



**HAL**  
open science

## The GH51 $\alpha$ -l-arabinofuranosidase from *Paenibacillus* sp. THS1 is multifunctional, hydrolyzing main-chain and side-chain glycosidic bonds in heteroxylans

Hanen Bouraoui, Marie-Laure Desrousseaux, Eleni Ioannou, Pablo Alvira, Mohamed Manaï, Caroline Rémond, Claire Dumon, Narcis Fernandez-Fuentes, Michael O'Donohue

### ► To cite this version:

Hanen Bouraoui, Marie-Laure Desrousseaux, Eleni Ioannou, Pablo Alvira, Mohamed Manaï, et al.. The GH51  $\alpha$ -l-arabinofuranosidase from *Paenibacillus* sp. THS1 is multifunctional, hydrolyzing main-chain and side-chain glycosidic bonds in heteroxylans. *Biotechnology for Biofuels*, 2016, 9 (1), pp.1-14. 10.1186/s13068-016-0550-x . hal-01595481

**HAL Id: hal-01595481**

**<https://hal.science/hal-01595481v1>**

Submitted on 26 Sep 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.


L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

RESEARCH

Open Access



# The GH51 $\alpha$ -L-arabinofuranosidase from *Paenibacillus* sp. THS1 is multifunctional, hydrolyzing main-chain and side-chain glycosidic bonds in heteroxylans

Hanan Bouraoui<sup>1,5†</sup>, Marie-Laure Desrousseaux<sup>2†</sup>, Eleni Ioannou<sup>2,4</sup>, Pablo Alvira<sup>2</sup>, Mohamed Manai<sup>1</sup>, Caroline Rémond<sup>3</sup>, Claire Dumon<sup>2</sup>, Narcis Fernandez-Fuentes<sup>4</sup> and Michael J. O'Donohue<sup>2\*</sup> 

## Abstract

**Background:** Conceptually, multi-functional enzymes are attractive because in the case of complex polymer hydrolysis having two or more activities defined by a single enzyme offers the possibility of synergy and reduced enzyme cocktail complexity. Nevertheless, multi-functional enzymes are quite rare and are generally multi-domain assemblies with each activity being defined by a separate protein module. However, a recent report described a GH51 arabinofuranosidase from *Alicyclobacillus* sp. A4 that displays both  $\alpha$ -L-arabinofuranosidase and  $\beta$ -D-xylanase activities, which are defined by a single active site. Following on from this, we describe in detail another multi-functional GH51 arabinofuranosidase and discuss the molecular basis of multifunctionality.

**Results:** THSAbf is a GH51  $\alpha$ -L-arabinofuranosidase. Characterization revealed that THSAbf is active up to 75 °C, stable at 60 °C and active over a broad pH range (4–7). THSAbf preferentially releases para-nitrophenyl from the L-arabinofuranoside ( $k_{\text{cat}}/K_M = 1050 \text{ s}^{-1} \text{ mM}^{-1}$ ) and to some extent from D-galactofuranoside and D-xyloside. THSAbf is active on 4-O-methylglucuronoxylans from birch and beechwood (10.8 and 14.4 U mg<sup>-1</sup>, respectively) and on sugar beet branched and linear arabinans (1.1 ± 0.24 and 1.8 ± 0.1 U mg<sup>-1</sup>). Further investigation revealed that like the *Alicyclobacillus* sp. A4  $\alpha$ -L-arabinofuranosidase, THSAbf also displays endo-xylanase activity, cleaving  $\beta$ -1,4 bonds in heteroxylans. The optimum pH for THSAbf activity is substrate dependent, but ablation of the catalytic nucleophile caused a general loss of activity, indicating the involvement of a single active center. Combining the  $\alpha$ -L-arabinofuranosidase with a GH11 endoxylanase did not procure synergy. The molecular modeling of THSAbf revealed a wide active site cleft and clues to explain multi-functionality.

**Conclusion:** The discovery of single active site, multifunctional enzymes such as THSAbf opens up exciting avenues for enzyme engineering and the development of new biomass-degrading cocktails that could considerably reduce enzyme production costs.

**Keywords:** Glycoside hydrolase, Xylanase, Wheat bran, Enzyme cocktails, Biomass

\*Correspondence: michael.odonohue@insa-toulouse.fr

<sup>†</sup>Hanan Bouraoui and Marie-Laure Desrousseaux contributed equally to this work

<sup>2</sup> CNRS, INRA, INSA, LISBP, Université de Toulouse, Toulouse, France

Full list of author information is available at the end of the article

## Background

Heteroxylans (commonly referred to as xylans) composed of  $\beta$ -1,4-linked D-xylosyl subunits constitute an important class of polysaccharides that are widespread throughout the plant kingdom, notably in flowering plants [1]. In both perennial and annual grassy species, the main chains of xylans are mainly decorated with L-arabinofuranosyl moieties. However, depending on the exact origin of the source material, the frequency of the arabinofuranosyl decorations and the bonds linking them to the xylan main chain are different [2].

One consequence of the structural complexity of heteroxylans is the diversity of the enzymes that are required to break them down into sugars. This is illustrated by the complex enzymatic arsenals deployed by microbial plant pathogens, plant saprophytes and gut microbiota [3]. Generally, the core enzyme activities produced by xylanolytic microorganisms are endo- $\beta$ -1,4-D-xylanases (EC 3.2.1.8) and  $\beta$ -1,4-D-xylosidases (EC 3.2.1.37), which act on  $\beta$ -1,4 bonds that link D-xylosyl moieties in polysaccharides and oligosaccharides, and  $\alpha$ -L-arabinofuranosidases (EC 3.2.1.55) that hydrolyze the 1,2 and 1,3 glycosidic that link  $\alpha$ -L-arabinofuranosyl side chains to the main chain, although the actual list of enzyme activities is much longer.

Regarding different  $\alpha$ -L-arabinofuranosidases (or Abfs), enzymes displaying this activity are quite widespread in the present CAZy classification system (<http://www.cazy.org>), although families GH51, 43 and 62 are the main families. Of these, GH51 is by far the largest family, with a majority of its members being of bacterial origin. Structural data obtained for several different members of the family confirm that all enzymes in GH51 possess a catalytic domain that displays ( $\beta/\alpha$ )<sub>8</sub>-barrel architecture, and additionally all bear a C-terminal  $\beta$ -sandwich domain of unknown function [4]. GH51 enzymes operate through a double displacement mechanism that leads to the overall retention of the anomeric configuration of the carbon at the scissile bond [5] and involves two conserved glutamic acid residues [6]. Regarding the finer structural details of GH51 enzymes, work on *Thermobacillus xylanilyticus* (TxAbf, UNIPROT O69262, PDB 2VRK and 2VRQ) has revealed two putative active site configurations. In the so-called open configuration, the  $\beta$ 2 $\alpha$ 2 loop bearing Trp99 is distant from the active site. However, in the closed configuration the same loop closes down onto and forms the active site, with Trp99 contributing to the topology of the -1 subsite that accommodates non-reducing L-arabinosyl moieties.

So far, the majority of characterized GH51 family members have been shown to be  $\alpha$ -L-arabinofuranosidases (EC 3.2.1.55), although a few display endoglucanase (EC 3.2.1.4) activity [7]. Previously, Hövel et al. [8] rationalized

the co-occurrence of these two activities within GH51 by considering the structural similarities between a GH51  $\alpha$ -L-arabinofuranosidase and a GH5 endoglucanase and specifically showing how D-glucosyl moieties can conceivably be accommodated in the -1 and +1 subsites of the GH51 enzyme. In addition to these two substrate specificities, studies have revealed that GH51 Abfs can often hydrolyze the glycosidic bond in aryl-glycosides, such as *para*-nitrophenyl- $\beta$ -D-xylopyranoside or (*p*NP-Xylp) or *para*-nitrophenyl- $\beta$ -D-galactofuranoside (*p*NP-Galf), the explanation for this flexibility being found in the stereochemical relatedness of  $\alpha$ -L-arabinosyl,  $\beta$ -D-xylopyranosyl and  $\beta$ -D-galactofuranosyl moieties [8, 9]. Nevertheless, apart from rare examples [10], these activities are often minor and are not evidenced when GH51 Abfs are used to hydrolyze heteroxylan or xylo-oligosaccharides [11], unlike the clearly bifunctional  $\alpha$ -L-arabinofuranosidases/ $\beta$ -D-xylopyranosidases belonging to other GH families, such as GH3 and 43 [12, 13].

Conceptually, bifunctional enzymes are attractive, because in the case of complex polymer hydrolysis, the two activities are likely to be complementary and catalyze consecutive reactions [14]. In the context of heteroxylan hydrolysis, it is easy to imagine that bifunctional systems of the type  $\alpha$ -L-arabinofuranosidase/ $\beta$ -D-xylopyranosidases or  $\alpha$ -L-arabinofuranosidase/ $\beta$ -D-xylanase would be particularly useful and, indeed, the literature abounds with examples of the former type. However, only two examples of naturally occurring  $\alpha$ -L-arabinofuranosidase/ $\beta$ -D-xylanases have been reported. The Xln23 from *Streptomyces chattanoogensis* is bimodular, containing two catalytic domains; thus the activities are displayed by two distinct active sites [15]. Similarly, artificial enzymes, such as the chimeras designed by Fan et al. [16, 17] are also bimodular. On the other hand, a recently reported GH51 arabinofuranosidase from *Alicyclobacillus* sp. A4 appears to possess a single catalytic domain that is responsible for both  $\alpha$ -L-arabinofuranosidase and  $\beta$ -D-xylanase activities, although no xylosidase activity was evidenced [18]. This novel finding is intriguing, because it is not easy to understand how the active site topology of typical GH51 arabinofuranosidases can accommodate heteroxylan in a productive manner leading to the hydrolysis of  $\beta$ -(1-4) bonds linking main-chain D-xylosyl moieties.

From a technological standpoint, Abfs,  $\beta$ -D-xylopyranosidases and  $\beta$ -D-xylanases are useful for a wide variety of commercial applications, including the pre-hydrolysis of non-digestible fibers in animal feed and the production of sugar syrups from lignocellulosic biomass, being particularly vital when targeting xylan-rich agricultural co-products, such as brans from corn and wheat, and hardwoods. Therefore, it is easy to appreciate that enzymes displaying several xylanolytic activities

could be useful to reduce the complexity and the cost of enzyme cocktails [19] or, alternatively, to limit the number of coding sequences needed to construct designer xylanolytic microorganisms [20].

Consistent with the recent report by Yang et al. [18], we describe herein another multifunctional GH51 arabinofuranosidase that is produced by a bacterium isolated from a Tunisian hot spring. The detailed study of this apparently trifunctional enzyme has provided us with some clues as to how all three substrate specificities can be defined by a single catalytic site.

## Methods

### General chemicals and reagents

Most chemicals and reagents used in enzyme assays were purchased from Sigma-Aldrich Chimie S.a.r.l. (Lyon, France) unless stated otherwise. However, *pNP-Galf* was prepared in-house using a published protocol [9]. Polysaccharides and XOS for calibrating HPAEC experiments were purchased from Megazyme International (Wicklow, Ireland), although the monosaccharides L-arabinose, D-xylose and oat spelt xylan were purchased from Sigma-Aldrich. Destarched wheat bran was a gift from ARD (Pomacle, France).

### Genomic DNA preparation

A moderately thermophilic strain was isolated from a soil sample collected in south Tunisia (Gafsa region) using a batchwise enrichment procedure. Briefly, bacteria were grown in aerobic conditions at 55 °C with shaking in a previously described XE medium [21, 22], which contains 0.3 g/L  $\text{KH}_2\text{PO}_4$ , 0.6 g/L NaCl, 12 g/L  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ , 0.08 g/L  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ , 1 g/L  $\text{NH}_4\text{Cl}$ , 2 g/L yeast extract, containing 1 % v/v vitamin and 1 % v/v minerals and 5 g/L oat spelt xylan as the sole carbon source. Genomic DNA (gDNA) was isolated from a fresh culture of this strain, essentially using the method described by Harwood and Cutting [23]. After thorough extraction using an aqueous phenol/chloroform/isoamyl alcohol (50:48:2 v/v) solution saturated with 10 mM Tris, 1 mM EDTA, pH 8.0, the gDNA was precipitated using isopropanol and recovered using a glass Pasteur pipette. While attached to the pipette, the gDNA was washed using 70 % (v/v) ethanol and then dissolved in 10 mM Tris, 1 mM EDTA, pH 8.0.

### 16s rDNA cloning and sequencing

Using the gDNA as template and two universal 16s rDNA primers, E8F and E154R (see below), the 16s rDNA of the bacterial isolate was amplified by PCR, employing the high-fidelity Pfu Turbo DNA polymerase (Promega, WI, USA). The amplicon was purified using the QIAquick PCR purification kit (Qiagen S.A., Courtaboeuf, France), ligated to the linear, blunt-ended vector

pST1Blue (Novagen, Merck, Darmstadt, Germany) and the reaction mixture was used to transform *E. coli* XL1Blue. After isolation of positive clones, the recombinant plasmid containing the 16s rDNA fragment was amplified and prepared as plasmid DNA using standard methods. This DNA was submitted for sequencing, which was performed by MWG Biotech (Germany). Finally, the sequence data (GENBANK accession number AM283040) was compared to all of the type strain accessions in the ribosomal database (<http://www.rdp.cme.msu.edu/>). Oligonucleotide primers used to isolate the 16s rDNA of the bacterial isolate were:

E8F 5' AGAGTTTGATCCTGGCTCAG 3',

E1541R 5' AAGGAGGTGATCCANCCRCA 3'.

### Isolation and cloning of the gene encoding THSAbf

Genomic DNA was digested in individual reactions using four different restriction endonucleases (*HaeIII*, *RsaI*, *AluI* et *DpnI*) that generate blunt DNA fragments, according to standard molecular biology protocols [24]. Following digestion, DNA fragments were size separated by electrophoresis on an agarose gel, employing low melting point agarose (0.8 % w/v). DNA fragments in the size range 2–7 kb were recovered by excising the relevant zones of the gel and fragments were purified using the Qiaquick gel extraction kit (Qiagen, Courtaboeuf, France), before being ligated to the *EcoRV*-linearized vector pSMART-LCKan. After an appropriate incubation period, aliquots of the ligation mixture were used to transform *E. coli* cloni<sup>®</sup> 10G *Escherichia coli* chemically competent cells (Lucigen Corp, Middleton, WI, USA), which were subsequently cultured on solid LB agar medium containing kanamycin (30  $\mu\text{g mL}^{-1}$ ) and 5-bromo-4-chloro-3-indolyl- $\beta$ -D-xylopyranoside (X-Xyl, Glycosynth Ltd, Cheshire, UK). The latter (40  $\mu\text{L}$  aliquot of a solution of 20 mg/mL in DMSO) was applied to the surface of the solid medium to enable the selection of clones expressing  $\beta$ -D-xylopyranosidase activity. The sequence of one positive clone encoding the family GH51 Abf, designated THSAbf, was amplified by PCR using the primers below and inserted between the *NdeI* and *EcoRI* sites of the expression vector pET21a (+), thus creating the plasmid pET21-THSAbf that was used to transform the *E. coli* strain JM109 DE3. This plasmid encodes a recombinant version of THSAbf bearing a C-terminal (His)<sub>6</sub> tag.

FwAbf (*NdeI* underlined) 5' GGATCAGCATATGTCATCAACAGCACCGCG 3',

RevAbf (*EcoRI* underlined) 5' CCATGAATTCGCAGCTCTCGCTTGACCCAG 3'.

### Creation of the mutant E296A

To create an inactive variant of THSAbf, site-directed mutagenesis was performed to replace the amino acid

E296 (predicted catalytic nucleophile) by an alanine. This was achieved using the QuickChange II site-directed mutagenesis kit (Agilent Technologies, Les Ulis, France) and forward and reverse mutagenic primers:

F-E296A 5' GTACCTCAGCTTCGACGCGTGGAAACG TTTGGTAC 3',

R-E296A 5' GTACCAAACGTTCCACGCGTCAAG CTGAGGTAC 3'.

Mutagenesis was performed using a thermocycler programmed to perform 16 cycles of annealing (95 °C, 30 s and 55 °C, 1 min) and primer elongation at 68 °C for 5 min. Afterward, the mutated plasmid was used to transform *E. coli* strain JM109 DE3. The mutant enzyme E296A was expressed and purified as described below.

### Protein expression

Protein expression was conducted using standard procedures and single colonies were grown at 37 °C in liquid LB medium containing ampicillin (100 µg mL<sup>-1</sup>). A non-standard procedure for pT7-driven expression was used to express THSAbf, since no IPTG was used to induce protein expression. Briefly, a pre-culture established using a single colony *E. coli* strain JM109 DE3 bearing pET21-THSAbf was grown for 16 h at 37 °C with shaking (140 rpm) in LB medium containing ampicillin. This culture was then diluted (1:100 v/v) with fresh LB/ampicillin medium and grown at 37 °C until the OD<sub>600nm</sub> of the culture reached 0.5 absorbance unit (generally 4 h). At this point, the culture was cooled on ice and then submitted to centrifugation (5000×g, 12 min, 4 °C), thus recovering the cells for protein purification.

### Purification of THSAbf

To purify recombinant THSAbf, bacterial cells were suspended in 20 mL of Talon<sup>®</sup> buffer (Tris–HCl 20 mM, NaCl 300 mM, pH 8) and frozen at –20 °C. After defrosting on ice, the cell suspension was sonicated (0.5 s pulse followed by 0.5 s pause, during a total of 4 min using a ¾-inch probe operating at 30 % amplitude), maintaining the cells on ice throughout. Solid cellular debris was separated from soluble cellular extract by centrifugation (20 min at 15,000×g, 4 °C), with the latter being filtered (0.45 µm Minisart syringe filter, Sartorius) before being applied to a gravity flow column containing 3 mL of Talon<sup>®</sup> metal affinity resin (Clontech, La Jolla, CA, USA). The column was washed successively with 20 mL of Talon<sup>®</sup> buffer, 7.5 mL of Talon<sup>®</sup> containing 5 mM imidazole, pH 8, and 7.5 mL of Talon<sup>®</sup> containing 10 mM imidazole, pH 8. Finally, THSAbf was eluted in 15 mL (applied in 7.5 mL aliquots) of Talon<sup>®</sup> containing 100 mM imidazole, pH 8, and the collected fractions were dialyzed against 250 volumes of 50 mM sodium phosphate (2 successive baths), pH 7, before being filtered using a

sterile Minisart Sartorius filter (0.22 µm) and stored at 4 °C. The final concentration of the THSAbf solutions was determined by measuring the absorbance at 280 nm and applying the Lambert–Beer relationship, using theoretical values for Mw (57,223 Da) and molar extinction coefficient of  $\epsilon = 112,230 \text{ M}^{-1} \text{ cm}^{-1}$ .

### Bioinformatics and structural modeling of THSAbf–ligand complexes

Routine DNA sequence analysis was performed using Bioedit 7.2.5, while the analysis of 16s rRNA was performed using SILVA, the rRNA database project (<http://www.arb-silva.de/>). To perform phylogenetic analysis, a set of complete sequences was extracted from family GH51 in the CAZy database (<http://www.cazy.org/GH51.html>), and then redundant sequences were removed using CD-HIT, fixing the sequence identity cutoff at 1.0. Subsequently, 984 non-redundant sequences were submitted to BLASTCLUST (part of the Bioinformatics Toolkit of the Max Planck Institute for Developmental Biology at <http://www.toolkit.tuebingen.mpg.de/blast-clust/>), using 80 % sequence identity and 70 % sequence coverage as run settings. Sequences representing 577 clusters were then aligned using CLUSTAL OMEGA (<http://www.toolkit.tuebingen.mpg.de/clustalw>) and the phylogenetic tree was plotted using Figtree 1.4.2 software (<http://www.tree.bio.ed.ac.uk/>). The presence of a signal peptide in THSAbf was checked using the SignalP4.1 server (<http://www.cbs.dtu.dk/services/SignalP/>), entering UNIPROT sequence B1A0T7 in Fasta format as the query.

For protein modeling, a structural model of THSAbf (GENBANK ABZ10760) was generated by homology modeling using M4T, which selects and optimally combines multiple template structures benefitting from the unique contribution of each template [25]. Selected templates were the GH51  $\alpha$ -L-arabinofuranosidases from *Geobacillus stearothermophilus* [8] (PDB code 1qw9) and *Thermobacillus xylanilyticus* [4] (PDB code 2vrq). The quality and stereochemistry of the models were assessed using ProSA-II (<http://www.prosa.services.came.sbg.ac.at/prosa.php>) and PROCHECK (<http://www.ebi.ac.uk/thornton-srv/software/PROCHECK/>), respectively. Prior to ligand docking, an array of representative protein conformations defining the conformational space of THSAbf was generated using CONCOORD ([http://www3.mpibpc.mpg.de/groups/de\\_groot/conCOORD/](http://www3.mpibpc.mpg.de/groups/de_groot/conCOORD/)) and GROMACS (<http://www.gromacs.org/>) as follows. Using the structural model of THSAbf as the starting conformation, geometrical constraints were defined using the van der Waals parameters from the OPLS-AA force field and CONCOORD's default parameters for atomic bonds and angles. However, chirality was assessed 'on the fly',

and the 'bump check' option was activated to avoid steric clashes. A total of 500 conformations were generated using the DISCO application in CONCOORD, adjusting the number of trials per run and iterations per structure to 100 and 4000, respectively. In the final stage, structural clustering of the conformers was performed using the *g\_cluster* application in GROMACS implementing the *gromos* clustering algorithm and a root mean square deviation cutoff of 2 Å. This resulted in 25 clusters, whose centroids were used as representative conformations. The final structural models of the different substrates-bound THSAbf were derived using AutoDock Vina (ADV1.1.2, <http://www.vina.scripps.edu/>) as follows. The structures of D-xylose, xylobiose and xylotriose, and the branched pentasaccharide XA<sup>3</sup>XX and xylopentose were obtained from X-ray structures deposited in the PDB databank (1px8, 1gor, 1b30, 2vrq and 3wn2). To adapt these to the requirements of ADV, polar hydrogen atoms and Gasteiger–Marsili charges were added to the ligands using AutoDock tools implemented in MGL Tools version 1.5.6 (<http://www.mgltools.scripps.edu/>), before docking them onto all modeled conformers of THSAbf using a docking grid box centered on the two catalytic glutamic acid residues (E177, and E296) with a size of 45 × 40 × 60 Å<sup>3</sup>. The docking poses were ranked by their predicted binding energy (as per ADV implementation) and the top ten poses were considered for each pair THSAbf–ligand representative conformers. Finally, poses were visually inspected and selected based on the correct and putative productive geometry between catalytic residues and substrates.

#### Enzyme activity measurements

Routinely, activity measurements were performed in 50 mM sodium phosphate buffer, pH 6.5, containing BSA (1 mg/mL) and *p*NP-Araf (4 mM final concentration). Before adding THSAbf, the mixture was pre-incubated for 5 min at 60 °C in a water bath, and then 100 µL of THSAbf solution was added and incubation pursued. The total volume of the reaction mix was 1 mL. Aliquots (100 µL) of the reaction mixture were removed at regular intervals over a 14-min incubation period and immediately added to 500 µL of cold (0 °C) 1 M Na<sub>2</sub>CO<sub>3</sub>. To read absorbance values (at 401 nm), a 200 µL aliquot of the mixture was transferred to a microtiter plate that was then placed in a Sunrise UV/Vis microspectrophotometer (Tecan Ltd, Männedorf, Switzerland). Reaction rates, measured in terms of µmoles released per min, were determined by comparing absorbance values with a standard curve prepared using known concentrations of *p*NP-OH (over the range 0 to 1 mM) and 1 U of activity was defined as the quantity of THSAbf required to release 1 µmol *p*NP-OH per min.

To determine the effect of pH on THSAbf activity, the above method was performed at 37 °C using *p*NP-Araf and *p*NP-Xylp as substrates, or the DNS method was employed to measure the release of reducing sugar from LVWAX. Reactions were performed at different pH values using either sodium citrate (pH 3–6), sodium phosphate (pH 6–8), or citrate–phosphate (pH 5.5–6.5) buffer to cover the whole range from pH 3 to 8. Four independent measurements were performed for each pH value and standard deviations were calculated. Similarly, to determine the effect of different temperatures on activity, measurements were performed as described, but covering a range of temperatures, from 30 to 85 °C, collecting four independent measurements for each temperature and calculating standard deviations. Regarding enzyme thermostability, this was determined by incubating samples of THSAbf in 50 mM sodium phosphate, pH 6.5, containing BSA (1 mg/mL) at different temperatures (50, 60, 65, 70 °C). Aliquots (100 µL) were removed at regular intervals and activity was assayed as described above (at 60 °C). Thermostability measurements were performed in duplicate.

#### Determination of kinetic parameters

Fourteen different substrates, including *p*NP-α-L-arabinofuranoside, *p*NP-Araf; *p*NP-β-D-xylopyranoside (*p*NP-Xylp), *p*NP-β-D-galactofuranoside (*p*NP-Galf), *p*NP-β-D-galactopyranoside, *p*NP-β-D-glucopyranoside, *p*NP-α-L-glucopyranoside, *p*NP-β-D-fucopyranoside, *p*NP-α-L-fucopyranoside, *p*NP-β-L-fucopyranoside, *p*NP-α-D-mannopyranoside, *p*NP-β-D-mannopyranoside, *p*NP-α-L-arabinopyranoside, *p*NP-β-L-arabinopyranoside, and *p*NP-α-L-rhamnopyranoside, were used to perform the initial tests. Based on the results, the Michaelis–Menten parameters,  $k_{cat}$  and  $K_M$ , were determined for the hydrolysis of three different *p*NP-Araf, *p*NP-Xylp and *p*NP-Galf by THSAbf. All reactions were performed in triplicate at 60 °C in 50 mM sodium phosphate, pH 6.5, containing BSA (1 mg/mL). To investigate the relationship between substrate concentration, [S] and reaction velocity,  $v$ , the following substrate concentrations were employed: *p*NP-Araf, 0–5 mM; *p*NP-Xylp, 0–40 mM; *p*NP-Galf, 0–14 mM. Afterward, data were analyzed using the enzyme kinetics module of Sigma Plot 10.0 software and the Michaelis–Menten model.

#### Assessing the activity of THSAbf on complex substrates

The activity of THSAbf was measured on several polymeric substrates. These included branched sugar beet and linear arabinans (Megazyme), and xylans from birch and beechwood. Reactions contained the substrate (10 mg/mL) dissolved, or partially suspended, in 900 µL of 50 mM sodium phosphate, pH 6.5, containing

BSA (1 mg/mL). THSAbf-mediated hydrolysis was performed at 60 °C and the time-dependant progress of the reaction was assessed in a discontinuous assay, in which aliquots (100 µL) were removed at regular intervals to measure the quantity of soluble reducing sugars using the DNS method [26]. To account for different substrate specificities, the number of units (defined based on the release of *p*NP-OH from *p*NP-Araf) of THSAbf used in the reactions was different (0.85 and 1.7 U for birch and beechwood xylans, respectively, and 400 and 1200 U for the branched and linear arabinans, respectively). Results, obtained by measuring absorbance at 540 nm, were compared to standard curves prepared using either D-xylose or L-arabinose (solubilized in reaction buffer containing complex substrate at the initial concentration). One unit of THSAbf activity was defined as the quantity of enzyme required to release 1 µmol arabinose or xylose equivalent per minute. To determine the kinetic parameters,  $k_{cat}$  and  $K_{M(app)}$  for THSAbf-catalyzed hydrolysis of arabinans, substrate concentrations were varied, either in the range 10–140 mg/mL (sugar beet arabinan) or in the range 2–19 mg/mL (linear debranched arabinan). The low solubility of linear arabinan required the use of a Thermomixer Comfort (Eppendorf, Hamburg, Germany) operating at 60 °C and 1500 rpm.

#### Hydrolysis of wheat arabinoxylan and bran using hemicellulolytic cocktails

First, to study the action of THSAbf and other enzymes on low viscosity wheat arabinoxylan, or LWAX (Megazyme) and destarched wheat bran, or DWB (ARD, Pomacle, France), the monosaccharide composition of these substrates was determined using acid hydrolysis, followed by separation and quantification of the monosaccharides by HPAEC-PAD, using a DIONEX system (Fisher Scientific, Strasburg, France) equipped with a CarboPac SA10 column and pre-column equilibrated in an aqueous solution of 1 mM sodium hydroxide. Analyses were performed in triplicate ( $n = 3$ ), over a 30-min period, using a flow rate of 1.2 mL min<sup>-1</sup> and an eluent composed of an aqueous solution of 100 mM sodium hydroxide. Quantitative data analysis was performed referring to linear concentration–response reference curves prepared using the same column and five different monosaccharide solutions in the range 0 to 5 mg L<sup>-1</sup> (Table 1).

Afterward, the action of THSAbf on LWAX and DWB was compared with that of the GH51 Abf and the GH11 β-D-xylanase from *Thermobacillus xylanilyticus* (*Tx*Abf and *Tx*Xyn, respectively) and the possible synergy between THSAbf and these enzymes was investigated. To achieve this, eight reactions were set up for each substrate, mixing a substrate suspension (in 50 mM sodium phosphate, pH 6.5,

containing 1 mg/mL BSA and pre-incubated at 60 °C) with the following: (1) water (control), (2) THSAbf, (3) *Tx*Xyn, (4) *Tx*Abf, (5) THSAbf + *Tx*Xyn, (6) THSAbf + *Tx*Abf, (7) *Tx*Abf + *Tx*Xyn and finally (8) THSAbf + *Tx*Xyn + *Tx*Abf. Regarding enzyme concentrations, THS and *Tx*Abfs were used at 15 U/mL (activity determined on *p*NP-Araf), while *Tx*Xyn was used at 6.25 U/mL (activity determined on birchwood xylan). To conduct the reactions, substrates (20 mg) were thoroughly rehydrated in sodium phosphate, pH 6.5, for 16 h at 60 °C under stirring, then the enzyme was added. Aliquots were removed at several intervals during a 48 h period and the supernatants were immediately isolated by centrifugation at 20,000×g and 4 °C for 5 min, and then stored frozen (−20 °C) while awaiting analysis. To identify and quantify the reaction products, samples were injected onto a CarboPac PA100 column running on an HPAEC-PAD system (Dionex). Briefly, having equilibrated the column in 150 mM sodium hydroxide containing sodium acetate (5 mM), elution of the different reaction products was achieved using a gradient of increasing sodium acetate concentration (5–85 mM over 30 min, then 500 mM for 3 min, followed by a return to 5 mM for 17 min). To calibrate the analysis, internal standards (D-xylose, L-arabinose, xylobiose, xylotriose, xylo-tetraose, xylopentaose, xylohexaose) were injected at concentrations ranging from 0 to 25 mg L<sup>-1</sup>. Results were analyzed using Chromeleon 6.8 software to perform peak integration of the chromatogram.

## Results and discussion

### Strain isolation, gene discovery, cloning, and analysis

A microbial collection campaign performed in Tunisia in 2004 allowed the isolation of a bacterial species from a soil sample (collected in the region of Gafsa) on the

**Table 1 A. Monosaccharide composition of LWAX and B. monosaccharide composition of DWB**

	Quantity mmol/g	% wt	Xyl/Ara ratio
A. Destarched wheat bran			
Xylose	1.95 ± 0.2	47.21	1.74
Arabinose	1.12 ± 0.22	27.04	
Glucose	0.07 ± 0.01	1.79	
Galactose	0.96 ± 0.16	23.33	
Mannose	0.03 ± 0.01	0.64	
B. Low-viscosity wheat arabinoxylan			
Xylose	2.89 ± 0.17	57.28	1.31
Arabinose	2.20 ± 0.10	42.02	
Glucose	0.02	0.32	
Galactose	0.02	0.38	
Mannose	ND	0	

$n = 3$  and SD values are shown

basis of its ability to grow at 55 °C in the culture medium. Sequencing of the rDNA of the bacterial isolate and comparison with entries in the SILVA ribosomal database revealed that it was closely related to several *Paenibacillus* strains, with *Paenibacillus* sp. MK17 (EF173331) being its nearest neighbor. Accordingly, the isolate was designated *Paenibacillus* sp. *HanTHS1*.

The creation and subsequent screening of a clone library constructed using *Paenibacillus* sp. *HanTHS1* genomic DNA revealed seven individual isolates that formed a blue color at the colony level in the presence of X-Xyl. Plasmid extraction from one of these and DNA sequence analysis revealed the presence of a DNA insert composed of 4464 bp. Further analysis disclosed three ORFs, one of which (1521 bp) was identified by BLASTX (<http://www.blast.ncbi.nlm.nih.gov/Blast.cgi>) as a family 51 glycoside hydrolase (<http://www.cazy.org/GH51>). The sequence of this ORF, designated *abfA*, was deposited in GENBANK (accession number ABZ10760). Further analysis of the BLASTX results revealed that the amino acid sequence encoded by *abfA* is 100 % identical to that of the Abf (GH51) from *Paenibacillus* sp. oral taxon 786 str D14 (UniProt C6IZP3) and displays >80 % homology with GH51 Abfs from *Paenibacillus lactis* 154 (G4HDR3), *Paenibacillus* sp. HGF5 (F3M674), *Paenibacillus* sp. Y412MC10 (D3ELR9), and *Paenibacillus vortex* V453 (E5YTZ2), results that further support the conclusion that the Tunisian bacterial isolate belongs to the genus *Paenibacillus*. Moreover, analysis of the translated *abfA* sequence using the SignalP 4.1 server failed to reveal the presence of an N-terminal signal peptide, indicating that the native protein is possibly located in the cytoplasm.

Phylogenetic analysis of non-redundant sequences in family GH51, including that of THSAbf, revealed an unrooted tree in which Abfs of fungal, plant and viral origin are clustered into three relatively well-defined regions. Moreover, the GH51 members that are annotated in the CAZy database as cellulases (EC 3.2.1.4) are localized in a part of the tree that is very distinct from that harboring THSAbf (Additional file 1). Indeed, the latter forms part of a different subgroup that contains the Abf from *Paenibacillus* sp. Y412MC10 (D3ELR9) and the sequence of the well-characterized *GsAbf* (B3EYN4).

Cloning of *abfA* in pET21 under the control of the T7 promoter facilitated the expression of recombinant THSAbf in *E. coli*. Unusually, THSAbf could be expressed without induction by IPTG, even though no lactose or other analog was present in the culture medium. This advantageous feature was not further investigated, but it is possible that auto-induction is of the leakiness of the Lac-UV5 lac promoter. Working in this way, the average final pure protein yields were in the range of

20–25 mg L<sup>-1</sup> culture, whereas IPTG-induced protein production led to approximately twofold lower yields.

#### Activity of THSAbf and E296A on aryl-monosaccharides and polymeric substrates

Rapid screening of THSAbf activity, using 14 different aryl-monosaccharides, representing both furanose and pyranose conformations and sugars from D- and L-series, confirmed that THSAbf is specific for sugars displaying related conformations, notably *pNP-Araf*, *pNP-Galf* and *pNP-Xylp*. THSAbf was most active on *pNP-Araf*, with its specific activity on this substrate being 27- and 2300-fold higher than that on *pNP-Galf* and *pNP-Xylp*, respectively (Table 2). These results reveal that like the majority of GH51 glycoside hydrolases [27], THSAbf can best be described as an Abf and that the subsite -1 has a clear preference for the furanose conformation. Further testing on polymeric substrates revealed that, like the recently described Abfs from *Cellulomonas fimi* ATCC 484 and *Alicyclobacillus* sp. A4 [18, 28], THSAbf could release reducing sugars from a variety of polysaccharides, including branched and linear sugar beet arabinans (SA of 1.1 ± 0.24 and 1.8 ± 0.1 U mg<sup>-1</sup>, respectively), and xylans from birch and beechwood (SA of 10.8 ± 0.1 and 14.5 ± 1.3 U mg<sup>-1</sup> respectively). While it was unsurprising to detect activities on sugar beet arabinan ( $k_{\text{cat}}/K_M$  of 0.194 s<sup>-1</sup> mg<sup>-1</sup> mL), which bears L-arabinosyl ramifications, it was more surprising to detect activity on linear debranched arabinan ( $k_{\text{cat}}/K_M$  of 0.196 s<sup>-1</sup> mg<sup>-1</sup> mL), because this suggests that THSAbf can hydrolyze both α-1,3 bonds linking side-chain arabinosyl moieties and α-1,5 bonds linking main-chain arabinosyl moieties. Moreover, it is noteworthy that the detection of activity on 4-O-methylglucuronoxylans (i.e., birch and beechwood xylans) was unexpected for a GH51 Abf, since these polysaccharides contain very few α-L-linked arabinosyl moieties. Significantly, unlike wild-type THSAbf, the mutant E296A displayed only very weak (approximately, 1600-fold lower than the wild-type enzyme) on *pNP-Araf* and displayed no measurable activity on *pNP-Xylp* and beechwood xylan, consistent with the fact that the mutant is crippled by the lack of a catalytic nucleophile

**Table 2 Kinetics of hydrolysis of aryl-monosaccharides by THSAbf**

Substrate	$K_M$ (mM)	$k_{\text{cat}}$ (s <sup>-1</sup> )	$k_{\text{cat}}/K_M$ (s <sup>-1</sup> mM <sup>-1</sup> )	SA (U mg <sup>-1</sup> )
<i>pNP-Araf</i>	0.31 ± 0.1	328 ± 23	1050	426 ± 36
<i>pNP-Xylp</i>	23.8 ± 4.1 <sup>a</sup>	10.6 ± 0.9	0.44	1.70 ± 5
<i>pNP-Galf</i>	8.6 ± 0.9 <sup>a</sup>	327 ± 20	38	103 ± 2

<sup>a</sup> These values are subject to caution since the maximum substrate concentration used in each case was less than 2 ×  $K_M$ ,  $n = 3$  and SD values are shown



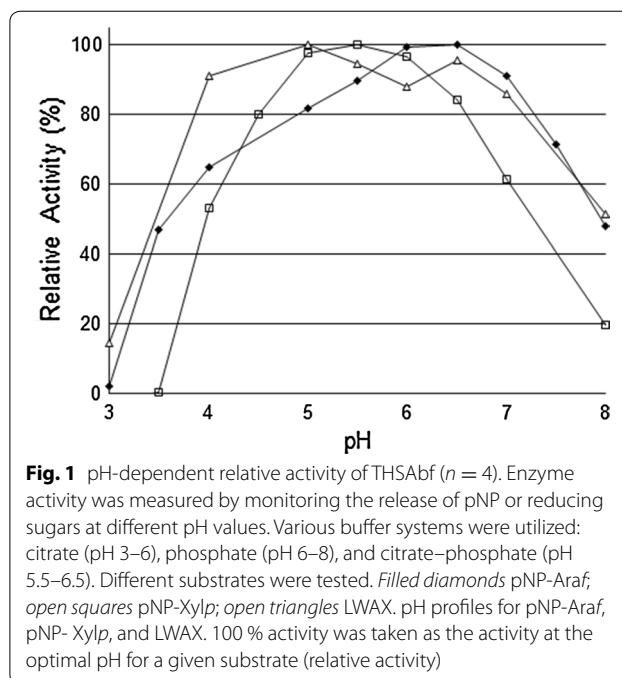
and that all of the activities attributed herein to THSAbf described are defined by a single active site.

In the light of the above results, it is worth recalling that THSAbf was first revealed by functional screening using a 5-bromo-4-chloro-3-indolyl- $\beta$ -D-xylopyranoside and that the subsequent discovery of an Abf appeared to be anomalous. However, the fact that THSAbf appears to be multi-functional rationalizes this preliminary finding. The discovery that THSAbf displays an endo-mode of action, hydrolyzing internal  $\beta$ -1,4 bonds between D-xylosyl subunits in heteroxylans and perhaps internal  $\alpha$ -1,5 bonds in arabinan is unusual, even if the results also reveal that THSAbf is not a potent endoxylanase. Indeed, when compared with GH11 *TxXyn* (approximately, 1500 U mg<sup>-1</sup> on birchwood), the specific activities of THSAbf on birch and beechwood heteroxylans (10–15 U mg<sup>-1</sup>) are quite low. While this indicates the intrinsic limit of describing THSAbf as an endoxylanase, the results described below indicate that it is rather a question of arabinose substitution levels.

#### Influence of pH and temperature on activity

The measurement of THSAbf activity on three different substrates (*p*NP-Araf, *p*NP-Xylp and LWAX) at different pH values revealed that activities are near maximal over a relatively wide pH range from approximately 4 to 7, but the actual pH curves are not identical. In particular, the bell-shaped curve representing the pH dependence of the reaction involving *p*NP-Xylp is narrower with maximal activity occurring in the range 5–6, centered on pH 5.5 (Fig. 1), while the curves describing the pH dependence of reactions involving LWAX and *p*NP-Araf are broader, the latter displaying a rather irregular form. Overall, these results are unexpected and apparently inconsistent with the notion of a single active site in THSAbf. However, it is noteworthy that similar results have been described for other glycoside hydrolases [29, 30]. Moreover, considering that the binding of structurally different substrates within the active site of THSAbf probably induces topological modifications of the active site, it is likely that the distance separating the catalytic residues (nucleophile and acid/base) is altered. This is significant for the optimum pH for activity, because this mainly results from the distance-dependent Coulomb interaction between the catalytic residues and the interaction of these with neighboring residues that participate in an often complex protonic network.

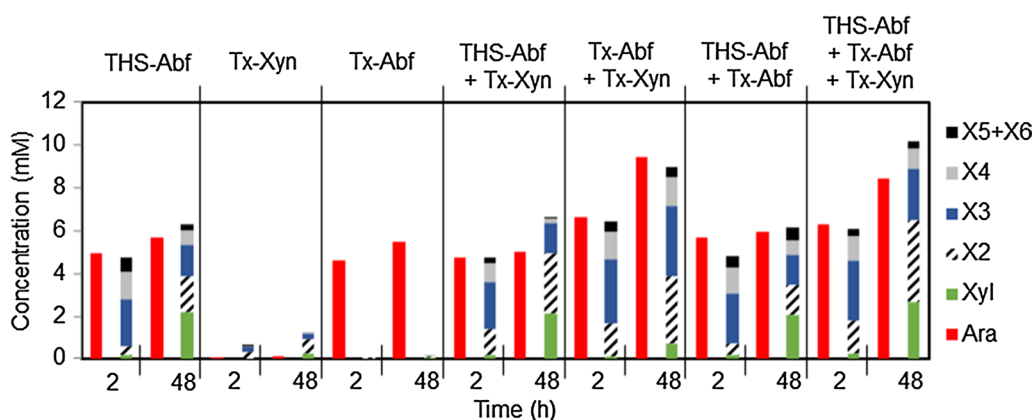
Studying the influence of temperature on THSAbf activity revealed that optimal activity was achieved at 75 °C, although our data also revealed that THSAbf is relatively unstable at this temperature (Additional file 2, A). At 60 °C, the stability of THSAbf was outstanding, since 100 % activity could still be measured after 24 h



(Additional file 2, B). Consequently, in subsequent studies, 60 °C was chosen as the preferred operational temperature for THSAbf.

#### Activity of THSAbf on isolated wheat arabinoxylan and wheat bran

Since THSAbf is thermophilic, the activity of this enzyme on complex substrates was studied alone and combined with two other thermostable enzymes *TxAbf* and *TxXyn*. Operating alone, THSAbf, *TxAbf* and *TxXyn*, clearly displayed hydrolytic activity on LVWAX (Fig. 2; Table 3; Additional file 3). Typical of an Abf, *TxAbf* only released arabinose as the reaction product (25.9 % w/w total arabinose), which is consistent with previous results for this enzyme [11]. Arabinose release by *TxAbf* was rapid, with almost maximal yields being achieved after only 2 h. Likewise, the hydrolytic behavior of *TxXyn* on LVWAX was consistent with the known properties of this enzyme, with xylobiose and xylotriose being the main reaction products, though only 8.9 % w/w total xylose was solubilized. Regarding THSAbf, its behavior on LVWAX was most untypical of a GH51 Abf, because in addition to the presence of soluble arabinose (26.8 % w/w total arabinose), analysis of the reaction products revealed soluble xylose (48.8 % w/w total xylose), which was mainly in the form of xylo-oligosaccharides (xylotriose, xylotetraose and xylobiose in that order). Analysis of the kinetics of product formation revealed that, like arabinose, xylotriose and xylotetraose were formed very rapidly (maximal yields achieved after 1 h), but xylobiose appeared more



**Fig. 2** Hydrolysis of LVWAX using single enzymes or enzyme combinations. Histograms were prepared using results from experiments performed in triplicate

**Table 3** Yield of soluble sugars obtained upon incubating LVWAX with different enzymes and enzyme combinations

Enzymes	% w/w yield soluble sugars		
	Arabinose	Xylose (mono) <sup>a</sup>	Xylose (equivalent) <sup>b</sup>
THSAbf	26.8	7.6	48.8
TxXyn	0.6	1.0	8.9
TxAbf	25.9	0.2	0.6
THSAbf + TxXyn	23.7	7.5	45.1
TxAbf + TxXyn	44.6	2.5	85.2
THSAbf + TxAbf	28.0	7.1	51.7
THSAbf + TxAbf + TxXyn	39.8	9.3	79.9
Control (no enzyme)	0.2	0.0	0.1

<sup>a</sup> Soluble monosaccharide

<sup>b</sup> Total soluble xylose;  $n = 3$

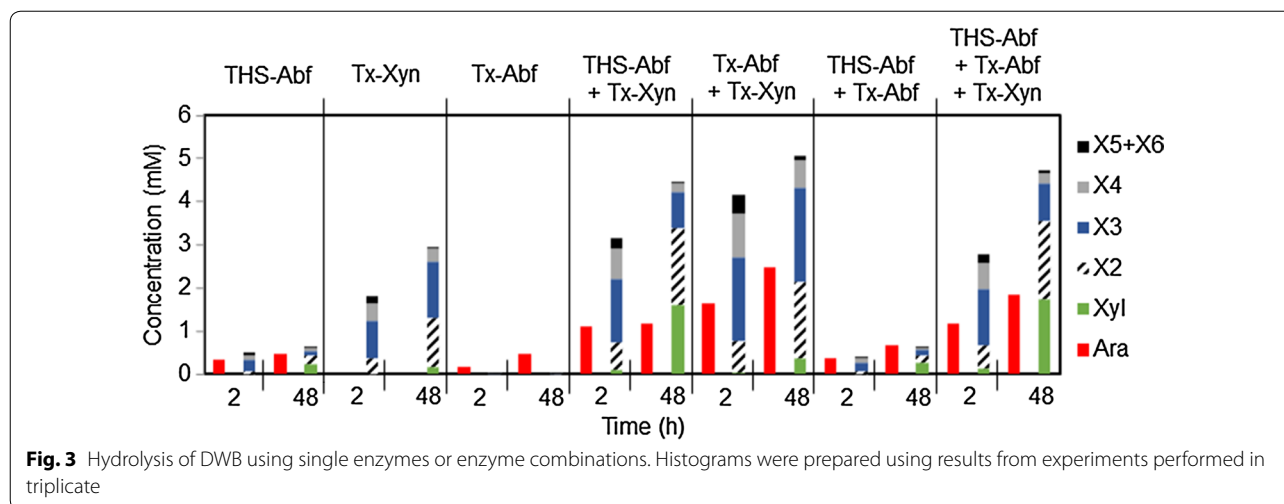
progressively, possibly being the result of the hydrolysis of higher DP xylo-oligosaccharides.

Combining the action of the enzymes (THSAbf, TxAbf and TxXyn), either pairwise or all together, revealed some logical catalytic interplay on LVWAX (Fig. 2; Additional file 4). For example, when TxAbf was deployed simultaneously with TxXyn, the yields of soluble arabinose and xylose were increased (44.6 % w/w total arabinose and 85.2 % w/w total xylose), with the quantity of soluble xylose equivalents being almost tenfold higher (Fig. 2; Table 3). In sharp contrast, when THSAbf was used with TxXyn, no benefits were evidenced, since the yields of soluble products were almost identical to those obtained using THSAbf alone. Similarly, using the two Abfs (THSAbf and TxAbf) together yielded results similar to those obtained with THSAbf alone, although the overall yield of soluble arabinose and xylose equivalents

was slightly higher. Finally, using all three enzymes simultaneously, quite high solubilization yields were procured (39.8 % w/w total arabinose and 79.9 % w/w total xylose), although these values were lower than those obtained using TxAbf and TxXyn together. Nevertheless, a significant increase in the yield of soluble monomeric xylose (9.3 % w/w total xylose) could be clearly attributed to the presence of THSAbf.

The treatment of DWB with TxAbf led to the release of arabinose as the sole product, which is consistent with the fact that this enzyme displays typical Abf activity (Fig. 3; Additional file 5). After 48 h, approximately 4.3 % (w/w) of arabinose had been released, slightly higher than previous data [11]. Regarding TxXyn, its impact on DWB was consistent with previous results [31], since approximately 40 % (w/w) xylose was released, mostly as xylobiose and xylotriose, with a smaller amount of xylose being evidenced. Xylotetraose and xylopentaose were also visible among the products, but logically no free arabinose was formed. Most importantly, the hydrolysis of DWB with THSAbf confirmed the previous results obtained on LVWAX. Compared to TxAbf, although THSAbf solubilized an almost identical amount of arabinose (4.3 % after 48 h), the reaction was much faster (3.2 % solubilized after only 2 h) and was accompanied by the equally fast release of xylotriose and xylotetraose (together representing 5.2 % w/w xylose) and the more progressive solubilization of xylose (1.3 % w/w), thus providing a clear indication that THSAbf possesses the ability to hydrolyze the  $\beta$ -1,4 bonds linking both internal and terminal D-xylosyl moieties.

Combining the action of THSAbf with that of TxXyn provided overall gains in the solubilization of arabinose and xylose from DWB (Fig. 3; Additional file 6). For example, the quantity of arabinose released into



the reaction supernatant was 2.5-fold greater than that obtained using THSAbf alone, and the amount of total xylose released was approximately 18 % higher, with 8.2 % w/w total xylose being released in monosaccharide form (Table 4). Nevertheless, combining *TxAbf* and *TxXyn* procured an even better performance, since 22 % w/w total arabinose and 69.6 % w/w total xylose were released in soluble form, although the amount of soluble xylose in monosaccharide form represented a little less than 2 % w/w total xylose. Predictably, the association of the two Abfs, THSAbf and *TxAbf*, for the treatment of DWB yielded a composite product profile, with approximately 6 % w/w total arabinose and 7 % w/w total xylose being recovered in soluble form, the latter represented by xylo-oligosaccharides and a small amount of xylose (1.5 % w/w total xylose). Finally, the simultaneous deployment of all

three enzymes did not procure added value when compared with the deployment of *TxAbf* and *TxXyn*. Indeed, the overall yields of soluble xylose and arabinose were lower (approximately, 47 % w/w total xylose and 16.5 % w/w total arabinose, respectively) when compared with the dual action of *TxAbf* and *TxXyn* (which solubilized approximately 69 % w/w total xylose and 22 % w/w total arabinose), though the soluble monosaccharide (xylose) production was significantly higher (8.93 % w/w, instead of 1.9 % w/w total xylose).

Previous work has shown that the heteroxylans in DWB are differentially organized within its component layers. Heteroxylans in the aleurone layer display moderate arabinose substitution (Xyl/Ara ratio of 2.07), while those in the middle and outer layers of DWB are characterized by lower Xyl/Ara ratios [32]. *TxXyn* very efficiently solubilizes the aleurone arabinoxylans, an observation that correlates well with this enzyme’s preference for heteroxylans displaying lower levels of arabinose substitution [31]. Inversely, taking into account the poor activity of THSAbf on birch and beechwood heteroxylans, presumably the activity of THSAbf on the aleurone layer arabinoxylans is also lower than that of *TxXyn*. A further indication of this clear distinction between endoxylanase activities of THSAbf and *TxXyn* is provided by the comparison of their action on LVWAX. On this substrate, THSAbf is clearly more efficient than *TxXyn*, which correlates well with the higher level of arabinose substitution (Xyl/Ara ratio of 1.3). Presumably, on LVWAX THSAbf benefits from the synergy between its arabinofuranosidase and endoxylanase activities. When considering the potential usefulness of THSAbf in xylanolytic cocktails, it is interesting to note that its interplay with other enzymes was not intuitive, in the sense that THSAbf appeared to compete with, rather than complement, the action of a GH11 xylanase. This

**Table 4** Yield of soluble sugars obtained incubating DWB with different enzymes and enzyme combinations

Enzymes	% w/w yield soluble sugars		
	Arabinose	Xylose (mono) <sup>a</sup>	Xylose (equivalent) <sup>b</sup>
THSAbf	4.29	1.32	6.47
<i>TxXyn</i>	0.00	0.98	39.68
<i>TxAbf</i>	4.31	0.04	0.46
THSAbf + <i>TxXyn</i>	10.74	8.24	44.73
<i>TxAbf</i> + <i>TxXyn</i>	22.14	1.92	69.58
THSAbf + <i>TxAbf</i>	5.95	1.52	6.94
THSAbf + <i>TxAbf</i> + <i>TxXyn</i>	16.53	8.93	47.25
Control (no enzyme)	0.18	0.08	0.34

<sup>a</sup> Soluble monosaccharide

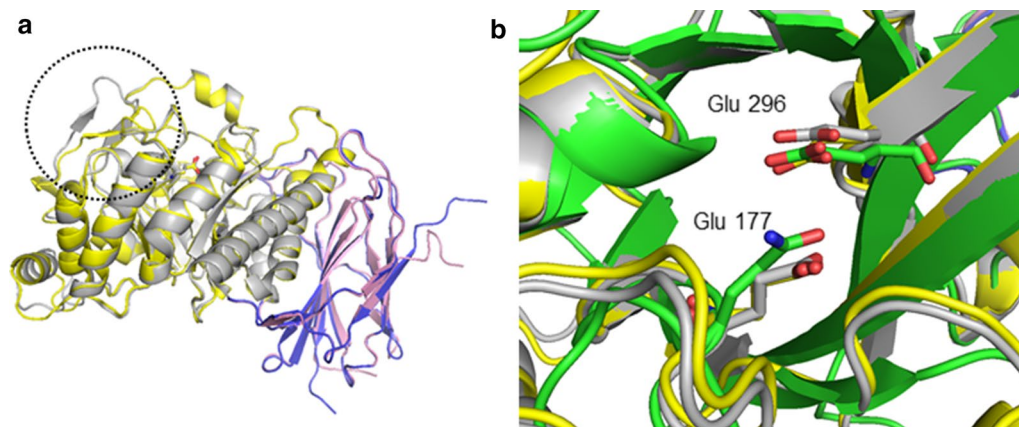
<sup>b</sup> Total soluble xylose

observation contrasts somewhat with the fact that *AlAbf51* was reported to complement the action of a GH10 xylanase. However, unlike GH10 xylanases, those from GH11 display rather narrow active site clefts (average 5.5 Å) [33] that preclude the binding of decorated D-xylosyl moieties in either subsite -1 or +1, and only L-arabinosyl (and not other substitutions) is generally tolerated in subsite -2. Therefore, the results presented in this study might imply that THSAbf and *TxXyn* vie for the undecorated regions of heteroxylans. In this respect, it is noteworthy that a similar competition phenomenon was reported between GH10 and GH11 xylanases, which led to a decrease in the release of XOS from wheat arabinoxylan [34]. In the case of THSAbf, primarily an Abf, this is more surprising because one would expect this enzyme to first act on available L-arabinosyl moieties. However, in our experiments, THSAbf was used in 20-fold molar excess when compared with *TxXyn*; thus in these conditions it is easy to appreciate why *TxXyn* was out-competed.

#### Protein modeling and investigation of different THSAbf–ligand complexes

Attempts to obtain high-quality X-ray diffraction data from THSAbf crystals have so far failed. Therefore, to gain insight into the possible structural determinants of the different activities displayed by THSAbf, structural models were generated and used for ligand docking. Two THSAbf models were built using comparative modeling and templates that display more than 70 % sequence identity with THSAbf, thus ensuring that high-quality structural models were obtained (i.e., Z-score of 6.9 using ProSA-II and absence of major stereo-chemical problems revealed by

PROCHECK). Inspection of the overall structure of the two THSAbf models reveals that like all other GH51 Abfs, THSAbf displays two domains, with the catalytic domain displaying  $(\beta/\alpha)_8$  architecture and the other a  $\beta$ -sandwich fold (Fig. 4a). Closer analysis of the two THSAbf models reveals that the principal differences between these are loop positions, notably the position of the  $\beta 2\alpha 2$  loop. In this respect, it is useful to recall that previous work on *TxAbf* [4, 35] has shown that the mobility of its  $\beta 2\alpha 2$  loop is likely to be an essential feature of catalysis. Nevertheless, when the  $\beta 2\alpha 2$  loop of THSAbf is in so-called open configuration, the overall position of the loop is dissimilar to that of *TxAbf* (2VRQ, chains C); residue Trp101 (borne on the  $\beta 2\alpha 2$  loop) is distant from the predicted subsite -1 and the overall active site topology of THSAbf resembles that of the wide *GsAbf* active site (Additional file 7). On the other hand, in the closed form, the  $\beta 2\alpha 2$  loop of THSAbf adopts a configuration that is similar to that of the so-called closed form of *TxAbf* (2VRQ, chains A and B), with Trp101 superposing well with its equivalent (Trp99) in *TxAbf* (Additional file 8). Nevertheless, the active site topology remains wider than that of *TxAbf*, because of the presence of a shortened  $\beta 6\alpha 6$  loop. Moreover, like *GsAbf*, THSAbf displays a long  $\beta 7\alpha 7$  loop that is positioned in much the same way as its analog in *GsAbf*. Likewise, superposition of the structures of *GsAbf* and *TxAbf* on the models of THSAbf reveals that the spatial position of the catalytic dyad (E177 and E296) coincides well with those of *GsAbf* (E175 and E294 in 1QW8,) and *TxAbf* (E176 and E298, not shown), and is localized just above a clearly defined cavity, which presumably constitutes the subsite -1 (Fig. 4b).



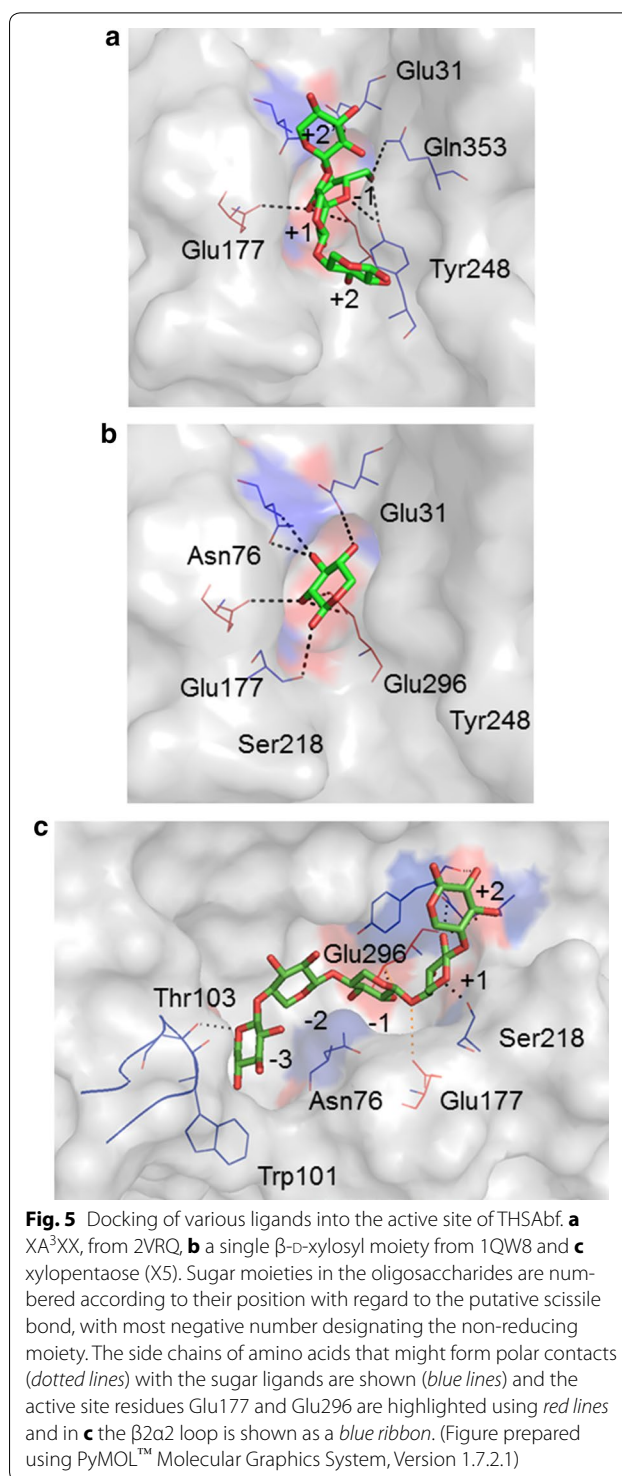
**Fig. 4** Modeled structures of THSAbf. **a** The overall two domain structure composed of a  $(\beta/\alpha)_8$  folded catalytic domain and a C-terminal domain displaying  $\beta$ -sandwich architecture. Two structural conformers of THSAbf (open gray and violet, closed yellow and pink) are superposed and the  $\beta 2\alpha 2$  loop in two different positions is encircled. **b** Zoom on the catalytic site of the two modeled THSAbf conformers (gray and yellow) and that of *GsAbf* (1QW8, green). The side chains of the two catalytic residues (Glu 177 and 296 in THSAbf) are shown as sticks. (Figure prepared using PyMOL™ Molecular Graphics System, Version 1.7.2.1)

The docking of different putative ligands, including  $XA^3XX$ , a single  $\beta$ -D-xylosyl moiety and xylopentaose (X5) revealed that it was possible to generate satisfactory models of ligand-bound enzymes harboring either an  $\alpha$ -L-arabinofuranosyl or a  $\beta$ -D-xylosyl moiety in subsite -1 (Fig. 5). In the case of the oligosaccharide  $XA^3XX$ , the  $\alpha$ -L-arabinofuranosyl moiety adopted a position highly similar to that already observed in ligand-bound *TxA*bF, with the scissile (O4-C1) bond being in a potentially productive position with respect to the two catalytic residues. Importantly, docking revealed that a single  $\beta$ -D-xylosyl moiety could also enter into subsite -1, with its O-5, O-2 and O-3 superposing well with those of a similarly bound  $\alpha$ -L-arabinofuranosyl moiety, a result that is consistent with previous findings for *GsAbf* and with the biochemical data presented here. Even more satisfyingly, ligand docking using X5 provided a rational basis for ligand binding in endo-mode. The combination in the model of THSAbf of a short  $\beta 6\alpha 6$  loop and an 'open'  $\beta 2\alpha 2$  loop creates a rather long and wide active site that can readily accommodate X5. Moreover, the absence of Trp101 in the vicinity of subsite -1 allows the  $\beta$ -D-xylosyl moiety in position 3 to sink quite low into the active, thus properly occupying, subsite -1 and positioning the scissile bond within an almost plausible distance of the catalytic dyad (i.e., 4.1 Å for  $Glu177-O^{e1} \rightarrow O4\beta$ -Xyl and 4.3 Å for  $Glu296-O^{e1} \rightarrow C1\beta$ -Xyl). Finally, it is noteworthy that the model of the THSAbf:X5 complex predicts that Asn76 N $\delta$  is well positioned to form H-bonds with both Xyl3(O3) and Xyl2(O5), and O $^{\nu 1}$  of Thr103, which is on the  $\beta 2\alpha 2$  loop, is predicted to form a H-bond with Xyl5(O3), the moiety forming the non-reducing end of X5.

Overall, modeling reveals that the position of the  $\beta 2\alpha 2$  loop could be a major factor to explain the multi-functionality of THSAbf, especially because a key function has already been attributed to this loop in other GH51  $\alpha$ -L-arabinofuranosidases. Apparently, the mobility of the  $\beta 2\alpha 2$  loop could form the basis for the absence of the bulky Trp101 in the active site of THSAbf and thus the ability of the enzyme to productively bind xylooligosaccharides. In this respect, it is perhaps noteworthy that the replacement of Tyr166 by a less bulky amino acid in an unrelated GH43 arabinoxyylan hydrolase (*AXHd3* from *Humicola insolens*) also provoked a significant modification of active topology and facilitated the productive binding of xylo-oligosaccharides (DP  $\geq$  5).

## Conclusion

This work follows the discovery of the bifunctional *Ali-cyclobacillus* GH51 arabinofuranosidase. Like the former, THSAbf differs from most GH51 Abfs, because it catalyzes the hydrolysis of external and internal glycosidic



bonds. However, THSAbf is multifunctional, since it also possesses xylosidase and possibly endo-arabinanase activities. Importantly, the discovery of *AlAbf51* and THSAbf reveals that current knowledge of glycoside hydrolases is incomplete, especially because only a small

fraction of GHs have been characterized and many studies are insufficiently thorough. From an industrial standpoint, we expect that discovery of THSAbf will facilitate the design of simpler, lower-cost xylanolytic cocktails.

## Additional files

**Additional file 1.** Phylogenetic analysis of GH51 sequences. The figure shows an unrooted phylogram built using 577 GH51 sequences. The approximate locations in the phylogram of THSAbf and other characterized GH51 Abfs (*GsAbf* and *TxAbf*) are shown.

**Additional file 2.** Thermoactivity and thermostability of THSAbf. Figure S2A shows THSAbf activity as a function of temperature and Figure S2B shows thermostability plots.

**Additional file 3.** Hydrolysis of LVWAX by various enzymes. Figure S3A, B and C show the progress of hydrolysis LVWAX by *TxAbf*(B), *TxXyn* and THSAbf respectively. Several reactions products are monitored, including xylose, arabinose and xylooligosaccharides (X2-X6).

**Additional file 4.** Hydrolysis of LVWAX using different enzyme combinations. Figure S4A, B, C and D show the progress of hydrolysis of LVWAX by (A) *TxAbf* + *TxXyn*, (B) THSAbf + *TxXyn*, (C) THSAbf + *TxAbf* and (D) THSAbf + *TxAbf* + *TxXyn*. Several reactions products are monitored, including xylose, arabinose and xylooligosaccharides (X2-X6).

**Additional file 5.** Hydrolysis of DWB by various enzymes. Figure S5A, B and C show the progress of hydrolysis of DWB by *TxAbf*(B), *TxXyn* and THSAbf respectively. Several reactions products are monitored, including xylose, arabinose and xylooligosaccharides (X2-X6).

**Additional file 6.** Hydrolysis of DWB using different enzyme combinations. Figure S6A, B, C and D show the progress of hydrolysis of DWB by (A) *TxAbf* + *TxXyn*, (B) THSAbf + *TxXyn*, (C) THSAbf + *TxAbf* and (D) THSAbf + *TxAbf* + *TxXyn*. Several reactions products are monitored, including xylose, arabinose and xylooligosaccharides (X2-X6).

**Additional file 7.** Modelling alternative active site topologies of THSAbf. Figure S7A and B show proposed models for the so-called open and closed conformers of THSAbf respectively. These are compared to models of active sites (Figure S7C) *GsAbf* (1QW8 chain B) and (Figure S7D) *TxAbf* (2VVRQ chain C).

**Additional file 8.** Comparison of the position of the  $\beta$ 2a2 loop in the closed and open forms of THSAbf and *TxAbf*. Figure S8A and B show two views (rotation of 90° through X-axis) of the active site with a zoom on the  $\beta$ 2a2 loop.

## Abbreviations

Abf:  $\alpha$ -L-arabinofuranosidase; Ara: arabinose; DWB: destarched wheat bran; GH: glycoside hydrolase; HPAEC-PAD: high-performance anion-exchange chromatography with pulsed amperometric detection; LB: Luria Bertoni; LVWAX: low-viscosity wheat arabinoxylan; pNP: 4-nitrophenol; pNP- $\alpha$ -L-arabinofuranoside: pNP-Araf; pNP- $\beta$ -D-galactofuranoside: pNP-Galf; pNP- $\beta$ -D-xylopyranoside: pNP-Xylp; SA: specific activity; XOS: xylooligosaccharides; Xyl: xylose; X-Xyl: 5-bromo-4-chloro-3-indolyl- $\beta$ -D-xylopyranoside.

## Authors' contributions

Together, HB, MLD, EI, NFF and PA, performed all of the experiments described in the manuscript, though HB and MLD were the major contributors. PA created and tested the active site mutant and EI participated in the determination of pH optima on different substrates, and with NFF did the molecular modeling work. CR and CD provided significant experimental supervision and participated in writing the manuscript, while MJOD designed and supervised the whole study and wrote most of the manuscript. MM supervised HB during her Ph.D. thesis, in particular the isolation of the *Paenibacillus* strain. All authors agreed to the publication of the manuscript. All authors read and approved the final manuscript.

## Author details

<sup>1</sup> UBMB, Université de Tunis El Manar, BP 94, 1068 Rommana, Tunisia. <sup>2</sup> CNRS, INRA, INSA, LISBP, Université de Toulouse, Toulouse, France. <sup>3</sup> INRA, FARE, Université de Reims Champagne Ardenne, 2, Esplanade Roland Garros, 51100 Reims, France. <sup>4</sup> Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth SY23 3DA, Ceredigion, UK. <sup>5</sup> Present Address: Laboratoire des Ressources Sylvo-Pastorales, Institut Sylvo-Pastoral de Tabarka, Institution de la Recherche et de l'Enseignement Supérieur Agricoles, Université de Jendouba, Jendouba, Tunisia.

## Acknowledgements

The authors would like to thank Béatrice Hermant and Nathalie Aubry for technical support and Lannig Richard and Emilie Mauré, both of whom contributed during internships that formed part of their undergraduate and BTS courses, respectively.

## Competing interests

All authors declare that they have no competing interests.

## Availability of supporting data

The 16s rDNA is available at GENBANK Accession Number AM283040.

The DNA encoding THSAbf, designated *abfA*, is available at GENBANK Accession Number ABZ10760.

Additional files accompany this manuscript and are available online.

## Funding

This work was funded in part by the French Agence Nationale de la Recherche (Contract No ANR-08-CP2D-17-02), which importantly provided a stipend to MLD. HB benefited from grants from the Comité Mixte de Coopération Universitaire (France-Tunisia) and the Tunisian Ministry of Higher Education and Scientific Research, and Eleni Ioannou is funded by a PhD Grant from Climate KIC.

Received: 22 April 2016 Accepted: 23 June 2016

Published online: 08 July 2016

## References

- Carpita NC. Structure and biogenesis of the cell walls of grasses. *Annu Rev Plant Physiol Plant Mol Biol.* 1996;47:445–76.
- Ebringerova A, Heinze T. Xylan and xylan derivatives—biopolymers with valuable properties, 1 Naturally occurring xylans structures, isolation procedures and properties. *Macromol Rapid Commun.* 2000;21:542–56.
- Shallom D, Shoham Y. Microbial hemicellulases. *Curr Opin Microbiol.* 2003;6:219–28. <http://www.linkinghub.elsevier.com/retrieve/pii/S1369527403000560>. Accessed 3 Nov 2012.
- Paës G, Skov LK, O'Donohue MJ, Rémond C, Kastrup JS, Gajhede M, et al. The structure of the complex between a branched pentasaccharide and *Thermobacillus xylanilyticus* GH-51 arabinofuranosidase reveals xylan-binding determinants and induced fit. *Biochemistry.* 2008;47:7441–51.
- McCarter JD, Withers SG. Mechanisms of enzymatic glycoside hydrolysis. *Curr Opin Struct Biol.* 1994;4:885–92.
- Shallom D, Belakhov V, Solomon D, Shoham G, Baasov T, Shoham Y. Detailed kinetic analysis and identification of the nucleophile in  $\alpha$ -L-arabinofuranosidase from *Geobacillus stearothermophilus* T-6, a family 51 glycoside hydrolase. *J Biol Chem.* 2002;277:43667–73. <http://www.ncbi.nlm.nih.gov/pubmed/12221104>. Accessed 17 Nov 2012.
- Malburg SRC, Malburg LM, Liu T, Iyo AH. Catalytic properties of the cellulose-binding endoglucanase F from fibrobacter succinogenes S85. *Appl Environ Microbiol.* 1997;63:2449–53.
- Hövel K, Shallom D, Niefind K, Belakhov V, Shoham G, Baasov T, et al. Crystal structure and snapshots along the reaction pathway of a family 51  $\alpha$ -L-arabinofuranosidase. *EMBO J.* 2003;22:4922–32.
- Rémond C, Plantier-Royon R, Aubry N, O'Donohue MJ. An original chemoenzymatic route for the synthesis of  $\beta$ -D-galactofuranosides using an  $\alpha$ -L-arabinofuranosidase. *Carbohydr Res.* 2005;340:637–44.

10. Minic Z, Rihouey C, Do CT, Lerouge P, Jouanin L. Purification and characterization of enzymes exhibiting arabinanase A from *Pseudomonas fluorescens* subsp cellulosa exhibits both an endo- and an exo- mode of action. *Plant Physiol*. 2004;135:867–78.
11. Rémond C, Boukari I, Chambat G, O'Donohue M. Action of a GH 51  $\alpha$ -L-arabinofuranosidase on wheat-derived arabinoxylans and arabinoxylooligosaccharides. *Carbohydr Polym*. 2008;72:424–30. <http://www.linkinghub.elsevier.com/retrieve/pii/S0144861707004870>. Accessed 17 Nov 2012.
12. Lee RC, Hrmova M, Burton RA, Lahnstein J, Fincher GB. Bifunctional family 3 glycoside hydrolases from barley with  $\alpha$ -L-arabinofuranosidase and  $\beta$ -D-xylosidase activity. Characterization, primary structures, and COOH-terminal processing. *J Biol Chem*. 2003;278:5377–87. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12464603>. Accessed 17 Nov 2012.
13. Ravanal MC, Callegari E, Eyzaguirre J. Novel bifunctional  $\alpha$ -L-arabinofuranosidase/xylobiohydrolase (ABF3) from *Penicillium purpogenum*. *Appl Environ Microbiol*. 2010;76:5247–53. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2916492&tool=pmcentrez&rendertype=abstract>. Accessed 4 Feb 2013.
14. Khandeparker R, Numan MT. Bifunctional xylanases and their potential use in biotechnology. *J Ind Microbiol Biotechnol*. 2008;35:635–44. <http://www.ncbi.nlm.nih.gov/pubmed/18365260>. Accessed 3 Nov 2012.
15. Hernández A, Copa-Patiño JL, Soliveri J. xln23 from *Streptomyces chatanoogensis* UAH23 encodes a putative enzyme with separate xylanase and arabinofuranosidase catalytic domains. *DNA Seq*. 2001;12:166–77.
16. Fan Z, Wagschal K, Lee CC, Kong Q, Shen KA, Maiti IB, et al. The construction and characterization of two xylan-degrading chimeric enzymes. *Biotechnol Bioeng*. 2009;102:684–92. <http://www.ncbi.nlm.nih.gov/pubmed/18828173>. Accessed 7 Nov 2012.
17. Fan Z, Werkman JR, Yuan L. Engineering of a multifunctional hemicellulase. *Biotechnol Lett*. 2009;31:751–7. <http://www.ncbi.nlm.nih.gov/pubmed/19169889>. Accessed 17 Nov 2012.
18. Yang W, Bai Y, Yang P, Luo H, Huang H, Meng K, et al. A novel bifunctional GH51 exo- $\alpha$ -L-arabinofuranosidase/endo-xylanase from *Alicyclobacillus* sp. A4 with significant biomass-degrading capacity. *Biotechnol Biofuels*. *BioMed Central*. 2015;8:197. <http://www.biotechnologyforbiofuels.com/content/8/1/197>.
19. Sørensen HR, Pedersen S, Jørgensen CT, Meyer AS. Enzymatic hydrolysis of wheat arabinoxylan by a recombinant "minimal" enzyme cocktail containing beta-xylosidase and novel endo-1,4-beta-xylanase and alpha-L-arabinofuranosidase activities. *Biotechnol Prog*. 2007;23:100–7. <http://www.ncbi.nlm.nih.gov/pubmed/17269676>.
20. Bokinsky G, Peralta-Yahya PP, George A, Holmes BM, Steen EJ, Dietrich J, et al. Synthesis of three advanced biofuels from ionic liquid-pretreated switchgrass using engineered *Escherichia coli*. *Proc Natl Acad Sci USA*. 2011;108:19949–54. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3250153&tool=pmcentrez&rendertype=abstract>. Accessed 2 Nov 2012.
21. Touzel JP, Donohue MO, Debeire P, Samain E, Breton C. A new aerobic thermophilic xylan-degrading bacterium isolated from farm soil. *Int J Syst Evol Microbiol*. 2000;50:315–20.
22. Bouraoui H, Rebib H, Aissa MB, Touzel JP, O'Donohue MJ. *Paenibacillus marinum* sp. nov., a thermophilic xylanolytic bacterium isolated from a marine hot spring in Tunisia. *J Basic Microbiol*. 2013;53:877–83.
23. Harwood CR, Archibald AR. Growth, maintenance and general techniques. In: Harwood CR, Cutting SM, editors. *Molecular biological methods for Bacillus*. Chichester: Wiley; 1990. p. 1–26.
24. Sambrook J, Maniatis T, Fritsch EF. *Molecular cloning: a laboratory manual*. New York: Cold Spring Harbor Laboratory Press; 1987.
25. Fernandez-Fuentes N, Madrid-Aliste CJ, Rai BK, Fajardo JE, Fiser A. M4T: a comparative protein structure modeling server. *Nucleic Acids Res*. 2007;35:363–8.
26. Miller GL. Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal Chem*. 1959;31:426–8.
27. Saha BC. Alpha-L-arabinofuranosidases: biochemistry, molecular biology and application in biotechnology. *Biotechnol Adv*. 2000;18:403–23.
28. Yang Y, Zhang L, Guo M, Sun J, Matsukawa S, Xie J, et al. Novel  $\alpha$ -L-arabinofuranosidase from *Cellulomonas fimi* ATCC 484 and its substrate-specificity analysis with the aid of computer. *J Agric Food Chem*. 2015;63:3725–33. <http://www.pubs.acs.org/doi/abs/10.1021/jf5059683>.
29. Keating L, Kelly C, Fogarty W. Mechanism of action and the substrate-dependent pH maximum shift of the alpha-amylase of *Bacillus coagulans*. *Carbohydr Res*. 1998;309:311–8.
30. Jeon EJ, Jung JH, Seo DH, Jung DH, Holden JF, Park CS. Bioinformatic and biochemical analysis of a novel maltose-forming alpha-amylase of the GH57 family in the hyperthermophilic archaeon *Thermococcus* sp. CL1. *Enzyme Microb Technol*. 2014;60:9–15. doi:10.1016/j.enzmictec.2014.03.009.
31. Beaugrand J, Chambat G, Wong VWK, Goubet F, Rémond C, Paës G, et al. Impact and efficiency of GH10 and GH11 thermostable endoxylanases on wheat bran and alkali-extractable arabinoxylans. *Carbohydr Res*. 2004;339:2529–40. <http://www.ncbi.nlm.nih.gov/pubmed/15476714>. Accessed 17 Nov 2012.
32. Benamrouche S, Crônier D, Debeire P, Chabbert B. A chemical and histological study on the effect of (1  $\rightarrow$  4)- $\beta$ -endo-xylanase treatment on wheat bran. *J Cereal Sci*. 2002;36:253–60. <http://linkinghub.elsevier.com/retrieve/pii/S0733521001904278>. Accessed 17 Nov 2012.
33. Paës G, Berrin J-G, Beaugrand J. GH11 xylanases: Structure/function/properties relationships and applications. *Biotechnol Adv*. 2012;30:564–92. <http://www.ncbi.nlm.nih.gov/pubmed/22067746>. Accessed 3 Mar 2013.
34. Goldbeck R, Damásio AR, Gonçalves TA, Machado CB, Paixão DA, Wolf LD, et al. Development of hemicellulolytic enzyme mixtures for plant biomass deconstruction on target biotechnological applications. *Appl Microbiol Biotechnol*. 2014;98:8513–25. <http://www.link.springer.com/10.1007/s00253-014-5946-6>.
35. Arab-Jaziri F, Bissaro B, Barbe S, Saurel O, Débat H, Dumon C, et al. Functional roles of H98 and W99 and  $\beta$ 2a2 loop dynamics in the  $\alpha$ -L-arabinofuranosidase from *Thermobacillus xylanilyticus*. *FEBS J*. 2012;279:3598–611. <http://www.ncbi.nlm.nih.gov/pubmed/22823427>. Accessed 17 Nov 2012.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

