Novel Ethanol- and 5-Hydroxymethyl Furfural-Stimulated β-Glucosidase Retrieved From a Brazilian Secondary Atlantic Forest Soil Metagenome

Luana de Fátima Alves[1], Luana Parras Meleiro[2], Roberto N. Silva[1], Cauã Antunes Westmann[3] and María-Eugenia Guazzaroni[4]*

[1] Department of Biochemistry and Immunology, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, São Paulo, Brazil, [2] Department of Chemistry, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, São Paulo, Brazil, [3] Department of Cellular and Molecular Biology, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, São Paulo, Brazil, [4] Department of Biology, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, São Paulo, Brazil

Beta-glucosidases are key enzymes involved in lignocellulosic biomass degradation for bioethanol production, which complete the final step during cellulose hydrolysis by converting cellobiose into glucose. Currently, industry requires enzymes with improved catalytic performance or tolerance to process-specific parameters. In this sense, metagenomics has become a powerful tool for accessing and exploring the biochemical biodiversity present in different natural environments. Here, we report the identification of a novel β-glucosidase from metagenomic DNA isolated from soil samples enriched with decaying plant matter from a Secondary Atlantic Forest region. For this, we employed a functional screening approach using an optimized and synthetic broad host-range vector for library production. The novel β-glucosidase – named Lfa2 – displays three GH3-family conserved domains and conserved catalytic amino acids D283 and E487. The purified enzyme was most active in pH 5.5 and at 50°C, and showed hydrolytic activity toward several pNP synthetic substrates containing β-glucose, β-galactose, β-xylose, β-fucose, and α-arabinopyranose, as well as toward cellobiose. Lfa2 showed considerable glucose tolerance, exhibiting an IC_{50} of 300 mM glucose and 30% of remaining activity in 600 mM glucose. In addition, Lfa2 retained full or slightly enhanced activity in the presence of several metal ions. Further, β-glucosidase activity was increased by 1.7-fold in the presence of 10% (v/v) ethanol, a concentration that can be reached in conventional fermentation processes. Similarly, Lfa2 showed 1.7-fold enhanced activity at high concentrations of 5-hydroxymethyl furfural, one of the most important cellulase inhibitors in pretreated sugarcane bagasse hydrolysates. Moreover, the synergistic effect of Lfa2 on *Bacillus subtilis* GH5-CBM3 endoglucanase activity

was demonstrated by the increased production of glucose (1.6-fold). Together, these results indicate that β-glucosidase Lfa2 is a promissory enzyme candidate for utilization in diverse industrial applications, such as cellulosic biomass degradation or flavor enhancement in winemaking and grape processing.

INTRODUCTION

Conversion of lignocellulosic biomass into biofuels is a promising alternative to replace fossil fuels derived from non-renewable energy sources. Lignocellulose consists primarily of cellulose (40–60%), hemicelluloses (20–40%) and lignin (10–25%) (McKendry, 2002). Cellulose is a linear polymer of D-glucose subunits bound by β-1,4-glycosidic linkages with a polymerization degree ranging from few hundreds of D-glucose units to more than thousands. Hydroxyl groups of D-glucose units interact by hydrogen bonds resulting in a compact crystalline structure. As a result, although cellulose is the most recalcitrant material present in plant cell walls, it also has a substantial potential as a bioethanol production source (Sun et al., 2016).

The classical cellulose degradation framework for fungi includes: (i) the endo-1,4-β-glucanases (commonly found in GH families 5, 9, 44, and 45) which cleave the inner bonds of cellulose chain; (ii) exo-1,4-β-glucanases or cellobiohydrolases (most often from GH families 6, 7, 9, and 48) which release cellobiose molecules from either the reducing or non-reducing ends of a cellulose polymer, and (iii) the β-glucosidases (for example, belonging to the GH1 and GH3 families), which hydrolyze cellobiose in glucose molecules (Kubicek, 1992; Nidetzky et al., 1994; van den Brink and de Vries, 2011; Glass et al., 2013; Lombard et al., 2014). A limiting step in enzymatic cellulose saccharification is the conversion of oligosaccharides and cellobiose, which are products of endoglucanases and cellobiohydrolases hydrolysis, into glucose, and it has been demonstrated that the reaction products inhibit activities of most cellobiohydrolases and endoglucanases (Gruno et al., 2004; Xin et al., 2015; Yamamoto and Tamaru, 2016). In this context, β-glucosidases reduce cellobiose inhibition by hydrolyzing cellobiose into glucose, playing an important role in cellulose degradation process and acting as a key rate-limiting enzyme (Wojtusik et al., 2017). Further, synergism between cellulases plays an important role in biomass degradation and it has been reported for many cellulolytic enzyme systems, including free enzymes and also those that are part of cellulosomes (Beukes et al., 2008; Doi, 2008; Hu et al., 2013; Yang M. et al., 2016; Zhang et al., 2017). Quantitatively, when the combination of two enzymes is more efficient than the sum of each enzyme activities acting separately, the two enzymes show synergy. Due to the synergistic effect, one enzyme is able to accelerate the action of the other, with a consequent increase in hydrolysis yield (Saini et al., 2014).

Currently, one of the bottlenecks for second-generation bioethanol production is the high cost (Brijwani et al., 2010;

Koppram et al., 2014) and the low efficiency of enzymes required for the hydrolysis of cellulosic materials into fermentable sugars (Li et al., 2016). It means that there is an increasing demand for new strategies to reduce process costs and for new biocatalysts with improved properties for industrial applications, such as high catalytic efficiency, increased stability at high temperatures and certain pHs, biocatalysts that are not inhibited by the product, as well as not inhibited by toxic compounds resulting from lignocelluloses pretreatments (Ximenes et al., 2010, 2011; Sun et al., 2016). According to Ramos et al. (2016), the second generation ethanol production expenses with enzymes can be about 15% of the total alcohol production cost. In this way, many efforts have been made in order to enhance the efficiency of enzyme production and the activity of these enzymes, as well as in finding new enzymes with such features (Ramos et al., 2016). In this context, metagenomics allows the identification of new enzymes with specific activities without the need of previous isolation and cultivation of microorganisms, which opens the door to the huge biochemical potential of most of the microbial life existing in environments of interest (Guazzaroni et al., 2015; Alves et al., 2017). Several examples reported in literature showed the identification of new cellulases and other enzymes from different environments through metagenomic approaches (Mori et al., 2014; Montella et al., 2015; Yang C. et al., 2016; Zhou et al., 2016; Lewin et al., 2017; Tiwari et al., 2017, 2018; Dadheech et al., 2018; Lee et al., 2018).

Enzymes presenting tolerance to organic solvents, such as ethanol, are preferred candidates for enzymatic hydrolysis in processes such as simultaneous saccharification and fermentation (SSF), used in the production of second generation bioethanol, in which ethanol concentrations may reach up to 8% (w/v) (Liu et al., 2016; Qin et al., 2018). In this context, ethanol-tolerant cellulases are interesting enzymes, once they could remain active in these conditions. Furthermore, enzymes that are tolerant to toxic compounds, such as 5-hydroxymethyl furfural (5-HMF), produced during biomass pretreatments, are preferably desired.

Here, we report the identification and biochemical characterization of a novel ethanol-and 5-HMF-stimulated β-glucosidase, using an activity-based metagenomic screening strategy. For this, we used DNA from a microbial community inhabitant of soil samples enriched with decaying plant matter from a Brazilian Secondary Atlantic Forest region. The rationale behind this choice was that environmental biomass-rich soil samples should also be enriched in genes encoding for cellulases. Additionally, a synthetic broad host-range vector was used for library generation and for subsequently confirming, in a straightforward manner, the enzyme activity in other hosts more suitable than *Escherichia coli* for industrial processes. In

this sense, we showed Lfa activity against several substrates in *Pseudomonas putida*, a natural occurring soil bacterium with high metabolic versatility, robustness and significant potential for biotechnological applications (Wierckx et al., 2005; Craig et al., 2010; Beuttler et al., 2011; Van Duuren et al., 2011; Yu et al., 2016). Moreover, Lfa showed enhanced activity in harsh operating conditions, such as in presence of one of the most toxic inhibitory by-products from pretreated lignocellulose and also in 10% ethanol. Finally, we applied Lfa2 to the hydrolysis of a commercial polymeric substrate (carboxymethyl cellulose), in combination with GH5-CBM3 endoglucanase from *Bacillus subtilis*, to demonstrate its efficiency for enhancing the production of glucose.

MATERIALS AND METHODS

Metagenomic Library Construction and β-Glucosidase Screening

The soil sample was obtained on July 2015 from a Secondary Atlantic Forest region at University of São Paulo (Ribeirão Preto, São Paulo, Brazil; 21°09′58.4′′S, 47°51′20.1′′W, at an altitude of 540 m) which was naturally enriched with decaying plant matter and was stored at −20°C until DNA extraction. For the library construction, metagenomic DNA was extracted using the UltraClean^{TM} Soil DNA isolation kit (Mo Bio, United States), partially digested using *Sau*3AI and fragments ranging from 2 to 7 kb were recovered directly from an agarose gel and ligated into a dephosphorylated and *Bam*HI-digested pSEVA232 vector (Silva-Rocha et al., 2013) (for vector map, see **Supplementary Figure S4A**). The resulting plasmids were transformed by electroporation in *E. coli* DH10B cells (for detailed workflow, see **Supplementary Figure S3**). The resulting metagenomic library (named LFA-USP3) was analyzed for the percentage of plasmids bearing metagenomic DNA and average insert size.

Screening for β-glucosidase activity was performed by plating LFA-USP3 metagenomic library into LB-agar plates supplemented with kanamycin (50 μg/ml) and streptomycin (50 μg/ml) and containing 0.15% (w/v) esculin hydrate and 0.03% (w/v) ferric chloride as described by Eberhart in 1964 (Eberhart et al., 1964). *E. coli* clones were incubated at 37°C for 24 h and further incubation at 30°C for 4 days was performed. Positive clones were identified by the appearance of dark halos surrounding the colonies. Plasmids of potential positive colonies were extracted from individual clones, retransformed into *E. coli* DH10B cells and plated on LB-agar plates containing 0.15% esculin hydrate and 0.03% ferric chloride to confirm the phenotype. Thus, one positive clone was confirmed, herein named pLFA2 (for vector map, see **Supplementary Figure S4B**), and its DNA insert was fully sequenced by using an Applied Biosystems 3500xL Genetic Analyzer sequencer (Applied Biosystems, United States).

In silico Analysis and 3D-Structure Model Generation

The putative ORFs were predicted using ORF Finder program available in https://www.ncbi.nlm.nih.gov/orffinder/. ORF

nucleotide sequences were translated to amino acid sequences using Translate tools program[1] and putative amino acid sequences were analyzed using BlastP against non-redundant databases. Amino acid sequences were further analyzed for protein domains using Pfam database[2], for signal peptide sequences using SignalP server[3] and for physical and chemical parameters using ProtParam tool[4]. The 3D-structure was generated by SWISS-MODEL server using the automated mode[5] (Schwede et al., 2003; Benkert et al., 2011; Waterhouse et al., 2018) and the catalytic site prediction in complex with glucose was generated by the ITASSER server[6] (Zhang, 2008; Roy et al., 2010; Yang et al., 2014). All structures were analyzed using the PyMol software[7]. Multiple alignment of β-glucosidase with other GH3 family members was performed using Clustal Omega server[8] and the multiple alignment printing was generated by the BoxShade server[9]. Twenty-seven amino acid sequences from characterized β-glucosidases were recovered from CAZy database[10] to perform a phylogenetic reconstruction analysis along with the amino acid sequence of the Lfa2 using MEGA 6.0 software (Kumar et al., 1994; Tamura et al., 2013). Neighbor-joining statistical method was applied and accuracy of phylogenetic analysis was predicted by using 1,000 bootstrapping replications.

Cloning, Gene Expression and Protein Purification

The putative β-glucosidase gene, named *lfa2*, was amplified by PCR using the primers: 5′ GACT**GGATCC**AT GAAATCCAGACTCGTAGCC 3′ containing a *Bam*HI restriction site (in bold) and 5′ AGTC**CTCGAG**TCA CTTTGACGACGATAGCTC 3′ containing a *Xho*I restriction site (in bold) and using the pLFA2 plasmid as template. PCR conditions were as follows: initial denaturation at 98°C for 30 s, followed by 30 cycles of 98°C for 10 s, 65°C for 20 s, 72°C for 60 s and a final extension step of 72°C for 5 min. Then, the PCR product was digested with *Bam*HI and *Xho*I restriction enzymes and ligated into a pET28a vector previously digested by the same enzymes. The resulting clone was fully sequenced to check for the absence of mutations and transformed by electroporation into *E. coli* Rosetta (DE3) cells.

For expression and purification experiments, a single colony of *E. coli* Rosetta (DE3) cell harboring the resulting plasmid pET28a-lfa2 was grown in LB containing kanamycin (50 μg/ml) and chloramphenicol (36 μg/ml) at 37°C with shaking until to an $A_{600\ nm}$ of ∼0.8. Then, protein expression was induced by 0.1 mM IPTG followed by 18 h incubation at 18°C. Cells

---

[1]https://web.expasy.org/translate/
[2]http://pfam.xfam.org/
[3]http://www.cbs.dtu.dk/services/SignalP/
[4]https://web.expasy.org/protparam/
[5]https://swissmodel.expasy.org/
[6]https://zhanglab.ccmb.med.umich.edu/I-TASSER/
[7]https://pymol.org/2/
[8]https://www.ebi.ac.uk/Tools/msa/clustalo/
[9]https://embnet.vital-it.ch/software/BOX_faq.html
[10]http://www.cazy.org/

were harvest by centrifugation at 4,000 rpm for 30 min and cell pellet was resuspended in 6 mL of lysis buffer [30 mM HEPES pH 7.4, 300 mM NaCl, 1 mM PMSF, 1% (v/v) Triton-X and 20 mM imidazole]. After disruption by sonication, cell debris were collected by centrifugation at 10,000 rpm for 30 min and the supernatant was loaded on a nickel-charged HisTrap HP affinity column (GE Healthcare, United Kingdom). Protein purification was performed by using 30 mM HEPES buffer pH 7.4, 300 mM NaCl and increasing imidazole concentrations (40 and 80 mM). Finally, protein was eluted with the same buffer containing 300 mM imidazole. All fractions were collected and samples were analyzed by 12% SDS-PAGE (Laemmli, 1970). Purified protein concentration measurements were performed according to Read and collaborators using bovine serum albumin as standard (Read and Northcote, 1981).

Biochemical Characterization

Biochemical properties of the β-glucosidase were determined using the purified enzyme and all experiments were performed in triplicate with suitable controls. Specific activity (U/mg) of β-glucosidase was determined at 50°C using 3 mM of pNP-based substrates including pNP-β-D-glucopyranoside (pNPβGlu), pNP-β-D-galactopyranoside (pNPβGal), pNP-β-D-xylopyranoside (pNPβXyl), pNP-α-D-xylopyranoside (pNPαXyl), pNP-β-D-fucopyranoside (pNPβFuc), pNP-β-D-mannopyranoside (pNPβMan), pNP-α-L-arabinopyranoside (pNPαAra) in 50 mM of citrate-phosphate buffer pH 5.5 and 0.3 μg of enzyme in a final volume of 30 μL. Reactions were stopped by addiction of 30 μL of a $Na_2[B_4O_5(OH)_4]$ saturated solution and the amount of pNP released were measured in a 96-well plate at 405 nm in a xMark Microplate Spectrophotometer (Bio-Rad). One unit (U) was defined as the amount of β-glucosidase required to release 1 μmol of pNP per minute under assay conditions. For cellobiose, β-D-lactose, maltose and sucrose the amount of glucose released was measured using a glucose oxidase-based kit (Labtest, BR), following the manufacturer recommendations. Kinetic parameters (K_M, V_{max}, and k_{cat}) were determined at 50°C in a substrate range of 0 to 15 mM for pNP-β-D-glucopyranoside and of 0 to 150 mM for cellobiose in 50 mM of citrate-phosphate buffer pH 5.5 and 0.5 μg of β-glucosidase. The data were analyzed with the Graph Prism 5.0 software (Graph-Pad Prism Software, United States).

Effect of temperature on enzyme activity was evaluated in temperatures ranging from 25 to 70°C in 50 mM of citric acid-phosphate buffer pH 5.5. For optimal pH determination, we used the range 2.0–8.0 at 50°C, using 50 mM of citric acid-phosphate buffer. Results were expressed in relative activity percentage according to either the optimum pH or optimum temperature.

The effects of several different compounds on β-glucosidase (0.2 μg) activity were analyzed using 4 mM of pNP-β-D-glucopyranoside as substrate. In these assays, the chelating agent ethylenediaminetetraacetic acid (EDTA) was tested at the final concentration of 10 mM. Each divalent metal ions Mg^{2+}, Co^{2+}, Mn^{2+}, Cu^{2+}, Ca^{2+}, Ni^{2+}, Zn^{2+}, Fe^{2+} (using their respective salts: $MgCl_2$, $CoCl_2$, $MnSO_4$, $CuSO_4$, $CaCl_2$,

$NiSO_4$, $ZnSO_4$, and $FeSO_4$) were tested at the final concentration of 10 mM, as well. Ethanol and DMSO were assayed each one at 10, 25, and 50% (v/v) and NaCl was added at final concentrations of 250 mM and 1 M. Ethanol assays were performed in tightly closed tubes, in order to avoid ethanol evaporation. All reactions were performed in 50 mM of citric acid-phosphate buffer pH 5.5 at 50°C in a tabletop heater. Control assays were performed using 4 mM of pNP-β-D-glucopyranoside in 50 mM of citric acid-phosphate buffer pH 5.5 at 50°C without any addictives.

Similarly, the inhibitory effect of different concentrations of vanillin (0–2%), 4-hydroxybenzoic acid (4-HBA) (0–2%), furfural (0–1%) and 5-hydroxymethyl furfural (5-HMF) (0–1%) and the inhibitory effect of different concentrations of glucose (0–1.35 M) were determined by incubating the purified β-glucosidase (0.2 μg) with 4 mM of pNP-β-D-glucopyranoside in 50 mM of citric acid-phosphate buffer pH 5.5 at 50°C. Likewise, control assays were performed using 4 mM of pNP-β-D-glucopyranoside in 50 mM of citric acid-phosphate buffer pH 5.5 at 50°C without any addictives. The IC50 is defined as the concentration of glucose that inhibited 50% of enzymatic activity when compared with the control (without glucose). All experiments were analyzed in percentage of relative activities according to the control assay (without addictives).

Hydrolysis of Commercial Polymeric Substrate

Hydrolysis of carboxymethyl cellulose sodium salt (CMC) medium viscosity (Sigma, United States) by a GH5-CBM3 endoglucanase (BsCel5A) from B. subtilis 168 (in crude enzymatic extract) and β-glucosidase Lfa2 was performed in a 200 μL total volume containing 0.5% of CMC (w/v) in 50 mM of citric acid-phosphate buffer. Plasmid pSEVA242-BsCel5A was transferred to E. coli DH10B and a single colony was grown overnight in LB medium containing kanamycin (50 μg/ml). Overnight culture was used to inoculate 100 mL of LB medium and the culture was incubated at 37°C, 120 rpm for 24 h. Cells were collected by centrifugation and resuspended in 1 mL of 0.1 M citric acid-phosphate buffer pH 7.4 and disrupted by sonication. Cell debris were collected by centrifugation and 10 μL (10 U of BsCel5A endoglucanase per gram of CMC) of the crude protein extract was added to the reaction mixture which was incubated at 40°C per 45 min. Then, the reaction was supplemented with 20 μL (85 U of β-glucosidase per gram of CMC) of the purified β-glucosidase and incubated for further 1h. Glucose released was measured using a glucose oxidase-based kit (Labtest, BR). The synergism degree was defined as the ratio of the glucose amount released by the action of both enzymes to the sum of the glucose amount released by each enzyme independently.

Evaluation of β-Glucosidase Activity in Crude Protein Extracts From Different Hosts

In order to evaluate β-glucosidase activity expressed in different hosts, pLFA2 plasmid was transferred to P. putida KT2440 and E. coli DH10B. Overnight cultures of each bacteria harboring pLFA2 plasmid were used to inoculate 100 mL of LB medium.

Cultures were incubated at 30°C, 120 rpm for 48 h. Cultures of *P. putida* KT2440 and *E. coli* DH10B harboring pSEVA232 empty vector were used as negative controls (experiments with *p*NPβGlu, *p*NPβXyl, and *p*NPβFuc). When experiments were performed toward *p*NPβGal, cells of both bacteria without plasmid were used as negative controls, since pSEVA232 empty vector harbor the lacZα gene fragment, which encodes a β-galactosidase that could interfere in results interpretation. Cells were collected by centrifugation, resuspended in 1 mL of 0.1 M citric acid-phosphate buffer pH 7.4 and disrupted by sonication. Cell debris were collected by centrifugation and 5 μL of the crude protein extracts were used to compare specific activity of β-glucosidase expressed by two different hosts in 2 mM of different *p*NP-based substrates.

Statistical Analysis

Tukey test at 5% significance level was accomplished by using the software Statistic 7.0.

RESULTS

Generation and Evaluation of LFA-USP3 Metagenomic Library and β-Glucosidase Screening

A metagenomic library, herein named LFA-USP3, was generated from an environmental soil sample collected at South America, from a Secondary Atlantic Forest at the University of São Paulo (Ribeirão Preto, São Paulo, Brazil) enriched in decaying leaves from *Anadenanthera* sp. Metagenomic DNA was partially digested and cloned into the pSEVA232 vector (Silva-Rocha et al., 2013). The library comprised around 257 Mb of environmental DNA distributed into approximately 63,000 clones (100,000 clones – total library – which 63% carrying inserts) carrying insert fragments with an average size of 4.1 kb, thus presenting a number of genomes estimated in about 57; assuming 4.5 Mb per genome (Raes et al., 2007). In total, approximately 216,000 clones from LFA-USP3 library were screened for β-glucosidase activity using an esculin-based plate assay. We identified some dark halo-forming colonies which were selected for DNA plasmid extraction and retransformation into *E. coli* DH10B cells for verifying phenotype maintenance. After retransformation, one colony recovered the phenotype observed in the initial screening. This colony was selected for DNA plasmid extraction and its metagenomic insert was fully sequenced using a primer walking approach. The insert sequence of approximately 5.8 kb presented three open reading frames, designated as Lfa1, encoding for a protein displaying a conserved domain for the AXE1 superfamily (Acetyl Xylan Esterase superfamily), Lfa2, which encodes a potential GH3 glycosyl hydrolase, and Lfa3, encoding a truncated carboxypeptidase (**Figure 1A** and **Table 1**). Lfa2 consists of a 2,358-bp fragment encoding 785 amino acids. It shares the highest amino acid sequence identity (59%) with a non-characterized GH3 glycosyl hydrolase from *Pyrinomonas methylaliphatogenesis* (Genbank No. WP_041974888.1), a thermophilic and acidophilic Gram-negative bacteria from soil. The new sequence is available at Genbank with accession number MH397474.

Phylogenetic analysis of β-glucosidase Lfa2 (**Figure 1B**) with other 27 β-glucosidase amino acid sequences showed the presence of three different groups in the phylogenetic tree: (i) GH3 β-glucosidases from bacteria and archaea, (ii) GH3 β-glucosidases from fungi and (iii) GH1 β-glucosidases from bacteria, archaea and fungi. In accordance with previous bioinformatic analyses, Lfa2 is located in the first group, indicating that it is possibly sharing some functional and structural features with proteins positioned in this group.

Sequence and 3D-Structure Analysis of the β-Glucosidase Lfa2

Multiple amino acid sequence alignment of Lfa2 with other members of the GH3 family is shown in **Figure 2**. All proteins used in the alignment have shown activity against cellobiose in previous works and shared highly conserved sequences (Pozzo et al., 2010; Yoshida et al., 2010; Pozo et al., 2012; McAndrew et al., 2013). From this alignment, we found that the catalytic nucleophile (D283) and the acid/base amino acid (E487) residues were conserved (see **Supplementary Material** for full amino acid alignment, **Supplementary Figure S1**), which are also fully conserved in members of the GH3 family (Dan et al., 2000; Pozzo et al., 2010; Yoshida et al., 2010; McAndrew et al., 2013; Thongpoo et al., 2013), consistent with their key roles as catalytic amino acids.

The predicted 3D-structure model of the Lfa2 (89% of coverage, residues 34–762) was obtained from Swiss Model server based on the crystal structure of the β-glucosidase 3B from *Thermotoga neapolitana* (PDB: 2x42), which shares 49.2% of amino acid identity and 44% of amino acid similarity with Lfa2. The resulting GMQE and QMEAN scores were 0.7 and −1.38, respectively (Benkert et al., 2011; Waterhouse et al., 2018). As shown in **Figure 3A**, the predicted 3D-structure is composed by an α/β triose phosphate isomerase (TIM) barrel-like domain, an α/β sandwich domain and a C-terminal fibronectin type III domain, according to other β-glucosidases from GH3 family (Pozzo et al., 2010; Suzuki et al., 2013). Side chains of the important functional amino acids involved in substrate recognition and hydrolysis D98, R171, N204, R215, Y251, and S404 are highlighted in red. The TIM barrel-like domain and the α/β sandwich domains play an important role in the active-site organization; the last domain may be involved with substrate interaction and could affect enzyme stability (Pozzo et al., 2010). Additionally, the interaction analysis between a molecule of glucose in complex with Lfa2 showed that D283 (located in the TIM barrel-like domain) and E487 (positioned in the α/β sandwich domain) were conserved (**Figure 3B**).

Biochemical Characteristics of the β-Glucosidase Lfa2

Lfa2 was overexpressed and purified from cell-free enzymatic extract by Ni^{2+} affinity chromatography. Molecular mass of the enzyme was estimated to be 84 kDa using the
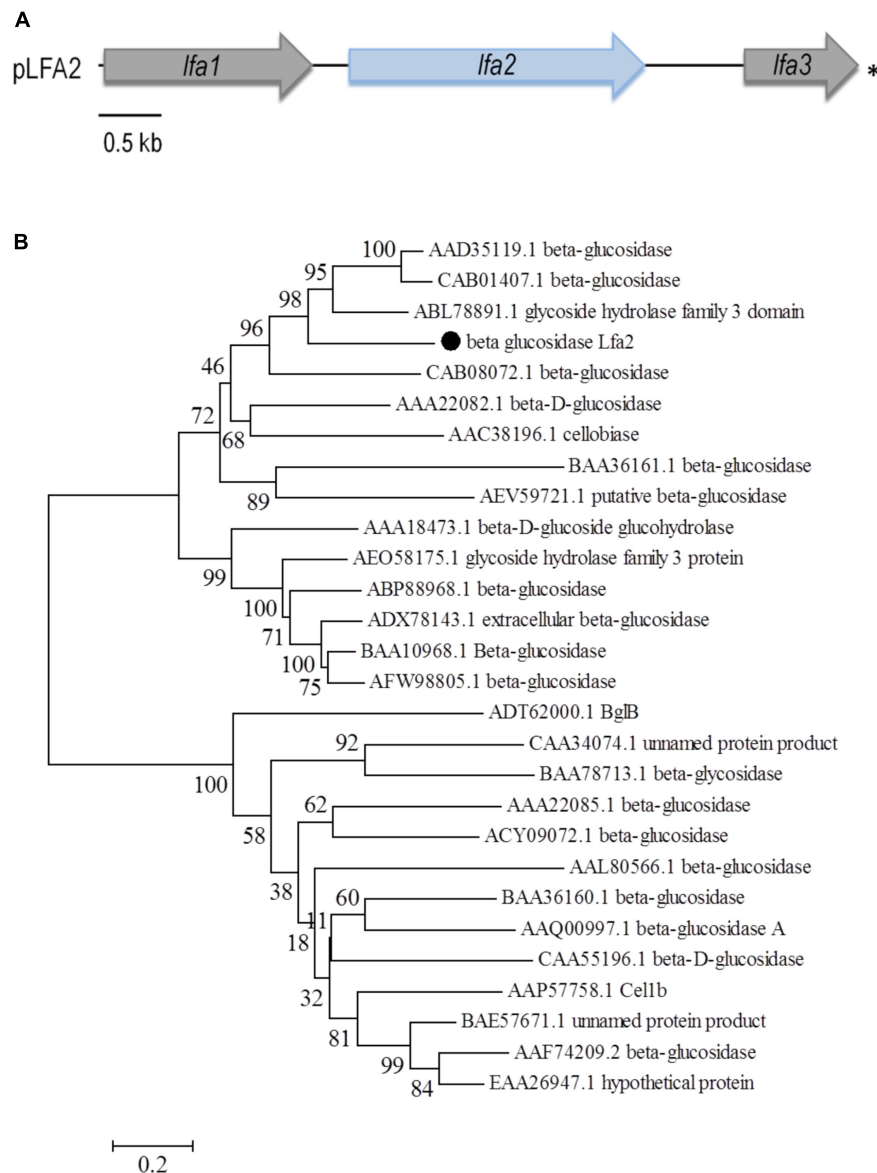
FIGURE 1 | Schematic representation of the ORFs identified in the pLFA2 plasmid and phylogenetic analysis of β-glucosidase Lfa2. **(A)** The complete metagenomic insert is shown. Arrows display the location, length and transcriptional orientation of the ORFs. Lfa2 ORF is showed as a light blue arrow. Asterisk indicates incomplete ORFs. **(B)** Phylogenetic tree of 28 candidate sequences using the Neighbor-Joining (NJ) method. Amino acid sequences were recovered from GenBank as indicate by accession numbers as follows: AAD35119.1 (*Thermotoga maritima* MSB8), CAB01407.1 (*Thermotoga neapolitana*), ABL78891.1 (*Thermofilum pendens* Hrk 5), MH397474 (beta glucosidase Lfa2), CAB08072.1 (*Clostridium stercorarium*), AAA22082.1 (*Agrobacterium tumefaciens*), AAC38196.1 (*Cellulomonas biazotea*), BAA36161.1 (*Bacillus* sp.), AEV59721.1 (uncultured bacterium), AAA18473.1 (*Trichoderma reesei*), AEO58175.1 (*Thermothelomyces thermophila* ATCC 42464), ABP88968.1 (*Penicillium brasilianum*), ADX78143.1 (*Aspergillus fumigatus*), BAA10968.1 (*Aspergillus aculeatus*), AFW98805.1 (*Aspergillus niger*), ADT62000.1 (bacterium enrichment culture), CAA34074.1 (*Sulfolobus solfataricus*), BAA78713.1 (*Thermococcus kodakarensis* KOD1), AAA22085.1 (*Agrobacterium* sp.), ACY09072.1 (uncultured bacterium [2]), AAL80566.1 (*Pyrococcus furiosus* DSM 3638), BAA36160.1 (*Bacillus* sp.), AAQ00997.1 (*Clostridium cellulovorans*), CAA55196.1 (*Avena sativa*), AAP57758.1 (*Trichoderma reesei*), BAE57671.1 (*Aspergillus oryzae* RIB40), AAF74209.2 (*Aspergillus niger*), EAA26947.1 (*Neurospora crassa*). Numbers on nodes correspond to percentage bootstrap values for 1,000 replicates.

ProtParam server[11], with a theoretical pI of 9.24. The purified recombinant protein appeared on SDS-PAGE electrophoresis gel as a band (see **Supplementary Figure S2**) between 66.2 and 116 kDa bands. Biochemical characterization of

[11]https://web.expasy.org/protparam/

the Lfa2 was investigated using the purified recombinant enzyme.

The optimal temperature for Lfa2 was determined to be 50°C, while the enzyme activity was approximately 50% of the maximum activity at a range from 40 to 55°C, maintaining more than 20% of the maximum activity at a range from 35 to 60°C

TABLE 1 | Description of the ORFs contained in pLFA2 plasmid and their sequence similarities.

Plasmid name [insert (bp)]	% G + C	ORF[a]	Length (aa)[b]	Closest similar protein (length in aa)	Organism	Identity (%)	E-value	Putative function	No. Of TMH[c]	Signal Peptide
pLFA2 (5,868)	53,2	1	577	hypothetical protein A2W03_04630 [OGD23027.1] (641)	Candidatus Aminicenantes bacterium RBG_16_63_16	56	<5.00E-133	Acetyl xylan esterase	0	No
		2	785	Glycosyl hydrolase family 3 [WP_041974888.1] (787)	Pyrinomonas methylaliphatogenes	59	<5.00E-133	Glycosyl hydrolase family 3	0	Yes[d]
		3*	385**	Carboxypeptidase [OGU07865.1] (461)	Gemmatimonadetes bacterium RBG 16668	62	5.00E-133	Carboxypeptidase	1	No

[a]Open reading frame. [b]Amino acids. [c]Number of transmembrane helices (TMH) predicted by using the TMHMM program (http://www.cbs.dtu.dk/services/TMHMM-2.0/). [d]According to SignalP software (http://www.cbs.dtu.dk/services/SignalP/), the clivage site is between aminoacids 22 and 23; considering an analysis for Gram-negative bacteria. *Asterisks indicate incomplete ORFs and **proteins truncated at C-terminal end.

(**Figure 4A**). Lfa2 was most active at pH 5.5, retaining more than 60% of the maximum activity in a pH range from 5 to 6.5 (**Figure 4B**).

Effect of Glucose on the Enzymatic Activity of Lfa2

In order to evaluate the effect of increasing glucose concentrations on Lfa2 activity, catalytic activity was measured using pNP-Glu as a substrate in presence of 0–1.35 M of glucose. As shown in **Figure 4C**, the enzyme presented considerable tolerance to glucose inhibition, presenting an IC_{50} of 300 mM glucose. When glucose concentrations were further increased, the enzyme activity was gradually inhibited, with 20% of remaining activity at higher glucose concentrations as 800 mM.

Effects of Metal Ions, Organic Solvents, EDTA and Salt on Lfa2 Activity

The effect of divalent metal ions, organic solvents, chelating agent and salt on the enzymatic activity of Lfa2 were analyzed (**Table 2**). The enzymatic activity of Lfa2 was slightly increased in the presence of Mg^{2+}, Co^{2+}, and Mn^{2+} and strongly inhibited in the presence of Cu^{2+} and Fe^{2+}. The relative enzymatic activity improved to $130.8 \pm 4.3\%$ (Mg^{2+}) $134.1 \pm 11.2\%$ (Co^{2+}) and $109.4 \pm 0.3\%$ (Mn^{2+}), but it was reduced to 5.8 ± 1.3 and $17.4 \pm 4.1\%$ in presence of Cu^{2+} and Fe^{2+}, respectively. The effects of Ca^{2+}, Ni^{2+}, and Zn^{2+} on the enzymatic activity were not significant. Chelating agent EDTA (ethylenediaminetetraacetic acid) reduced the relative activity to $79.3 \pm 1.6\%$. On the other hand, relative activity was increased approximately 30% in presence of NaCl. Enzymatic activity was inhibited in presence of increasing concentrations of the organic solvent DMSO (dimethyl sulfoxide). Instead, in presence of 10% ethanol the relative enzymatic activity was improved in 70% relative to the control (without additives). Additionally, increasing concentrations of ethanol (25 and 50%) reduced enzyme activity to approximately 59 and 35%, respectively.

Substrate Specificity and Kinetic Parameters of Lfa2

Under optimal pH and temperature parameters, the hydrolytic activity of Lfa2 toward various substrates was measured and the substrate specificity was determined (**Table 3**). Among the chromogenic substrates, β-glucosidase Lfa2 showed activity toward pNPβGlu, pNPβXyl, pNPαAra, pNPβFuc, and pNPβGal, while no activity was observed upon pNPαXyl and pNPβMan. Lfa2 showed activity toward cellobiose, but no detectable activity was observed toward other oligosaccharides tested. The enzyme showed to be more active toward pNPβGlu, followed by pNPβXyl and cellobiose which were approximately 25 and 10% of the maximum activity observed toward pNPβGlu. These results indicate that Lfa2 is a β-glucosidase with a broad specificity substrate range.

Enzymatic reaction rate regarding substrate concentration of pNPβGlu and cellobiose were tested and it followed the

FIGURE 2 | Multiple alignment of sequences from β-glucosidase Lfa2 and other GH3 family β-glucosidases. Amino acid sequence alignment of the metagenomic β-glucosidase Lfa2 was carried out using the Clustal Omega server. Results were shaded using BOXSHADE (black background = strictly conserved; gray or white background = conservatively substituted or non-conserved). Asterisks indicate the nucleophile residue D283 and the acid/base amino acid residue E487. Full species names and Genbank, PDB or NCBI IDs of the β-glucosidases sequences are shown as follow: *Kluyveromyces marxianus* PDB ID: 3ABZ (3ABZ); metagenomic β-glucosidases from cow rumen NCBI accession numbers: JX163905, JX163906, and JX163904 (AGH13475.1, AGH13484.1, and AGH13451.1); *Thermotoga neapolitana* PDB ID: 2X41 (2X41) and metagenomic β-glucosidaseLfa2 from Secondary Atlantic Forest soil (this study) Genbank accession number: MH397474 (Lfa2).
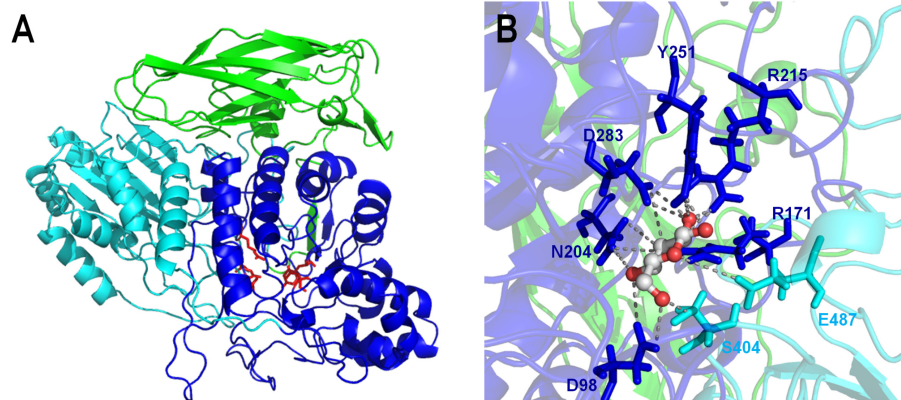
FIGURE 3 | 3D-structure model of the β-glucosidase Lfa2 characterized in this work. **(A)** Overall 3D-structure representation showing the three domains of the β-glucosidase: TIM barrel-like domain (in blue), α/β sandwich domain (in cyan) and the C-terminal fibronectin type III domain (in green). Side chains of the catalytic amino acids and amino acids involved in substrate binding are shown in red. **(B)** Catalytic site representation in complex with glucose. Amino acid side chains are represented by sticks, indicating the conserved catalytic amino acids D283 and E487 and other amino acids involved in substrate recognition and stabilization (D98, R171, N204, R215, Y251, and S404).

Michaelis–Menten kinetics. Then, half-saturation coefficient (K_M), catalytic rate constant (k_{cat}) and catalytic efficiency (k_{cat}/K_M) values were determined (**Table 4**). According to k_{cat}/K_M values, Lfa2 showed greater efficiency for pNPβGlu (17.4×10^3 s^{-1} M^{-1}) when compared with cellobiose (3.02×10^2 s^{-1} M^{-1}).

FIGURE 4 | Effects of temperature, pH and glucose concentration on the activity of the β-glucosidase Lfa2. **(A)** Effect of temperature on the activity of Lfa2 was determined in pH 5.5 (citric acid-phosphate buffer 50 mM) in a range of temperature from 25 to 70°C. **(B)** Effect of pH on the activity of Lfa2 was determined in the range of pH 2.0–8.0 at 50°C using citric acid-phosphate buffer (50 mM). **(C)** Influence of increasing glucose concentrations (0–1.35 M) on Lfa2 activity was performed in citric acid-phosphate buffer 50 mM pH 5.5 at 50°C. All enzymatic properties were determined using pNPβG as substrate. All measurements were done in triplicates. Error bars show SD. Activity at 100% refers to a specific activity of 9.6 μmol min^{-1} mg^{-1}.

TABLE 2 | Influence of metal ions, chelating agent, organic solvents and salt on Lfa2 activity.

Compound	Relative activity (%)
Control (no addictives)	100 ± 6.1
10 mM MgCl$_2$	130.8 ± 4.3
10 mM CoCl$_2$	134.1 ± 11.2
10 mM MnSO$_4$	109.4 ± 0.3*
10 mM CuSO$_4$	5.8 ± 1.3
10 mM CaCl$_2$	105.7 ± 5.5*
10 mM NiSO$_4$	92.1 ± 3*
10 mM ZnSO$_4$	94.6 ± 8.8*
10 mM FeSO$_4$	17.4 ± 4.1
10 mM EDTA	79.3 ± 1.6
250 mM NaCl	137.4 ± 6.5
1 M NaCl	135.9 ± 6.1
10% DMSO	86.6 ± 3.1*
25% DMSO	65.5 ± 3
50% DMSO	21.3 ± 1.2
10% EtOH	170.1 ± 1.7
25% EtOH	59.1 ± 2.4
50% EtOH	35.6 ± 0.8

*No significant difference when compared with the control (no addictives) as revealed by Tukey test, $p < 0.05$. Activity at 100% refers to a specific activity of 7.9 μmol min^{-1} mg^{-1} using pNPβGlu as substrate.

TABLE 3 | Substrate specificity of Lfa2 for chromogenic substrates and oligossacarides.

Substrate (3 mM)	Specific activity (U/mg)
Chromogenic substrates	
pNPβGlu	8.09 ± 0.01[a]
pNPαAra	0.63 ± 0.09[cd]
pNPβXyl	2.1 ± 0.2[b]
pNPβGal	0.26 ± 0.04[e]
pNPβFuc	0.38 ± 0.01[de]
pNPαXyl	n.d*
pNPβMan	n.d*
Oligosaccharides	
Cellobiose	0.73 ± 0.03[c]
Maltose	n.d*
Lactose	n.d*
Sucrose	n.d*

*n.d – activity not detected. [a,b,c,d,e]Indicate significant difference between different specific activities for each substrate, as revealed by Tukey test, $p < 0.05$.

Effect of Lignocellulose-Derived Compounds on Lfa2 Activity

Effects of four potential cellulase inhibitors, furfural, 5-hydroxymethyl furfural (5-HMF), vanillin and 4-hydroxybenzoic acid (4-HBA) on Lfa2 activity were analyzed using 4 mM pNPβGlu as substrate. Different concentrations of furfural (0–1%), 5-HMF (0–1%), vanillin (0–2%), and 4-HBA (0–2%) showed contrasting effects on Lfa2 activity (**Figure 5**). For instance, Lfa2 activity was strongly inhibited in presence of 0.5 and 1% 4-HBA (enzymatic activity decreased approximately 40 and 90%, respectively), while 2% 4-HBA produced practically 100% inhibition. Additionally, 0.1% vanillin resulted in 30% of Lfa2 inhibition, whereas in the presence of 2% vanillin enzyme activity showed 70% of Lfa2 inhibition. Furfural displayed a slight effect on Lfa2 activity, inhibiting enzymatic activity no more than

30%, even in higher concentrations, as 1%. On the other hand, 5-HMF did not affect enzymatic activity in lower concentrations as 0.05 and 0.1%, whereas in higher concentrations, such as 0.5 and 1%, unexpectedly, 5-HMF activated Lfa2 activity, increasing pNPβGlu hydrolysis in 60 and 70%, respectively.

Synergistic Effect of β-Glycosidase Lfa2 on *Bacillus subtilis* GH5-CBM3 Endoglucanase Activity

The synergistic effect of purified β-glycosidase Lfa2 and GH5-CBM3 (BsCel5A) endoglucanase of *B. subtilis* (in crude extract) was investigated using CMC as substrate. For this, enzymatic crude extract containing *Bs*Cel5A endoglucanase expressed from pSEVA242-*Bs*Cel5 plasmid and the purified Lfa2 were used. As shown in **Figure 6**, Lfa2 and crude enzymatic extract from empty plasmid pSEVA242 were unable to produce glucose from CMC. On the other hand, when combining endoglucanase activity of *Bs*Cel5A with β-glucosidase activity of Lfa2, the amount of glucose released was about 1.6-fold higher compared with endoglucanase alone. This result indicates that Lfa2 has positive effects on cellulose hydrolysis, acting in synergy with *Bs*Cel5A. The degree of synergism was estimated in 1.53-fold which was calculated according to the methodology in Section "Hydrolysis of Commercial Polymeric Substrate."

Evaluation of Lfa2 Activity in Crude Protein Extract Generated From Different Hosts

As previously mentioned, a synthetic broad host-range vector was used for library production. Plasmid pSEVA232 has an origin of replication that can be recognized by a large number of Gram-negative hosts (Silva-Rocha et al., 2013). Thus, activity of Lfa2 was verified in crude extracts from *E. coli* DH10B and *P. putida* KT2440 harboring pLFA2 plasmid. As shown in **Figure 7**, different specific activities (U/mL) toward pNP-derived substrates were observed among both bacteria, being lower in crude extract from *P. putida* when compared with *E. coli*, probably due to differences in the protein expression levels in both hosts (Troeschel et al., 2012).
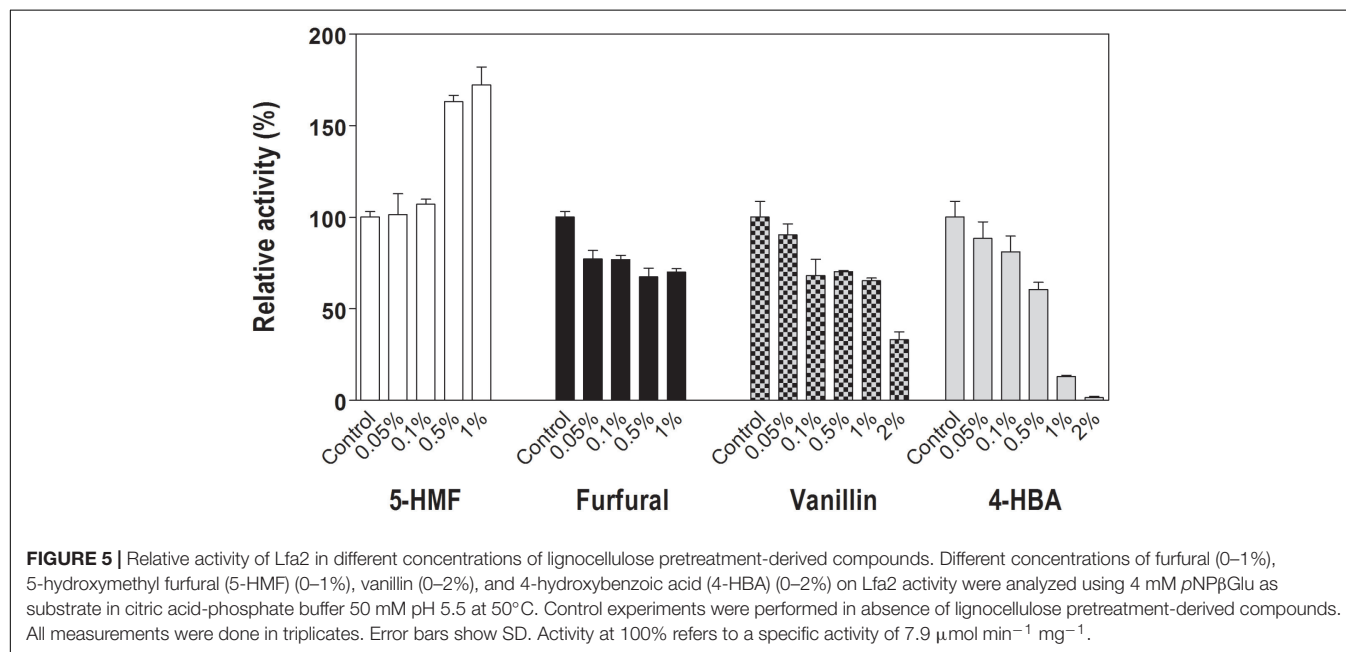
DISCUSSION

Metagenomic approaches have been successfully used in biotechnological fields, mining new biocatalysts for industrial purposes. In particular, several efforts have been made to find cellulases with relevant characteristics for specific parameters used in ethanol industry (Rees et al., 2003; Ferrer et al., 2005; Schröder et al., 2014; Garg et al., 2016; Lewin et al., 2017; Dadheech et al., 2018; Duque et al., 2018; Lee et al., 2018; Tiwari et al., 2018). Here, we report the identification and biochemical characterization of a novel ethanol- and 5-HMF-stimulated β-glucosidase, using a functional-based metagenomic screening strategy. For metagenomic library generation, the pSEVA232 synthetic broad host-range vector was used (Silva-Rocha et al., 2013). Thus, library construction in a wide-ranging host vector

allowed us to confirm in a straightforward manner the enzyme activity in other hosts more suitable than *E. coli* for industrial applications. In this sense, we tested Lfa2 activity in *P. putida*, a metabolically robust bacterium with high reducing power output and low cellular maintenance demand (Wierckx et al., 2005; Craig et al., 2010; Beuttler et al., 2011; Van Duuren et al., 2011; Ng et al., 2015; Yu et al., 2016). We observed that Lfa2 was active against several substrates in both bacteria (**Figure 7**), but displayed different specific activities, probably due to particular expression levels of heterologous proteins in different hosts. Differences in plasmid copy number, codon usage, promoter and ribosome binding site recognition may lead to different expression levels in distinct hosts and it is difficult to predict which host strain may be the best choice for an heterologous protein expression (Terpe, 2006). In a recent study, expression of two reporter proteins was compared when expressed in *E. coli* BL21 and *P. putida* KT2440. Using a broad-host range shuttle vector based on PBBR1 origin of replication (same origin that plasmid pSEV232 used in this study), authors showed that plasmid copy number was similar in *E. coli* and *P. putida*, but expression of the reporter proteins in *P. putida* resulted in decreased enzymatic activities when compared with *E. coli* (Troeschel et al., 2012). Thus, a metagenomic library constructed in a vector that allows straightforward plasmid transfer to other bacterial hosts should simplify the process of identifying a proper host for ideal heterologous protein expression.

β-glucosidases are key enzymes involved in lignocellulosic biomass degradation for bioethanol production (Koppram et al., 2014; Liu and Cotta, 2015; Treebupachatsakul et al., 2016) which complete the final step during cellulose hydrolysis by converting cellobiose into glucose (Bhatia et al., 2002), a key rate-limiting step in cellulose degradation. Classification of glycosyl hydrolases based on amino-acid sequence similarities was proposed some years ago (Henrissat et al., 1991; Henrissat and Bairoch, 1996), and β-glucosidases were divided among two different families: GH1 and GH3. The amino acid sequence of Lfa2 displayed the highest sequence identity (59%) with a GH3 glycosyl hydrolase from *P. methylaliphatogenes* (Genbank No. WP_041974888.1). Furthermore, phylogenetic analysis showed that Lfa2 has a close relationship with GH3 β-glucosidases from bacteria and archaea. Thus, *in silico* and experimental results indicated that Lfa2 is a β-glucosidase that belongs to the GH3 family, presenting conserved domains and conserved catalytic amino acids of this family. The 3D-structure model revealed three conserved domains: α/β triose phosphate isomerase (TIM) barrel-like domain, an α/β sandwich domain and a C-terminal fibronectin-like type III domain, also present in other β-glucosidases from GH3 family (Pozzo et al., 2010; McAndrew et al., 2013; Suzuki et al., 2013). The two former domains play an important role in the active site formation (which is positioned at the interface of this two domains), since the catalytic residues (the nucleophile residue D283 and the acid/base amino acid residue E487) as well as other amino acids involved in substrate recognition, binding and/or stabilization (D98, R171, N204, R215, Y251, and S404) are conserved. The fibronectin-like type III domain possibly does not play a role in small substrates recognition, due to its position, but it could

TABLE 4 | Kinetic parameters of the purified β-glucosidase Lfa2.

Substrate	$V_{máx}$ (μmol min^{-1} mg^{-1})	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} M^{-1})
pNPβGlu	9.4 ± 0.2	0.76 ± 0.06	13.2 ± 0.2	17.4 × 10^3
Cellobiose	16.2 ± 1.2	75.3 ± 11.6	22.8 ± 1.7	3.02 × 10^2

FIGURE 5 | Relative activity of Lfa2 in different concentrations of lignocellulose pretreatment-derived compounds. Different concentrations of furfural (0–1%), 5-hydroxymethyl furfural (5-HMF) (0–1%), vanillin (0–2%), and 4-hydroxybenzoic acid (4-HBA) (0–2%) on Lfa2 activity were analyzed using 4 mM pNPβGlu as substrate in citric acid-phosphate buffer 50 mM pH 5.5 at 50°C. Control experiments were performed in absence of lignocellulose pretreatment-derived compounds. All measurements were done in triplicates. Error bars show SD. Activity at 100% refers to a specific activity of 7.9 μmol min^{-1} mg^{-1}.

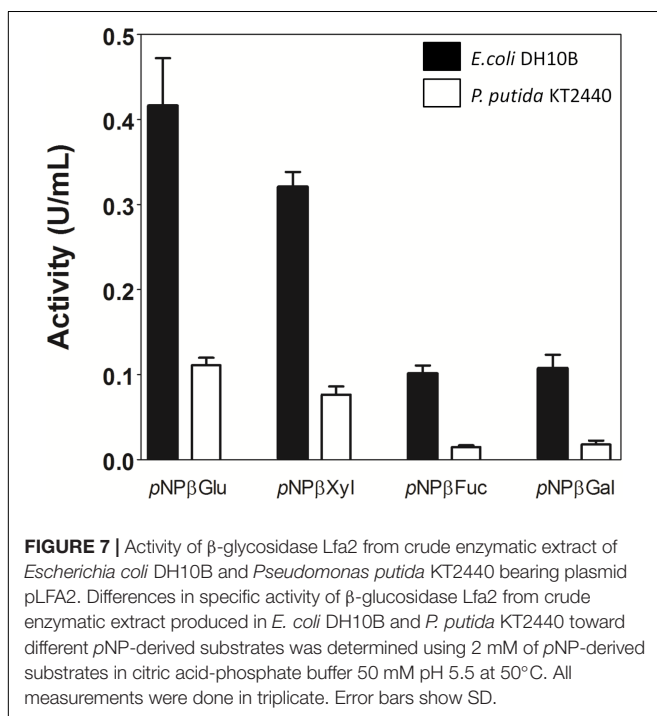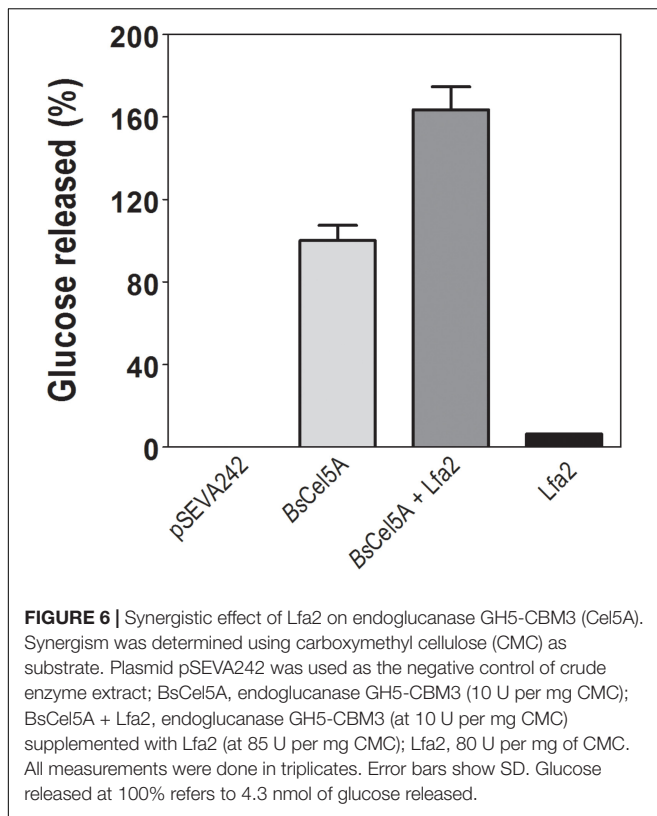be involved in thermostability and binding in long polymeric substrates.

Regarding substrate specificity, β-Glucosidases can be divided into three different groups, namely, (1) aryl-β-glucosidases, which display strong affinity to aryl-β-glucosides; (2) true cellobiases, which hydrolyze only oligosaccharides and (3) broad substrate-range, that show activity on a large range of substrates (Rojas et al., 1995; Riou et al., 1998). The β-glucosidase identified in this study, which showed activity on a wide range of pNP substrates, such as pNPβGlu, pNPβXyl, pNPαAra, pNPβFuc, pNPβGal, and cellobiose, but presented no activity upon pNPαXyl, pNPβMan, sucrose, maltose, and lactose, belongs to the first group. Kinetic parameters of Lfa2 were measured and the enzyme exhibited catalytic efficiency coefficients (k_{cat}/K_M) of 17.4 × 10^3 s^{-1} M^{-1} for pNPβGlu and 3.02 × 10^2 s^{-1} M^{-1} for cellobiose. This values of catalytic efficiency are between some values of catalytic efficiency already reported for β-glucosidases from bacteria, fungi and metagenomes (we found values ranging from 2.9 × 10^4 s^{-1} M^{-1} to 1.6 s^{-1} M^{-1} for pnpβGlu) (Zanoelo et al., 2004; Karnchanatat et al., 2007; Joo et al., 2009; Yapi Assoi Yapi et al., 2009; Beloqui et al., 2010; Cobucci-Ponzano et al., 2010; Jiang et al., 2010, 2011; Naz et al., 2010; Fan et al., 2011; Kim et al., 2011, 2006; Bao et al., 2012; Haq et al., 2012; Zhao et al., 2012) (see **Supplementary Table S2**). On the other hand, these values are lower than catalytic efficiencies determined to the mainly commercial β-glucosidases (Seidle et al., 2004; Chauve et al., 2010). Additionally, Lfa2 showed considerable glucose tolerance capacity with an IC$_{50}$ of 300 mM. Most microbial

β-glycosidases are inhibited by glucose, which is an important limitation for its use in industry (Leite et al., 2008). High glucose concentrations may interfere directly or indirectly with substrate binding at the active site of the enzyme, reducing reaction rates (Singhania et al., 2013). It is worth mentioning that some glucose tolerant β-glucosidases were already isolated and characterized (Zanoelo et al., 2004; Pei et al., 2012; Zhao et al., 2013; Matsuzawa and Yaoi, 2017), presenting a value of IC$_{50}$ higher than found for Lfa2 (Pei et al., 2012; Lu et al., 2013). For instance, Matsuzawa and Yaoi (2017) described a β-glucosidase stimulated by saccharides that presented more than sixfold of increased activity in presence of 250 mM glucose. Pei and collaborators, showed a β-glucosidase with an IC$_{50}$ of 600 mM glucose. In other study, Lu and collaborators showed a β-glucosidase-tolerant enzyme with an IC$_{50}$ of 1500 mM glucose. Many studies have suggested that glucose tolerance and stimulation of β–glucosidases properties are associated with transglycosylation activity (Uchiyama et al., 2013, 2015; Yang et al., 2015), so enzymes with these features could overcome the product inhibition by transglycosylation activity during the cellulose hydrolysis (Wang et al., 2017).

Notably, effect of ethanol on enzymatic activity is essential for β-glycosidase characterizations, since these enzymes are exposed to substantial concentrations of ethanol in a number of applications, such as the simultaneous saccharification and fermentation process (Garcia et al., 2015; Liu et al., 2016; Sun et al., 2016). Ethanol concentrations of 10% increased the enzymatic activity toward pNPβGlu by 1.7-fold compared with the activity without ethanol addition (**Table 2**). Moreover,

FIGURE 6 | Synergistic effect of Lfa2 on endoglucanase GH5-CBM3 (Cel5A). Synergism was determined using carboxymethyl cellulose (CMC) as substrate. Plasmid pSEVA242 was used as the negative control of crude enzyme extract; BsCel5A, endoglucanase GH5-CBM3 (10 U per mg CMC); BsCel5A + Lfa2, endoglucanase GH5-CBM3 (at 10 U per mg CMC) supplemented with Lfa2 (at 85 U per mg CMC); Lfa2, 80 U per mg of CMC. All measurements were done in triplicates. Error bars show SD. Glucose released at 100% refers to 4.3 nmol of glucose released.

FIGURE 7 | Activity of β-glycosidase Lfa2 from crude enzymatic extract of *Escherichia coli* DH10B and *Pseudomonas putida* KT2440 bearing plasmid pLFA2. Differences in specific activity of β-glucosidase Lfa2 from crude enzymatic extract produced in *E. coli* DH10B and *P. putida* KT2440 toward different *p*NP-derived substrates was determined using 2 mM of *p*NP-derived substrates in citric acid-phosphate buffer 50 mM pH 5.5 at 50°C. All measurements were done in triplicate. Error bars show SD.

increased ethanol concentrations were not able to completely inactivate Lfa2 activity, which remained with almost 60 and 35% of activity under 25 and 50% ethanol, respectively. It is worth mentioning that these activities are still high in comparison to

other enzymes described as tolerant and/or activated by ethanol (Gueguen et al., 1996; Manzanares Rojas et al., 2000; Gosset and Martinez, 2012; Karnaouri et al., 2013; Garcia et al., 2015). Most β-glucosidases are inhibited in presence of ethanol, even in lower concentrations (Zhou et al., 2016; Matsuzawa and Yaoi, 2017). Though, some β-glucosidases activated by ethanol and/or ethanol-tolerant have been described (Parry et al., 2001; Muñoz-Gutiérrez et al., 2012; Karnaouri et al., 2013; Fang et al., 2016; Wu et al., 2018; Xue et al., 2018) (as showed in **Supplementary Table S1**). Uchiyama et al. (2013) described a metagenomic β-glucosidase activated 1.16-fold by ethanol 10% whereas 25% ethanol dramatically reduced enzyme activity. Parry et al. (2001) described a thermostable β-glucosidase from *Thermoascus aurantiacus* which was activated by 40% at 30% ethanol and inactivated at concentrations higher than 50% ethanol. A GH3 β-glucosidase with remarkable stimulation and tolerance to ethanol was characterized from *Myceliophthora thermophila* by Karnaouri et al. (2013). Enzyme activity was increased twofold in a range from 10 to 30% of ethanol. Considering that the final concentration of ethanol in a conventional fermentation process is approximately 10–15% (Koppram et al., 2014), enzymes with tolerance toward ethanol may be useful for application in saccharification processes for bioethanol production (Baffi et al., 2013).

Another interesting feature of Lfa2 is its behavior in presence of lignocellulose pretreatment-derived compounds. Some conditions used during lignocellulose pretreatment may lead to the generation of certain compounds that could affect the subsequent enzymatic hydrolysis (inhibiting enzymes) and fermentation (inhibiting microorganisms growth and survival) (Lavarack et al., 2002; Ximenes et al., 2011; Quéméneur et al., 2012). Generation of these compounds depends on the lignocellulose source, the type of pretreatment used and the conditions employed in pretreatment (Du et al., 2010; van der Pol et al., 2014). The main compounds generated from pentoses and hexoses degradation are furfural and 5-hydroxymethylfurfural (5-HMF) and, among phenolic compounds, 4-hydroxybenzoic acid (4-HBA) and vanillin should be highlighted among the most toxic ones (Tomás-Pejó et al., 2011). Here, we tested the enzymatic activity of Lfa2 toward *p*NPβGlu in the presence of different concentrations of furfural (0–1%), 5-HMF (0–1%), 4-HBA (0–2%) and vanillin (0–2%). Furfural and vanillin were not able to strongly inhibit Lfa2 activity, even at higher concentrations as 1 and 2%, Lfa2 activity has remained approximately 70 and 33%, respectively, when compared with the control, whereas in presence of 4-HBA 2%, Lfa2 activity was dramatically decreased. On the other hand, in the presence of higher concentrations of 5-HMF (1%), the Lfa2 activity was enhanced in more than 70%. Most cellulases are inhibited in the presence of this lignocellulose-derived compounds (Ximenes et al., 2010, 2011; Maruthamuthu and Van Elsas, 2017), however, cellulases are found to be more sensitive to inhibition by some phenolic compounds than β-glucosidases (Berlin et al., 2006; Ximenes et al., 2010). Furfural and 5-HMF concentrations in sugarcane bagasse hydrolysates pretreated by acid solution are 0.1–1.25% (w/v) and less than 0.1%, respectively (Chen et al., 2012; Tizazu and Moholkar, 2018), whereas in corn stover the

levels of 5-HMF and furfural were 0.001–0.004% and 0.0026–0.02% (w/v), respectively (Du et al., 2010). Vanillin and 4-HBA appears in lower concentrations in corn stover hydrolysates, being vanillin ranging from 2.8 to 4.0 mg/L and 4-HBA around 0.028 mg/L. Interestingly, Lfa2 has remained active in higher concentrations of lignocellulose-derived compounds than the ones reported for sugarcane bagasse and corn stover hydrolysates. To the best of our knowledge, Lfa2 is the first β-glucosidase described in the literature as stimulated in presence of 5-HMF. Until now, two native β-glucosidases from the fungus *Clavispora* have been described as resistant to 5-HMF, but its activity was not enhanced in presence of the compound. Wang et al. (2016) showed that 0.12% furfural and 5-HMF did not affect the activity of these two β-glucosidases.

During the process of enzymatic lignocellulose degradation, β-glucosidases play an important role hydrolyzing cello-oligosaccharides – such as cellobiose and cellotrioses – into glucose (Singhania et al., 2013). Cello-oligosaccharides are produced during the degradation of cellulose by cellulases such as endo-β-glucanases and cellobiohydrolases, thus, β-glucosidase activity is essential for efficient plant biomass saccharification process, since cello-oligosaccharides may inhibit the activity of endo-β-glucanases and cellobiohydrolases. Chauve et al. (2010) demonstrated that a β-glucosidase from *Aspergillus niger* may be used to achieve complete hydrolysis of oligosaccharides produced by endoglucanases. Other studies have shown that the addition of β-glucosidases in hydrolysis of some lignocellulosic matters improved hydrolysis in around 20% when compared with non-supplemented reactions (Han and Chen, 2008; Pallapolu et al., 2011; Zhang et al., 2017). Here, we demonstrated that β-glucosidase Lfa2 acts in synergy with endoglucanase *Bs*Cel5A of *B. subtilis* hydrolyzing CMC. We showed in a straightforward experiment that addition of Lfa2 increased in 60% the glucose release compared to the non-supplemented reaction (*Bs*Cel5A alone).

CONCLUSION

In summary, we have found and characterized a novel ethanol- and 5-hydroxymethyl furfural-stimulated β-glucosidase recovered from a Brazilian Secondary Atlantic Forest soil metagenome. The combined properties of Lfa2, including its increased activity in the presence of ethanol; medium tolerance level to glucose; high tolerance to 5-HMF; low responsiveness

to several metal ions; synergistic effect when coupled to endoglucanase activity and broad substrate specificity, supported this enzyme as a very promising candidate for utilization in a wide range of industrial applications, such as cellulosic biomass degradation or flavor enhancement in winemaking and grape processing.

DATA AVAILABILITY

Nucleotide sequence obtained for the plasmid insert have been deposited in the GenBank database under the accession number (MH397474), as described in Results session.

AUTHOR CONTRIBUTIONS

LA and M-EG designed the experiments. LA, LM, RS, and CW performed the experiments. LA and LM analyzed the data. LA and M-EG wrote the manuscript. All authors reviewed the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.02556/full#supplementary-material

REFERENCES

Alves, L. D. F., Silva-Rocha, R., and Guazzaroni, M.-E. (2017). "Enhancing metagenomic approaches through synthetic biology," in *Functional Metagenomics: Tools and Applications*María-Eugenia Guazzaroni, eds T. C. Charles, M. R. Liles, and A. Sessitsch (Berlin: Springer International Publishing), 1–14. doi: 10.1007/978-3-319-61510-3

Baffi, M. A., Martin, N., Tobal, T. M., Ferrarezi, A. L., Lago, J. H. G., Boscolo, M., et al. (2013). Purification and characterization of an ethanol-tolerant β-glucosidase from sporidiobolus pararoseus and its potential for hydrolysis of wine aroma precursors. *Appl. Biochem. Biotechnol.* 171, 1681–1691. doi: 10.1007/s12010-013-0471-0

Bao, L., Huang, Q., Chang, L., Sun, Q., Zhou, J., and Lu, H. (2012). Cloning and characterization of two β-glucosidase/xylosidase enzymes from yak rumen metagenome. *Appl. Biochem. Biotechnol.* 166, 72–86. doi: 10.1007/s12010-011-9405-x

Beloqui, A., Nechitaylo, T. Y., López-Cortés, N., Ghazi, A., Guazzaroni, M. E., Polaina, J., et al. (2010). Diversity of glycosyl hydrolases from cellulose-depleting communities enriched from casts of two earthworm species. *Appl. Environ. Microbiol.* 76, 5934–5946. doi: 10.1128/AEM.00902-10

Benkert, P., Biasini, M., and Schwede, T. (2011). Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* 27, 343–350. doi: 10.1093/bioinformatics/btq662

Berlin, A., Balakshin, M., Gilkes, N., Kadla, J., Maximenko, V., Kubo, S., et al. (2006). Inhibition of cellulase, xylanase and β-glucosidase activities by softwood lignin preparations. *J. Biotechnol.* 125, 198–209. doi: 10.1016/j.jbiotec.2006.02.021

Beukes, N., Chan, H., Doi, R. H., and Pletschke, B. I. (2008). Synergistic associations between Clostridium cellulovorans enzymes XynA, ManA and EngE against sugarcane bagasse. *Enzyme Microb. Technol.* 42, 492–498. doi: 10.1016/j.enzmictec.2008.01.010

Beuttler, H., Hoffmann, J., Jeske, M., Hauer, B., Schmid, R. D., Altenbuchner, J., et al. (2011). Biosynthesis of zeaxanthin in recombinant *Pseudomonas putida*. *Appl. Microbiol. Biotechnol.* 89, 1137–1147. doi: 10.1007/s00253-010-2961-0

Bhatia, Y., Mishra, S., and Bisaria, V. S. (2002). Microbial β-Glucosidases: cloning, properties, and applications. *Crit. Rev. Biotechnol.* 22, 375–407. doi: 10.1080/07388550290789568

Brijwani, K., Oberoi, H. S., and Vadlani, P. V. (2010). Production of a cellulolytic enzyme system in mixed-culture solid-state fermentation of soybean hulls supplemented with wheat bran. *Process Biochem.* 45, 120–128. doi: 10.1016/j.procbio.2009.08.015

Chauve, M., Mathis, H., Huc, D., Casanave, D., Monot, F., and Ferreira, N. L. (2010). Comparative kinetic analysis of two fungal β-glucosidases. *Biotechnol. Biofuels* 3, 1–8. doi: 10.1186/1754-6834-3-3

Chen, W. H., Ye, S. C., and Sheen, H. K. (2012). Hydrolysis characteristics of sugarcane bagasse pretreated by dilute acid solution in a microwave irradiation environment. *Appl. Energy* 93, 237–244. doi: 10.1016/j.apenergy.2011.12.014

Cobucci-Ponzano, B., Aurilia, V., Riccio, G., Henrissat, B., Coutinho, P. M., Strazzulli, A., et al. (2010). A new archaeal β-glycosidase from *Sulfolobus solfataricus*: seeding a novel retaining β-glycan-specific glycoside hydrolase family along with the human non-lysosomal glucosylceramidase GBA. *J. Biol. Chem.* 285, 20691–20703. doi: 10.1074/jbc.M109.086470

Craig, J. W., Chang, F. Y., Kim, J. H., Obiajulu, S. C., and Brady, S. F. (2010). Expanding small-molecule functional metagenomics through parallel screening of broad-host-range cosmid environmental DNA libraries in diverse proteobacteria. *Appl. Environ. Microbiol.* 76, 1633–1641. doi: 10.1128/AEM.02169-09

Dadheech, T., Shah, R., Pandit, R., Hinsu, A., Chauhan, P. S., Jakhesara, S., et al. (2018). Cloning, molecular modeling and characterization of acidic cellulase from buffalo rumen and its applicability in saccharification of lignocellulosic biomass. *Int. J. Biol. Macromol.* 113, 73–81. doi: 10.1016/j.ijbiomac.2018.02.100

Dan, S., Marton, I., Dekel, M., Bravdo, B. A., He, S., Withers, S. G., et al. (2000). Cloning, expression, characterization, and nucleophile identification of family 3, *Aspergillus niger* -glucosidase. *J. Biol. Chem.* 275, 4973–4980. doi: 10.1074/jbc.275.7.4973

Doi, R. H. (2008). Cellulases of mesophilic microorganisms: Cellulosome and noncellulosome producers. *Ann. N. Y. Acad. Sci.* 1125, 267–279. doi: 10.1196/annals.1419.002

Du, B., Sharma, L. N., Becker, C., Chen, S. F., Mowery, R. A., van Walsum, G. P., et al. (2010). Effect of varying feedstock-pretreatment chemistry combinations on the formation and accumulation of potentially inhibitory degradation products in biomass hydrolysates. *Biotechnol. Bioeng.* 107, 430–440. doi: 10.1002/bit.22829

Duque, E., Daddaoua, A., Cordero, B. F., Udaondo, Z., Molina-santiago, C., Roca, A., et al. (2018). Brief report Ruminal metagenomic libraries as a source of relevant hemicellulolytic enzymes for biofuel production. *Microb. Biotechnol.* 11, 781–787. doi: 10.1111/1751-7915.13269

Eberhart, B., Cross, D. F., and Chase, L. R. (1964). B-glcosidase system of *Neurospora crassa*. *J. Bacteriol.* 87, 761–770.

Fan, H. X., Miao, L. L., Liu, Y., Liu, H. C., and Liu, Z. P. (2011). Gene cloning and characterization of a cold-adapted β-glucosidase belonging to glycosyl hydrolase family 1 from a psychrotolerant bacterium *Micrococcus antarcticus*. *Enzyme Microb. Technol.* 49, 94–99. doi: 10.1016/j.enzmictec.2011.03.001

Fang, W., Yang, Y., Zhang, X., Yin, Q., Zhang, X., Wang, X., et al. (2016). Improve ethanol tolerance of β-glucosidase Bgl1A by semi-rational engineering for the hydrolysis of soybean isoflavone glycosides. *J. Biotechnol.* 227, 64–71. doi: 10.1016/j.jbiotec.2016.04.022

Ferrer, M., Golyshina, O. V., Chernikova, T. N., Khachane, A. N., Reyes-Duarte, D., Santos, V. A., et al. (2005). Novel hydrolase diversity retrieved from a metagenome library of bovine rumen microflora. *Environ. Microbiol.* 7, 1996–2010. doi: 10.1111/j.1462-2920.2005.00920.x

Garcia, N. F. L., da Silva Santos, F. R., Gonçalves, F. A., da Paz, M. F., Fonseca, G. G., and Leite, R. S. R. (2015). Production of β-glucosidase on solid-state fermentation by *Lichtheimia ramosa* in agroindustrial residues: characterization and catalytic properties of the enzymatic extract. *Electron. J. Biotechnol.* 18, 1–6. doi: 10.1016/j.ejbt.2015.05.007

Garg, R., Srivastava, R., Brahma, V., Verma, L., Karthikeyan, S., and Sahni, G. (2016). Biochemical and structural characterization of a novel halotolerant cellulase from soil metagenome. *Sci. Rep.* 6, 1–15. doi: 10.1038/srep39634

Glass, N. L., Schmoll, M., Cate, J. H. D., and Coradetti, S. (2013). Plant cell wall deconstruction by *Ascomycete Fungi*. 67, 477–498. doi: 10.1146/annurev-micro-092611-150044

Gosset, G., and Martinez, A. (2012). Cell surface display of a b -glucosidase employing the type V secretion system on ethanologenic *Escherichia coli* for the fermentation of cellobiose to ethanol. *J. Ind. Microbiol. Biotechnol.* 39, 1141–1152. doi: 10.1007/s10295-012-1122-0

Gruno, M., Väljamäe, P., Pettersson, G., and Johansson, G. (2004). Inhibition of the *Trichoderma reesei* cellulases by cellobiose is strongly dependent on the nature of the substrate. *Biotechnol. Bioeng.* 86, 503–511. doi: 10.1002/bit.10838

Guazzaroni, M. E., Silva-Rocha, R., and Ward, R. J. (2015). Synthetic biology approaches to improve biocatalyst identification in metagenomic library screening. *Microb. Biotechnol.* 8, 52–64. doi: 10.1111/1751-7915.12146

Gueguen, Y., Chemardin, P., Janbon, G., Arnaud, A., and Galzy, P. (1996). A very efficient β-Glucosidase catalyst for the hydrolysis of flavor precursors of wines and fruit juices. *J. Agric. Food Chem.* 44, 2336–2340. doi: 10.1021/jf950360j

Han, Y., and Chen, H. (2008). Characterization of β-glucosidase from corn stover and its application in simultaneous saccharification and fermentation. *Bioresour. Technol.* 99, 6081–6087. doi: 10.1016/j.biortech.2007.12.050

Haq, I. U., Khan, M. A., Muneer, B., Hussain, Z., Afzal, S., Majeed, S., et al. (2012). Cloning, characterization and molecular docking of a highly thermostable β-1,4-glucosidase from *Thermotoga petrophila*. *Biotechnol. Lett.* 34, 1703–1709. doi: 10.1007/s10529-012-0953-0

Henrissat, B., and Bairoch, A. (1996). Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.* 316, 695–696. doi: 10.1042/bj3160695

Henrissat, B., Vegetales, M., and Grenoble, F. (1991). A classification of glycosyl hydrolases based sequence similarities amino acid. *Biochem. J.* 280, 309–316. doi: 10.1007/s007920050009

Hu, J., Arantes, V., Pribowo, A., and Saddler, J. N. (2013). The synergistic action of accessory enzymes enhances the hydrolytic potential of a "cellulase mixture" but is highly substrate specific. *Biotechnol. Biofuels* 6:112. doi: 10.1186/1754-6834-6-112

Jiang, C., Hao, Z. Y., Jin, K., Li, S. X., Che, Z. Q., Ma, G. F., et al. (2010). Identification of a metagenome-derived β-glucosidase from bioreactor contents. *J. Mol. Catal. B Enzym.* 63, 11–16. doi: 10.1016/j.molcatb.2009.11.009

Jiang, C., Li, S. X., Luo, F. F., Jin, K., Wang, Q., Hao, Z. Y., et al. (2011). Biochemical characterization of two novel β-glucosidase genes by metagenome expression cloning. *Bioresour. Technol.* 102, 3272–3278. doi: 10.1016/j.biortech.2010.09.114

Joo, A. R., Jeya, M., Lee, K. M., Sim, W. I., Kim, J. S., Kim, I. W., et al. (2009). Purification and characterization of a β-1,4-glucosidase from a newly isolated strain of *Fomitopsis pinicola*. *Appl. Microbiol. Biotechnol.* 83, 285–294. doi: 10.1007/s00253-009-1861-7

Karnaouri, A., Topakas, E., Paschos, T., Taouki, I., and Christakopoulos, P. (2013). Cloning, expression and characterization of an ethanol tolerant GH3 β-glucosidase from *Myceliophthora thermophila*. *PeerJ* 1:e46. doi: 10.7717/peerj.46

Karnchanatat, A., Petsom, A., Sangvanich, P., Piaphukiew, J., Whalley, A. J. S., Reynolds, C. D., et al. (2007). Purification and biochemical characterization of an extracellular β-glucosidase from the wood-decaying fungus *Daldinia eschscholzii* (Ehrenb.:Fr.) Rehm. *FEMS Microbiol. Lett.* 270, 162–170. doi: 10.1111/j.1574-6968.2007.00662.x

Kim, B. J., Singh, S. P., and Hayashi, K. (2006). Characteristics of chimeric enzymes constructed between *Thermotoga maritima* and *Agrobacterium tumefaciens* β-glucosidases: role of C-terminal domain in catalytic activity. *Enzyme Microb. Technol.* 38, 952–959. doi: 10.1016/j.enzmictec.2005.08.038

Kim, Y. S., Yeom, S. J., and Oh, D. K. (2011). Characterization of a GH3 family -glucosidase from Dictyoglomus turgidum and its application to the hydrolysis of isoflavone glycosides in spent coffee grounds. *J. Agric. Food Chem.* 59, 11812–11818. doi: 10.1021/jf2025192

Koppram, R., Tomás-Pejó, E., Xiros, C., and Olsson, L. (2014). Lignocellulosic ethanol production at high-gravity: challenges and perspectives. *Trends Biotechnol.* 32, 46–53. doi: 10.1016/j.tibtech.2013.10.003

Kubicek, C. P. (1992). The cellulase proteins of *Trichoderma reesei*: structure, multiplicity, mode of action and regulation of formation. *Enzym. Prod. Bact. Fungi Plant Cells* 45, 1–27. doi: 10.1007/BFb0008754

Kumar, S., Tamura, K., and Nei, M. (1994). MEGA: molecular evolutionary genetics analysis software for microcomputers. *Comput. Appl. Biosci.* 10, 189–191.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685. doi: 10.1038/227680a0

Lavarack, B. P., Griffin, G. J., and Rodman, D. (2002). The acid hydrolysis of sugarcane bagasse hemicellulose to produce xylose, arabinose, glucose and other products. *Biomass Bioenergy* 23, 367–380. doi: 10.1016/S0961-9534(02)00066-1

Lee, J. P., Lee, H. W., Na, H. B., Lee, J. H., Hong, Y. J., Jeon, J. M., et al. (2018). Characterization of truncated endo-β-1,4-glucanases from a compost metagenomic library and their saccharification potentials. *Int. J. Biol. Macromol.* 115, 554–562. doi: 10.1016/j.ijbiomac.2018.04.102

Leite, R. S. R., Alves-Prado, H. F., Cabral, H., Pagnocca, F. C., Gomes, E., and Da-Silva, R. (2008). Production and characteristics comparison of crude β-glucosidases produced by microorganisms *Thermoascus aurantiacus* e *Aureobasidium pullulans* in agricultural wastes. *Enzyme Microb. Technol.* 43, 391–395. doi: 10.1016/j.enzmictec.2008.07.006

Lewin, A., Zhou, J., Pham, V. T. T., Haugen, T., Zeiny, M. E., Aarstad, O., et al. (2017). Novel archaeal thermostable cellulases from an oil reservoir metagenome. *AMB Express* 7:183. doi: 10.1186/s13568-017-0485-z

Li, Y., Sun, Z., Ge, X., and Zhang, J. (2016). Effects of lignin and surfactant on adsorption and hydrolysis of cellulases on cellulose. *Biotechnol. Biofuels* 9:20. doi: 10.1186/s13068-016-0434-0

Liu, D., Zhang, H., Lin, C.-C., and Xu, B. (2016). Optimization of rice wine fermentation process based on the simultaneous saccharification and fermentation kinetic model. *Chinese J. Chem. Eng.* 24, 1406–1412. doi: 10.1016/j.cjche.2016.05.037

Liu, Z. L., and Cotta, M. A. (2015). Technical assessment of cellulosic ethanol production using β-Glucosidase producing yeast clavispora NRRL Y-50464. *Bioenergy Res.* 8, 1203–1211. doi: 10.1007/s12155-014-9575-9

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, 490–495. doi: 10.1093/nar/gkt1178

Lu, J., Du, L., Wei, Y., Hu, Y., and Huang, R. (2013). Expression and characterization of a novel highly glucose-tolerant beta-glucosidase from a soil metagenome. *Acta Biochim. Biophys. Sin.* 45, 664–673. doi: 10.1093/abbs/gmt061

Manzanares Rojas, V., Genoves, S., and Valles, S. P. (2000). A preliminary search for anthocyanin-ß-D-glucosidase activity in non-*Saccharomyces* wine yeasts. *Int. J. Food Sci. Technol* 35, 95–103. doi: 10.1046/j.1365-2621.2000.00364.x

Maruthamuthu, M., and Van Elsas, J. D. (2017). Molecular cloning, expression, and characterization of four novel thermo-alkaliphilic enzymes retrieved from a metagenomic library. *Biotechnol. Biofuels* 10, 1–17. doi: 10.1186/s13068-017-0808-y

Matsuzawa, T., and Yaoi, K. (2017). Screening, identification, and characterization of a novel saccharide-stimulated β-glycosidase from a soil metagenomic library. *Appl. Microbiol. Biotechnol.* 101, 633–646. doi: 10.1007/s00253-016-7803-2

McAndrew, R. P., Park, J. I., Heins, R. A., Reindl, W., Friedland, G. D., D'haeseleer, P., et al. (2013). From soil to structure, a novel dimeric β-glucosidase belonging to glycoside hydrolase family 3 isolated from compost using metagenomic analysis. *J. Biol. Chem.* 288, 14985–14992. doi: 10.1074/jbc.M113.458356

McKendry, P. (2002). Energy production from biomass (part 1): overview of biomass. *Bioresour. Technol.* 83, 37–46. doi: 10.1016/S0960-8524(01)00118-3

Montella, S., Amore, A., and Faraco, V. (2015). Metagenomics for the development of new biocatalysts to advance lignocellulose saccharification for bioeconomic development. *Crit. Rev. Biotechnol.* 8551, 1–12. doi: 10.3109/07388551.2015.1083939

Mori, T., Kamei, I., Hirai, H., and Kondo, R. (2014). Identification of novel glycosyl hydrolases with cellulolytic activity against crystalline cellulose from metagenomic libraries constructed from bacterial enrichment cultures. *Springerplus* 3:365. doi: 10.1186/2193-1801-3-365

Muñoz-Gutiérrez, I., Oropeza, R., Gosset, G., and Martinez, A. (2012). Cell surface display of a β-glucosidase employing the type v secretion system on ethanologenic *Escherichia coli* for the fermentation of cellobiose to ethanol. *J. Ind. Microbiol. Biotechnol.* 39, 1141–1152. doi: 10.1007/s10295-012-1122-0

Naz, S., Ikram, N., Rajoka, M. I., Sadaf, S., and Akhtar, M. W. (2010). Enhanced production and characterization of a beta-glucosidase from *Bacillus halodurans* expressed in *Escherichia coli*. *Biochem. Biokhimiia* 75, 513–525.

Ng, C. Y., Farasat, I., Maranas, C. D., and Salis, H. M. (2015). Rational design of a synthetic Entner-Doudoroff pathway for improved and controllable NADPH regeneration. *Metab. Eng.* 29, 86–96. doi: 10.1016/j.ymben.2015.03.001

Nidetzky, B., Steiner, W., Hayn, M., and Claeyssens, M. (1994). Cellulose hydrolysis by the cellulases from *Trichoderma reesei*: a new model for synergistic interaction. *Biochem. J.* 298(Pt 3), 705–710. doi: 10.1042/bj2980705

Pallapolu, V. R., Lee, Y. Y., Garlock, R. J., Balan, V., Dale, B. E., Kim, Y., et al. (2011). Effects of enzyme loading and β-glucosidase supplementation on enzymatic hydrolysis of switchgrass processed by leading pretreatment technologies. *Bioresour. Technol.* 102, 11115–11120. doi: 10.1016/j.biortech.2011.03.085

Parry, N. J., Beever, D. E., Owen, E., Vandenberghe, I., Van Beeumen, J., and Bhat, M. K. (2001). Biochemical characterization and mechanism of action of a thermostable beta-glucosidase purified from *Thermoascus aurantiacus*. *Biochem. J.* 353, 117–127.

Pei, J., Pang, Q., Zhao, L., Fan, S., and Shi, H. (2012). *Thermoanaerobacterium thermosaccharolyticum* β-glucosidase: a glucose-tolerant enzyme with high specific activity for cellobiose. *Biotechnol. Biofuels* 5:31. doi: 10.1186/1754-6834-5-31

Pozo, M. V., Del, Fernández-arrojo, L., Gil-martínez, J., Montesinos, A., Chernikova, T. N., Nechitaylo, T. Y., et al. (2012). Microbial β-glucosidases from cow rumen metagenome enhance the saccharification of lignocellulose in combination with commercial cellulase cocktail. *Biotechnol. Biofuels* 5:73.

Pozzo, T., Pasten, J. L., Karlsson, E. N., and Logan, D. T. (2010). Structural and functional analyses of -Glucosidase 3B from *Thermotoga neapolitana*: a Thermostable three-domain representative of glycoside hydrolase 3. *J. Mol. Biol.* 397, 724–739. doi: 10.1016/j.jmb.2010.01.072

Qin, L., Zhao, X., Li, W. C., Zhu, J. Q., Liu, L., Li, B. Z., et al. (2018). Biotechnology for Biofuels Process analysis and optimization of simultaneous saccharification and co-fermentation of ethylenediamine-pretreated corn stover for ethanol production. *Biotechnol. Biofuels* 11:118. doi: 10.1186/s13068-018-1118-8

Quéméneur, M., Hamelin, J., Barakat, A., Steyer, J. P., Carrre, H., and Trably, E. (2012). Inhibition of fermentative hydrogen production by lignocellulose-derived compounds in mixed cultures. *Int. J. Hydrogen Energy* 37, 3150–3159. doi: 10.1016/j.ijhydene.2011.11.033

Raes, J., Korbel, J. O., Lercher, M. J., von Mering, C., and Bork, P. (2007). Prediction of effective genome size in metagenomic samples. *Genome Biol.* 8:R10. doi: 10.1186/gb-2007-8-1-r10

Ramos, J. L., Valdivia, M., García-Lorente, F., and Segura, A. (2016). Benefits and perspectives on the use of biofuels. *Microb. Biotechnol.* 9, 436–440. doi: 10.1111/1751-7915.12356

Read, S. M., and Northcote, D. H. (1981). Minimization of variation in the response to different proteins of the Coomassie blue G dye-binding assay for protein. *Anal. Biochem.* 116, 53–64. doi: 10.1016/0003-2697(81)90321-3

Rees, H. C., Grant, S., Jones, B., Grant, W. D., and Heaphy, S. (2003). Detecting cellulase and esterase enzyme activities encoded by novel genes present in environmental DNA libraries. *Extremophiles* 7, 415–421. doi: 10.1007/s00792-003-0339-2

Riou, C., Salmon, J. M., Vallier, M. J., Günata, Z., and Barre, P. (1998). Purification, characterization, and substrate specificity of a novel highly glucose-tolerant beta-glucosidase from *Aspergillus oryzae*. *Appl. Environ. Microbiol.* 64, 3607–3614. doi: 10.1002/adsc.200700082

Rojas, A., Arola, L., and Romeu, A. (1995). beta-Glucosidase families revealed by computer analysis of protein sequences. *Biochem. Mol. Biol. Int.* 35, 1223–1231.

Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: A unified platform for automated protein structure and function prediction. *Nat. Protoc.* 5, 725–738. doi: 10.1038/nprot.2010.5

Saini, J. K., Saini, R., and Tewari, L. (2014). Lignocellulosic agriculture wastes as biomass feedstocks for second-generation bioethanol production: concepts and recent developments. *3 Biotech* 5, 337–353. doi: 10.1007/s13205-014-0246-5

Schröder, C., Elleuche, S., Blank, S., and Antranikian, G. (2014). Enzyme and microbial technology characterization of a heat-active archaeal $^{N}_{L}$-glucosidase

from a hydrothermal spring metagenome. *Enzyme Microb. Technol.* 57, 48–54. doi: 10.1016/j.enzmictec.2014.01.010

Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C. (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.* 31, 3381–3385. doi: 10.1093/nar/gkg520

Seidle, H. F., Marten, I., Shoseyov, O., and Huber, R. E. (2004). Physical and kinetic properties of the family 3 β-glucosidase from *Aspergillus niger* which is important for cellulose breakdown. *Protein J.* 23, 11–23. doi: 10.1023/B: JOPC.0000016254.58189.2a

Silva-Rocha, R., Martínez-García, E., Calles, B., Chavarría, M., Arce-Rodríguez, A., De Las Heras, A., et al. (2013). The standard European vector architecture (SEVA): A coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res.* 41, 666–675. doi: 10.1093/nar/gks1119

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

Sun, S., Sun, S., Cao, X., and Sun, R. (2016). The role of pretreatment in improving the enzymatic hydrolysis of lignocellulosic materials. *Bioresour. Technol.* 199, 49–58. doi: 10.1016/j.biortech.2015.08.061

Suzuki, K., Sumitani, J.-I., Nam, Y.-W., Nishimaki, T., Tani, S., Wakagi, T., et al. (2013). Crystal structures of glycoside hydrolase family 3 β-glucosidase 1 from *Aspergillus aculeatus*. *Biochem. J.* 452, 211–221. doi: 10.1042/BJ20130054

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* 72, 211–222. doi: 10.1007/s00253-006-0465-8

Thongpoo, P., McKee, L. S., Araújo, A. C., Kongsaeree, P. T., and Brumer, H. (2013). Identification of the acid/base catalyst of a glycoside hydrolase family 3 (GH3) β-glucosidase from *Aspergillus niger* ASKU28. *Biochim. Biophys. Acta* 1830, 2739–2749. doi: 10.1016/j.bbagen.2012.11.014

Tiwari, R., Nain, L., Labrou, N. E., and Shukla, P. (2017). Bioprospecting of functional cellulases from metagenome for second generation biofuel production: a review. *Crit. Rev. Microbiol.* 7828, 1–14. doi: 10.1080/1040841X.2017.1337713

Tiwari, R., Nain, L., Labrou, N. E., and Shukla, P. (2018). Bioprospecting of functional cellulases from metagenome for second generation biofuel production: a review. *Crit. Rev. Microbiol.* 44, 244–257. doi: 10.1080/1040841X.2017.1337713

Tizazu, B. Z., and Moholkar, V. S. (2018). Kinetic and thermodynamic analysis of dilute acid hydrolysis of sugarcane bagasse. *Bioresour. Technol.* 250, 197–203. doi: 10.1016/j.biortech.2017.11.032

Tomás-Pejó, E., Alvira, P., Ballesteros, M., and Negro, M. J. (2011). "Pretreatment technologies for lignocellulose-to-bioethanol conversion," in *Biofuels*, eds A. Pandey, C. Larroche, S. C. Ricke, C. G. Dussap, and E. Gnansounou (Amsterdam: Academic Press), 149–176. doi: 10.1016/B978-0-12-385099-7.00007-3

Treebupachatsakul, T., Nakazawa, H., Shinbo, H., Fujikawa, H., Nagaiwa, A., Ochiai, N., et al. (2016). Heterologously expressed *Aspergillus aculeatus* β-glucosidase in *Saccharomyces cerevisiae* is a cost-effective alternative to commercial supplementation of β-glucosidase in industrial ethanol production using *Trichoderma reesei* cellulases. *J. Biosci. Bioeng.* 121, 27–35. doi: 10.1016/j.jbiosc.2015.05.002

Troeschel, S. C., Thies, S., Link, O., Real, C. I., Knops, K., Wilhelm, S., et al. (2012). Novel broad host range shuttle vectors for expression in *Escherichia coli*, *Bacillus subtilis* and *Pseudomonas putida*. *J. Biotechnol.* 161, 71–79. doi: 10.1016/j.jbiotec.2012.02.020

Uchiyama, T., Miyazaki, K., and Yaoi, K. (2013). Characterization of a novel β-Glucosidase from a compost microbial metagenome with strong transglycosylation activity. *J. Biol. Chem.* 288, 18325–18334. doi: 10.1074/jbc.M113.471342

Uchiyama, T., Yaoi, K., and Miyazaki, K. (2015). Glucose-tolerant β-glucosidase retrieved from a Kusaya gravy metagenome. *Front. Microbiol.* 6:548. doi: 10.3389/fmicb.2015.00548

van den Brink, J., and de Vries, R. P. (2011). Fungal enzyme sets for plant polysaccharide degradation. *Appl. Microbiol. Biotechnol.* 91, 1477–1492. doi: 10.1007/s00253-011-3473-2

van der Pol, E. C., Bakker, R. R., Baets, P., and Eggink, G. (2014). By-products resulting from lignocellulose pretreatment and their inhibitory effect on fermentations for (bio)chemicals and fuels. *Appl. Microbiol. Biotechnol.* 98, 9579–9593. doi: 10.1007/s00253-014-6158-9

Van Duuren, J. B., Wijte, D., Leprince, A., Karge, B., Puchałka, J., Wery, J., et al. (2011). Generation of a catR deficient mutant of *P. putida* KT2440 that produces cis, cis-muconate from benzoate at high rate and yield. *J. Biotechnol.* 156, 163–172. doi: 10.1016/j.jbiotec.2011.08.030

Wang, X., Liu, Z. L., Weber, S. A., and Zhang, X. (2016). Two new native β-glucosidases from *Clavispora* NRRL Y-50464 confer its dual function as cellobiose fermenting ethanologenic yeast. *PLoS One* 11:e0151293 doi: 10.1371/journal.pone.0151293

Wang, X., Wu, Y., and Zhou, Y. (2017). Transglycosylation, a new role for multifunctional cellulase in overcoming product inhibition during the cellulose hydrolysis. *Bioengineered* 8, 129–132. doi: 10.1080/21655979.2016.1215787

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46:gky427. doi: 10.1093/nar/gky427

Wierckx, N. J. P., Ballerstedt, H., De Bont, J. A. M., and Wery, J. (2005). Engineering of Solvent-Tolerant. *Appl. Environ. Microbiol.* 71, 8221–8227. doi: 10.1128/AEM.71.12.8221

Wojtusik, M., Villar, J. C., Zurita, M., Ladero, M., and Garcia-Ochoa, F. (2017). Study of the enzymatic activity inhibition on the saccharification of acid pretreated corn stover. *Biomass Bioenergy* 98, 1–7. doi: 10.1016/j.biombioe.2017.01.010

Wu, J., Geng, A., Xie, R., Wang, H., and Sun, J. (2018). Characterization of cold adapted and ethanol tolerant β-glucosidase from *Bacillus cellulosilyticus* and its application for directed hydrolysis of cellobiose to ethanol. *Int. J. Biol. Macromol.* 109, 872–879. doi: 10.1016/j.ijbiomac.2017.11.072

Ximenes, E., Kim, Y., Mosier, N., Dien, B., and Ladisch, M. (2010). Enzyme and microbial technology inhibition of cellulases by phenols. *Biotechnol. Bioeng.* 46, 170–176. doi: 10.1016/j.enzmictec.2009.11.001

Ximenes, E., Kim, Y., Mosier, N., Dien, B., and Ladisch, M. (2011). Deactivation of cellulases by phenols. *Enzyme Microb. Technol.* 48, 54–60. doi: 10.1016/j.enzmictec.2010.09.006

Xin, D., Sun, Z., Viikari, L., and Zhang, J. (2015). Role of hemicellulases in production of fermentable sugars from corn stover. *Ind. Crops Prod.* 74, 209–217. doi: 10.1016/j.indcrop.2015.05.017

Xue, D., Zeng, X., Gong, C., Lin, D., and Yao, S. (2018). A cold adapt and ethanol tolerant endoglucanase from a marine *Bacillus subtilis*. *Chin. J. Chem. Eng.* (in press). doi: 10.1016/j.cjche.2018.02.007

Yamamoto, K., and Tamaru, Y. (2016). Synergistic properties of cellulases from *Clostridium cellulovorans* in the presence of cellobiose. *AMB Express* 6, 1–6. doi: 10.1186/s13568-015-0169-5

Yang, C., Xia, Y., Qu, H., Li, A.-D., Liu, R., Wang, Y., et al. (2016). Discovery of new cellulases from the metagenome by a metagenomics-guided strategy. *Biotechnol. Biofuels* 9, 1–12. doi: 10.1186/s13068-016-0557-3

Yang, M., Zhang, K., Di Zhang, P. Y., Zhou, X., Ma, X. Q., and Li, F. L. (2016). Synergistic cellulose hydrolysis dominated by a multi-modular processive endoglucanase from *Clostridium cellulosi*. *Front. Microbiol.* 7:1–8. doi: 10.3389/fmicb.2016.00932

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2014). The I-TASSER suite: protein structure and function prediction. *Nat. Methods* 12, 7–8. doi: 10.1038/nmeth.3213

Yang, Y., Zhang, X., Yin, Q., Fang, W., Fang, Z., Wang, X., et al. (2015). A mechanism of glucose tolerance and stimulation of GH1 β-glucosidases. *Sci. Rep.* 5, 1–12. doi: 10.1038/srep17296

Yapi Assoi Yapi, D., Gnakri, D., Lamine Niamke, S., and Patrice Kouame, L. (2009). Purification and biochemical characterization of a specific β-Glucosidase from the digestive fluid of larvae of the palm weevil, *Rhynchophorus palmarum*. *J. Insect Sci.* 9:4. doi: 10.1673/031.009.0401

Yoshida, E., Hidaka, M., Fushinobu, S., Koyanagi, T., Minami, H., Tamaki, H., et al. (2010). Role of a PA14 domain in determining substrate specificity of a glycoside

hydrolase family 3 β-glucosidase from *Kluyveromyces marxianus*. *Biochem. J.* 431, 39–49. doi: 10.1042/BJ20100351

Yu, S., Plan, M. R., Winter, G., and Krömer, J. O. (2016). Metabolic engineering of *Pseudomonas putida* KT2440 for the production of *para*-hydroxy benzoic acid. *Front. Bioeng. Biotechnol.* 4:90. doi: 10.3389/fbioe.2016.00090

Zanoelo, F. F., Polizeli Mde, L., Terenzi, H. F., and Jorge, J. A. (2004). β-Glucosidase activity from the thermophilic fungus *Scytalidium thermophilum* is stimulated by glucose and xylose. *FEMS Microbiol. Lett.* 240, 137–143. doi: 10.1016/j.femsle.2004.09.021

Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9:40. doi: 10.1186/1471-2105-9-40

Zhang, Z., Wang, M., Gao, R., Yu, X., and Chen, G. (2017). Synergistic effect of thermostable β-glucosidase TN0602 and cellulase on cellulose hydrolysis. *3 Biotech* 7:54. doi: 10.1007/s13205-017-0672-2

Zhao, L., Pang, Q., Xie, J., Pei, J., Wang, F., and Fan, S. (2013). Enzymatic properties of *Thermoanaerobacterium thermosaccharolyticum* β-glucosidase fused to Clostridium cellulovorans cellulose binding domain and its application in hydrolysis of microcrystalline cellulose. *BMC Biotechnol.* 13:101. doi: 10.1186/1472-6750-13-101

Zhao, W., Peng, R., Xiong, A., Fu, X., Tian, Y., and Yao, Q. (2012). Expression and characterization of a cold-active and xylose-stimulated β-glucosidase from Marinomonas MWYL1 in *Escherichia coli*. *Mol. Biol. Rep.* 39, 2937–2943. doi: 10.1007/s11033-011-1055-0

Zhou, Y., Wang, X., Wei, W., Xu, J., Wang, W., Xie, Z., et al. (2016). A novel efficient β-glucanase from a paddy soil microbial metagenome with versatile activities. *Biotechnol. Biofuels* 9:36. doi: 10.1186/s13068-016-0449-6

Conflict of Interest Statement: 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.