




Paenibacillus sp. A59 GH10 and GH11 Extracellular Endoxylanases: Application in Biomass Bioconversion

Silvina Ghio¹ · Ornella Ontañón^{2,3} · Florencia E. Piccinni^{2,3} · Rubén Marrero Díaz de Villegas² · Paola Talia^{2,3} · Daniel H. Grasso^{1,4} · Eleonora Campos^{2,3} 

© Springer Science+Business Media, LLC, part of Springer Nature 2017

Abstract

The cost-efficient degradation of xylan to fermentable sugars is of particular interest in second generation bioethanol production, feed, food, and pulp and paper industries. Multiple potentially secreted enzymes involved in polysaccharide deconstruction are encoded in the genome of *Paenibacillus* sp. A59, a xylanolytic soil bacterium, such as three endoxylanases, seven GH43 β -xylosidases, and two GH30 glucuronoxylanases. In secretome analysis of xylan cultures, ten glycoside hydrolases were identified, including the three predicted endoxylanases, confirming their active role. The two uni-modular xylanases, a 32-KDa GH10 and a 20-KDa GH11, were recombinantly expressed and their activity on xylan was confirmed (106 and 85 IU/mg, respectively), with differences in their activity pattern. Both endoxylanases released mainly xylobiose (X2) and xylotriose (X3) from xylan and pre-treated biomasses (wheat straw, barley straw, and sweet corn cob), although only rGH10XynA released xylose (X1). rGH10XynA presented optimal conditions at pH 6, with thermal stability at 45–50 °C, while rGH11XynB showed activity in a wider range of pH, from 5 to 9, and was thermostable only at 45 °C. Moreover, GH11XynB presented sigmoidal kinetics on xylan, indicating possible cooperative binding, which was further supported by the structural model. This study provides a detailed analysis of the complete set of carbohydrate-active enzymes encoded in *Paenibacillus* sp. A59 genome and those effectively implicated in hemicellulose hydrolysis, contributing to understanding the mechanisms necessary for the bioconversion of this polysaccharide. Moreover, the two main free secreted xylanases, rGH10XynA and rGH11XynB, were fully characterized, supporting their potential application in industrial bioprocesses on lignocellulosic biomass.

Keywords *Paenibacillus* · Endoxylanases · GH10 · GH11 · Pre-treated biomass

Introduction

Increasing the need for renewable fuels highlights the requirement of suitable technology development from sustainable

energy sources. In the last decades, scientific and engineering efforts have been conducted to improve the bioconversion of polysaccharides from wood materials, forest derived lignocellulose, agricultural residues, or energy crops into fermentable sugars for biofuels and added value by-products [1]. Also, increasing knowledge in biomass structural components and its efficient hydrolysis resulted in the upgrading of other processes using lignocellulose as raw materials, such as paper, textile, food, and animal feed industries [2, 3].

Lignocellulose is the complex structural material of all plant cell walls, composed mainly from cellulose (40–50%), hemicellulose (25–30%), and lignin (15–20%) [4]. The recalcitrance of these structural components is based on a complex polymeric network that provides plants with strength and resistance. For this reason, enzyme-mediated hydrolysis of lignocellulose, to release soluble and fermentable sugars, is the central step for lignocellulosic based biofuels [5]. Most studies have focused on hydrolysis of cellulose (the main homogeneous non-branched polysaccharide) into its repeating unit—cellobiose—or glucose (monomer) by cellulases [6, 7]. On the other side,

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s12155-017-9887-7>) contains supplementary material, which is available to authorized users.

✉ Eleonora Campos
campos.eleonora@inta.gob.ar; eleonoracampos@yahoo.com

¹ Instituto de Suelos, CIRN, Instituto Nacional de Tecnología Agropecuaria (INTA), Hurlingham, Buenos Aires, Argentina

² Instituto de Biotecnología, CICVyA, Instituto Nacional de Tecnología Agropecuaria (INTA), Los Reseros y N. Repetto, Hurlingham B1686, Buenos Aires, Argentina

³ Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Ciudad Autónoma de Buenos Aires, Argentina

⁴ Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, Bernal, Buenos Aires, Argentina

hemicelluloses are branched heteropolymers, with variable composition among different plants; while xylans are more abundant in grasses and angiosperm hardwoods, mannans are major components of gymnosperm softwood species [8]. Xylan, the most abundant hemicellulose in nature, is formed by a backbone of β -1,4-xylopyranosyl units and a diversity of substituted groups (arabinose, acetyl, glucuronic acids, feruloyl acids). Xylanases (EC 3.2.1.8) and β -xylosidases (EC 3.2.1.37) are the main hemicellulases involved in the depolymerization of the xylan backbone, while branch-point degrading enzymes include α -L-arabinofuranosidases (EC 3.2.1.55), acetylxylan esterases (EC 3.1.1.72), and α -D-glucuronidases (EC 3.2.1.131) [9]. Xylanases are usually classified into glycoside hydrolase (GH) families GH 5, 7, 8, 10, 11, 26, 30, and 43, according to carbohydrate-active enzymes (CAZy) database [10]. Family GH10 includes a few enzymes with endo- β -1,3-xylanase activity (EC 3.2.1.32), although the majority are endo- β -1,4-xylanases (EC 3.2.1.8). Some of the latter display limited activity on aryl cellobiosides, but not on cellulose. They exhibit greater catalytic versatility and lower substrate specificity than enzymes from family GH11, and they are generally able to act on the xylan backbone, close to substituted regions. On the other hand, the glycoside hydrolases of family GH11 are only endo- β -1,4-xylanases (EC 3.2.1.8), monospecific, strict to xylan, and act on unsubstituted regions of the substrate. While family GH10 generally includes xylanases of high molecular weight/low isoelectric point, family GH11 includes mainly low molecular weight/ high isoelectric point enzymes [11, 12].

In order to hydrolyze a complex substrate like heteroxylan, microorganisms secrete the majority of xylanases into the extracellular environment since the large size of this polysaccharide prevents its entrance into the cell. Besides, some intracellular xylanases that may act on small xylooligosaccharides have also been described [13]. Because of their optimum activities at neutral or alkaline pHs or over a wide range of temperatures, bacterial xylanases can act as a good complement of fungal xylanases in the industrial field [14]. Several bacterial genera such as *Bacillus*, *Cellulomonas*, *Micrococcus*, *Staphylococcus*, *Paenibacillus*, *Arthrobacter*, *Microbacterium*, *Pseudoxanthomonas*, and *Rhodothermus* have been reported to produce xylanases [15]. In particular, members of the *Paenibacillus* genus—facultative anaerobe bacteria—have been described as producers of several enzymes for industrial applications, including lignocellulose-degrading enzymes [16, 17].

Microbial bioprospection from diverse natural sources and proteogenomic analysis of novel cellulolytic and xylanolytic strains can lead to the identification of glycoside hydrolases that are secreted under different culture conditions, such as growth on diverse substrates, and contribute to the understanding of the complex interactions of polysaccharide degrading systems as

well as the discovery of novel enzymes [18–20]. Thorough information about key enzymes involved in a particular bacterial lignocellulolytic system is important to understand the mechanisms of bioconversion in nature. In this context, we aimed to describe the predicted carbohydrate-active enzymes of the recently isolated hemicellulolytic *Paenibacillus* sp. A59 strain [21] based on genome sequence analysis, along with the study of the main secreted glycoside hydrolases by bacterial growth in xylan. Also, we have recombinantly expressed two of the major xylanases identified, predicted to belong to GH10 and GH11 families, and we have characterized their activity on xylan and lignocellulosic biomass, for further application in industrial bioconversion processes.

Materials and Methods

Carbohydrate-Active Enzyme Prediction

CAZymes analysis was performed using dbCAN annotation web server (<http://csbl.bmb.uga.edu/dbCAN/annotate.php>) [22] based on the protein family classification of CAZy database (<http://www.cazy.org/>) [23]. Protein-coding sequences of *Paenibacillus* sp. A59, previously annotated in GenBank [24], were used for the search. Signal peptides of each protein were predicted by Signal P v4.1 [25]. Protein Blast (BLASTP) (<http://blast.ncbi.nlm.nih.gov>) and Expasy-Prosites (<http://prosite.expasy.org/>) were used to identify conserved domains.

Secretome Protein Extraction and Mass Spectrometry Analysis

Paenibacillus sp. A59 culture was grown for 96 h in 20 ml of minimal medium [26] with 0.1% (w/v) yeast extract (Bacto) and 0.5% (w/v) beechwood xylan (Sigma) at 30 °C. It was filtered twice through 1.2 μ m glass-fiber discs (Schleicher and Schuell) and then with 0.2 μ m, followed by centrifugation (10,000 \times g, 20 min). Protease inhibitor (Thermo Fisher) was added to the supernatant to prevent protein degradation. Total protein content was quantified using Micro-BCA (bicinchoninic acid) Colorimetric Assay Kit (Thermo Fisher), with bovine serum albumin (BSA) as standard. Twenty micrograms of protein sample were loaded into a 12% (w/v) SDS-polyacrylamide gel for electrophoresis. After that, gel was stained with colloidal Coomassie G-250, and the gel lane was divided into four fractions of different molecular sizes. Gel fractions were excised for mass spectrometry analysis. Protein digestion and mass spectrometry analysis were performed at the Proteomics Core Facility CEQUIBIEM, at the Faculty of Exact Sciences, University of Buenos Aires/CONICET (National Research Council) as follows: gel fragments were reduced with DTT (10 mM) for 45 min (min) at

56 °C, alkylated with iodoacetamide (55 mM) for 45 min in the dark, and digested with trypsin (200 ng) overnight at 37 °C. Peptides were eluted from gel with 50% acetonitrile (ACN)-0.5% trifluoroacetic acid, further concentrated by speed-vacuum dry and desalted with Zip-Tip C18 (Merck Millipore). Samples were eluted in 10 µl of water/ACN/formic acid 40:60:0.1%. The digests were analyzed by nano LC-MS/MS in a Thermo Scientific Q-Exactive Mass Spectrometer coupled to a nano HPLC EASY-nLC 1000 (Thermo Scientific). Peptides were loaded onto a C18 Easy-Spray Accucore column and eluted for 120 min at a flow of 33 nl/min, using a gradient of water/ACN. Full-scan mass spectra were acquired in the Orbitrap analyzer. The scanned mass range was 400–1800 m/z , at a resolution of 70,000 at 400 m/z . Q-Exactive raw data was processed using Proteome Discoverer software (version 1.4 Thermo Scientific) and searched against *Paenibacillus* spp. protein sequences database (including *Paenibacillus* sp. A59) with trypsin specificity and a maximum of one missed cleavage per peptide. Carbamido methylation of cysteine residues was set as a fixed modification, and oxidation of methionine was set as variable modifications. Only the proteins identified by peptides with high confidence were used for analysis. Results correspond to one biological replicate.

Protein Sequence Analysis and Molecular Modeling

Deduced amino acid sequences of GH10XynA (WP_053782506.1) and GH11XynB (WP_053781844.1) were analyzed by BLASTP against reference protein sequences (refseq) available in NCBI database (<http://blast.ncbi.nlm.nih.gov/>). The 15 most related full-length amino acid sequences (identity and coverage over 90%) were retrieved, and multiple sequence alignments were performed using ClustalW program from MEGA software v6.0 [27]. Phylogenetic analysis was inferred by neighbor-joining method and the evolutionary distances were computed using the p-distance method. The robustness of the trees in the analysis was assessed by 1000 bootstrap replications. *Clostridium termitidis* GH10 (WP_004625225.1) and GH11 (WP_004629994.1) were included as outgroup.

The molecular structures of the mature GH10XynA and GH11XynB proteins (without signal peptide) were generated using the modeling software Iterative Threading ASSEMBLY Refinement (I-TASSER) (<http://zhanglab.ccmb.med.umich.edu/I-TASSER/>) [28]. To construct the model, structure of proteins with similar folds was retrieved from a PDB library (<http://www.rcsb.org>) by LOMETS (combination of ten threading programs: MUSTER, FFAS3D, SPARKSX, HHSEARCH2 5, HHSEARCH I, NeffPPAS, HHSEARCH, pGenTHREADER, cdPPAS, PROSPECT2). The fragments from templates with the highest significance (measured by

the Z-score) in the threading alignments were re-assembled into a full-length model. The following PDB templates were selected to construct the model, for GH10: 3nyd-A, 1nq6-A, 4pmx-A, 1r85-A; and for GH11: 1axk-B, 1axk-A, 3lgr-A, 3wp3-A, 3mfa-A, 3wp3-A, 2z79-A. The final I-TASSER model resulted from a second simulation round which removed the steric clashes and refined the global topology. The confidence of each model was quantitatively measured by C-score, TM-score, and root-mean-square deviation (RMSD). C-score was calculated based on the significance of threading template alignments and the convergence parameters of the structure assembly simulations. It is typically in the range of $[-5, 2]$, where a higher value signifies a model with a higher confidence and vice versa (a C-score > -1.5 indicates a model of correct topology). TM-score is a metric for measuring the structural similarity of two protein models. It has the value in $(0, 1)$, where 1 indicates a perfect match between two structures (a TM-score > 0.5 indicates a model of correct topology and a TM-score < 0.17 means a random similarity). RMSD is the root-mean-square deviation of atomic positions, the average distance between the atoms of superimposed proteins.

Functional analysis was conducted by COACH (combined results from COFACTOR, TMSITE, and SSITE programs), within I-TASSER, which included prediction of the protein-ligand-binding site. The confidence of the predictions was determined by a C-score, with values $(0, 1)$, where a higher score indicated a more reliable prediction, and the cluster size, which is the total number of templates used. The full I-TASSER output results, including secondary structure prediction, solvent accessibility, model construction and prediction of ligand-binding sites, active sites, and GO ontology for GH10XynA and GH11XynB can be accessed in Online Resources 1 and 2, respectively. Model manipulation, alignment, and imaging were performed using Chimera software v1.11 [29].

Strains, Plasmids, and Media

Total genomic DNA from *Paenibacillus* sp. A59 was obtained using Wizard genomic DNA purification kit (Promega) and used as template for gene amplification. pGemT Easy vector (Promega) was used for cloning and pET28a (Novagen) and pJExpress (DNA 2.0) vectors for protein expression. Chemically competent *Escherichia coli* cells were used as hosts for gene cloning (DH5 α strain, Thermo Fisher Scientific) and protein expression (Rossetta DE3, Novagen and XL1 Blue, Stratagene). *E. coli* cells containing recombinant plasmids were grown in Luria Bertani (LB) medium supplemented with either 100 µg/ml ampicillin or 50 µg/ml kanamycin, depending on the plasmid used.

Gene Expression and Protein Purification

Open-reading frames of *Paenibacillus* sp. A59 extracellular endoxylanases GH10XynA (WP_053782506.1) and GH11XynB (WP_053781844.1), without signal peptides, were amplified using specific oligonucleotides: GH11F1Pae: 5'AAACATATGCATCATCATCCATCACCACGCAACAGATTATTGGCAAAT3', GH11R2Pae 5'ATCTAGATTAC CACACCGTTACGTTGGA3', GH10-2DF 5'AAAGGAT CCAAAGGAAGCAAGTTTCTGGGTAAT3', and GH10-2DR 5'AAACTCGAGTTAGGGATTGTTGGCAAGATAATT3', for N-terminal fusion to 6xHis (restriction sites underlined). GH10XynA amplification product was first cloned in pGemT Easy vector using *E. coli* DH5- α competent cells. Then, plasmids from selected colonies were cloned in pET28a vector (*Bam*HI/*Xho*I) and transformed into competent *E. coli* Rossetta cells. GH11XynB amplification product was directly cloned in pJExpress (*Nde*I/*Xba*I) using *E. coli* XL1-blue cells for protein expression. GH10XynA protein expression was induced with 1 mM IPTG for 4 h at 37 °C and GH11XynB with 0.5 mM IPTG for 16 h at 28 °C. After cell lysis and sonication (six pulses of 10 s, 28% amplitude), recombinant proteins were purified in the soluble fraction by IMAC with Ni-NTA agarose resin (Qiagen), using 50 mM NaH₂PO₄, 300 mM NaCl, 250 mM imidazole pH 8 as elution buffer for GH10XynA and 20 mM Tris-HCl, 20 mM KCl, 1 mM EDTA, and 0.1% (v/v) Igepal (Sigma) pH 7 for GH11XynB. Usually, 20 mg of rGH10XynA and 10 mg of rGH11XynB recombinant proteins were obtained from 50 ml *E. coli* induced cultures.

Enzyme Activity Assays

Enzyme activity assays were performed using each enzyme (diluted in appropriate buffer) in a final concentration of 5 μ g/ml. Xylanase activity was assayed on 1% (w/v) beechwood xylan (Sigma), in a final reaction volume of 0.2 ml at 400 rpm for 15 min in a Thermomixer (Eppendorf). Reducing sugars released from xylan hydrolysis were measured using the 3,5-dinitrosalicylic acid (DNS) assay [30] with xylose as standard. The optimal pH (from 3 to 10) and temperature (from 30 to 70 °C) of xylanase activity were assayed as previously described [21]. Thermal stability was performed by preincubating the enzymes at 45, 50, and 60 °C from 0 to 20 min. Also, kinetic parameters were determined under optimal assay conditions using 0–20 mg/ml of beechwood xylan as substrate, by fitting the models to data with the software GraphPad Prism v 6.0. Other substrates were also used for testing enzyme specificity: 1% (w/v) wheat arabinoxylan (medium viscosity, 38% arabinose substitution, Megazyme), 2% (w/v) carboxymethylcellulose (CMC low viscosity, Sigma), 1% (w/v) Avicel (Fluka), 1% (w/v) bacterial microcrystalline cellulose (BMCC), and 1% (w/v) phosphoric acid swollen

cellulose (PASC), by DNS quantification of reducing sugars. *p*NP-arabinofuranoside, *p*NP-cellobioside, *p*NP-xylopyranoside, and *p*NP-glucopyranoside (Sigma) (5 mM) were used for testing arabinofuranosidase, cellobiosidase, β -xylosidase, and β -glucosidase activities, respectively, using a *p*-nitrophenol curve as standard. Enzyme activities were expressed as international units per milligram of protein. One international unit was defined as the amount of enzyme that released 1 μ mol of product per minute under the assay conditions. All experiments were conducted with two biological replicates in independent assays and three replicates for each condition. Equivalent results were obtained in both biological replicates. The average shown corresponds to the triplicates of one biological replicate.

Hydrolysis of Xylan and Pre-treated Biomasses

Hydrolysis reactions on beechwood xylan (1%) (w/v), wheat arabinoxylan (1%) (w/v), non-branched xylooligosaccharides (xylotetraose, X4, and xylotriose, X3) (Megazyme) (1.5 mg/ml), and xylobiose (X2) (Sigma) (1.5 mg/ml) were carried out using each recombinant enzyme (5 IU/mg of substrate) at pH 6 and 45 °C, for 20 min for xylan and 2 h for X2, X3, X4, and wheat arabinoxylan. Biomass hydrolysis was performed using wheat straw, sweet corn cob, and barley straw (2%) pre-treated by extrusion and each recombinant enzyme (2.5 IU/mg of biomass) in a final volume of 1 ml, pH 6 at 45 °C for 24 h. Also, combination of enzymes for wheat straw hydrolysis was prepared in a 1:1 proportion (1.25 IU/mg substrate). The commercial hemicellulase cocktail NS22002 (Novozymes) 1% (v/v) was also tested on pre-treated wheat straw, at the same assay conditions as the recombinant xylanases. Xylan, wheat arabinoxylan, X2, X3, and X4 hydrolysis reactions were carried out in a Thermomixer (Eppendorf) at 400 rpm, and biomass hydrolysis reactions were conducted in an orbital shaker at 250 rpm to allow the proper mixing of the components. After that, reactions were stopped at 100 °C for 10 min and then centrifuged and filtered by 0.2 μ m nylon filters. Xylan and wheat straw hydrolysis patterns were qualitatively analyzed by thin layer chromatography (TLC) in silica gel plates, using butanol/acetic acid/water (2:1:1) as solvents and revealed by water/ethanol/sulfuric acid (20:70:3) with 1% (v/v) orcinol solution, over flame. Soluble sugars from all the reactions were quantified by HPLC (Agilent 1100), using a Rezex RPM-Monosaccharide column (Phenomenex) (80 °C, flow 0.6 ml/min), with a RI detector at 35 °C. X4, X3, X2, and xylose were used as chromatography standards. Two biological replicates were conducted in independent assays, with three replicates for each condition. Results from only one of the biological replicates is shown, although the results obtained were equivalent.

Results

CAZymes Prediction in *Paenibacillus* sp. A59 Genome

Paenibacillus sp. A59, a novel hemicellulolytic strain, was isolated from decaying forest soil [21]. In order to study the xylanolytic and cellulolytic system of this isolate, its genome was sequenced by Illumina MiSeq [24] and coding sequences for proteins potentially involved in biomass deconstruction were identified. NCBI Prokaryotic Genome Annotation pipeline revealed 5958 protein-coding sequences that were then analyzed using dbCAN carbohydrate-active enzymes (CAZy) annotation algorithm. A total of 504 CAZy domains were identified, including 208 glycoside hydrolases (GHs) domains, 73 carbohydrate-binding domains (CBMs), 13 polysaccharide lyases (PLs), 57 glycoside transferases (GTs), 67 carbohydrate esterases (CEs), 78 S-layer homology domains (SLHs), and 8 domains with auxiliary activities (AAs). By detailed manual analysis, we assigned these domains to 396 proteins. Although most frequently one domain was assigned to a single protein, in some cases, a combination of a catalytic domain with one or more CBMs or SLH domains was predicted in the same protein. No multiple-activity GHs—a combination of two or more catalytic domains—were predicted. In order to define the potential extracellular glycoside hydrolases involved in cellulose or hemicelluloses (xylans, heteroxylans, heteromannans, and xyloglucans) hydrolysis, the output results obtained from dbCAN were subjected to BLASTP homology search and signal peptide prediction analysis. We identified 70 GHs in *Paenibacillus* sp. A59 genome for biomass hydrolysis, belonging to GH families associated with hydrolysis of cellulose (GH1, GH2, GH3, GH5, GH6, GH8, GH9, GH12, GH16, GH30, GH48) and xylan-based hemicellulose (GH10, GH11, GH30, GH31, GH43, GH51, GH52, GH74) (Table 1). Among them, enzymes with potential endoglucanase, β -glucosidase, exoglucanase, endoxylanase, β -xylosidase, β -glucuronidase, α -xylosidase, xyloglucanase, arabinofuranosidase, and glucuronoxylase activities were recognized. Interestingly, we found only one coding sequence for proteins classified as GH1, 6, 8, 9, 11, 12, 48, and 52, predominantly associated with endoglucanase, exoglucanase, or endo-xylanase activities. Other GH families, mainly related to β -xylosidase, β -glucosidase, and arabinofuranosidase activities (GH2, 3, 5, 16, 30, 43, and 51), were more abundant. In all cases, the coding sequence for each protein was found in single copy along the genome. Furthermore, other coding sequences for proteins with relevant activities on non-lignocellulosic carbohydrates such as chitin and starch were identified in *Paenibacillus* sp. A59 genome. Six protein-coding sequences with potential chitinase activity (GH18) and 14 with amylase (GH13) activity were predicted (data not shown). By signal peptide prediction, 35 of all the GHs mentioned above were inferred to be extracellular and GHs from families 1, 2, 31,

51, and 52 were predicted to be exclusively intracellular (Table 1). These proteins have predicted functions related to short oligosaccharides and disaccharides hydrolysis, which typically occur inside the bacterial cell at the final hydrolysis steps and therefore correlate with their potential intracellular localization. However, we have identified two GH3 β -glucosidases (one of them with three SLH domains) and seven GH43 β -xylosidases/ α -arabinofuranosidases (six of which have associated CBMs) that present signal peptide sequences, which may indicate that these proteins could be either extracellular or surface exposed.

About half of the CAZy catalytic modules identified in *Paenibacillus* sp. A59 were associated with one or more CBM, although a CBM9 and a CBM35 (generally found associated to xylanases) were found in proteins with not known catalytic domain, suggesting that these proteins may have a non-hydrolytic role or activities not yet described. All predicted CBMs had been previously described to bind cellulose, xylan, chitin, or mannan polymers [31]. Remarkably, proteins with similar function from the same GH family showed different CBMs associations. For example, in family GH43, diverse types of CBMs (6, 13, 35, 40, 42, 66) were associated with the catalytic domain. In other cases, from the enzymes identified corresponding to the same GH family (GH5, GH10, GH16, GH30), only one of the proteins of each family was predicted to have associated CBMs. S-layer homology domains (SLH) were observed in some proteins, generally by way of three subsequent subunits. Multiple carbohydrate esterases (CEs) protein sequences, including acetyl xylan and feruloyl esterases, as well as a lytic polysaccharide monooxygenase (LPMO) enzymes (auxiliary activity AA10) were also identified in *Paenibacillus* sp. A59 (data not shown). These activities had been described as relevant for hemicellulose and chitin debranching and cellulose deconstruction. Most species of aerobic bacteria growing on insoluble substrates have genes encoding one or two AA10 proteins, which catalyze oxidative cleavage of polysaccharides, but the majority of these proteins have not been characterized yet [32]. Noteworthy, proteins, annotated as hypothetical or with unknown function by NCBI prokaryotic genome annotation pipeline, were predicted to have putative CAZy domains using dbCAN, which could give a hint on potential activity of these still uncharacterized proteins. This also highlights the importance of using more than one annotation pipeline for searching conserved features in cellulose and hemicellulose hydrolytic enzymes.

Paenibacillus sp. A59 Extracellular Proteins Induced by Xylan

Maximal xylanase activity was achieved in culture supernatant when *Paenibacillus* sp. A59 was grown on xylan, compared to lignocellulosic biomass or CMC [21]. By mass

Table 1 Carbohydrate-active enzymes (CAZymes) involved in cellulose and xylan-based hemicellulose bioconversion predicted in *Paenibacillus* sp. A59 genome

Predicted activity ^a	EC number ^b	Accession number ^c	CAZy domains ^d	Signalpeptide ^e
Endo- β -1,4-xylanase	3.2.1.8	WP_053779305.1	GH10 + CBM9 + (CBM22) ₂ + (SLH) ₃	Y
		WP_053780820.1	GH10	N
		WP_053782506.1	GH10	Y
		WP_053781844.1	GH11	Y
β -Xylosidase/ α -L-arabinofuranosidase	3.2.1.37/3.2.1.55	WP_053780819.1	GH43	N
		WP_053780821.1	GH43	N
		WP_053781471.1	GH43 + CBM42	Y
		WP_053781573.1	GH43	N
		WP_053782335.1	GH43 + CBM35	Y
		WP_053782339.1	GH43	N
		WP_053782398.1	GH43 + CBM40	Y
		WP_053782497.1	GH43 + CBM32 + CBM42	Y
		WP_053782893.1	GH43 + CBM13	Y
		WP_053783386.1	GH43	N
		WP_053783510.1	GH43	N
		WP_053783509.1	GH43	Y
		WP_053781832.1	GH43 + (CBM6) ₂ + CBM66	Y
		WP_053780607.1	GH52	N
		WP_053782397.1	GH51	N
		WP_053782407.1	GH51	N
		WP_053783366.1	GH51	N
WP_053779045.1	GH51	N		
Xyloglucanase	3.2.1.151	WP_053779044.1	GH74	Y
		WP_053781908.1	GH74	Y
Glucuronoxylanase	3.2.1.136	WP_053781470.1	GH30 + CBM13	Y
		WP_053781633.1	GH30	N
		WP_053782376.1	GH30	N
		WP_053782507.1	GH30	N
		WP_053779434.1	GH30	Y
α -Xylosidase	3.2.1.177	WP_053783720.1	GH31	N
		WP_053781003.1	GH31	N
β -Glucosidase	3.2.1.21	WP_053780132.1	GH1	N
		WP_053780488.1	GH3	N
		WP_053781619.1	GH3	N
		WP_053782189.1	GH3	Y
		WP_053782759.1	GH3 + CBM6	N
		WP_053782986.1	GH3	N
		WP_053782991.1	GH3	N
		WP_053783318.1	GH3	N
		WP_053783722.1	GH3	N
		WP_053784388.1	GH3	N
		WP_053779370.1	GH3 + (SLH) ₃	Y
β -Glucuronidase	3.2.1.31	WP_053780053.1	GH3	N
		WP_053781895.1	GH2	N
		WP_053782272.1	GH2	N
		WP_053782273.1	GH2	N
		WP_053782417.1	GH2	N
WP_053782906.1	GH2	N		

Table 1 (continued)

Predicted activity ^a	EC number ^b	Accession number ^c	CAZy domains ^d	Signalpeptide ^e
		WP_053783756.1	GH2	N
		WP_053784098.1	GH2 + CBM6	N
		WP_053779343.1	GH2	N
		WP_053779956.1	GH2	N
Endo- β -1,4-glucanase	3.2.1.4	WP_053783226.1	GH5	N
		WP_053782127.1	GH5	Y
		WP_053782136.1	GH5 + (CBM46)2	Y
		WP_053782765.1	GH5	Y
		WP_053779698.1	GH8	Y
		WP_053783505.1	GH9 + (CBM3)2 + CBM37	Y
		WP_053779894.1	GH12	Y
Cellobiohydrolase (no-reducing end)	3.2.1.91	WP_053784024.1	GH6 + CBM3	Y
Cellobiohydrolase (reducing-end)	3.2.1.176	WP_053783364.1	GH48 + CBM3	Y
Endo- β -1,3-glucanase	3.2.1.39	WP_053781078.1	GH16	Y
		WP_053781865.1	GH16	Y
		WP_053782420.1	GH16 + CBM13	Y
Endo- β -1,6-glucanase	3.2.1.75	WP_053782296.1	GH30	Y
Cellulose binding		WP_053783378.1	CBM9	N
		WP_053784172.1	CBM9	N
		WP_053781829.1	CBM9 (SLH)3	Y
Xylan binding		WP_053781007.1	CBM35	Y
		WP_053783673.1	(CBM35)2 + (SLH)3	Y

^a According to CAZy database classification and sequence homology comparison using BLASTP (NCBI)

^b Enzyme commission number

^c *Paenibacillus* sp. A59 NCBI functional annotation

^d CAZy domains from dbCAN annotation

^e Presence (Y) or absence (N) of signal peptide (Signal P v4.1)

spectrometry of xylan-induced whole secretome, previously separated by one-dimensional polyacrylamide gel electrophoresis, a total of 23 extracellular proteins were identified (Table 2), from which ten were glycoside hydrolases (GHs), nine were ABC type transporter proteins, one amidase, one peptidase, and two uncharacterized proteins. Thirty-six intracellular proteins without predicted signal peptide were also detected, probably as a result of cell lysis during secretome extraction; thus, they were not considered for the analysis. Among GHs, proteins from diverse families were identified: three endo-xylanases (two GH10 and a GH11), two chitinases (GH18), one cellobiohydrolase (GH6), three endoglucanases (β -1,3 GH16, β -1,4 GH8, and β -1,6 GH30), and one amylase (GH13). Besides, many of them exhibited CBMs associated to the catalytic domains. In the multidomain GH10 endoxylanase (160 kDa), three xylan-binding domains (two CBMs22 and one CBM9) as well as three C-terminal SLH domains were associated to the catalytic domain. The two GH18 chitinases identified in the secretome presented one CBM12, chitin-binding domain, each. Also, a CBM20

starch-binding domain was associated to a GH13 amylase domain, although no signal peptide was predicted for this sequence. GH6 and GH16 exo and endoglucanases were linked to CBM3 and CBM13 cellulose-binding domains, respectively. Noteworthy, three high score peptides identified in the mass spectrometry analysis did not match with any predicted protein from *Paenibacillus* sp. A59. However, they matched to other *Paenibacillus* species, probably due to gaps in *Paenibacillus* sp. A59 genome sequence as it is still a draft version [24]. When the sequences were compared to *Paenibacillus* sp. A59 by NCBI BLASTP, the corresponding proteins were identified. These proteins corresponded to a GH9 endoglucanase from *Paenibacillus* sp. FSL H7-0357 (62% homology with WP_053783505.1) and two GH5 endoglucanases from *Paenibacillus* sp. FSL R7-0331 and *Paenibacillus* sp. A1 (63 and 95% homology with WP_053782136.1 and WP_053782127.1, respectively).

The three predicted extracellular endoxylanases encoded in *Paenibacillus* sp. A59 genome (two from family GH10 and one from family GH11) were found in xylan-induced

Table 2 Secreted proteins identified by mass spectrometry in xylan-culture supernatant of *Paenibacillus* sp. A59

Protein description	GenBank accession number	Score	Coverage	Unique peptides	Theoretical protein mass (kDa)/ isoelectric point	CAZy domains
Endo- β -1,4-xylanase	WP_053782506.1	103.9	25.0	5	32/8.92	GH10
Endo- β -1,4-xylanase	WP_053779305.1	2.8	0.5	1	160.1/4.93	GH10 + (CBM22) ₂ + CBM9 + (SLH) ₃
Endo- β -1,4-xylanase	WP_053781844.1	61.8	12.0	2	22/9.38	GH11
Chitinase	WP_053782577.1	92.0	25.6	13	73.6/6.35	GH18 + CBM12
Chitinase	WP_053782576.1	6.2	3.5	3	55.5/6.35	GH18 + CBM12
Cellobiohydrolase	WP_053784024.1	5.1	4.3	2	75.7/5.47	GH6 + CBM3
Endo- β -1,3-glucanase	WP_053782420.1	4.5	2.2	1	44.3/5.39	GH16 + CBM13
Endo- β -1,4-glucanase	WP_053779698.1	4.2	6.0	2	44.8/8.35	GH8
Endo- β -1,6-glucanase	WP_053782296.1	2.7	2.0	1	55.4/8.4	GH30
α -amylase	WP_053784175.1	34.9	21.0	9	76.1/6.27	GH13 + CBM20
Peptidase M4	WP_053783199.1	67.5	18.6	6	55.9/5.02	–
ABC transporter substrate-binding protein	WP_053780420.1	36.9	27.5	7	36.8/4.87	–
ABC transporter substrate-binding protein	WP_053781964.1	28.0	16.8	8	67.2/4.84	–
ABC transporter substrate-binding protein	WP_053780095.1	21.2	34.7	6	29.8/4.91	–
ABC transporter substrate-binding protein	WP_053781518.1	4.8	2.0	2	76.4/4.7	–
ABC transporter substrate-binding protein	WP_053783256.1	7.2	15.1	2	30.4/4.9	–
ABC transporter substrate-binding protein	WP_053783994.1	5.5	5.2	2	67.1/5.02	–
ABC transporter substrate-binding protein	WP_053779742.1	2.4	3.0	1	47.8/4.91	–
DNA-binding protein	WP_053780115.1	9.7	6.8	3	52.5/5.36	–
Peptide-binding protein	WP_053781233.1	24.1	15.9	9	62.8/5.06	–
<i>N</i> -acetylmuramoyl-L-alanine amidase	WP_053780743.1	4.7	2.0	1	49.9/7.61	–
Uncharacterized protein	WP_053782877.1	9.7	4.8	2	52.2/4.67	–
Uncharacterized protein	WP_053781861.1	1.9	1.0	1	54.4/5.36	–

secretome. This may indicate the requirement of the complete xylanase repertoire for efficient hydrolysis of substrate.

Sequence Analysis and Molecular Modeling of Main Secreted Non-modular Endoxylanases

The two uni-modular secreted endoxylanases (without CBMs or SLH domains) belonging to families GH10 and GH11 were selected for further studies, in order to evaluate their role in the xylanolytic system of *Paenibacillus* A59. We named them GH10XynA and GH11XynB, respectively. Both proteins were analyzed based on their amino acid sequences and conserved domains by NCBI and ExPasy-Prosite databases. GH10XynA showed high sequence identity (up to 96%, in more than 90% coverage) with several sequences annotated as endoxylanases from other isolates classified as *Paenibacillus* sp. Neighbor-joining phylogenetic analysis was also conducted, and GH10XynA showed a close

phylogenetic relationship with *Paenibacillus* sp. AD87 (WP_064641585.1) and *Paenibacillus* sp. Ov031 (WP_072733866.1) endoxylanases, as they were placed in the same cluster supported by a bootstrap value of 70% (Fig. 1a). GH11XynB presented high identity (97% with 100% coverage) with an endoxylanase from *Paenibacillus pabuli* (WP_076288544.1), confirmed by phylogenetic analysis as they were grouped in the same cluster along with the sequences *Paenibacillus* sp. AD87 (WP_064638409.1) and *Paenibacillus* sp. O199 (WP_063565023.1), with a bootstrap support value of 50% (Fig. 1b). All these protein sequences were inferred from whole genome shotgun studies, except for *Paenibacillus* sp. O199 GH11 endoxylanase that was also identified in proteome studies (when bacteria were grown on microcrystalline cellulose and wheat straw). However, to the best of our knowledge, no recombinant expression or activity characterization of any of these enzymes had been undertaken so far.

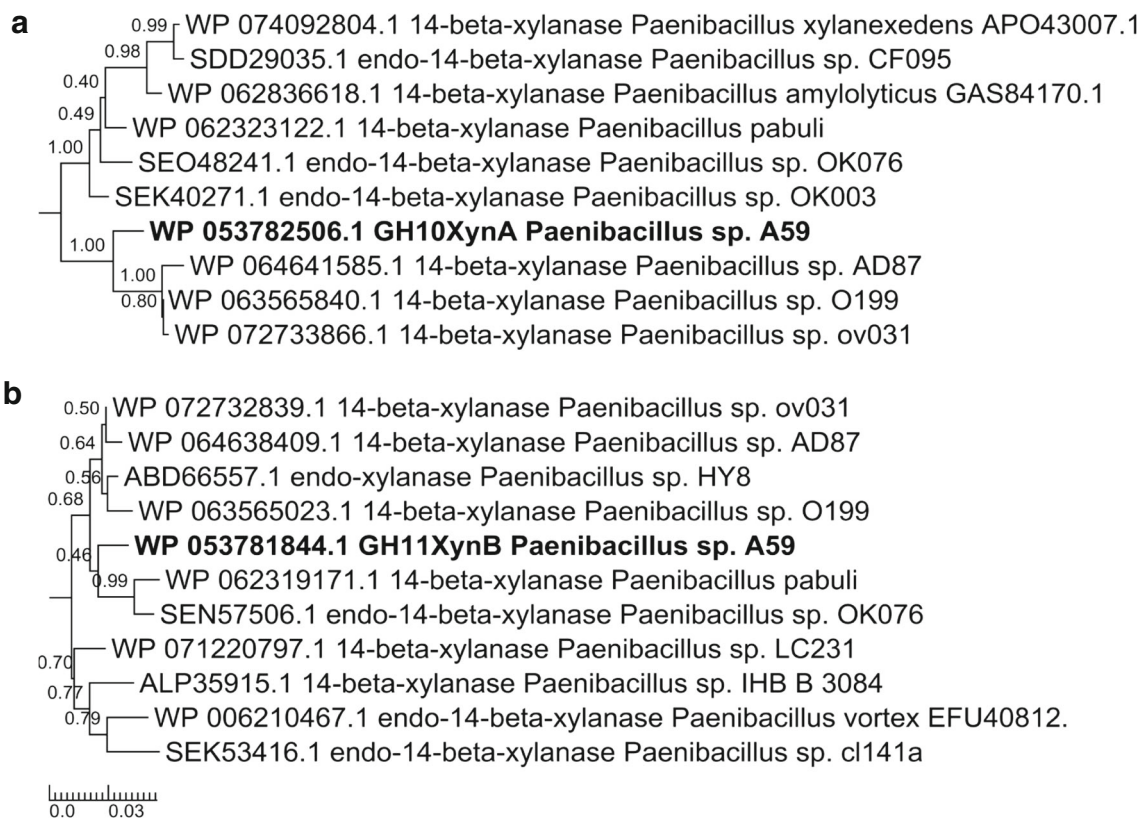


Fig. 1 Molecular phylogenetic analysis. The evolutionary history of *Paenibacillus* sp. A59 GH10XynA (full square) (a) and GH11XynB (full circle) (b) was inferred using the neighbor-joining method with p-distance analysis (MEGA6). The tree is drawn to scale, with branch lengths units measured in number of amino acid differences per site.

Additionally, we developed a molecular model for the GH10XynA and the GH11XynB mature proteins using the automated modeling server I-TASSER (Iterative Threading ASSEmblY Refinement) [28]. A model with very high confidence was obtained for GH10XynA (*C*-score 1.2, estimated TM-score 0.88 ± 0.07 , estimated RMSD 3.7 ± 2.5 Å) based on structural homology to previously crystallized xylanases. The final predicted model presented top structural homology (with a TM-score of 0.969) with a GH10 xylanase from *Thermotoga petrophila* (PDB: 3niy) (Fig. 2a) (for full analysis details please refer to Online Resources 1). This analysis showed that GH10XynA protein sequence fitted with the classical catalytic domain of GH10 enzymes from the clan A, a (β/α) 8-barrel fold, also called a TIM-barrel fold, which resembles the shape of a salad bowl (Fig. 2a). Although the general structure is highly conserved, the differences observed between the two aligned models could be related to the different characteristics of both proteins, as GH10 from *T. petrophila* is a hyperthermophilic enzyme [33]. Prediction of the protein-ligand-binding site was conducted by COACH analysis, obtaining a top-ranked prediction with a *C*-score of 0.61, based on 135 templates, being its main representative a GH10 xylanase crystallized as a complex with xylobiose (XynGR40, PDB:

Clostridium termitidis GH10 and GH11 were used as outgroup, for each tree. The percentage of replicate trees, in which the associated taxa clustered together in the bootstrap test (1000 replicates), is shown next to the branches (only bootstrap values over 50% are indicated)

5d4y). By alignment of GH10XynA structure model with the top ranked structural representatives, we identified the catalytic residues E111 and E218 (mature protein) with an atomic distance of 5 Å between them, which had been indicated to be a structural requirement for the catalytic retaining mechanism of other GH10 xylanases (Fig. 2a; Online Resources 3a) [34].

A high confidence molecular model of the mature GH11XynB protein was also obtained by I-TASSER, with a *C*-score of 1.64, an estimated TM-score of 0.94 ± 0.05 , and estimated RMSD of 2.0 ± 1.6 Å. The predicted model presented top structural homology, with a TM-score of 0.993, with the 1,4-beta-xylanase domain from a *Bacillus* engineered protein (PDB: 1axkA1) [35], with an excellent alignment (Fig. 2b) (for full analysis details please refer to Online Resources 2). GH11XynB structural model presented a putative catalytic domain with the typical β -jelly roll structure that resembles the shape of a partially closed hand and consists of two twisted antiparallel β -sheets (the “fingers”) and a single α -helix (Fig. 2b) [36, 37]. The predicted active site, deduced from COACH analysis followed by alignment with top structural hits, contained two glutamic acid residues (E77 and E170) situated on the

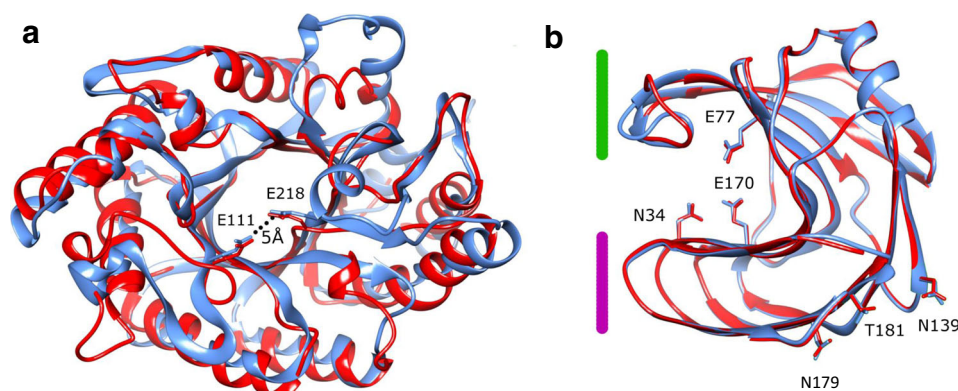


Fig. 2 Molecular modeling of GH10XynA and GH11XynB. Molecular models of GH10XynA (a) and GH11XynB (b) were generated by I-TASSER. Superimposed models of GH10XynA and GH11XynB (light blue) with their closest structural analog (red) are shown: PDB, 3niy (GH10 xylanase from *Thermotoga petrophila*), and PDB, 1axk (engineered xylanase catalytic module from *Bacillus* sp), respectively.

Catalytic glutamate residues are shown for each model. For GH11XynB, the thumb and fingers domains are indicated by green and violet lines respectively, as well as the residues of the predicted secondary binding site (N139, N179, T181) with a yellow line. pH relevant N34 residue is also shown

concave side of the palm, previously described in other GH11 endo-xylanases as nucleophile and proton donor, respectively [38]. The loop between β -strands B7 and B8 made the “thumb” containing the classical motif PSIXG (Fig. 2b; Online Resources 3b). The structural model of GH11XynB also displayed other relevant features, as the conformational relevant P89 residue at the cord as well as the N34, a residue strongly correlated with the functionality of the enzyme over pH 5, in so-called alkaline xylanases [39, 40]. Also, amino acids N139, N179, and T181, potentially involved in a secondary binding site for long substrates were identified [41, 42].

Protein sequence and structural analysis described above supported the potential endo-xylanolytic activity of GH10XynA and GH11XynB, which correlated with the enzymatic activity features observed (see below), and in accordance to the assigned GH family. Recombinant expression and enzyme activity characterization were performed in order to confirm these hypothesis and to evaluate their potential industrial applications.

Enzymatic Activity Characterization of GH10 and GH11 Recombinant Endoxylanases

In order to characterize the enzymatic activity of the main secreted *Paenibacillus* sp. A59 non-modular endoxylanases, we have cloned and recombinantly expressed GH10XynA and GH11XynB without their signal peptides and as His6X N-terminal fusion proteins. rGH10XynA had a molecular weight of 32 kDa and rGH11XynB 20 kDa, as observed in a SDS-PAGE of purified proteins. Both displayed endoxylanase activity in zymograms using beechwood xylan as substrate (Fig. 3) and were observed as monomeric proteins in non-denaturing gels (not shown). Substrate specificity of both enzymes was assayed by using different hemicellulosic and

cellulosic substrates (Table 3). rGH11XynB was only active on beechwood xylan (85.04 ± 0.02 IU/mg) and wheat arabinoxylan (131.56 ± 0.08 IU/mg), confirming the true xylanase feature of the enzymes of family GH11 [43]. In contrast, although the main activity of rGH10XynA was on beechwood xylan (105.96 ± 0.05 IU/mg) and wheat arabinoxylan (99.58 ± 0.14 IU/mg), some activity was also observed on *p*NP-xylopyranose (0.06 ± 0.02 IU/mg) and *p*NP-cellobioside (1.25 ± 0.01 IU/mg). However, no activity on polymeric cellulosic substrates (either soluble or crystalline) was observed. Optimal temperature and pH of each xylanase were tested on beechwood xylan to further characterize their activities under different reaction conditions (Fig. 4). rGH10XynA showed the highest xylanase activity at pH 6 and 60 °C, with more than 70% of activity in a pH range from 5.5 to 7.5 and at temperatures from 45 to 60 °C. However, the enzyme lost almost all activity after 60 min at 60 °C, while it remained completely stable for 120 min at 45 and 50 °C. Thus, pH 6 and 45 to 50 °C were selected as the best reaction conditions for further assays. rGH11XynB showed the highest xylanase activity at pH 9 and 45 °C, with more than 60% of activity in a wide pH range from 5 to 9 and from 40 to 55 °C. At the optimal temperature of 45 °C, the enzyme retained more than 75% of xylanase activity after 120 min. Thus, 45 °C, pH 9, and pH 6 were selected to be used for further assays. Kinetic analysis determined on beechwood xylan as substrate showed differences between both endoxylanases: while rGH10XynA presented a hyperbolic kinetic profile (fitting Michaelis-Menten function), rGH11XynB had a sigmoidal one (non-Michaelis-Menten behavior), indicating cooperative binding (Fig. 4e, f). The K_M and V_{max} of rGH10XynA were 14 mg/ml and 570 μ mol/min/mg, respectively. Also, the turnover number for this enzyme ($V_{max}/[Et]$) was calculated as 304/seg and thus the catalytic efficiency (k_{cat}/K_M) resulted 21.7 ml/s/mg, which is in the

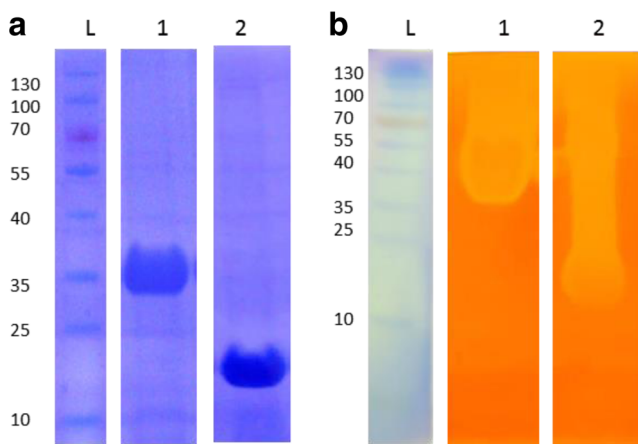


Fig. 3 Analysis of purified rGH10XynA and rGH11XynB. Soluble IMAC purification (10% SDS-PAGE) (**a**) and xylan zymography (12% SDS-PAGE-0.1% beechwood xylan) (**b**) of rGH10XynA (lane 1) and rGH11XynB (lane 2). Pre-stained protein marker (L)

order of other reported GH10 xylanases [44, 45]. For rGH11XynB, the K_{half} was 6.3 mg/ml and the V_{max} was 192 $\mu\text{mol}/\text{min}/\text{mg}$, with a Hill coefficient (η_{H}) of 4.06. These results suggest the possible presence of secondary binding sites (SBSs) in rGH11XynB, as recently observed in other two GH11 endoxylanases (from *Bacillus circulans* and *Pseudozyma brasiliensis*) [41, 42]. A relatively flat region flanked by asparagine and threonine residues, along with tryptophan, could act as SBSs, and the key amino acids identified by mutational studies [42] are highly conserved in rGH11XynB (Fig. 2).

Xylan and Pre-treated Biomass Bioconversion

With the objective of understanding the mechanism of action of rGH10XynA and rGH11XynB, we performed hydrolysis reactions using non-branched xylo-oligosaccharides (X3 and X4) and xylobiose (X2). As expected, none of the enzymes

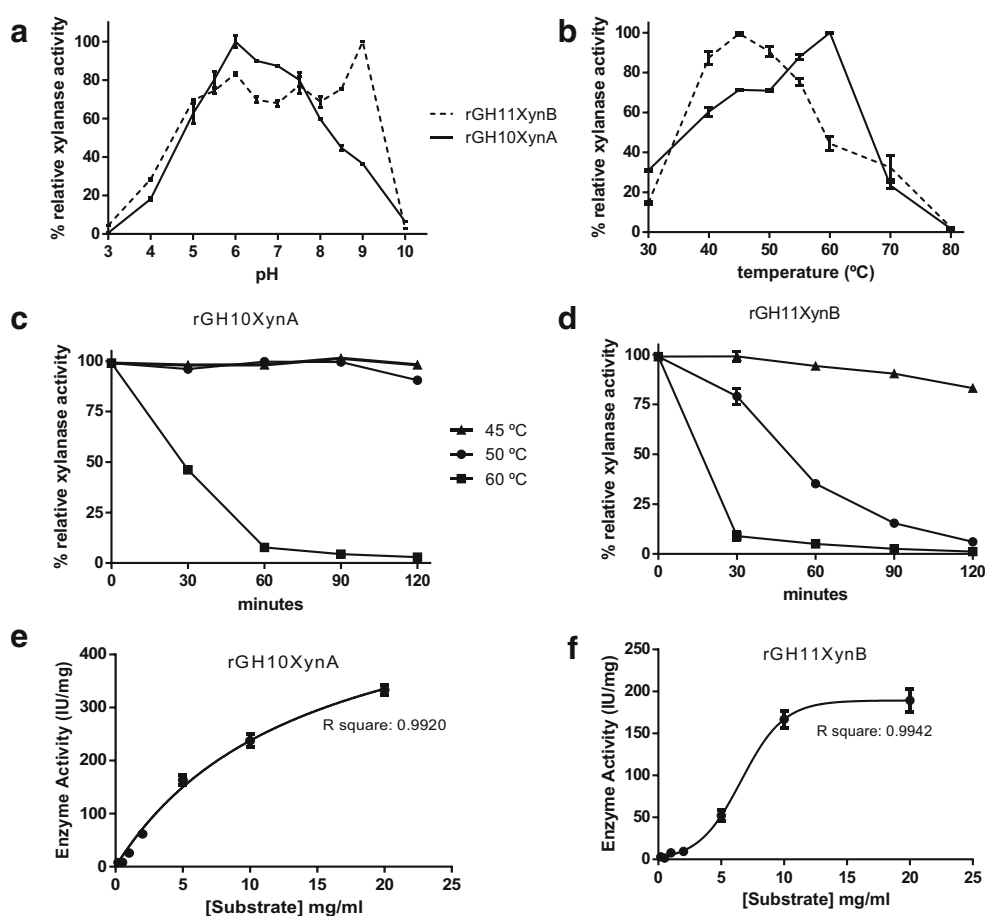
hydrolyzed X2. rGH10XynA hydrolyzed X4 and X3, releasing mainly X2 and X1, which could indicate the hydrolysis of terminal glycosidic bonds of X3 and X4 in a processive way (therefore releasing xylose), as the main mechanism (Table 4; Online Resources 4). In contrast, rGH11XynB was active on X4, releasing X3 and X2 as the main products, and presented very low activity on X3. While rGH10XynA was able to hydrolyze substrates with $\text{DP} \geq 3$, rGH11XynB showed the main activity on substrates with $\text{DP} \geq 4$. Furthermore, in order to study, the xylanolytic activity of the two recombinant enzymes, analysis of soluble sugars released from xylan was conducted. The profile of sugars released from this substrate was found to be different for the two enzymes. rGH10XynA released xylobiose (X2) as the main sugar and also xylose (X1) and xylotriose (X3), while rGH11XynB did not release quantifiable amounts of X1, releasing only X2, X3 (at similar proportion), and xylo-oligosaccharides (XOS) of higher DP. In both cases, X3, X4, and XOS of higher DP were observed in TLC but they were not resolved by HPLC (Fig. 5a, c, e). These results could be explained by the diverse substrate accession of enzymes, based on the different structural conformations of their active sites. In order to quantify sugars released from xylan, different reaction conditions were tested and also a combination of both enzymes was assayed to evaluate if they had additive effect (Table 5). As previously determined by DNS, at the best reaction conditions for rGH10XynA (pH 6 and 45 or 50 $^{\circ}\text{C}$), equal amounts of reducing sugars were measured (400 $\mu\text{mol}/\text{min}$ of reaction); however, different hydrolysis products were observed by HPLC for sugars with DP higher than 2, with practically no X3 detected at 50 $^{\circ}\text{C}$ when rGH10XynA was used. Similarly, comparing rGH11XynB activity at the two optimal pH conditions, pH 6 and pH 9 (at 45 $^{\circ}\text{C}$), almost twice the concentration of X2 was obtained at pH 6 (Table 5). This can be explained by the fact that pH dependence activity profiles of glycoside hydrolases reflect the ionization state of the two carboxylic catalytic residues in the active site [46]. These results

Table 3 Substrate specificity of rGH10XynA and rGH11XynB endoxylanases

	rGH10XynA (IU/mg)	rGH11XynB (IU/mg)
Beechwood xylan	105.96 \pm 0.05	85.04 \pm 0.02
Wheat arabinoxylan (medium viscosity)	99.58 \pm 0.14	131.56 \pm 0.08
Avicel	ND	ND
Carboxymethyl cellulose	ND	ND
Bacterial cellulose	ND	ND
PASC (phosphoric acid swollen cellulose)	ND	ND
pNP-arabinofuranoside	ND	ND
pNP-xylopyranoside	0.06 \pm 0.02	ND
pNP-glucopyranoside	ND	ND
pNP-cellobioside	1.25 \pm 0.01	ND

ND not detected, IU international units (μmol of product/min of reaction)

Fig. 4 Xylanase profile activity of rGH10XynA and rGH11XynB. Optimal pH condition (a), temperature (b) thermal stability (c, d) and kinetic analysis (e, f) of endoxylanases rGH10XynA and rGH11XynB. Results correspond to mean and standard deviations of technical triplicates. Two independent biological replicate assays were undertaken with equivalent results. IU international units: $\mu\text{mol}/\text{min}$



highlighted the importance of analyzing the detailed pattern of hydrolysis of the substrate in study to select the most suitable reaction conditions for different industrial applications where sugars with different DP are needed. The combination of both enzymes at equal proportion resulted in no additive or synergic effect, since the observed results were equivalent to those resulting from the action of rGH10XynA.

Analysis of hydrolysis products released from wheat arabinoxylan, a highly arabinose-branched xylan (38%

arabinose substitution), showed a low amount of xylobiose as the main product (0.54 ± 0.07 mg/ml and 0.22 ± 0.00 mg/ml for rGH10XynA and rGH11XynB, respectively) and non-quantifiable xylose (<0.1 mg/ml) (Table 5). These results indicated a very low conversion of highly branched arabinoxylans by both recombinant enzymes. In order to evaluate the activity on arabinoxylan-rich biomass, extruded wheat straw (3.5% arabinose substitution) was used as substrate. Different hydrolysis patterns were

Table 4 Activity of rGH10XynA and rGH11XynB on short non-branched xylooligosaccharides

	Substrate	Hydrolysis products (mg/ml)			
		X4	X3	X2	X1
rGH10XynA	X4	ND	ND	1.66 ± 0.90 (86%)	0.28 ± 0.01 (14%)
	X3	ND	ND	1.72 ± 0.06 (77%)	0.51 ± 0.00 (23%)
	X2	ND	ND	1.48 ± 0.11 (100%)	ND
rGH11XynB	X4	ND	0.37 ± 0.04 (27%)	1.01 ± 0.06 (73%)	NQ
	X3	ND	1.41 ± 0.17 (94%)	0.10 ± 0.00 (6%)	NQ
	X2	ND	ND	1.25 ± 0.14 (100%)	ND

The percentage (%) of each reaction product is shown between parentheses

X4 xylotetraose, X3 xylotriose, X2 xylobiose, X1: xylose, ND not detected, NQ not quantifiable (<0.1 mg/ml)

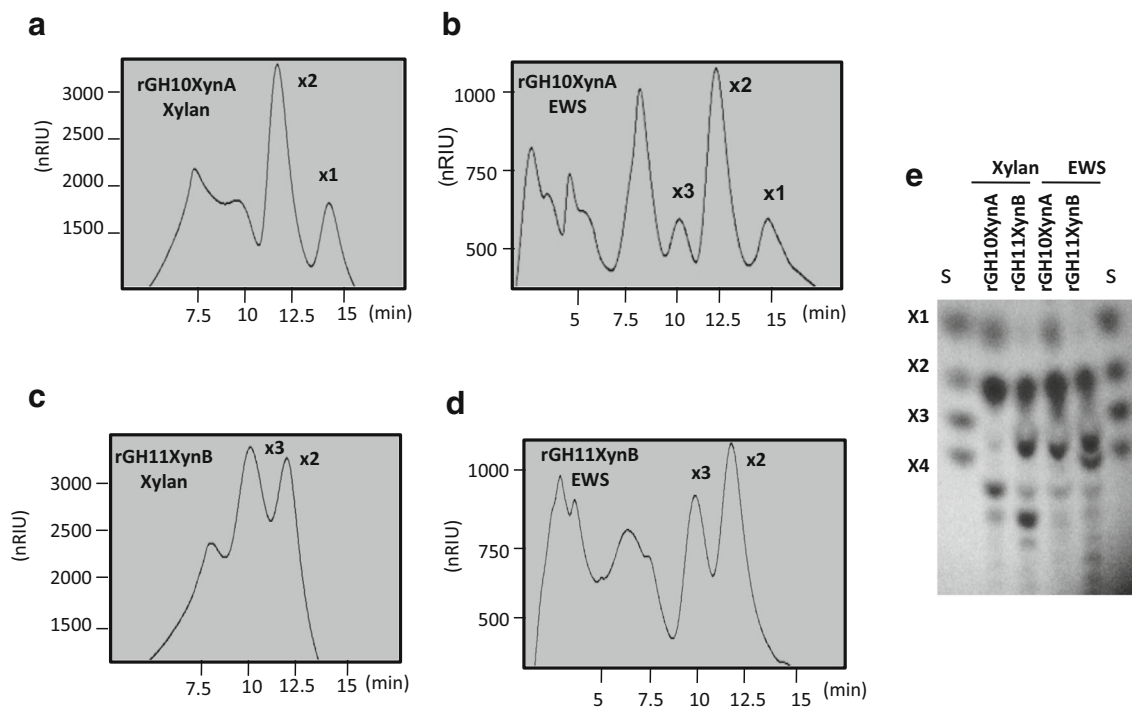


Fig. 5 Hydrolysis products released from xylan and extruded wheat straw. HPLC determination of soluble sugars released by rGH10XynA and rGH11XynB at pH 6 and 45 °C from xylan (a, c) and extruded wheat straw (EWS) (b, d), respectively. X-axis corresponds to retention time in minutes (min) and Y-axis to signal intensity (nRIU). Results correspond to

mean and standard deviations of technical triplicates. Two independent biological replicate assays were undertaken with equivalent results. TLC confirmation of hydrolysis products (e). Abbreviation code: EWS extruded wheat straw, S standards, X1 xylose, X2 xylobiose, X3 xylotriose, X4 xylotetraose

observed for both endoxylanases by TLC and HPLC (Fig. 5b, d, e; Table 5). While XOS of DP > 2 and no arabinose were released by both enzymes, X1 was only detected by the action of rGH10XynA (0.45 ± 0.10 mg/ml), and the highest concentration of X3 was released by rGH11XynB (0.46 ± 0.10 mg/ml). No additive effect or synergism was observed with the combination of rGH11XynB and

rGH10XynA at the same proportion, similarly to the results obtained on xylan. Similar results were achieved on extruded sweet corn cob (SCC) and barley straw (BS) (with 7.1 and 5.1% of arabinose substitutions respectively), obtaining xylobiose (1.29 ± 0.27 mg/ml for SCC and 0.24 ± 0.05 mg/ml for BS) and xylose (0.24 ± 0.05 mg/ml for SCC and 0.32 ± 0.05 mg/ml for BS) as the main

Table 5 Soluble sugars released from beechwood xylan and extruded wheat straw and conversion rates, determined by HPLC

	Assay condition	Xylotriose (mg/ml)	Xylobiose (mg/ml)	Xylose (mg/ml)	Arabinose (mg/ml)
Beechwood xylan					
rGH10XynA	45 °C, pH 6	NR	2.54 ± 0.50 (54.25%)	0.90 ± 0.20 (15.84%)	ND
rGH10XynA	50 °C, pH 6	ND	1.80 ± 0.20 (38.44%)	0.77 ± 0.60 (13.55%)	ND
rGH11XynB	45 °C, pH 6	0.76 ± 0.60	1.30 ± 0.20 (27.70%)	ND	ND
rGH11XynB	45 °C, pH 9	0.80 ± 0.30	0.60 ± 0.20 (12.80%)	ND	ND
rGH10XynA + rGH11XynB	45 °C, pH 6	NR	2.90 ± 0.34 (61.90%)	0.76 ± 0.10 (13.37%)	ND
Extruded wheat straw					
rGH10XynA	45 °C, pH 6	0.12 ± 0.10	0.82 ± 0.10 (14.50%)	0.45 ± 0.10 (6.58%)	ND
rGH11XynB	45 °C, pH 6	0.46 ± 0.10	0.69 ± 0.10 (12.30%)	NQ	ND
rGH10XynA + rGH11XynB	45 °C, pH 6	ND	1.11 ± 0.10 (19.70%)	0.61 ± 0.10 (8.92%)	ND
NS22002 (1% v/v)	45 °C, pH 6	ND	2.70 ± 0.13 (48%)	1.65 ± 0.11 (24.14%)	0.76 ± 0.05 (94%)

Percentages (%) of conversion to xylobiose, xylose, and arabinose are shown between parentheses

NR not resolved, ND not detected, NQ not quantifiable (≤ 0.1 mg/ml)

products of rGH10XynA and xylotriose (0.47 ± 0.27 mg/ml for SCC and 0.25 ± 0.04 mg/ml for BS) and xylobiose (1.09 ± 0.26 mg/ml for SCC and 0.58 ± 0.12 mg/ml for BS) as the main products of rGH11XynB. The highly optimized commercial hemicellulase cocktail NS22002 (at 1% v/v, equivalent to 0.015 mg protein/mg biomass) was used as benchmark on wheat straw hydrolysis (Table 5), resulting in high conversion to X2, X1, and arabinose as main products, in accordance to its beta-xylosidase and arabinofuranosidase additional activities.

Discussion

Paenibacillus sp. A59, a Gram-positive, facultative anaerobic bacterium capable of hydrolyzing a wide variety of carbohydrate and protein based substrates, was previously isolated as a mainly hemicellulolytic strain from a soil microbial consortium [21]. In the present work, we analyzed the xylanolytic system of *Paenibacillus* sp. A59 by studying the diversity of catalytic and substrate-binding domains of glycoside hydrolases encoded in the genome as well as identifying which of these proteins were actually being secreted. By genome sequence analysis of *Paenibacillus* sp. A59, we identified about 90 glycoside hydrolases potentially active on abundant polymers in nature such as cellulose, hemicellulose, chitin and starch, showing the great potential of this genus for prospection of biotechnologically important enzymes [19]. From the 35 potentially secreted glycoside hydrolases, 10 were identified in xylan culture supernatant (approximately one third of total encoded GHs). We had previously assayed sugar cane straw and carboxymethylcellulose culture supernatants and had identified only a GH10 and a GH48, respectively, in accordance to the higher xylanolytic activity achieved when *Paenibacillus* sp. A59 was grown on xylan than on any other substrate evaluated so far [21]. Surprisingly, not only xylanases or proteins with xylan-binding CBMs were induced, as several other enzymes acting on polysaccharides were found (endoglucanases, chitinases and amylases), suggesting a cross-induction of the cellulose/hemicellulose deconstruction system. The potentially secreted xylanolytic system encoded in the genome included three endoxylanases (two from GH10 and one from GH11 families), seven GH43 β -xylosidases/ α -arabinofuranosidase, and two GH30 glucuronoxylanases. However, only the two GH10s and the GH11 were detected in the supernatant, which also correlated with our previous findings of high xylanase activity but poor β -xylosidase activity in culture supernatant [21]. Therefore, expression assays of the potentially secreted beta-xylosidases would help to fully understand their role in xylan deconstruction.

One of the GH10 endoxylanases (160 kDa) identified in xylan culture supernatant contained a CBM9, two CBM22, and three SLH domains. Related GH10 endoxylanases from

Paenibacillus sp. W61 and *Paenibacillus* sp. JDR-2, with the same modular structure described above had been previously characterized as active on methyl-glucuronoxylan, by anchoring to the cell-surface with the SLH domains [47, 48]. The other two endoxylanases identified in this study by xylan secretome analysis presented only the catalytic module. The secreted enzymes on xylan hydrolysis had also been studied in other bacteria by proteomic analysis. Other Gram-positive strains secreted different endoxylanases as part of their hemicellulolytic secretome: the GH10-CBM9-CBM22-SLH previously mentioned from *Paenibacillus* JDR2 was described as the main endoxylanase involved in efficient xylan utilization [49], *Bacillus subtilis* secreted two endoxylanases (a GH11 and a GH30) after growth in glucuronoxylan [50], the thermophilic anaerobe *Caldicellulosiruptor kronotskyensis* [51] secreted a GH10-CBM22 endoxylanase, and the hyperthermophile *Thermotoga maritima* presented a membrane bound multidomain GH10 endoxylanase as main component of its xylanolytic system [52]. Based on our findings and existing data, a possible hypothesis is that the modular GH10-CBM-SLH endoxylanase could remain adsorbed to the bacterial cell wall by the SLH domains while the CBMs would assist in binding to xylan [53]. This protein, along with the free endoxylanases GH10XynA and GH11XynB, could act concertedly to generate short xylooligosaccharides, xylobiose, and some xylose to be incorporated into the cell. The secreted sugar transporters and substrate-binding proteins, identified in the culture supernatant, could have a role in this mechanism. In a previous study, we had also identified an uncharacterized protein with SLH domains as one of the main proteins secreted by *Paenibacillus* sp. A59 after growth in sugarcane agricultural residue (sugarcane straw) [21]. These results suggested an important role of SLH domains in *Paenibacillus* polysaccharide degradation systems.

Whole shotgun genome sequencing of (hemi) cellulolytic bacteria, combined with annotation platforms and CAZY module prediction, can reveal potential enzymes not previously described. However, the true proof of enzymatic activity is only achieved by characterization of the enzymes involved. In this regard, we have recombinantly expressed the free endoxylanases, GH10 and GH11 of *Paenibacillus* sp. A59, and also developed a structural model for both of them.

Molecular modeling of GH10XynA and GH11XynB allowed us confirmation of their tertiary structure and the identification of the key active site residues as well as other sites that can lead to enzyme improving by site-directed mutagenesis. Replacement or variations of residues from other xylanases were previously undertaken in order to improve thermal stability, activity at alkaline pH and substrate selectivity [54–56].

Regarding activity, rGH10XynA presented higher activity values (~ 100 IU/mg) than that reported for some purified

bacterial xylanases commercially available, such as a GH10 endo-1,4- β -xylanase from *Aeromonas punctata* (8.4 IU/mg) or a GH10 endo-1,4- β -xylanase from *Cellvibrio japonicus* (25 IU/mg) (Megazyme, www.megazyme.com), and in the same range as other GH10 endoxylanases from *Paenibacillus* previously described (GH10 from *Paenibacillus* sp. HC1, 12.8 IU/mg; GH10 from *Paenibacillus* DG22, 454.4 IU/mg) [57–60].

Similarly, rGH11XynB presented activity of about 85 IU/mg on xylan and 130 IU/mg on wheat arabinoxylan, which is in the range of some commercial GH11 xylanases, such as a GH11 endo-1,4- β -xylanase M4 from *Aspergillus niger* (80 IU/mg), although there are some with higher activity levels, such as GH11 endo-1,4- β -xylanase from *Neocallimastix patriciarum* (1500 IU/mg).

rGH11XynB showed a wide range of pH tolerance and released soluble xylo-oligosaccharides of higher degree of polymerization than rGH10XynA. Similarly, the GH11 endoxylanase from *Paenibacillus* sp. HY8 (isolated from the digestive tract of an herbivorous longicorn beetle), with 96% sequence homology to GH11XynB, presented the same pattern of xylan hydrolysis and range of endoxylanase activity (147.8 IU/mg, pH 6) but at a narrower range of pH [61]. This suggests that GH11XynB could be of use, based on its alkaline stability and cellulase-free properties, for paper industry applications, where hemicellulose needs to be selectively removed from cellulose under alkaline pre-treatment during the pulping process.

Although activity from recombinant xylanases from *Cellulomonas*, *Bacillus* and other *Paenibacillus* species had been previously studied [62–64], information on activity of purified proteins on pre-treated biomass and conversion to monomeric sugars is not highly abundant in the literature [65, 66].

In this work, we performed a thorough study of the xylanolytic system of *Paenibacillus* sp. A59 and have fully characterized two endoxylanases that can be used for industrial applications, either by themselves or as part of enzymatic cocktails requiring xylanase activity. GH10XynA and GH11XynB can act on pre-treated biomass, efficiently releasing soluble sugars from the hemicellulose fraction and can therefore be used to study XOS production for prebiotics development. In biofuel industry, in order to improve conversion rates, additional enzymes with arabinofuranosidase and beta-xylosidase activities would be also required.

Acknowledgments SG and FP are Ph.D. students from the Department of Biological Chemistry (QB), FCEN, Univ. of Buenos Aires, Argentina. SG holds an INTA doctoral fellowship. PF and OO have doctoral and post-doctoral fellowships, respectively, from the Argentinean National Council of Research (CONICET). EC and PT acknowledge CONICET as Research Career members. Pre-treated wheat straw and barley straw were gently supplied by Dr. Mercedes Ballesteros from CIEMAT, Madrid, Spain. Pre-treated sweet corn cob was supplied by INPT,

France. NS22002 was kindly provided by Novozymes. We also acknowledge Dr. Supriya Ratnaparkhe for her support on TLC assays by means of Bi-national Project MINCYT IN/14/04 (Arg)/ DST/INT/AGR/P-4/2015 (India).

Funding Information This work was supported by project grants MINCYT PICT2014-0791 and INTA PNAIyAV 1130034, from Argentina and the project BABETREAL5, grant agreement no. 654365 funded by the European Commission under the program H2020.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interests.

References

- Sorek N, Yeats TH, Szemenyei H, Youngs H, Somerville CR (2014) The implications of lignocellulosic biomass chemical composition for the production of advanced biofuels. *Bioscience* 64(3): 192–201. <https://doi.org/10.1093/biosci/bit037>
- Juturu V, Wu JC (2012) Microbial xylanases: engineering, production and industrial applications. *Biotechnol Adv* 30(6):1219–1227. <https://doi.org/10.1016/j.biotechadv.2011.11.006>
- Juturu V, Wu JC (2014) Microbial cellulases: engineering, production and applications. *Renew Sust Energy Rev* 33:188–203. <https://doi.org/10.1016/j.rser.2014.01.077>
- Gray KA, Zhao L, Emptage M (2006) Bioethanol. *Curr Opin Chem Biol* 10(2):141–146. <https://doi.org/10.1016/j.cbpa.2006.02.035>
- Brethauer S, Studer MH (2015) Biochemical conversion processes of lignocellulosic biomass to fuels and chemicals—a review. *Chim Int J Chem* 69(10):572–581. <https://doi.org/10.2533/chimia.2015.572>
- Brethauer S, Wyman CE (2010) Review: continuous hydrolysis and fermentation for cellulosic ethanol production. *Bioresour Technol* 101(13):4862–4874. <https://doi.org/10.1016/j.biortech.2009.11.009>
- Kim IJ, Lee HJ, Choi IG, Kim KH (2014) Synergistic proteins for the enhanced enzymatic hydrolysis of cellulose by cellulase. *Appl Microbiol Biotechnol* 98(20):8469–8480. <https://doi.org/10.1007/s00253-014-6001-3>
- Scheller HV, Ulvskov P (2010) Hemicelluloses. *Annu Rev plan Biol* 61(1):263–289. <https://doi.org/10.1146/annurev-arplant-042809-112315>
- Moreira LRS, Filho EXF (2016) Insights into the mechanism of enzymatic hydrolysis of xylan. *Appl Microbiol Biotechnol* 100(12):5205–5214. <https://doi.org/10.1007/s00253-016-7555-z>
- Cantarel BI, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res* 37(Database):233–238. <https://doi.org/10.1093/nar/gkn663>
- Pastor FIJ, Gallardo Ó, Sanz-Aparicio J, Díaz P (2007) Xylanases: molecular properties and applications. In: Polaina J, Mac Cabe AP (eds) *Ind. Enzym. Struct. Funct. Appl.* Springer, Netherlands, pp 65–82. <https://doi.org/10.1007/1-4020-5377-0>
- Beg QK, Kapoor M, Mahajan L, Hoondal GS (2001) Microbial xylanases and their industrial applications: a review. *Appl Microbiol Biotechnol* 56(3-4):326–338. <https://doi.org/10.1007/s002530100704>
- Collins T, Gerday C, Feller G (2005) Xylanases, xylanase families and extremophilic xylanases. *FEMS Microbiol Rev* 29(1):3–23. <https://doi.org/10.1016/j.femsre.2004.06.005>

14. Chen C-C, Ko T-P, Huang J-W, Guo R-T (2015) Heat-and alkaline-stable xylanases: application, protein structure and engineering. *Chem Bioengineering Rev* 2(2):95–106. <https://doi.org/10.1002/cben.201400035>
15. Chakdar H, Kumar M, Pandiyan K, Singh A, Nanjappan K, Kashyap PL, Srivastava AK (2016) Bacterial xylanases: biology to biotechnology. *3 Biotech* 150(2):1–15. <https://doi.org/10.1007/s13205-016-0457-z>
16. Subramaniyan S, Prema P (2000) Cellulase-free xylanases from *Bacillus* and other microorganisms. *FEMS Microbiol Lett* 183(1): 1–7. [https://doi.org/10.1016/S0378-1097\(99\)00620-5](https://doi.org/10.1016/S0378-1097(99)00620-5)
17. Orencio-Trejo M, De la Torre-Zavala S, Rodriguez-Garcia A, Aviles-Arnaut H, Gastelum-Arellanez A (2016) Assessing the performance of bacterial cellulases: the use of *Bacillus* and *Paenibacillus* strains as enzyme sources for lignocellulose saccharification. *Bioenergy Res* 9(4):1023–1033. <https://doi.org/10.1007/s12155-016-9797-0>
18. Talamantes D, Biabini N, Dang H, Abdoun K, Berlemont R (2016) Natural diversity of cellulases, xylanases, and chitinases in bacteria. *Biotechnol Biofuels* 9(133):1–11. <https://doi.org/10.1186/s13068-016-0538-6>
19. López-Mondéjar R, Zühlke D, Větrovský T, Becher D, Riedel K, Baldrian P (2016) Decoding the complete arsenal for cellulose and hemicellulose deconstruction in the highly efficient cellulose decomposer *Paenibacillus* O199. *Biotechnol Biofuels* 9(104):1–12. <https://doi.org/10.1186/s13068-016-0518-x>
20. Ma J, Zhang K, Liao H, Hector SB, Shi X, Li J, Liu B, Xu T, Tong C, Liu X, Zhu Y (2016) Genomic and secretomic insight into lignocellulolytic system of an endophytic bacterium *Pantoea ananatis* Sd-1. *Biotechnol Biofuels* 9(25):1–15. <https://doi.org/10.1186/s13068-016-0439-8>
21. Ghio S, Insani EM, Piccinni FE, Talia PM, Grasso DH, Campos E (2016) GH10 XynA is the main xylanase identified in the crude enzymatic extract of *Paenibacillus* sp. A59 when grown on xylan or lignocellulosic biomass. *Microbiol Res* 186:16–26. <https://doi.org/10.1016/j.micres.2016.02.006>
22. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y (2012) DbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 40(W1):445–451. <https://doi.org/10.1093/nar/gks479>
23. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42(D1):490–495. <https://doi.org/10.1093/nar/gkt1178>
24. Ghio S, Martínez Cáceres AI, Talia P, Grasso DH, Campos E (2015) Draft genome sequence of cellulolytic and xylanolytic *Paenibacillus* sp. A59, isolated from decaying forest soil from Patagonia, Argentina. *Genome Announc* 3(5):2013–2014. <https://doi.org/10.1128/genomeA.01233-15>
25. Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) Signal P 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8(10):785–786. <https://doi.org/10.1038/nmeth.1701>
26. Anagnostakis L, Hankin S (1977) Solid media containing carboxymethylcellulose to detect CX cellulose activity of microorganisms. *J Gen Microbiol* 98(3):108–115. <https://doi.org/10.2307/3758395>
27. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30(12):2725–2729. <https://doi.org/10.1093/molbev/mst197>
28. Roy A, Kucukural A, Zhang Y (2011) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* 5:725–738. <https://doi.org/10.1038/nprot.2010.5.I-TASSER>
29. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrinte (2004) UCSF chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25(13): 1605–1612. <https://doi.org/10.1002/jcc.20084>
30. Miller GL (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal Chem* 31(3):426–430. <https://doi.org/10.1021/ac60147a030>
31. Boraston AB, Kwan E, Chiu P, Warren RA, Kilburn DG (2003) Recognition and hydrolysis of noncrystalline cellulose. *J Biol Chem* 278(8):6120–6127. <https://doi.org/10.1074/jbc.M209554200>
32. Krueger-Zerhusen N, Wilson DB (2015) Bacterial AA10 lytic polysaccharide monooxygenases enhance the hydrolytic degradation of recalcitrant substrates. In: Himmel ME (ed) *Direct microbial conversion of biomass to advanced biofuels*. Elsevier, Amsterdam, pp 91–110. <https://doi.org/10.1016/B978-0-444-59592-8.01001-0>
33. Santos CR, Meza AN, Hoffmann ZB, Silva JC, Alvarez TM, Ruller R, Giesel GM, Verli H, Squina FM, Prade RA, Murakami MT (2010) Thermal-induced conformational changes in the product release area drive the enzymatic activity of xylanases 10B: crystal structure, conformational stability and functional characterization of the xylanase 10B from *Thermotoga petrophila* RKU-1. *Biochem Biophys Res Commun* 403(2):214–219. <https://doi.org/10.1016/j.bbrc.2010.11.010>
34. Solomon V, Teplitsky A, Shulami S, Zolotnitsky G, Shoham Y, Shoham G (2007) Structure-specificity relationships of an intracellular xylanase from *Geobacillus stearothermophilus*. *Acta Crystallogr Sect D Biol Crystallogr* 63(8):845–859. <https://doi.org/10.1107/S0907444907024845>
35. Ay J, Götz F, Borriss R, Heinemann U (1998) Structure and function of the *Bacillus* hybrid enzyme GluXyn-1: native-like jellyroll fold preserved after insertion of autonomous globular domain. *Proc Natl Acad Sci U S A* 95(12):6613–6618. <https://doi.org/10.1073/pnas.95.12.6613>
36. Törrönen A, Harkki A, Rouvinen J (1994) Three-dimensional structure of endo-1,4-beta-xylanase II from *Trichoderma reesei*: two conformational states in the active site. *EMBO J* 13(11):2493–2501
37. Törrönen A, Rouvinen J (1995) Structural comparison of two major endo-1,4-xylanases from *Trichoderma reesei*. *Biochemistry* 34(3): 847–856. <https://doi.org/10.1021/bi00003a019>
38. Paës G, Berrin JG, Beaugrand J (2012) GH11 xylanases: structure/function/properties relationships and applications. *Biotechnol Adv* 30(3):564–592. <https://doi.org/10.1016/j.biotechadv.2011.10.003>
39. Törrönen A, Rouvinen J (1997) Structural and functional properties of low molecular weight endo-1,4-beta-xylanases. *J Biotechnol* 57(1-3):137–149. [https://doi.org/10.1016/S0168-1656\(97\)00095-3](https://doi.org/10.1016/S0168-1656(97)00095-3)
40. Sapag A, Wouters J, Lambert C, de Loannes P, Eyzaguirre J, Depiereux E (2002) The endoxylanases from family 11: computer analysis of protein sequences reveals important structural and phylogenetic relationships. *J Biotechnol* 95(2):109–131. [https://doi.org/10.1016/S0168-1656\(02\)00002-0](https://doi.org/10.1016/S0168-1656(02)00002-0)
41. Cuyvers S, Dornez E, Rezaei MN, Pollet A, Delcour JA, Courtin CM (2011) Secondary substrate binding strongly affects activity and binding affinity of *Bacillus subtilis* and *Aspergillus niger* GH11 xylanases. *FEBS J* 278(7):1098–1111. <https://doi.org/10.1111/j.1742-4658.2011.08023.x>
42. Borges TA, Tieppo de Souza A, Squina FM, Riaño-Pachón DM, Corrêa dos Santos RA, Machado E, Velasco de Castro Oliveira J, Damásio ARL, Goldman GH (2014) Biochemical characterization of an endoxylanase from *Pseudozyma brasiliensis* sp. nov. strain GHG001 isolated from the intestinal tract of *Chrysomelidae* larvae associated to sugarcane roots. *Process Biochem* 49(1):77–83. <https://doi.org/10.1016/j.procbio.2013.10.004>

43. Pollet A, Delcour JA, Courtin CM (2010) Structural determinants of the substrate specificities of xylanases from different glycosyl hydrolase families. *Crit Rev Biotechnol* 30(3):176–191. <https://doi.org/10.3109/07388551003645599>
44. Guo B, Chen XL, Sun CY, Zhou BC, Zhang YZ (2009) Gene cloning, expression and characterization of a new cold-active and salt-tolerant endo-beta-1,4-xylanase from marine *Glaciicola mesophila* KMM 241. *Appl Microbiol Biotechnol* 84(6):1107–1115. <https://doi.org/10.1007/s00253-009-2056-y>
45. Liao H, Zheng H, Li S, Wei Z, Mei X, Ma H, Shen Q, Xu Y (2015) Functional diversity and properties of multiple xylanases from *Penicillium oxalicum* GZ-2. *Sci Rep* 5(1):12631. <https://doi.org/10.1038/srep12631>
46. Ly HD, Withers SG (1999) Mutagenesis of glycosidases. *Annu Rev Biochem* 68(1):487–522. <https://doi.org/10.1146/annurev.biochem.68.1.487>
47. Ito Y, Tomita T, Roy N, Nakano A, Sugawara-Tomita N, Watanabe S, Okai N, Abe N, Kamio Y (2003) Cloning, expression, and cell surface localization of *Paenibacillus* sp. strain W-61 xylanase 5, a multidomain xylanase. *Appl Environ Microbiol* 69(12):6969–6978. <https://doi.org/10.1128/AEM.69.12.6969-6978.2003>
48. St John FJ, Rice JD, Preston JF (2006) *Paenibacillus* sp. strain JDR-2 and XynA1: a novel system for methyl glucuronoxylan utilization. *Appl Environ Microbiol* 72(2):1496–1506. <https://doi.org/10.1128/AEM.72.2.1496-1506.2006>
49. Sawhney N, Crooks C, Chow V, Preston JF, St John FJ (2016) Genomic and transcriptomic analysis of carbohydrate utilization by *Paenibacillus* sp. JDR-2: systems for bioprocessing plant polysaccharides. *BMC Genomics* 17(1):131. <https://doi.org/10.1186/s12864-016-2436-5>
50. Rhee MS, Wei L, Sawhney N, Rice JD, St. John FJ, Hurlbert JC, Preston JF (2014) Engineering the xylan utilization system in *Bacillus subtilis* for production of acidic xylooligosaccharides. *Appl Environ Microbiol* 80(3):917–927. <https://doi.org/10.1128/AEM.03246-13>
51. Jia X, Qiao W, Tian W, Peng X, Mi S, Su H, Han Y (2016) Biochemical characterization of extra- and intracellular endoxylanase from thermophilic bacterium *Caldicellulosiruptor kronotskyensis*. *Sci Rep* 6(1):21672. <https://doi.org/10.1038/srep21672>
52. Liebl W, Winterhalter C, Baumeister W, Armbrecht M, Valdez M (2008) Xylanase attachment to the cell wall of the hyperthermophilic bacterium *Thermotoga maritima*. *J Bacteriol* 190(4):1350–1358. <https://doi.org/10.1128/JB.01149-07>
53. Ozdemir I, Blumer-Schuette SE, Kelly RM (2012) S-layer homology domain proteins Csac_0678 and Csac_2722 are implicated in plant polysaccharide deconstruction by the extremely thermophilic bacterium *Caldicellulosiruptor saccharolyticus*. *Appl Environ Microbiol* 78(3):768–777. <https://doi.org/10.1128/AEM.07031-11>
54. Moers K, Bourgois T, Rombouts S, Beliën T, Van Campenhout S, Volckaert G, Robben J, Brijs K, Delcour JA, Courtin CM (2007) Alteration of *Bacillus subtilis* XynA endoxylanase substrate selectivity by site-directed mutagenesis. *Enzym Microb Technol* 41(1-2):85–91. <https://doi.org/10.1016/j.enzmictec.2006.12.019>
55. Xie J, Song L, Li X, Yi X, Xu H, Li J, Qiao D, Cao Y (2011) Site-directed mutagenesis and thermostability of xylanase XYNB from *Aspergillus niger* 400264. *Curr Microbiol* 62(1):242–248. <https://doi.org/10.1007/s00284-010-9697-z>
56. Li F, Xie J, Zhang X, Zhao L (2015) Improvement of the optimum pH of *Aspergillus niger* xylanase towards an alkaline pH by site-directed mutagenesis. *J Microbiol Biotechnol* J25(1):11–17. <https://doi.org/10.4014/jmb.1402.02055>
57. Lee SH, Lee HE (2014) Cloning and characterization of a multidomain GH10 xylanase from *Paenibacillus* sp. DG-22. *J Microbiol Biotechnol* 24(11):1525–1535. <https://doi.org/10.4014/jmb.1407.07077>
58. Harada KM, Tanaka K, Fukuda Y, Hashimoto W, Murata K (2008) *Paenibacillus* sp. strain HC1 xylanases responsible for degradation of rice bran hemicellulose. *Microbiol Res* 163(3):293–298. <https://doi.org/10.1016/j.micres.2006.05.011>
59. Waonukul R, Kyu KL, Sakka K, Ratanakhanokchai K (2009) Isolation and characterization of a multienzyme complex (cellulosome) of the *Paenibacillus curdlanolyticus* B-6 grown on Avicel under aerobic conditions. *J Biosci Bioeng* 107(6):610–614. <https://doi.org/10.1016/j.jbiosc.2009.01.010>
60. Valenzuela SV, Díaz P, Javier Pastor FI (2010) Recombinant expression of an alkali stable GH10 xylanase from *Paenibacillus barcinonensis*. *J Agric Food Chem* 58(8):4814–4818. <https://doi.org/10.1021/jf9045792>
61. Heo S, Kwak J, Oh HW, Park DS, Bae KS, Shin DH, Park HY (2006) Characterization of an extracellular xylanase in *Paenibacillus* sp. HY-8 isolated from an herbivorous longicorn beetle. *J Microbiol Biotechnol* 16:1753–1759
62. Liu MQ, Liu GF (2008) Expression of recombinant *Bacillus licheniformis* xylanase A in *Pichia pastoris* and xylooligosaccharides released from xylans by it. *Protein Expr Purif* 57(2):101–107. <https://doi.org/10.1016/j.pep.2007.10.020>
63. Amaya-Delgado L, Mejia-Castillo T, Santiago-Hernández A, Vega-Estrada J, Amelia FG, Xoconostle-Cázares B, Ruiz-Medrano R, Montes-Horcasitas Mdel C, Hidalgo-Lara ME (2010) Cloning and expression of a novel, moderately thermostable xylanase-encoding gene (Cflxyn11A) from *Cellulomonas flavigena*. *Bioresour Technol* 101(14):5539–5545. <https://doi.org/10.1016/j.biortech.2010.02.057>
64. Zheng HC, Sun MZ, Meng LC, Pei HS, Zhang XQ, Yan Z, Zeng WH, Zhang JS, Hu JR, Lu FP, Sun JS (2014) Purification and characterization of a thermostable xylanase from *Paenibacillus* sp. NF1 and its application in xylooligosaccharides production. *J Microbiol Biotechnol* 24(4):489–496. <https://doi.org/10.4014/jmb.1312.12072>
65. Lequart C, Nuzillard JM, Kurek B, Debeire P (1999) Hydrolysis of wheat bran and straw by an endoxylanase: production and structural characterization of cinnamoyl-oligosaccharides. *Carbohydr Res* 319(1-4):102–111. [https://doi.org/10.1016/S0008-6215\(99\)00110-X](https://doi.org/10.1016/S0008-6215(99)00110-X)
66. Liavoga AB, Bian Y, Seib PA (2007) Release of D-xylose from wheat straw by acid and xylanase hydrolysis and purification of xylitol. *J Agric Food Chem* 55(19):7758–7766. <https://doi.org/10.1021/jf070862k>