et al. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallographica*, *D67*, 235–242.
19. Langer, G., Cohen, S. X., Lamzin, V. S., & Perrakis, A. (2008). Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nature Protocols*, *3*, 1171–1179.
20. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., et al. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica*, *D67*, 355–367.
21. Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of coot. *Acta Crystallographica*, *D66*, 486–501.
22. Chen, V. B., Arendall 3rd, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., et al. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica*, *D66*, 12–21.
23. Krissinel, E., & Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology*, *372*, 774–797.
24. Holm, L., & Rosenström, P. (2010). Dali server: conservation mapping in 3D. *Nucleic Acids Research*, *38*, W545–W549.
25. Pons, T., Naumoff, D. G., Martínez-Fleites, C., & Hernández, L. (2004). Three acidic residues are at the active site of a β -propeller architecture in glycoside hydrolase families 32, 43, 62, and 68. *Proteins*, *54*, 424–432.
26. Ladeveze, S., Cioci, G., Roblin, P., Mourey, L., Tranier, S., & Potocki-Veronese, G. (2015). Structural bases for N-glycan processing by mannoside phosphorylase. *Acta Crystallographica*, *D71*, 1335–1346.
27. Ficko-Blean, E., Duffieux, D., Rebuffet, E., Larocque, R., Groisillier, A., Michel, G., et al. (2015). Biochemical and structural investigation of two paralogous glycoside hydrolases from *Zobellia galactanivorans*: novel insights into the evolution, dimerization plasticity and catalytic mechanism of the GH117 family. *Acta Crystallographica*, *D71*, 209–223.

28. Fujimoto, Z., Ichinose, H., Maehara, T., Honda, M., Kitaoka, M., & Kaneko, S. (2010). Crystal structure of an exo-1,5- α -L-arabinofuranosidase from *Streptomyces avermitilis* provides insights into the mechanism of substrate discrimination between exo- and endo-type enzymes in glycoside hydrolase family 43. *Journal of Biological Chemistry*, 285, 34134–34143.
29. Cartmell, A., McKee, L. S., Pena, M. J., Larsbrink, J., Brumer, H., Kaneko, S., et al. (2011). The structure and function of an arabinan-specific α -1,2-arabinofuranosidase identified from screening the activities of bacterial GH43 glycoside hydrolases. *Journal of Biological Chemistry*, 286, 15483–15495.
30. Brüx, C., Ben-David, A., Shallom-Shezifi, D., Leon, M., Niefind, K., Shoham, G., et al. (2006). The structure of an inverting GH43 β -xylosidase from *Geobacillus stearothermophilus* with its substrate reveals the role of the three catalytic residues. *Journal of Molecular Biology*, 359, 97–109.
31. Wilkens, C., Andersen, S., Petersen, B. O., Li, A., Busse-Wicher, M., Birch, J., et al. (2016). An efficient arabinoxylan-debranching α -L-arabinofuranosidase of family GH62 from *Aspergillus nidulans* contains a secondary carbohydrate binding site. *Applied Microbiology and Biotechnology*, 100, 6265–6277.
32. Laskowski, R. A., & Swindells, M. B. (2011). LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *Journal of Chemical Information and Modeling*, 51, 2778–2786.
33. Pérez, R., & Eyzaguirre, J. (2016). *Aspergillus fumigatus* produces two arabinofuranosidases from glycosyl hydrolase family 62: comparative properties of the recombinant enzymes. *Applied Biochemistry and Biotechnology*, 179, 143–154.