An efficient arabinoxylan-debranching α-L-arabinofuranosidase of family GH62 from *Aspergillus nidulans* contains a secondary carbohydrate binding site

Casper Wilkens[1,A], Susan Andersen[1], Bent O. Petersen[2,B], An Li[3], Marta Busse-Wicher[3], Johnny Birch[1], Darrell Cockburn[1,C], Hiroyuki Nakai[1,D], Hans E. M. Christensen[4], Birthe B. Kragelund[5], Paul Dupree[3], Barry McCleary[6], Ole Hindsgaul[2], Maher Abou Hachem[1] and Birte Svensson[1,*]

[1]Enzyme and Protein Chemistry, Department of Systems Biology, Technical University of Denmark, Elektrovej, Building 375, DK-2800 Kgs. Lyngby, Denmark

[2]Carbohydrate Chemistry Group, Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-1799 Copenhagen V, Denmark

[3]Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK

[4]Metalloprotein Chemistry and Engineering, Department of Chemistry, Technical University of Denmark, Kemitorvet, Building 207, DK-2800 Kgs. Lyngby, Denmark

[5]Structural Biology and NMR Laboratory, Department of Biology, University of Copenhagen, Ole Maaloes Vej 5, DK-2200 Copenhagen N, Denmark

[6]Megazyme International, Bray Business Park, Bray, Co. Wicklow, Ireland

[*]To whom correspondence should be addressed: Enzyme and Protein Chemistry, Department of Systems Biology, Technical University of Denmark, Elektrovej, Building 375, DK-2800 Kgs. Lyngby, Denmark, Tel.: +45 4525 2740, E-mail: bis@bio.dtu.dk

[A]Present address: Department of Chemical and Biochemical Engineering, Technical University of Denmark, Søltofts Plads, Building 227, DK-2800 Kgs. Lyngby, Denmark

[B]Present address: Biophysics and Biotechnology, Novo Nordisk A/S, Novo Nordisk Park, DK-2760, Måløv, Denmark

[C]Present address: Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, MI 48109, Michigan, U.S.A.

[D]Permanent address: Graduate School of Science and Technology, Niigata University, 8050 Ikarashi, Nishi-ku, Niigata 950-2181, Japan

Abstract: An α-L-arabinofuranosidase of GH62 from *Aspergillus nidulans* FGSC A4 (*An*Abf62A-m2,3) has unusually high activity towards wheat arabinoxylan (WAX) (67 U/mg; k_{cat} = 178 s^{-1}, K_m = 4.90 mg/ml) and arabinoxylooligosaccharides (AXOS) with degree of polymerisation (DP) 3–5 (37—80 U/mg), but about 50 times lower activity for sugar beet arabinan and 4-nitrophenyl-α-L-arabinofuranoside. α-1,2- and α-1,3-linked arabinofuranose is released from mono-, but not from disubstituted xylose in WAX and different AXOS as demonstrated by NMR and polysaccharide analysis by carbohydrate gel electrophoresis (PACE). Mutants of the predicted general acid (Glu^{188}) and base (Asp^{28}) catalysts, and the general acid pK_a modulator (Asp^{136}) lost 1700-, 165- and 130-fold activity for WAX. WAX, oat spelt xylan, birchwood xylan and barley β-glucan retarded migration of *An*Abf62A-m2,3 in affinity electrophoresis (AE) although the two latter are neither substrates nor inhibitors. Trp^{23} and Tyr^{44}, situated about 30 Å from the catalytic site as seen in an *An*Abf62A-m2,3 homology model generated using *Streptomyces thermoviolaceus Sth*Abf62A as template, participate in carbohydrate binding. Compared to wild-type, W23A and W23A/Y44A mutants are less retarded in AE, maintain about 70 % activity towards WAX with K_i of WAX

substrate inhibition increasing 4–7 fold, but lost 77–96 % activity for the AXOS. The Y44A single mutant had less effect suggesting Trp^{23} is a key determinant. *An*Abf62A-m2,3 seems to apply different polysaccharide-dependent binding modes and Trp^{23} and Tyr^{44} belong to a putative surface binding site which is situated at a distance of the active site and has to be occupied to achieve full activity.

Keywords: Glycoside hydrolase family 62, Inverting mechanism, Arabinoxylan, Arabinoxylooligosaccharides, Affinity gel electrophoresis, Surface binding site

Introduction

Plants supply the most abundant biomass on earth and sustainable utilisation of this renewable resource is very important for society. Plant cell walls are rich in L-arabinofuranose (Ara*f*) found in arabinan main chains, pectin side chains and as decorations of arabinoxylan (AX), arabinogalactan and gum arabic. Removal of Ara*f* residues constitutes a bottleneck in plant biomass conversion (Jordan et al. 2012) and efficient α-L-arabinofuranosidases (ABFs) (EC 3.2.1.55) are needed for various industrial processes such as bioethanol production (Numan and Bhosle 2006).

ABFs occur in glycoside hydrolase families GH3, 43, 51, 54 and 62 of the Carbohydrate Active Enzymes database (CAZy) (Lombard et al. 2014) and are distinguished by the ability to release 1,2- and/or 1,3-linked Ara*f* from singly or doubly substituted Xyl*p* residues (Van Laere et al. 1999; Sakamoto et al. 2013). Only GH62 contains exclusively ABFs and it constitutes glycoside hydrolase clan F (GH-F) with GH43 (Lombard et al. 2014) that comprises ABF and several other specificities. GH62 is predicted to be inverting similar to GH43 (Kellett et al. 1990; McKie et al. 1997; Kimura et al. 2000) as was here confirmed experimentally by using NMR, which also demonstrated that *An*Abf62A-m2,3 releases α-1,3-linked three times faster than α-1,2-linked Ara*f*. Currently 17 GH62 members have been functionally characterized and kinetic data are reported for nine (Poutanen 1988; Margolles-Clark et al. 1996; Ransom and Walton 1997; Vincent et al. 1997; Lange et al. 2006; De La Mare et al. 2013; Siguier et al. 2014; Maehara et al. 2014; Wang et al. 2014; Kaur et al. 2014), while substrate specificity was determined for the remaining eight enzymes (Kellett et al. 1990; Kimura et al. 2000; Hashimoto et al. 2011; Sakamoto et al. 2011). The first GH62 crystal structures – five in total – were published in 2014 (Siguier et al. 2014; Maehara et al. 2014; Wang et al. 2014; Kaur et al. 2014) and share a five-bladed β-propeller fold catalytic domain with the six GH43 ABF structures of one fungal and five bacterial enzymes (Nurizzo et al. 2002; Lombard et al. 2014). Single mutants support the role of invariant glutamic acid and aspartic acid residues as general acid and base catalyst and of another invariant aspartic acid residue as pK_a modulator of the acid catalyst (Pitson et al. 1996; Nurizzo et al. 2002; Siguier et al. 2014).

The present study concerns *An*Abf62A-m2,3, one of the two *Aspergillus nidulans* GH62 enzymes available in the seminal tool box of *Pichia pastoris* transformants encoding *A. nidulans* plant cell wall degrading enzymes (Bauer et al. 2006). *An*Abf62A-m2,3 has no carbohydrate binding module (CBM) therefore its strong binding to different cell wall polysaccharides motivated establishing a homology model in which Trp^{23} and Tyr^{44} were tentatively localized to a surface binding site (SBS). In the light of the number of GH62 sequences in CAZy which is very recently grown by 40 % we here divided GH62 in four phylogenetics subgroups (Supplementary Fig. S1) rather than just two (Siguier et al. 2014).

Materials and methods

Structural modelling and phylogenetic subgrouping

An *An*Abf62A-m2,3 model obtained using HHpred (Söding et al. 2005) and the structure of *Sth*Abf62A from *Streptomyces thermoviolaceus* (PDB ID 4O8O) as template was judged as "extremely good/very good" (LGscore 5.1 and MaxSub 0.54) by ProQ (Wallner and Elofsson 2003). Alignment with *Sth*Abf62A using PyMol 1.3 (Schrödinger, LLC, New York, NY, USA; also used for rendering structural models) showed similar secondary structural elements (prediction server PSIPRED (Buchan et al. 2010)) having two *An*Abf62A-m2,3 outliers (Thr^{202} and Asn^{287}) in the Ramachandran plot. The overall rmsd for Cα was 0.15 Å.

The catalytic domain (cl14647) was identified by Conserved Domain Database (Marchler-Bauer and Lu 2011) in 142 GH62 sequences (May 15 2015) retrieved from CAZy and a multiple alignment (ClustalW default settings within MEGA 6 (Tamura et al. 2013)) was generated for building a phylogenetic tree using the maximum likelihood algorithm with MEGA 6 (Tamura et al. 2013). Peptide Pattern Recognition (Busk and Lange 2013) identified unique sequence motifs for the subgroups with the following parameters (peptide length: 7; number of peptides: 70; cut-off: 10). The identities were calculated by aid of ClustalW 2.1 (Li et al. 2015).

Cloning, mutagenesis, expression and purification of *An*Abf62A-m2,3

P. pastoris X-33 transformants (FGSC database accession no. 10088 and 10106; www.fgsc.net) harbouring full-length *A. nidulans* FGSC A4 ABF (GenBank ID: AN7908.2) were purchased (Fungal Genetics Stock Centre, School of Biological Sciences, University of Missouri-Kansas City, MO, USA). A 22 residues predicted signal peptide (SignalP 3.0 (Emanuelsson et al. 2007)) was removed using PCR (Expand High Fidelity DNA polymerase; Roche Diagnostics, Rotkreuz, Switzerland) (for primers see Supplementary Table S1) and a C2A mutation was introduced to avoid intermolecular disulfide formation. The construct was cloned (using *Eco*RI and *Not*I; New England BioLabs, Ipswich, MA, USA) in-frame in pPICZαA (Invitrogen, Carlsbad, CA, USA) with the sequence for the *Saccharomyces cerevisiae* α-mating factor and a stop codon upstream of a C-terminal His-tag (QuickChange kit; Stratagene, San Diego, CA, USA; Supplementary Table S1). A pPICZαA-*An*Abf62A-m2,3 transformant (*Escherichia coli* DH5α selected on low salt LB medium with 25 µg/ml zeocin; Novagen, Nottingham, United Kingdom) was sequenced (Eurofins MWG Operon, Ebersberg, Germany), linearized (*Pme*I; New England BioLabs, Ipswich, MA, USA), transformed into *P. pastoris* X-33 (Micropulser; Bio-Rad, Hercules, CA, USA), and selected (30 °C, 3 d) on yeast peptone dextrose plates with 100 µg/ml zeocin (Invitrogen, Carlsbad, CA, USA). *An*Abf62A-m2,3 W23A, D28A, Y44A, D136A, E188A, and W23A/Y44A mutants were made using site-directed mutagenesis (for primers see Supplementary Table S1) according to the manufacturer's recommendations (QuickChange kit; Stratagene, San Diego, CA, USA). *P. pastoris* transformants were grown in shake flasks in buffered glycerol-complex medium (BMGY; 30 °C, 24 h), harvested (3000*g*, 10 min, 22 °C) and resuspended to 1/4 of the BMGY culture volume in buffered methanol-complex medium (BMMY; 22 °C, 96 h; methanol supplemented to 0.5 % (v/v) every 24 h). Supernatants were filtered (0.45 µm Durapore membrane filters; Millipore, Billerica, MA, USA), 10 fold concentrated and buffer-exchanged to 10 mM sodium acetate pH 5.5 (Pellicon ultra-filtration unit, 10 kDa cut-off filter; Millipore, Billerica, MA, USA), applied (5 ml/min) onto a 15 ml CaptoQ

column (GE Healthcare, Little Chalfont, United Kingdom) equilibrated with 10 mM sodium acetate pH 5.5 and eluted by a linear 0–500 mM NaCl gradient (20 CV) (5 ml/min). Fractions containing *An*Abf62A-m2,3 (monitored by SDS-PAGE) were pooled, concentrated (4000*g*; Amicon Ultra-15 centrifugal filter units, 10 kDa cut-off; Millipore, Billerica, MA, USA) and gel filtrated (Hiload 26/60 Superdex G75 column; GE Healthcare, Little Chalfont, United Kingdom) in 10 mM sodium acetate, 0.15 M NaCl, pH 5.5 (0.5 ml/min). Fractions containing *An*Abf62A-m2,3 were pooled, concentrated and buffer-exchanged to 10 mM HEPES pH 7.5, applied (2 ml/min) to a 6 ml ResourceQ column (GE Healthcare, Little Chalfont, United Kingdom) in this buffer and eluted by a linear 0–500 mM NaCl gradient (20 CV; 2 ml/min). Pure *An*Abf62A-m2,3 was pooled, concentrated (to 30–970 µM), buffer-exchanged to 10 mM sodium acetate pH 5.5, added sodium azide to 0.02 % and stored at 4 °C. All steps were carried out at 4 °C.

Protein analyses

*An*Abf62A-m2,3 wild-type and mutants were analysed by SDS-PAGE (4–12 %; Invitrogen). Molecular mass of wild-type was determined by ESI-MS (LCT Premier mass spectrometer; Waters, Milford, MA, USA). Briefly, *An*Abf62A-m2,3 was exchanged into 2.3 M ammonium acetate (Micro Bio-Spin P-6 size exclusion columns; Bio-Rad, Hercules, CA, USA), sprayed from nanoES capillaries (ES380; Proxeon, Odense, Denmark) using the parameters; capillary voltage: 900–1500 V; sample cone voltage: 25 V; source temperature: 30 °C; and cone gas flow: 20 L/h (N_2) and spectra were collected in positive ion mode. The instrument was calibrated with 100 mg/ml CsI in 50 % (v/v) isopropanol. Spectra were processed by smoothing followed by manual deconvolution (MassLynx V4.1 software; Waters, Milford, MA, USA). Protein concentration was determined by aid of amino acid analysis (Barkholt and Jensen 1989). The melting temperature (T_m) was determined by far-UV CD spectroscopy (see Supplementary Fig. S2). Deglycosylation by endoglycosidase H was attempted under native and denaturing conditions as recommended by the manufacturer (New England Biolabs, Ipswich, MA, USA).

Affinity gel electrophoresis

*An*Abf62A-m2,3 and mutants (4 µg in sample buffer, 0.25 M Tris base, 0.12 M boric acid, 40 % glycerol, 0.05 % Bromphenol Blue, pH 8.7) were applied on 12 % (w/v) polyacrylamide gel cast with 0.001–1 % (w/v) low viscosity wheat AX (WAX-LV) (Megazyme, Wicklow, Ireland), oat spelt xylan, birchwood xylan (both Carl Roth, Karlsruhe, Germany), larch arabinogalactan, sugar beet L-arabinan (both Megazyme, Wicklow, Ireland), acacia tree gum arabic, hydroxyethyl cellulose (both Sigma-Aldrich, St. Louis, MI, USA), or barley β-glucan (Novo Industries, Gentofte, Denmark), and run in 0.25 M Tris base, 0.12 M boric acid, pH 8.7 (4 °C, 50 V, 16 h, XCell SureLock© Mini-Cell system; Invitrogen, Carlsbad, CA, USA) with reference proteins (NativeMark; Invitrogen, Carlsbad, CA, USA) in the same tank as a control without polysaccharide. Proteins were visualized by Simpleblue SafeStain (Invitrogen, Carlsbad, CA, USA). WAX-LV was dissolved in water (50 °C) and kept for 1 h. Birchwood and oat spelt xylans were dissolved in water in a microwave oven, and sugar beet L-arabinan, acacia tree gum arabic, and larch arabinogalactan in water (RT). Barley β-glucan was wet with a minimum volume 95 % ethanol, suspended in cold water with stirring, heated to

boiling with stirring and stirred for 1 h. Hydroxyethyl cellulose was dissolved in 10 mM sodium phosphate pH 6 and adjusted to pH 8.

The relative retardation of migration (R_m) by the polysaccharide compared to the control was determined from the following equation: $R_m = R_{mi} / R_{mo}$, where R_{mi} and R_{mo} are migration distances of sample relative to reference protein in the presence and in the absence of polysaccharide, respectively.

Enzyme activity assays

4NP-glycosides: 10 mM 4NPA*f*, 4-nitrophenyl-β-D-xylopyranoside (4NPX) (Sigma-Aldrich, St. Louis, MI, USA) or 4-nitrophenyl-α-L-arabinopyranoside (4NPA*p*) (Sigma-Aldrich, St. Louis, MI, USA) in water (20 µl) was preincubated with 125 mM sodium acetate, 0.005 % Triton-X-100, pH 5.5 (20 µl; 2 min; 37°C) and added *An*Abf62A-m2,3 (10 µl; 12–24 µM final concentration). The reaction (10 min; 37 °C) was stopped by 1 M Na_2CO_3 (200 µl) and 4NP quantified spectrophotometrically at 410 nm (200 µl; microtiter plate reader; Bio-Tek Instrument Inc., Winooski, VT, USA) using 4NP (0–0.5 mM) as standard. One activity unit (U) was defined as the amount of enzyme releasing 1 µmol/min 4NP. Kinetic parameters were determined from initial rates of 4NPA*f* (0.05–4 mM) hydrolysis by *An*Abf62A-m2,3 (0.25–0.5 µM; monitored up to 16 min). pH activity optimum was determined using 2.5 mM 4NPA*f* in 40 mM Britton-Robinson universal buffers pH 2–10 (Britton and Robinson 1931) and 50 mM sodium acetate pH 5.0–6.0 (37 °C). The temperature optimum was determined at pH 5.5 for 25–75 °C.

Polysaccharides: Activity was tested on 0.9 % (w/v) linear L-arabinan (Megazyme Wicklow, Ireland), birchwood xylan, and barley β-glucan in 50 mM sodium acetate, 0.005 % Triton X-100 pH 5.5 incubated (30 min; 37 °C) with 97 µM *An*Abf62A-m2,3 and quantifying reducing sugar by adding 3,5-dinitrosalicylic acid reagent (600 µl; 1 % DNS, 0.2 % phenol, 0.05 % $NaSO_3$, 1 % NaOH, and 20 % NaK-tartrate), heated (95 °C; 15 min), cooled (on ice; 15 min) and measuring A_{540} (Mohun and Cook 1962) (200 µl; microtiter plate reader) using L-arabinose as standard. Specific activity for 0.9 % WAX-LV, rye AX (Megazyme, Wicklow, Ireland), oat spelt xylan, larch arabinogalactan, sugar beet L-arabinan, and acacia tree gum arabic was determined for 0.05 µM *An*Abf62A-m2,3 in the above buffer (10 min reaction) and L-arabinose quantified by the lactose/D-galactose (rapid) kit (Megazyme, Wicklow, Ireland) (see below). One activity unit (U) was defined as the amount of enzyme releasing 1 µmol/min arabinose. The effect on hydrolysis of 0.1 % WAX-LV and 0.4 mM or 6 mM 4NPA*f* by barley β-glucan or birchwood xylan (0.05–0.8 %; 0.1 % with 4NPA*f*) was measured assaying released L-arabinose by the lactose/D-galactose kit (see below).

Kinetic parameters were determined from initial rates of L-arabinose release (lactose/D-galactose (rapid) kit (Megazyme, Wicklow, Ireland)) from WAX-LV (0.28–9 mg/ml) and sugar beet L-arabinan (4–90 mg/ml) in the above buffer (37 °C). Reactions were initiated by adding enzyme (WAX-LV: 0.03–2 µM; sugar beet L-arabinan: 1–5 µM). Aliquots (50 µl) were removed during 16 min (60 min for catalytic site mutants), added to 1 M Tris-HCl pH 8.6 (200 µl) followed by incubation (40 min; RT) with lactose/D-galactose kit solution (880 µl) and quantified (200 µl; microtiter plate reader; NADH $\varepsilon_{M,340} = 6300\ M^{-1}\times cm^{-1}$) using L-arabinose (0–1.75 mM) as standard. k_{cat}, K_m, and K_i were obtained (SigmaPlot 9.01; SYSTAT software Inc., San Jose, CA, USA) by fitting either the classical Michaelis Menten $V = V_{max} / (1 + (K_m / [S])$ or the modified equation including a term for uncompetitive substrate inhibition $V_{i,s} = V_{max} / (1 + ((K_m / [S]) + ([S] / K_i))$ to initial rate data. V and $V_{i,s}$ are reaction rates, V_{max} maximum rate, [S] substrate concentration, and K_i

dissociation constant for inhibited ternary [substrate-enzyme]-substrate complex. Catalytic efficiency (k_{cat}/K_m) is reported for 4NPA*f*, as K_m is too high to be determined.

Specificity analysis was also done by polysaccharide analysis by carbohydrate gel electrophoresis (PACE) as described (Goubet et al. 2002) and visualised according to (Bromley et al. 2013). For PACE, WAX was treated with *Np*Xyn11A (Vardakou et al. 2008), *Hi*Abf43 (McKee et al. 2012) and *An*Abf62A-m2,3 to generate the xylooligosaccharides and AXOS labelled and used to analyse the specificity of *An*Abf62A-m2,3 essentially as described (McKee et al. 2012).

AXOS: Specific activity of *An*Abf62A-m2,3 (final concentration: 0.5 µM) was analysed on 53.7 mM (final concentration) A^3X [α -L-Araf-(1→3)- β -D-Xylp-(1→4)-D-Xylp], 40.7 mM A^2XX [α -L-Araf-(1→2)- β -D-Xylp-(1→4)- β -D-Xylp-(1→4)-D- β -Xylp], a mixture of 40.7 mM (final concentration) A^2XX [α -L-Araf-(1→2)- β -D-Xylp-(1→4)- β -D-Xylp-(1→4)-D- β -Xylp] (70 %) plus A^3XX [α -L-Araf-(1→3)- β -D-Xylp-(1→4)- β -D-Xylp-(1→4)-D- β -Xylp] (30 %), a mixture of 32.8 mM (final concentration) XA^3XX [β -D-Xylp-(α -L-Araf-(1→3)- β -D-Xylp-(1→4)- β -D-Xylp-(1→4)- β -D-Xylp (50%) plus XA^2XX [β -D-Xylp-(1→4)-[α -L-Araf-(1→2)]- β -D-Xylp-(1→4)- β -D-Xylp] (50%), and 32.8 mM (final concentration) $A^{2+3}XX$ [α -L-Araf-(1→2)]-[α -L-Araf-(1→3)]- β -D-Xylp-(1→4)- β -D-Xylp-(1→4)- β -D-Xylp] prepared in 33 mM sodium acetate pH 4.5 at 40 ºC and released L-arabinose was quantified using the lactose/D-galactose kit as described previously (McCleary et al. 2015).

Relative activities of wild-type (3.7 µM), W23A (4.4 µM), Y44A (3.3 µM) and W23A/Y44A (11 µM) were analysed as above using 2.5 mM AX^3, XA^2XX+XA^3XX and A^2XX [α -L-Ara*f*-(1→2)- β -D-Xyl*p*-(1→4)- β -D-Xyl*p*-(1→4)-D- β -Xyl*p*] (69.5 %), XA^3X [β -D-Xyl*p*-(1→4)-[α -L-Ara*f*-(1→3)]- β -D-Xyl*p*-(1→4)- β -D-Xyl*p*] (19 %) plus A^3XX [α -L-Ara*f*-(1→3)- β -D-Xyl*p*-(1→4)- β -D-Xyl*p*-(1→4)- β -D-Xyl*p*] (11.5 %) (Barry McCleary, in house collection).

Action pattern towards α-1,2- and α-1,3-Ara*f* decorated Xyl*p* and the stereochemical course were both determined by NMR. Hydrolysis of 1 mg/ml of XA^3XX+XA^2XX by *An*Abf62A-m2,3 (0.03 nM), $A^{2+3}X$ [[α -L-Ara*f*-(1→2)]-[α -L-Ara*f*-(1→3)]- β -D-Xyl*p*-(1→4)- β -D-Xyl*p*] (by 0.25 µM *An*Abf62A-m2,3), and $A^{2+3}XX$ [[α -L-Ara*f*-(1→2)]-[α -L-Ara*f*-(1→3)]- β -D-Xyl*p*-(1→4)- β -D-Xyl*p*-(1→4)- β -D-Xyl*p*] (by 0.25 µM *An*Abf62A-m2,3) in 10 mM sodium phosphate pH 6 was monitored (800 MHz Bruker Avance II NMR spectrometer equipped with a TCI cryoprobe; Bruker, Billerica, MA, USA) at 308 K and referenced relative to acetone (δ^1H=2.225 ppm; δ^{13}C=30.89 ppm). $A^{2+3}X$ and $A^{2+3}XX$ are kind gifts of Maija Tenkanen. For kinetic experiments a series of 1D proton spectra were recorded and for assignment a series of homo- and heteronuclear 2D spectra were recorded as DQF-COSY, NOESY with 600 ms mixing time, TOCSY with a spin lock field applied for 60 ms, a multiplicity edited ^{1}H-^{13}C HSQC and a ^{1}H-^{13}C HMBC. The stereochemical course of XA^2XX+XA^3XX hydrolysis was followed at 308 K by ^{1}H NMR with single scan 1D proton experiments of 11.5 s intervals. The first spectrum was recorded 23 s after enzyme addition.

Results

GH62 phylogenetic subgrouping

Phylogenetic analysis combined with a peptide pattern search using PPR (Busk and Lange 2013) of 142 GH62 sequences revealed four distinct subfamilies (Supplementary Fig. S1). GH62_2, the largest subfamily, contains 103 55–

100 % identical amino acid sequences and corresponds to the GH62_2 subfamily defined previously (Siguier et al. 2014). GH62_1 has 25 39–100 % identical sequences, GH62_3 and GH62_4 each have seven 29–100 % and 57–85 % identical sequences, respectively. *An*Abf62A-m2,3 belongs to subfamily GH62_2 (Supplementary Fig. S1). It remains to be uncovered if these subfamilies and the assigned unique sequence motifs (Supplementary Fig. S3) represent distinct enzymatic properties. Enzyme kinetics is reported for two GH62_1 (Couturier et al. 2011; Siguier et al. 2014; Kaur et al. 2014) and 12 GH62_2 members (Poutanen 1988; Vincent et al. 1997; Kimura et al. 2000; Tsujibo et al. 2002; Sakamoto et al. 2011; Hashimoto et al. 2011; De La Mare et al. 2013; Siguier et al. 2014; Maehara et al. 2014; Wang et al. 2014; Kaur et al. 2014; McCleary et al. 2015), whereas no GH62_3 member has been characterised and one from GH62_4 was shown to degrade oat spelt xylan (Kellett et al. 1990).

Structural model

The model of *An*Abf62A-m2,3 generated based on *Sth*Abf62A from *S. thermoviolaceus* (PDB ID 4O8O) of 73 % sequence identity (Wang et al. 2014) showed a five-bladed β-propeller fold domain. Overlays of arabinose and xylopentaose from structures of *Sth*Abf62A (PDB ID 4O8O) and *Streptomyces coelicolor Sc*Af62A (PDB ID 3WN2) (Maehara et al. 2014) complexes (Fig. 1) indicated possible substrate interactions in *An*Abf62A-m2,3 to involve at least three main chain binding subsites towards the non-reducing end (+2NR, +3NR, +4NR), one subsite towards the reducing end (+2R) and subsites −1 and +1 accommodating Ara*f* to be cleaved off and the Xyl*p* it decorates, respectively. Equivalent residues at these subsites in *An*Abf62A-m2,3 and the five GH62 crystal structures are shown (Fig. 2).

Purification and physico-chemical characterization

*An*Abf62A-m2,3 wild-type, three mutants at the catalytic site and three at the putative SBS were obtained in yields of 150–235 mg/l from *P. pastoris* culture supernatants and migrated in SDS-PAGE as two close bands of apparent molecular weights 34 and 36 kDa (Supplementary Fig. S4). ESI-MS of *An*Abf62A-m2,3 wild-type gave M_r of 33327.3 ± 0.3 and 33633 ± 1 differing by 306 for the lower band, while for the upper and minor band five M_r values in the range 35434–36067 differed by approximately 162 corresponding to one hexose residue. Mass deviations of 139 Da and 2.4–2.8 kDa from the theoretical M_r of 33188.5, presumably reflect misprocessing of the signal peptide and / or *O*-glycosylation, which was not eliminated by endoglycosidase H treatment. *An*Abf62A-m2,3 forms corresponding to either of the 34 and 36 kDa bands, were purified in extremely low yield (<1%) by gel filtration (Supplementary Fig. S5), and found to have the same specific activity towards WAX, therefore *An*Abf62A-m2,3 wild-type and mutants were characterised without being subjected to this inefficient purification of each form. The conformational stability of wild-type and mutants was assessed by aid of CD spectroscopy and T_m values were determined to 71.53 ± 0.28 °C (wild-type), 69.96 ± 0.19 (D28A), 70.11 ± 0.20 (E188A), 62.48 ± 0.18 (D136A), 60.83 ± 0.22 (W23A), 64.63 ± 0.20 (Y44A), and 55.41 ± 0.49 (W23A/Y44A) (Supplementary Fig. S2A–G).

Affinity for polysaccharides

*An*Abf62A-m2,3 interacted exceptionally strongly with 0.05 % WAX-LV in AE (Fig. 3A) and got still importantly retarded by 0.001 % WAX-LV (R_m=0.67) (Fig. 3C; Supplementary Table S2), oat spelt xylan (R_m=0.73) (Fig. 3D; Supplementary Table S2) or birchwood xylan (R_m=0.80) (Fig. 3E; Supplementary Table S2). *An*Abf62A-m2,3 thus recognises the xylan backbone as birchwood xylan has very few (< 1%) or no Ara*f* substituents (Kormelink and Voragen 1993;Li et al. 2000). Two closely migrating bands of the *An*Abf62A-m2,3 control (Fig. 3B) merged in AE indicating all *An*Abf62A-m2,3 forms bind polysaccharides. By contrast 1 % sugar beet L-arabinan (R_m=1) (Fig. 3G; Supplementary Table S2), acacia tree gum Arabic (R_m=1) (Fig. 3H), or larch arabinogalactan (R_m=1) (Fig. 3I; Supplementary Table S2) did not retard the enzyme in AE even though they are decorated by Ara*f* and L-arabinan is a substrate (Table 1). Notably, *An*Abf62A-m2,3 contains no CBM but clearly binds to 0.001 % barley β-glucan (R_m=0.9) (Fig. 3F; Supplementary Table S2) and hydroxyethyl cellulose (not shown), which are not substrates. This affinity for β-glucans may be reminiscent to the accommodation of cellotriose at the active site in the *Pa*Abf62A structure (Siguier et al. 2014).

Substrate specificity and mechanism of action

*An*Abf62A-m2,3 degraded WAX-LV with exceptional high activity of 67.42 U/mg (Table 1), k_{cat} = 178 s^{-1} and K_m = 2.3 mg/ml (Table 2, Fig. 4A). WAX-LV exerted uncompetitive substrate inhibition with K_i = 2.89 mg/ml (Table 2, Fig. 4A) and inhibited hydrolysis of 4NPA*f* by ~60 % (data not shown). *An*Abf62A-m2,3 has almost the same high activity on rye AX and oat spelt xylan (Table 1), but low activity without substrate inhibition for sugar beet L-arabinan of k_{cat} = 1.03 s^{-1} and K_m = 15.63 mg/ml (Table 2, Fig. 4A, B). Ara*f* substituted larch arabinogalactan and acacia tree gum arabic were extremely poor substrates and unsubstituted sugar beet linear arabinan was not degraded (Table 1). Birchwood xylan and barley β-glucan were neither substrates of *An*Abf62A-m2,3 nor inhibited its hydrolysis of WAX-LV and 4NPA*f*. *An*Abf62A-m2,3 showed moderate activity with 4NPA*f* and optimum at pH 5.5 and 50 °C (Tabel 1; Supplementary Fig. S6A–C); its activity towards 4NPA*p* and 4NPX was 2–3 % compared to 4NPA*f* (Table 1).

^{1}H NMR analyses demonstrated that *An*Abf62A-m2,3 hydrolysed 1,2- and 1,3-Ara*f* in XA^2XX+XA^3XX (1:1) in singly, but not from 1,2- / 1,3-Ara*f* doubly substituted Xyl*p* in $XA^{2+3}X$ and $XA^{2+3}XX$ and 1,3- was released three times faster than 1,2-linked Ara*f* (Table 1, Fig. 5, Supplementary Figs. S7 and S8). Additionally, ^{1}H-NMR showed that *An*Abf62A-m2,3 liberated β-furanose (assigned anomer resonance: 5.283 ppm) from XA^2XX+XA^3XX (Fig. 5, Supplementary Fig. S7). Due to fast mutarotation, however, the anomeric signal decreased considerably already after 1 min (Fig. 5, Supplementary Fig. S7). The same specificity was determined by PACE using AXOS and WAX as substrates (Supplementary Fig. S9). *An*Abf62A-m2,3 attacked A^3XX and XA^2XX, but not the doubly 1,2- / 1,3-Ara*f* substituted Xyl*p* in $XA^{2+3}XX$. Hydrolysis of WAX by *An*Abf62A-m2,3 followed by *Np*Xyn11A, predominantly released $XA^{2+3}XX$, xylobiose, xylose and arabinose, confirming the specificity of *An*Abf62A-m2,3 on the polysaccharide.

Finally, alanine mutants of the invariant catalytic site Asp^{28}, Glu^{188} and Asp^{136} retained 7.7×10^{-3}, 5.9×10^{-4} and 6.1×10^{-3} fold of wild-type activity for WAX-LV (Table 2, Fig. 4C). While D28A showed Michaelis-Menten kinetics on WAX-LV, D136A complied with the uncompetitive substrate inhibition found for wild-type, but K_i was doubled (Table 2, Fig. 4C). The activity of the general acid E188A mutant was too low for kinetic analysis. The results agreed with the roles in

catalysis of the three residues as general base, general acid catalysts and acid catalyst pK_a modulator, respectively, also supported by crystal structures of *Um*Abf62A, *Pa*Abf62A (Siguier et al. 2014) and *Sc*Araf62A (Maehara et al. 2014).

Interaction at a putative surface binding site

In the structural model of *An*Abf62A-m2,3 Trp^{23} and Tyr^{44} are situated near the active site cleft, at a distance of about 30 Å from the catalytic site in a shallow cleft that runs perpendicular to the active site cleft, and which is almost a continuation of this (Fig. 1; 6A, B; Supplementary 3D data). Trp^{23} is conserved in 71 % of the 142 GH62 sequences, which all belong to GH62_2 and six of seven GH62_3 sequences. Tyr^{44} is seen in 10 (7 %) GH62_2 sequences and all 10 have Trp^{23}. The interaction in AE with WAX-LV, oat spelt xylan, birchwood xylan and barley β-glucan clearly weakened for W23A and W23A/Y44A, but not for the Y44A mutant that displayed essentially wild-type retardation (Fig. 3C–E; Supplementary Table S2). While W23A/Y44A retained some binding with the AXs and birchwood xylan in AE, this is not the case for barley β-glucan (Fig. 3C–F; Supplementary Table S2). Thus substitution of two aromatic residues at a putative surface binding site (SBS) situated outside of the active site cleft differentially affected polysaccharide binding specificity of *An*Abf62A-m2,3.

Mutation of Trp^{23} and Tyr^{44} did not dramatically alter k_{cat} and K_m for WAX-LV, sugar beet L-arabinan and 4NPA*f* (Table 2, Fig. 4A, D). Remarkably, however, K_i of WAX-LV substrate inhibition increased 4–7 fold for the three mutants relative to wild-type (Fig. 4D, Table 2) suggesting significant AX interaction involving Trp^{23} and Tyr^{44} to be clearly diminished in the mutants accompanied by modest effect on activity (Table 2, Fig. 3A), which can be interpreted as an effect of lack of or reduced affinity for WAX at the SBS. Remarkably, depending on the mutant and size of AXOS (Table 3) only 4–23 % activity was the retained even though Trp^{23} and Tyr^{44} according to the *An*Abf62A-m2,3 model (Figs. 1 and 6) are not situated at subsites accommodating AXOS for productive binding.

Discussion

Knowledge on GH62s is important to provide guidance on ABFs best suited for specific applications. For example addition of *An*Abf62A-m2,3 to unhydrolysed oligosaccharides from switchgrass treated with commercial enzymes efficiently improved the extent of conversion (Bowman et al. 2015). While insights on structure, substrate specificity, and mechanism of action in a broader sense are gained from sequence based classification of ABFs into GH families (Lombard et al. 2014), understanding of substrate specificity details and linking of functional diversity with phylogenetics require experimental studies. Comparison of *A. nidulans An*Abf62A-m2,3 with other GH62 enzymes underscored its unusually high activity on both AXs and AXOS and disclosed a putative SBS implicated in activity and interaction with cell wall polysaccharides.

Activity and structure/function relationships

*An*Abf62A-m2,3 cleaves off 1,2- and 1,3-Ara*f* from mono-substituted Xyl*p* in AXOS and AX and the same specificity was reported for other GH62_2 members *Sth*Abf62A (Wang et al. 2014), *St*Abf62A (Kaur et al. 2014), *Penicillum chrysogenum* AXS5 (Sakamoto et al. 2011), *Penicillum funiculosum* ABF62a–c (De La Mare et al. 2013), *Penicillium*

capsulatum ABF (Lange et al. 2006), and *St*Abf62C of GH62_1 (Kaur et al. 2014). The rate of release analysed by ^{1}H NMR was three times faster for 1,3- than 1,2-Ara*f* probably reflecting that 1,3- and 1,2-linked Ara*f* residues bind productively in opposite directions (Maehara et al. 2014; Wang et al. 2014).

*An*Abf62A-m2,3 acts on WAX-LV with 67.42 compared to 0.15–13 U/mg reported for 13 other GH62s (Kellett et al. 1990; Vincent et al. 1997; Kimura et al. 2000; Hashimoto et al. 2011; Sakamoto et al. 2011; Couturier et al. 2011; De La Mare et al. 2013; Siguier et al. 2014; Maehara et al. 2014; Kaur et al. 2014). *S. thermoviolaceus Sth*Abf62A, however, shows ~30 U/mg with WAX-HV (HV = high viscosity) of Ara*f*:Xyl*p* ratio of 0.5, which is a superior substrate to WAX-LV with Ara*f*:Xyl*p* of 0.3 (Pitkänen et al. 2009) on which *Sth*Abf62A shows ~18 U/mg (Wang et al. 2014). *An*Abf62A-m2,3 has k_{cat} of 178 s^{-1} on WAX-LV (Table 2, Fig. 4A) compared to k_{cat} = 180 s^{-1} of *Sth*Abf62A determined with the superior substrate, WAX-HV (Wang et al. 2014). Other GH62s gave much lower k_{cat} of 0.3–1.5 s^{-1} against WAX-LV and WAX-HV (Vincent et al. 1997; De La Mare et al. 2013; Siguier et al. 2014; Maehara et al. 2014; Kaur et al. 2014). K_m of *An*Abf62A-m2,3 is 2.3 mg/ml for WAX-LV (Table 2, Fig. 4A), which is intermediate to K_m values of 1 mg/ml for AbfB from *Streptomyces lividans* (Vincent et al. 1997), ABF62b and ABF62c from *P. funiculosum* (De La Mare et al. 2013) and 7–12 mg/ml for *Sth*Abf62A from *S. thermophilum* (Wang et al. 2014), *Sc*Araf62A from *S. coelicolor* (Maehara et al. 2014) and ABF62a from *P. funiculosum* (De La Mare et al. 2013). *S. lividans* AbfB contains a putative CBM, for which the specificity has not been tested without the catalytic domain and it is possible therefore that the binding of xylan stems from the CBM but it cannot be excluded that the interaction is with the catalytic domain (Vincent et al. 1997). ABF62c from *P. funiculosum* has a cellulose binding CBM13 (De La Mare et al. 2013) perhaps contributing to substrate binding, while *St*Abf62A has a cellulose binding CBM1 (Wang et al. 2014). *S. thermophilum St*Abf62C has K_m = 3.7 mg/ml (Kaur et al. 2014) which is similar to *An*Abf62A-m2,3 having K_m = 4.9 mg/ml (Table 2). *An*Abf62A-m2,3 and *Sth*Abf62A are subject to substrate inhibition with K_i of 2.89 (Table 2, Fig. 4A) and 1.5 mg/ml for WAX-LV and WAX-HV (Wang et al. 2014), respectively.

*An*Abf62A-m2,3 is slightly more active on oat spelt xylan and rye AX than *Sth*Abf62A (Wang et al. 2014) and neither *An*Abf62A-m2,3 nor five other GH62s degraded birchwood xylan (Vincent et al. 1997; Tsujibo et al. 2002; Hashimoto et al. 2011; Sakamoto et al. 2011; Wang et al. 2014).

GH62s differ conspicuously in activity level for sugar beet L-arabinan and *An*Abf62A-m2,3 thus has 173- and 3-fold lower and higher k_{cat} and K_m, respectively, than on WAX-LV (Table 2, Fig. 4A,B), whereas *Pa*Abf62A and *Um*Abf62A have k_{cat} 3- and 8-fold higher than *An*Abf62A-m2,3 for sugar beet L-arabinan, but these k_{cat} values were similar to and only 3-fold higher, respectively, compared to their values obtained with WAX-LV (Siguier et al. 2014). *Sth*Abf62A, however, has a 30-fold lower k_{cat} of 6 s^{-1} for L-arabinan than WAX-HV. The ability to accommodate both AX and arabinan was reported to involve structural movements upon binding of the xylan main chain in *Sth*Abf62A (Wang et al. 2014). *An*Abf62A-m2,3 has 3–4 orders of magnitude lower activity for Ara*f* substituted larch arabinogalactan and acacia tree gum arabic than WAX (Table 1) and did not hydrolyse unsubstituted linear sugar beet arabinan. As for other GH62s α-L-1,5 linked Ara*f* was not a substrate (Vincent et al. 1997; Tsujibo et al. 2002; Hashimoto et al. 2011; De La Mare et al. 2013; Kaur et al. 2014). *Sc*Araf62A was unable to accommodate L-arabinan at the active site as deduced both from lack of activity and the crystal structure (Maehara et al. 2014). Apparently substrate interactions differ between *An*Abf62A-m2,3 and *Sc*Araf62A although comparison of the *An*Abf62A-m2,3 model and the *Sc*Araf62A structure did not reveal striking dissimilarities anticipated to result in different ability to act on arabinan. Overall we conclude that the GH62 family presents important quantitative, but little qualitative variation in substrate specificity.

Catalytic mechanism

The present study provides experimental evidence for GH62 of its expected inverting mechanism by the release of β-furanose from AXOS as monitored by ^{1}H NMR, which is in accordance with the known inverting mechanism for GH43 (Pitson et al. 1996) constituting clan GH-F with GH62 (Lombard et al. 2014). The very low residual activities for WAX-LV of catalytic site mutants D28A (general base); E188A (general acid); and D136A (pK_a modulator of the acid catalyst) confirmed their proposed roles in catalysis. In comparison *St*Abf62C and *Sc*Araf62A catalytic acid and base mutants lost activity completely against WAX-LV and 4NPA*f* (Maehara et al. 2014; Kaur et al. 2014), as did *Sth*Abf62A, for which, however, a mutant of the acid catalyst pK_a modulator retained 2.1 10^{-5} fold of wild-type activity (Wang et al. 2014). A stabilising effect of the pK_a modulator on the catalytic site previously proposed in case of GH43 (Nurizzo et al. 2002) may be reflected in the 9 ^{o}C loss in T_m of *An*Abf62A-m2,3 D136A (Supplementary Fig. S2E, H).

Possible importance of the non-reducing and reducing end subsites

At subsites +2R, +1, +1NR, +2NR and +3NR in GH62 structures the residues vary and no hint to the higher activity of *An*Abf62A-m2,3 and *Sth*Abf62A towards WAX can be deduced from the structures (Fig. 2). At subsite −1 both *An*Abf62A-m2,3 and *Sth*Abf62A have tryptophan and threonine that interact with the Ara*f* (Trp^{51} and Thr^{43}, *An*Abf62A-m2,3 numbering), whereas the other enzymes have tyrosine and threonine, respectively (Fig. 2). Because the two former enzymes *An*Abf62A-m2,3 and *Sth*Abf62A have higher activity for WAX than reported for any other GH62 member, we speculate that tryptophan at subsite −1 may be associated with their unusually high activity.

The level of activity of *An*Abf62A-m2,3 was 22−48-fold higher for different AXOS than for 4NPA*f* suggesting that subsites beyond −1 and +1 are important for a perpendicular orientation of the Xyl*p* ring at subsite +1 positioning Ara*f* into the subsite −1 pocket (Fig. 2) (Maehara et al. 2014) and offer extra backing for productive accommodation of Ara*f*. Furthermore, two-fold higher specific activity for A^2XX+A^3XX (7:3) and XA^2XX+XA^3XX (1:1) compared to A^3X possibly reflects importance of subsite +3NR in productive substrate binding.

Putative surface binding site

The substrate inhibition by WAX involved Trp^{23} and Tyr^{44} as the corresponding alanine mutants were less inhibited by WAX and also showed improved productive binding (Table 2, Fig. 3A). Thus harmful strain or adverse binding in the productive complex of WAX-LV and wild-type *An*Abf62A-m2,3 is relieved by these mutations (Table 2, Fig. 4A). Although modest changes in k_{cat}/K_m (65−104%) for 4NPA*f* supports retained functional integrity of subsites −1 and +1 remarkably, the activity of W23A, Y44A and W23A/Y44A *An*Abf62A-m2,3 for different AXOS was only 4–23 % of wild-type (Table 3), A^3X of DP3 being most affected. Activity improved with DP of both 4 ($A^2XX+XA^3X+A^3XX$) and 5 (XA^2XX+XA^3XX). Apparently occupation also of subsites towards the non-reducing end is needed for effective productive AXOS interaction (Table 2), altogether suggesting that interaction with distal subsites is significant, as demonstrated for *St*Abf62C by mutational analysis (Kaur et al. 2014). It may be speculated that carbohydrate binding

e.g. by AXOS at a secondary site in *An*Abf62A-m2,3 involving Trp^{23} and Tyr^{44} allosterically triggers stimulation of catalysis as known for SBSs in barley α-amylase (Oudjeriouat et al. 2003; Nielsen et al. 2012) and *Aspergillus niger* xylanase (Cuyvers et al. 2011). It is likely that 4NPA*f* is unable to bind at or has low affinity for the SBS and the W23A, Y44A and W23A/Y44A mutations therefore do not affect activity towards this substrate. As birchwood xylan and barley β-glucan interact with *An*Abf62A-m2,3, but are neither hydrolyzed nor inhibiting activity against WAX, we propose a polysaccharide binding mode exists distinct from the AX substrate complex and involves an SBS containing Trp^{23} and Tyr^{44} situated at a distance of the active site region. This is in agreement with the weakened substrate inhibition by WAX-LV for the three SBS mutants, and especially the weakened interaction for W23A/Y44A leads us to suggest that the SBS provides prominent interaction with the polysaccharide in conjunction with the active site.

In conclusion, *An*Abf62A-m2,3 is the most active WAX-LV and AXOS degrading GH62 member reported to date. AE showed *An*Abf62A-m2,3 interacts with the Ara*f* decorated WAX-LV and oat spelt xylan as well as birchwood xylan and barley β–glucan. In conjunction with mutations of aromatic residues situated ~30 Å from the catalytic site as guided by a structural model of *An*Abf62A-m2,3, activity on AXs and AXOS suggests this site is important, whether it constitutes an SBS or formally would be considered is a distal subsite. Important SBSs are recognised in certain xylan-degrading enzymes in which the SBSs form shallow clefts that are almost perpendicular to the active site cleft, and most often have a pair of aromatic residues located in the centre of the SBS cleft (Schmidt et al. 1999; De Vos et al. 2006; Ludwiczek et al. 2007; Vandermarliere et al. 2008). Trp^{23} and Tyr^{44} in the *An*Abf62A-m2,3 model are also located in a shallow cleft perpendicular to the active site (Fig. 6 and Supplementary 3D data), but in the xylanases the SBSs are typically found on the other side of the enzyme than the active site (Schmidt et al. 1999; De Vos et al. 2006; Ludwiczek et al. 2007; Vandermarliere et al. 2008) as opposed to *An*Abf62A-m2,3 where the shallow SBS cleft is almost a continuation of the active site cleft.

PACE and NMR specificity analysis showed that singly substituting α-1,2- and α-1,3-linked arabinofuranose residues in WAX-LV and AXOS are hydrolysed by *An*Abf62A-m2,3. The NMR experiments confirmed release of the β-arabinofuranose anomer in agreement with the inverting mechanism known for GH43 that forms GH clan-H with GH62, and further demonstrated that α-1,3- is released faster than α-1,2-linked arabinofuranose residues from AXOS.

Acknowledgements

Mette Pries is thanked for technical assistance and Anne Blicher for amino acid analysis. The 800 MHz NMR spectra were recorded at the Danish National Instrument Centre for NMR spectroscopy of Biological Macromolecules at the Carlsberg Laboratory. Maja Tenkanen (University of Helsinki) is thanked for doubly substituted AXOS.

Compliance with Ethical Standards

Funding: This work is supported by the Danish Council for Independent Research | Natural Sciences (FNU) [grant number 09-072151], by 1/3 PhD fellowship from the Technical University of Denmark (to CW) and by a Hans Christian Ørsted postdoctoral fellowship from DTU (to DC).

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of interest: Barry McCleary is CEO and founder of Megazyme International.

References

Barkholt V, Jensen AL (1989) Amino acid analysis: determination of cysteine plus half-cystine in proteins after hydrochloric acid hydrolysis with a disulfide compound as additive. Anal Biochem 177:318–322. doi: 10.1016/0003-2697(89)90059-6

Bauer S, Vasu P, Persson S, Mort AJ, Somerville CR (2006) Development and application of a suite of polysaccharide-degrading enzymes for analyzing plant cell walls. Proc Natl Acad Sci U S A 103:11417–11422. doi: 10.1073/pnas.0604632103

Bowman MJ, Dien BS, Vermillion KE, Mertens JA. (2015) Isolation and characterization of unhydrolyzed oligosaccharides from switchgrass (*Panicum virgatum*, L.) xylan after exhaustive enzymatic treatment with commercial enzyme preparations. Carbohydr Res 407:42–50. doi: 10.1016/j.carres.2015.01.018

Britton HTS, Robinson RA (1931) Universal buffer solutions and the dissociation constant of veronal. J Chem Soc 1456–1462. doi: 10.1039/jr9310001456

Bromley JR, Busse-Wicher M, Tryfona T, Mortimer JC, Zhang Z, Brown DM, Dupree P (2013) GUX1 and GUX2 glucuronyltransferases decorate distinct domains of glucuronoxylan with different substitution patterns. Plant J 74:423–434. doi: 10.1111/tpj.12135

Buchan DWA, Ward SM, Lobley AE, Nugent TCO, Bryson K, Jones DT (2010) Protein annotation and modelling servers at University College London. Nucleic Acids Res 38:563–568. doi: 10.1093/nar/gkq427

Busk PK, Lange L (2013) Function-based classification of carbohydrate-active enzymes by recognition of short, conserved peptide motifs. Appl Environ Microbiol 79:3380–3391. doi: 10.1128/AEM.03803-12

Couturier M, Haon M, Coutinho PM, Henrissat B, Lesage-Meessen L, Berrin J-G (2011) *Podospora anserina* hemicellulases potentiate the *Trichoderma reesei* secretome for saccharification of lignocellulosic biomass. Appl Environ Microbiol 77:237–246. doi: 10.1128/AEM.01761-10

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 278:1098–1111. doi: 10.1111/j.1742-4658.2011.08023.x

De La Mare M, Guais O, Bonnin E, Weber J, Francois JM (2013) Molecular and biochemical characterization of three GH62 α-L-arabinofuranosidases from the soil deuteromycete *Penicillium funiculosum*. Enzyme Microb Technol 53:351–358. doi: 10.1016/j.enzmictec.2013.07.008

De Vos D, Collins T, Nerinckx W, Savvides SN, Claeyssens M, Gerday C, Feller G, Van Beeumen J (2006) Oligosaccharide binding in family 8 glycosidases: crystal structures of active-site mutants of the β-1,4-xylanase pXyl from *Pseudoaltermonas haloplanktis* TAH3a in complex with substrate and product. Biochemistry 45:4797–4807. doi: 10.1021/bi052193e

Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. Nat Protoc 2:953–971. doi: 10.1038/nprot.2007.131

Goubet F, Jackson P, Deery MJ, Dupree P (2002) Polysaccharide analysis using carbohydrate gel electrophoresis: a method to study plant cell wall polysaccharides and polysaccharide hydrolases. Anal Biochem 300:53–68. doi: 10.1006/abio.2001.5444

Hashimoto K, Yoshida M, Hasumi K (2011) Isolation and characterization of CcAbf62A, a GH62 α-L-arabinofuranosidase, from the Basidiomycete *Coprinopsis cinerea*. Biosci Biotechnol Biochem 75:342–345. doi: 10.1271/bbb.100434

Jordan DB, Bowman MJ, Braker JD, Dien BS, Hector RE, Lee CC, Mertens JA, Wagschal K (2012) Plant cell walls to ethanol. Biochem J 442:241–252. doi: 10.1042/BJ20111922

Kaur AP, Nocek BP, Xu X, Lowden MJ, Leyva JF, Stogios PJ, Cui H, Di Leo R, Powlowski J, Tsang A, Savchenko A (2014) Functional and structural diversity in GH62 α-L-arabinofuranosidases from the thermophilic fungus *Scytalidium thermophilum*. Microbiol Biotechnol. 8:419–433 doi: 10.1111/1751-7915.12168

Kellett LE, Poole DM, Ferreira LM, Durrant AJ, Hazlewood GP, Gilbert HJ (1990) Xylanase B and an arabinofuranosidase from *Pseudomonas fluorescens* subsp. *cellulosa* contain identical cellulose-binding domains and are encoded by adjacent genes. Biochem J 272:369–376.

Kimura I, Yoshioka N, Kimura Y, Tajima S (2000) Cloning, sequencing and expression of an α-L-arabinofuranosidase from *Aspergillus sojae*. J Biosci Bioeng 89:262–266.

Lange L, Sørensen HR, Hamann T (2006) New *Penicillium* arabinofuranosidase, used in dough and useful ethanol process, mashing process, and for producing feed composition. WO2006/125438-A1

Li W, Cowley A, Uludag M, Gur T, McWilliam H, Squizzato S, Park YM, Buso N, Lopez R (2015) The EMBL-EBI bioinformatics web and programmatic tools framework. Nucleic Acids Res 43:580–584. doi: 10.1093/nar/gkv279

Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42:D490–495. doi: 10.1093/nar/gkt1178

Ludwiczek ML, Heller M, Kantner T, McIntosh LP (2007) A secondary xylan-binding site enhances the catalytic activity of a single-domain family 11 glycoside hydrolase. J Mol Biol 373:337–354. doi: 10.1016/j.jmb.2007.07.057

Maehara T, Fujimoto Z, Ichinose H, Michikawa M, Harazono K, Kaneko S (2014) Crystal structure and characterization of the glycoside hydrolase family 62 α-L-arabinofuranosidase from *Streptomyces coelicolor*. J Biol Chem 289:7962–7972. doi: 10.1074/jbc.M113.540542

Marchler-Bauer A, Lu S (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res 39:D225–D229. doi: 10.1093/nar/gkq1189

Margolles-Clark E, Tenkanen M, Nakari-Setälä T, Penttilä M (1996) Cloning of genes encoding α-L-arabinofuranosidase and β-xylosidase from *Trichoderma reesei* by expression in *Saccharomyces cerevisiae*. Appl Environ Microbiol 62:3840–3846.

McCleary BV, McKie VA, Draga A, Rooney E, Mangan D, Larkin J (2015) Hydrolysis of wheat flour arabinoxylan, acid-debranched wheat flour arabinoxylan and arabino-xylo-oligosaccharides by β-xylanase, α-L-arabinofuranosidase and β-xylosidase. Carbohydr Res 407:79–96 doi: 10.1016/j.carres.2015.01.017

McKee LS, Peña MJ, Rogowski A, Jackson A, Lewis RJ, York WS, Krogh KBRM, Viksø-Nielsen A, Skjøt M, Gilbert HJ, Marles-Wright J (2012) Introducing endo-xylanase activity into an exo-acting arabinofuranosidase that targets side chains. Proc Natl Acad Sci U S A 109:6537–6542. doi: 10.1073/pnas.1117686109

McKie VA, Black GW, Millward-Sadler SJ, Hazlewood GP, Laurie JI, Gilbert HJ (1997) Arabinase A from *Pseudomonas fluorescens* subsp. *cellulosa* exhibits both an endo- and an exo- mode of action. Biochem J 555:547–555. doi: 10.1042/bj3230547

Mohun AF, Cook IJ (1962) An improved dinitrosalicylic acid method for determining blood and cerebrospinal fluid sugar levels. J Clin Pathol 15:169–180. doi: 10.1136/jcp.15.2.169

Nielsen JW, Kramhøft B, Bozonnet S, Abou Hachem M, Stipp SLS, Svensson B, Willemoës M (2012) Degradation of the starch components amylopectin and amylose by barley α-amylase 1: role of surface binding site 2. Arch Biochem Biophys 528:1–6. doi: 10.1016/j.abb.2012.08.005

Numan MT, Bhosle NB (2006) α-L-arabinofuranosidases: the potential applications in biotechnology. J Ind Microbiol Biotechnol 33:247–260. doi: 10.1007/s10295-005-0072-1

Nurizzo D, Turkenburg JP, Charnock SJ, Roberts SM, Dodson EJ, McKie VA, Taylor EJ, Gilbert HJ, Davies GJ (2002) *Cellvibrio japonicus* α-L-arabinanase 43A has a novel five-blade beta-propeller fold. Nat Struct Biol 9:665–668. doi: 10.1038/nsb835

Oudjeriouat N, Moreau Y, Santimone M, Svensson B, Marchis-Mouren G, Desseaux V (2003) On the mechanism of α-amylase. Eur J Biochem 270:3871–3879. doi: 10.1046/j.1432-1033.2003.03733.x

Pitkänen L, Virkki L, Tenkanen M, Tuomainen P (2009) Comprehensive multidetector HPSEC study on solution properties of cereal arabinoxylans in aqueous and DMSO solutions. Biomacromolecules 10:1962–1969. doi: 10.1021/bm9003767

Pitson SM, Voragen AG, Beldman G (1996) Stereochemical course of hydrolysis catalyzed by arabinofuranosyl hydrolases. FEBS Lett 398:7–11. doi: http://dx.doi.org/10.1016/S0014-5793(96)01153-2

Poutanen K (1988) An α-L-arabinofuranosidase of *Trichoderma reesei*. J Biotechnol 7:271–281. doi: 10.1016/0168-1656(88)90039-9

Ransom RF, Walton JD (1997) Purification and characterization of extracellular β-xylosidase and α-arabinosidase from the plant pathogenic fungus *Cochliobolus carbonum*. Carbohydr Res 297:357-364. doi: 10.1016/S0008-6215(96)00281-9

Sakamoto T, Ogura A, Inui M, Tokuda S, Hosokawa S, Ihara H, Kasai N (2010) Identification of a GH62 α-L-arabinofuranosidase specific for arabinoxylan produced by *Penicillium chrysogenum*. Appl Microbiol Biotechnol 90:137–146. doi: 10.1007/s00253-010-2988-2

Sakamoto T, Inui M, Yasui K, Hosokawa S, Ihara H (2013) Substrate specificity and gene expression of two *Penicillium chrysogenum* α-L-arabinofuranosidases (AFQ1 and AFS1) belonging to glycoside hydrolase families 51 and 54. Appl Microbiol Biotechnol 97:1121–1130. doi: 10.1007/s00253-012-3978-3

Schmidt A, Gu GM, Kratky C (1999) Xylan binding subsite mapping in the xylanase from *Penicillium simplicissimum* using xylooligosaccharides as cryo-protectant. Biochemistry 38:2403–2412. doi: 10.1021/bi982108l

Siguier B, Haon M, Nahoum V, Marcellin M, Burlet-Schiltz O, Coutinho PM, Henrissat B, Mourey L, O Donohue MJ, Berrin J-G, Tranier S, Dumon C (2014) First structural insights into α-L-arabinofuranosidases from the two GH62 glycoside hydrolase subfamilies. J Biol Chem 289:5261–5273. doi: 10.1074/jbc.M113.528133

Söding J, Biegert A, Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res 33:244–248. doi: 10.1093/nar/gki408

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular evolutionary genetics analysis

version 6.0. Mol Biol Evol 30:2725–2729. doi: 10.1093/molbev/mst197

Tsujibo H, Takada C, Wakamatsu Y, Kosaka M, Tsuji A, Miyamoto K, Inamori Y (2002) Cloning and expression of an α-L-arabinofuranosidase gene (*stxIV*) from *Streptomyces thermoviolaceus* OPC-520, and characterization of the enzyme. Biosci Biotechnol Biochem 66:434–438. doi: http://doi.org/10.1271/bbb.66.434

Van Laere KMJ, Voragen CHL, Kroef T, Van den Broek LAM, Beldman G, Vorage, Paper O (1999) Purification and mode of action of two different arabinoxylan arabinofuranohydrolases from *Bifidobacterium adolescentis* DSM 20083. Appl Microbiol Biotechnol 51:606–613. doi: 10.1007/s002530051439

Vandermarliere E, Bourgois TM, Rombouts S, Van Campenhout S, Volckaert G, Strelkov SV, Delcour JA, Rabijns A, Courtin CM (2008) Crystallographic analysis shows substrate binding at the -3 to +1 active-site subsites and at the surface of glycoside hydrolase family 11 endo-1,4-β-xylanases. Biochem J 410:71–79. doi: 10.1042/BJ20071128

Vincent P, Shareck F, Dupont C, Morosoli R, Kluepfel D (1997) New α-L-arabinofuranosidase produced by *Streptomyces lividans*: cloning and DNA sequence of the *abfB* gene and characterization of the enzyme. 852:845–852. doi: 10.1042/bj3220845

Wallner B, Elofsson A (2003) Can correct protein models be identified? Protein Sci 12:1073–1086. doi: 10.1110/ps.0236803.a

Wang W, Mai-Gisondi G, Stogios PJ, Kaur A, Xu X, Cui H, Turunen O, Savchenko A, Master ER (2014) Elucidation of the molecular basis for arabinoxylan-debranching activity of a thermostable family GH62 α-L-arabinofuranosidase from *Streptomyces thermoviolaceus*. Appl Environ Microbiol 80:5317–5329. doi: 10.1128/AEM.00685-14

610 **TABLES**

611 Table 1: Specific activities for *An*Abf62A-m2,3

Substrate	Specific activity (U/mg)
Wheat arabinoxylan	67.42 ± 4.53 (1.00)
Rye arabinoxylan	64.24 ± 1.82 (0.95)
Oat spelt xylan	49.14 ± 1.19 (0.73)
Birchwood xylan	n.d.
Barley β-glucan	n.d.
Sugar beet L-arabinan	1.43 ± 0.14 (0.02)
Linear L-arabinan	n.d.
Larch wood arabinogalactan	0.08 ± 0.01 (0.001)
Acacia tree gum arabic	0.25 ± 0.02 (0.003)
4-nitrophenyl α-L-arabinofuranoside	1.66 ± 0.24 (0.02)
4-nitrophenyl β-D-xylopyranoside	0.03 ± 0.01 (0.0004)
4-nitrophenyl α-L-arabinopyranoside	0.04 ± 0.01 (0.001)
A^3X	37 ± 1.1 (0.55)
A^2XX	59 ± 0.5 (0.88)
A^2XX+A^3XX (7:3)	80 ± 2.1 (1.19)
XA^2XX+XA^3XX (1:1)	80 ± 3.5 (1.19)
$A^{2+3}XX$	n.d.

612 n.d., not detected. Relative values are in parentheses. All experiments were done in triplicates.

613 Table 2: Kinetic parameters for hydrolysis of wheat arabinoxylan, sugar beet L-arabinan and 4-nitrophenyl α-L-arabinofuranoside by *An*Abf62A-m2,3 wild-type and

614 catalytic site (D28A, D136A and E188A) and putative SBS (W23A, Y44A and W23A/Y44A) mutants.

Wheat arabinoxylan	Wild-type	D28A	D136A	E188A	W23A	Y44A	W23A/Y44A
k_{cat} (s^{-1})	178 ± 26 (1.00)	0.64 ± 0.06 (0.00)	0.63 ± 0.09 (0.00)	n.d	52.11 ± 9.25 (0.29)	46.54 ± 3.16 (0.26)	80.21 ± 10.53 (0.45)
K_m ($mg \times ml^{-1}$)	4.90 ± 0.91 (1.00)	2.62 ± 0.05 (0.53)	1.57 ± 0.32 (0.32)	n.d.	2.35 ± 0.63 (0.48)	1.02 ± 0.13 (0.21)	6.93 ± 1.15 (1.41)
k_{cat}/K_m ($s^{-1} \times mM^{-1}$)	36.37 (1.00)	0.24 (0.01)	0.40 (0.01)	n.d	22.18 (0.61)	45.76 (1.26)	11.56 (0.32)
Specific activity ($U \times mg^{-1}$)	67.42 ± 4.53 (1.00)	0.52 ± 0.02 (0.01)	0.41± 0.02 (0.01)	0.04 ± 0.00 (0.00)	48.29 ± 6.98 (0.72)	45.82 ± 2.05 (0.68)	50.63 ± 3.26 (0.75)
K_i ($mg \times ml^{-1}$)	2.89 ± 0.58 (1.00)	-	6.0 ± 1.50 (2.08)	-	16.32 ± 8.51 (5.64)	11.89 ± 2.19 (4.11)	19.71 ± 7.68 (6.82)
Sugar beet L-arabinan	Wild-type	D28A	D136A	E188A	W23A	Y44A	W23A/Y44A
k_{cat} (s^{-1})	1.03 ± 0.03 (1.00)	-	-	-	0.73 ± 0.02 (0.71)	0.96 ± 0.02 (0.93)	0.81 ± 0.04 (0.79)
K_m ($mg \times ml^{-1}$)	15.63 ± 1.25 (1.00)	-	-	-	20.60 ± 1.72 (1.32)	12.03 ± 0.79 (0.77)	33.78 ± 3.47 (2.16)
k_{cat}/K_m ($s^{-1} \times mM^{-1}$)	0.07 (1.00)	-	-	-	0.04 (0.57)	0.08 (1.14)	0.02 (0.29)
Specific activity ($U \times mg^{-1}$)	1.43 ± 0.14 (1.00)	-	-	-	1.06 ± 0.08 (0.74)	0.77 ± 0.08 (0.54)	1.07 ± 0.01 (0.75)
4-nitrophenyl α-L-arabinofuranoside	Wild-type	D28A	D136A	E188A	W23A	Y44A	W23A/Y44A
k_{cat}/K_m ($s^{-1} \times mM^{-1}$)	0.26 ± 0.01 (1.00)	-	-	-	0.17 ± 0.00 (0.65)	0.21 ± 0.02 (0.81)	0.27 ± 0.01 (1.04)
Specific activity ($U \times mg^{-1}$)	1.66 ± 0.24 (1.00)	-	-	-	1.28 ± 0.03 (0.77)	2.42 ± 0.08 (1.46)	1.99 ± 0.12 (1.20)

615 n.d. - not measured. Relative values are in parentheses. All experiments were in triplicates.

Table 3: Relative activities on arabinoxylooligosaccharides for *An*Abf62A-m2,3 wild-type and mutants of the putative SBS

AXOS*	Wild-type	W23A	Y44A	W23A/Y44A
A^3X	1.00	0.04	0.03	0.04
$A^2XX+XA^3X+A^3XX$	1.00	0.12	0.10	0.11
XA^2XX+XA^3XX	1.00	0.25	0.18	0.23

All experiments were in triplicates.

FIGURE LEGENDS

Fig. 1 Structural homology model of *An*Abf62A-m2,3 overlayed with xylopentaose (cyan) from *Sc*Abf62A (PDB ID 3WN2) and arabinose (orange) from *Um*Abf62A (PDB ID 4N2R). Subsites are labelled according to McKee et al. (2012). The catalytic residues are light purple and the putative surface binding site residues in dark purple.

Fig. 2 Subsites and side chains shown to interact with xylooligosaccharides in crystal structures of *Sth*Abf62A (PDB ID 4O8O) (pink), *St*Abf62C (PDB ID 4PVI) (brown) *Um*Abf62A (PDB ID 4N2R) (green), *Pa*Abf62A (PDB ID 4N2Z) (salmon) and *Sc*Abf62A (PDB ID 3WN2) (yellow). Only side chains that differ from *An*Abf62A-m2,3 (grey) are included for the above mentioned. Xylopentaose (cyan) from *Sc*Abf62A (PDB ID 3WN2) and arabinose (orange) from *Sth*Abf62A (PDB ID 4O8O) are shown. Numbering refers to *An*Abf62A-m2,3.

Fig. 3 Affinity gel electrophoresis of *An*Abf62A-m2,3. A) 17 h run with 0.05 % wheat arabinoxylan, B) control (without polysaccharide) and with 0.001 % C) wheat arabinoxylan, D) oat spelt xylan, E) birchwood xylan, F) barley β-glucan, G) sugar beet L-arabinan, H) acacia tree gum arabic and I) larch arabinogalactan. Lane 1: marker; lane 2: wild-type; lane 3: W23A; lane 4: Y44A; lane 5: W23A/Y44A. The lower vertical line shows the migration of *An*AbfGH62A-m2,3 wild-type in the control gel without polysaccharide and the upper one shows a marker used to align the gels.

Fig. 4 Substrate hydrolysis curves by *An*Abf62A-m2,3 of A) wheat arabinoxylan, B) sugar beet L-arabinan, C) wheat arabinoxylan by catalytic site mutants and D) 4-nitrophenyl α-L-arabinofuranoside. *An*Abf62A-m2,3 wild-type (●), W23A (■), Y44A (○), W23A/Y44A (□), D28A (▲) and D136A (▼).

Fig. 5 Time course of hydrolysis by *An*Abf62A-m2,3 of AXOS (1:1 molar ratio of β-D-Xyl*p*-(1→4)-[α-L-Ara*f*-(1→2)]-β-D-Xyl*p*-(1→4)-β-D-Xyl*p*-(1→4)-β-D-Xyl*p* (A^2XX) and β-D-Xyl*p*-(1→4)-[α-L-Ara*f*-(1→3)]-β-D-Xyl*p*-(1→4)-β-D-Xyl*p*-(1→4)-β-D-Xyl*p* (A^3XX) by *An*Abf62A-m2,3 monitored by ^{1}H NMR spectroscopy. Peak area integration values are shown for the signals from 1,3-linked arabinofuranose (○), 1,2-linked arabinofuranose (●), and arabinose on β-furanose (▼), α-furanose (Δ), α-pyranose (■) and β-pyranose (□) forms, respectively.

Fig. 6 Close-up surface representation of *An*Abf62A-m2,3 putative surface binding site (SBS) situated Trp^{23} and Tyr^{44} (dark purple) with xylopentaose (cyan) from *Sc*Abf62A (PDB ID 3WN2) and arabinose (orange) from *Sth*Abf62A (PDB ID 4O8O). A) End-view from subsite +3NR on the substrate binding crevice, B) Side-view on the substrate binding crevice.

FIGURES

Fig. 1

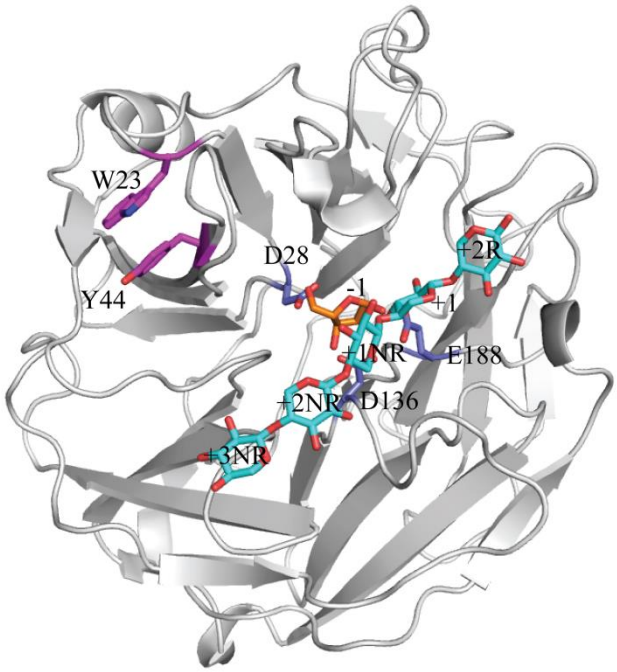


Fig. 2

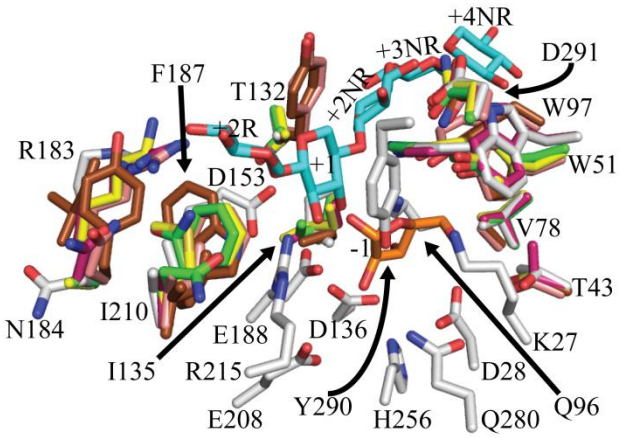
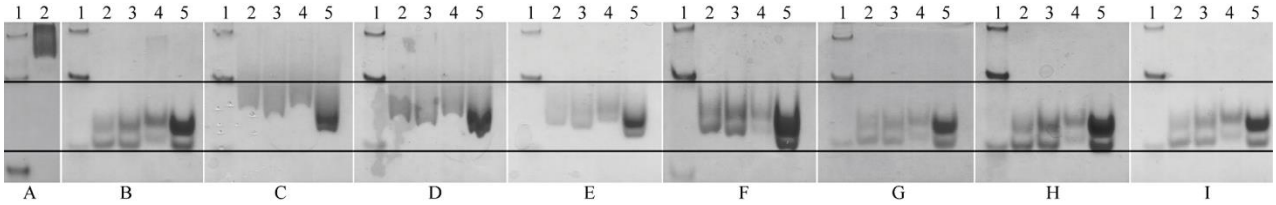
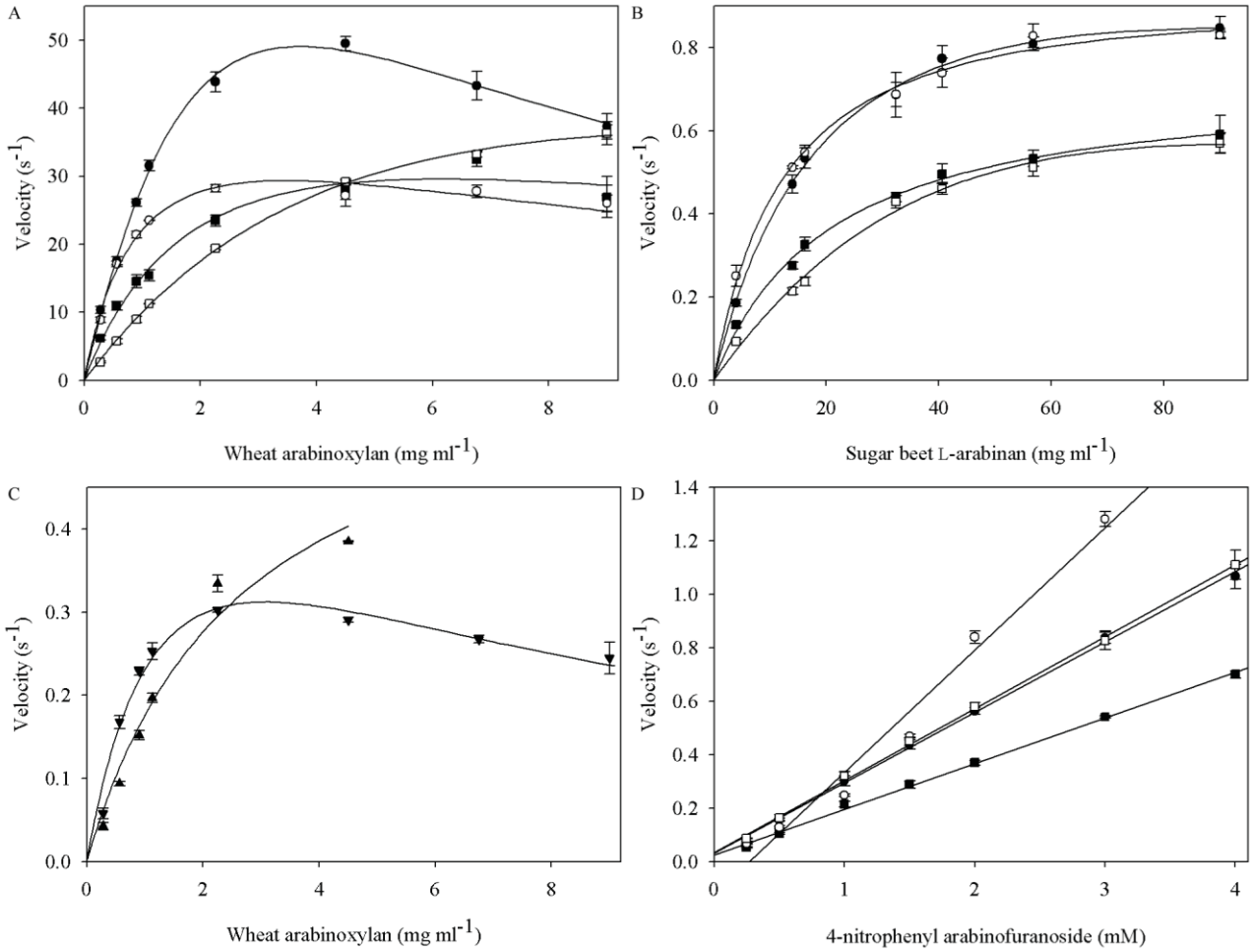

+4NR
+3NR
D291
F187
T132
+2NR
W97
+2R
R183
+1
W51
D153
V78
-1
T43
N184
I210
E188
D136
K27
I135
R215
D28
Y290
Q96
E208
H256
Q280

Fig. 3

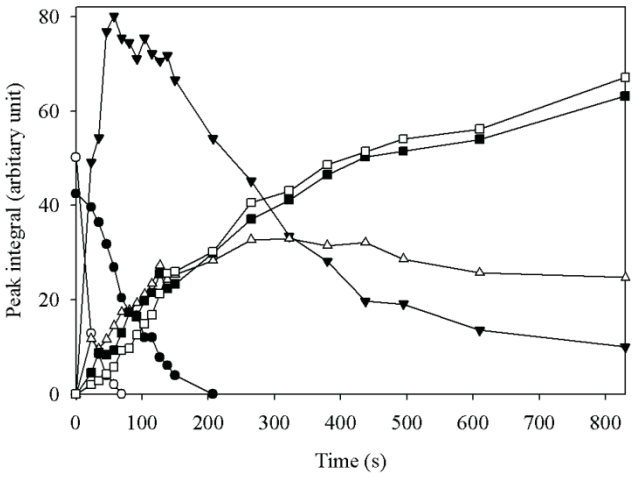

1 2 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5
A B C D E F G H I

780 Fig. 4

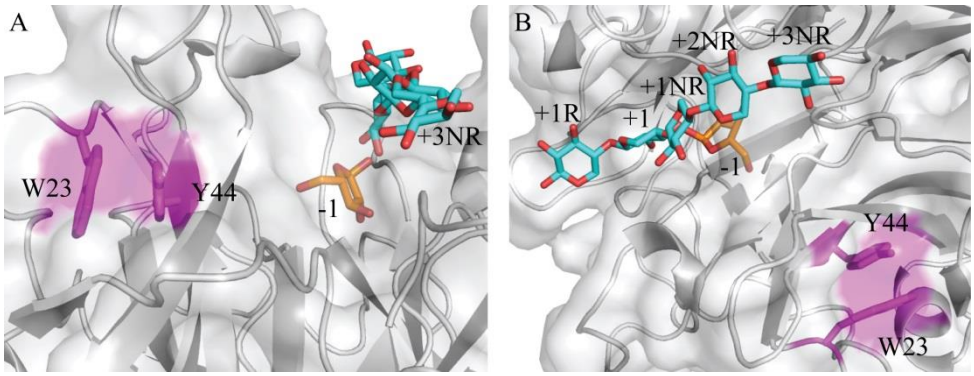


781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797

Fig. 5


Peak integral (arbitary unit)
Time (s)
0
20
40
60
80
0
100
200
300
400
500
600
700
800

827 Fig. 6



828
829
830
831
832