

ACCEPTED VERSION

This is the peer reviewed version of the following article:

Barbora Stratilová, Sergej Šesták, Jozef Mravec, Soňa Garajová, Zuzana Pakanová, Kristína Vadinová, Danica Kučerová, Stanislav Kozmon, Julian G. Schwerdt, Neil Shirley Eva Stratilová and Maria Hrmova

Another building block in the plant cell wall: Barley xyloglucan xyloglucosyl transferases link covalently xyloglucan and anionic oligosaccharides derived from pectin

The Plant Journal, 2020; 104(3):752-767

© 2020 Society for Experimental Biology and John Wiley & Sons Ltd

which has been published in final form at <http://dx.doi.org/10.1111/tpj.14964>

This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

PERMISSIONS

<https://authorservices.wiley.com/author-resources/Journal-Authors/licensing/self-archiving.html>

Wiley's Self-Archiving Policy

Accepted (peer-reviewed) Version

The accepted version of an article is the version that incorporates all amendments made during the peer review process, but prior to the final published version (the Version of Record, which includes; copy and stylistic edits, online and print formatting, citation and other linking, deposit in abstracting and indexing services, and the addition of bibliographic and other material.

Self-archiving of the accepted version is subject to an embargo period of 12-24 months. The standard embargo period is 12 months for scientific, technical, medical, and psychology (STM) journals and 24 months for social science and humanities (SSH) journals following publication of the final article. Use our [Author Compliance Tool](#) to check the embargo period for individual journals or check their copyright policy on [Wiley Online Library](#).

The accepted version may be placed on:

- the author's personal website
- the author's company/institutional repository or archive
- not for profit subject-based repositories such as PubMed Central

Articles may be deposited into repositories on acceptance, but access to the article is subject to the embargo period.

The version posted must include the following notice on the first page:

"This is the peer reviewed version of the following article: [FULL CITE], which has been published in final form at [Link to final article using the DOI]. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions."

The version posted may not be updated or replaced with the final published version (the Version of Record). Authors may transmit, print and share copies of the accepted version with colleagues, provided that there is no systematic distribution, e.g. a posting on a listserve, network or automated delivery.

There is no obligation upon authors to remove preprints posted to not for profit preprint servers prior to submission.

2 September 2021

<http://hdl.handle.net/2440/127393>

DR MARIA HRMOVA (Orcid ID : 0000-0002-3545-0605)

Article type : Original Article

Another building block in the plant cell wall: Barley xyloglucan xyloglucosyl transferases link covalently xyloglucan and anionic oligosaccharides derived from pectin

Barbora Stratilová^{1,2}, Sergej Šesták¹, Jozef Mravec³, Soňa Garajová^{1,4}, Zuzana Pakanová¹, Kristína Vadinová¹, Danica Kučerová¹, Stanislav Kozmon¹, Julian G. Schwerdt⁴, Neil Shirley⁴, Eva Stratilová¹ and Maria Hrmova^{4,5*}

¹ Institute of Chemistry, Centre for Glycomics, Slovak Academy of Sciences, Dúbravská cesta 9, SK-84538 Bratislava, Slovakia

² Faculty of Natural Sciences, Department of Physical and Theoretical Chemistry, Comenius University, Mlynská dolina, SK-842 15 Bratislava, Slovakia

³ Department of Plant and Environmental Sciences, University of Copenhagen, Thorvaldsensvej 40, 1871 Frederiksberg-C, Denmark

⁴ School of Agriculture, Food and Wine, University of Adelaide, Glen Osmond SA 5064, Australia

⁵ School of Life Sciences, Huaiyin Normal University, 223300 Huai'an, China

*To whom correspondence should be addressed; e-mail: maria.hrmova@adelaide.edu.au and maria.hrmova@hytc.edu.cn; Tel: +61 8 8313 7160

Running title: Substrate specificity of plant xyloglucan xyloglucosyl transferases.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/TPJ.14964](https://doi.org/10.1111/TPJ.14964)

This article is protected by copyright. All rights reserved

Key words: Clustered heatmap of expression profiles, evolution, glycoside hydrolase 16 family, homo- and hetero-transglycosylation, mutants, pectin fragment, RT-qPCR.

SUMMARY

We report on the homo- and hetero-transglycosylation activities of the HvXET3 and HvXET4 xyloglucan xyloglucosyl transferases (XET; EC 2.4.1.207) from barley (*Hordeum vulgare* L.), and the visualisation of these activities in young barley roots using Alexa Fluor 488-labelled oligosaccharides. We discover that these isozymes catalyse the transglycosylation reactions with the chemically defined donor and acceptor substrates, specifically with the xyloglucan donor and the penta-galacturonide [$\alpha(1-4)\text{GalAp}$]₅ acceptor – the homogalacturonan (pectin) fragment. This activity is supported by 3D molecular models of HvXET3 and HvXET4 with the docked XXXG donor and [$\alpha(1-4)\text{GalAp}$]₅ acceptor substrates at the -4 to +5 subsites in the active sites. Comparative sequence analyses of barley isoforms and seed-localised TmXET6.3 from nasturtium (*Tropaeolum majus* L.) permitted the engineering of mutants of TmXET6.3 that could catalyse the hetero-transglycosylation reaction with the xyloglucan/[$\alpha(1-4)\text{GalAp}$]₅ substrate pair, while wild-type TmXET6.3 lacked this activity. Expression data obtained by real-time quantitative PCR of *HvXET* transcripts and a clustered heatmap of expression profiles of the gene family revealed that *HvXET3* and *HvXET6* co-expressed but did not share the monophyletic origin. Conversely, *HvXET3* and *HvXET4* shared this relationship, when we examined the evolutionary history of 419 glycoside hydrolase 16 family members, spanning monocots, eudicots, and a basal Angiosperm. The discovered hetero-transglycosylation activity in HvXET3 and HvXET4 with the xyloglucan/[$\alpha(1-4)\text{GalAp}$]₅ substrate pair is discussed against the background of roles of xyloglucan-pectin heteropolymers and how they may participate in spatial patterns of cell wall formation and re-modelling, and affect the structural features of walls.

INTRODUCTION

Xyloglucans (XGs) were initially identified in nasturtium (*Tropaeolum majus*) and other plant seeds as reserve carbohydrates, although at that time little was known of their precise composition (Kooiman, 1957). In the following 60 years, intense research into XG chemistry and function has shown that XGs are the fundamental components of plant cell walls that form unique exoskeletal structures. XGs are hetero-polysaccharides composed of a (1,4)- β -glucan backbone substituted at C(O)6 with α -D-xylopyranosyl moieties or short chains of α -D-xylopyranosyl, β -D-

galactopyranosyl, α -L-fucopyranosyl residues (Hayashi, 1989; York *et al.*, 1990), or with galacturonic acid substituents replacing α -D-xylopyranosyl moieties (Peña *et al.*, 2012). XGs in *Poaceae* occur in lower levels in primary cell walls than in dicots (Vogel, 2008; Scheller and Ulvskov, 2010) but in both phyla, they share similar properties, in that they contain the repetitive blocks of the (1,4)- β -linked glucosyl residues substituted with α -D-xylopyranosyl residues, although they could include unbranched acetylated glucosyl residues. In *Poaceae*, the ratio of xylosyl:glucosyl residues equal around to 1:2 (Gibeaut *et al.*, 2005; Hsieh and Harris, 2009) and in sycamore wall XGs, galactosyl residues contain one or two *O*-acetyl groups *per* 2-linked galactosyl residues at C3, C4 or C6 carbons (Kiefer *et al.*, 1989).

XGs and other hemicelluloses form, along with cellulose and pectin (a linear chain of α -1,4-linked D-galacturonic acid is one of its components), the primary and conformationally flexible load-bearing XG-cellulose or XG-pectin networks that are believed to be pivotal for the primary (Keegstra *et al.*, 1973; Abasolo *et al.*, 2009) and secondary (Bourquin *et al.*, 2002) cell wall assembly. In the developing tensile wood, xyloglucan oligosaccharides are incorporated into new developing wall layers (Nishikubo *et al.*, 2007), and in tension wood fibers, xyloglucan linkages between the gelatinous layer of almost pure cellulose and secondary walls were implicated in the transmission of tensile stress (Gerttula *et al.*, 2015). While cellulose in the form of crystalline microfibrils cross-linked by hemicelluloses confers strength to cell walls, the cellulose-hemicellulose networks embedded in the pectin matrix glue the cell wall together and allow for its expansion (Nonogaki, 2019). It has been postulated that pectin may hinder the unfolding of XG chains and cause deformations of primary cell walls through the mechanical interactions of polymeric networks, as hypocotyl *Arabidopsis* cells elongate (Abasolo *et al.*, 2009). These XG-cellulose or XG-pectin network associations could be formed through backbone or side-chain residue hydrogen bonds, van der Waals and electrostatic interactions (Zykwinska *et al.*, 2008) or covalent linkages between XG and pectins (Cumming *et al.*, 2005; Popper and Fry, 2008; Cornuault *et al.*, 2014), cellulose or (1,3;1,4)- β -D-glucans (Hrmova *et al.*, 2007; Mohler *et al.*, 2013; Simmons *et al.*, 2015; Stratilová *et al.*, 2019; Viborg *et al.*, 2019). These more recent findings explain how the underlying mechanisms facilitate wall elongation and how expansins may contribute to these processes (Cosgrove, 2005; Abasolo *et al.*, 2009; Nonogaki, 2019).

The molecular mass and composition of XGs can be altered following their deposition into plant cell walls (Keegstra *et al.*, 1973) through the activities of xyloglucan endotransglycosylases (XETs), also known as xyloglucan xyloglucosyl transferases (EC 2.4.1.207) and xyloglucan

endohydrolases (XEHs, EC 3.2.1.151), forming the XTH group of enzymes. These widely distributed enzymes are catalogued in the glycoside hydrolase 16 (GH16) family of the Carbohydrate-Active enZYmes (CAZy) database (Lombart *et al.*, 2014) and can have either the XET activity, or both XET and XEH activities (Hayashi, 1989; Farkaš *et al.*, 1992; Fry *et al.*, 1992; Nishitani and Tominaga, 1992; Okazawa *et al.*, 1993). The reaction mechanism of XETs occurs in two steps. Firstly, the enzyme cleaves the (1,4)- β -D-glucan backbone of XG with the reducing-end fragment of the XG substrate diffusing away from the enzyme's surface. The non-reducing fragment of the polysaccharide remains covalently bound in the enzyme active site. In the next step, XET enzymes transfer the bound portion of the initial XG substrate onto the non-reducing end of another XG polysaccharide (Farkaš *et al.*, 1992; Fry *et al.*, 1992; Nishitani and Tominaga, 1992) or XG-derived or alternate XG-non-derived oligosaccharides (Ait Mohand and Farkaš 2006; Hrmova *et al.*, 2007, 2009; Simmons *et al.*, 2015; Shinohara *et al.*, 2017; Stratilová *et al.*, 2019). The initial polymer that is cleaved by the enzyme is referred to as a donor substrate, while the polysaccharide or oligosaccharide to which the cleaved product is transferred, is known as an acceptor substrate.

In this work, we focus on the barley HvXET3 and HvXET4 isozymes (Kaewthai *et al.*, 2010; Vaaje-Kolstad *et al.*, 2010a and 2010b) to define their homo- and hetero-transglycosylation activities, and extend our previous data on nasturtium TmXET6.3 (Stratilová *et al.*, 2019). We discover that the barley isoforms exhibited significant hetero-transglycosylation activities with chemically defined donor and acceptor substrates, explicitly with the fragment of homogalacturonan, the penta-galacturonide [α (1-4)GalAp]₅ as the acceptor. Based on structural analyses of the barley isoforms, we designed mutants of TmXET6.3 that catalysed the hetero-transglycosylation reaction with [α (1-4)GalAp]₅, and compare these activities with wild-type (WT) TmXET6.3 lacking this activity. The ability of barley isoforms to catalyse hetero-transglycosylations is reconciled with the structural models containing the docked donor XXXG hepta-saccharide/acceptor [α (1-4)GalAp]₅ substrate pair.

Moreover, we use real-time quantitative PCR (RT-qPCR) and the global barley RNA-sequence datasets to reveal the expression profiles of *HvXET3*, *HvXET4*, and *HvXET6* transcripts in various tissues and generate a clustered heatmap of the *HvXET* gene family that defines their co-expression patterns. We reveal that *HvXET3* and *HvXET6* expression coincided with the hetero-transglycosylation activities of the XG/[α (1-4)GalAp]₅ substrate pair, but not with HvXET5 that lacked this activity. These data suggested that certain barley XET enzymes with the latter substrate

specificity may co-evolved. These findings are supported by the phylogenetic analyses of 419 sequences classified in the GH16 family across monocots, eudicots, and a basal Angiosperm. In conclusion, the discovery of the hetero-transglycosylation activity of the XG/[α (1-4)GalAp]₅ substrate pair catalysed by HvXET3 and HvXET4 is discussed in the context of the XG-pectin networks thought to be involved in spatial patterns of the cell wall assembly, and how they may affect the structure of the cell wall.

RESULTS

The localisation of non-specific XET enzymes in cell walls of young barley roots using Alexa Fluor 488-labelled oligosaccharides

To localise homo- or hetero-transglycosylation XET activities in young barley roots obtained from 4-days old seedlings, we dissected root hair apices and incubated with the Alexa Fluor 488 dye (A488)-labelled nona-saccharide XXFG (XXFG-A488) and the mixture of [α (1-4)GalAp]₇-A488/[α (1-4)GalAp]₈-A488 substrates (Figure 1). After around 30-min incubations of root tips with the fluorescently-labelled oligosaccharides, we imaged the elongation zones of these tissues, which showed the incorporation of XXFG-A488 and [α (1-4)GalAp]₇-A488/[α (1-4)GalAp]₈-A488 in the epidermal cell walls of young roots (Figures 1c, 1d). In control incubations, lacking A488-labelled oligosaccharides (Figure 1a) or containing GlcN-A488 (Figure 1b) or when the α (1-4)GalAp]₇-A488/[α (1-4)GalAp]₈-A488 mixture was added to the Proteinase K pre-treated roots (Figure 1e), the fluorescence signals were compromised. The UV-autofluorescence emanating from cell wall components coincided with that of XXFG-A488 and the α (1-4)GalAp]₇-A488/[α (1-4)GalAp]₈-A488 mixture (Figure 1f - overlay) indicating the fluorescently labelled acceptors were incorporated in cell wall polymers.

N-glycosylation patterns of HvXET3, HvXET4 and HvXET6

Mass spectrometry analyses of N-glycosylation patterns of HvXET3, HvXET4, and HvXET6 deglycosylated by the peptide-N-glycosidase F indicated that the three XET isoforms expressed in *Pichia pastoris* (*Komagataella phaffii*) contained the Hex₈HexNAC₂, Hex₉HexNAC₂ and Hex₁₀HexNAC₂ N-glycans as the main determinants, and other poly-hexosylated HexNAC₂ species (Supplemental Figure S2; top three panels); these are presumably attached to the sole N85 (HvXET3 and HvXET4) or N86 (HvXET6) sites (numbering omits signal peptides). The material

secreted by *Pichia* cells with an empty vector contained poly-hexosylated species (Supplemental Figure S2; bottom panel).

Physicochemical parameters of HvXET3 and HvXET4

We determined thermal and pH optima of HvXET3 and HvXET4, their thermal stability in the 1-40 °C temperature ranges, and pH-dependent activity tolerance in the pH range of 4-8. We established that the pH optima were 6.0 for both XET isoforms and the thermal activity optima were at 28 °C (Supplemental Figures S3a, 3b), while thermal stabilities and the pH-dependent activity tolerance of HvXET3 and HvXET4 in buffers varied (Supplemental Figures S4a-4d). Here, 50% of thermal unfolding occurred after around 1.5 and 1.25 hours for HvXET3 and HvXET4 at 40 °C, respectively, while pH-dependent activity tolerance was considerably higher for HvXET4 than that for HvXET3 in the pH ranges 4-8; we observed that at pH 4, HvXET3 lost the activity after two days, while HvXET4 was no longer active after three days (Supplemental Figures S4c, S4d). We also noticed that after four freeze/thaw cycles HvXET3 lost approximately 50% of the activity, while nearly 90% of the activity remained in HvXET4 (Supplemental Figure S4e).

Homo- and hetero-transglycosylation activities, and substrate specificity of HvXET3, HvXET4, and HvXET6

We used affinity-purified barley XET isoforms (Supplemental Figure S1) and high-performance liquid chromatography (HPLC) with an inert size-exclusion chromatography column coupled to fluorescent and refractive index detectors to analyse transglycosylation activities. Reactions were monitored *via* a gradual increase of fluorescence in the XG or hydroxyethyl cellulose (HEC) regions (Stratilová *et al.*, 2019), where XG and HEC we used as donors, and a variety of oligosaccharides with defined chemistries as acceptors (Figure 2, Supplemental Scheme S1).

The HvXET3 and HvXET4 isoforms behaved as non-specific XET enzymes, and catalysed effective homo- and hetero-transglycosylation reactions with varying rates using XG and HEC fragments and a range of neutral and charged sulforhodamine (SR)-labelled acceptors, composed of tri-, tetrosyl and pentosyl oligosaccharides or other acceptors with a range of chemistries (Figures 2, 3, S5, Supplemental Scheme S1). Control reactions with enzymes inactivated by boiling did not show any activity with tested substrate pairs (not shown).

During investigations of the substrate specificity of HvXET3 and HvXET4 with XG or HEC donors and XG-derived and alternate XG-non-derived oligosaccharides (Figures 3, S5), we noticed that the activities of HvXET4 were always significantly higher than those of HvXET3. The highest transglycosylation activity was observed with the XG/XG-derived oligosaccharide substrate pairs in both isoforms (Supplemental Figure S5a), while those with the HEC/XG-derived or XG/XG-non-derived oligosaccharide substrate pairs were lower (Figures 3b, S5b-S5d). The highest hetero-transglycosylation rates catalysed by HvXET3 and HvXET4 with the XG donor were with $[\alpha(1-4)\text{GalAp}]_5$, the second-fastest with $[\beta(1-4)\text{GlcP}]_4$, while the weakest rates were detected with the arabino-galacto-oligosaccharides $\{[\beta(1-4)\text{GalP}]_4[\text{Araf}]_5\}_n$ (Figure 3a). The hetero-transglycosylation rates of HvXET3 and HvXET4 using HEC and a variety of XG-non-derived acceptors were the most effective with $[\beta(1-4)\text{GlcP}]_4$, while a negligible rate was observed with $\{[\beta(1-4)\text{GalP}]_4[\text{Araf}]_5\}_n$ (Figure 3b). We identified that the key parameter that affected the hetero-transglycosylation rates of HvXET3 and HvXET4 with the XG/(1,4)- β -D-linked oligosaccharides pairs were the degree of polymerisation of acceptors (Supplemental Figure S5c), whereas the key parameter with the XG/mixed-linkage oligoglucoside pairs was the position of the (1,3)- β -linkage in acceptors (Supplemental Figure S5d).

The data of HvXET3 and HvXET4 contrasted with those of HvXET6, which overall showed substantially faster homo- or hetero-transglycosylation rates than the former two isoforms (Figures 3, S5; insets), but notably HvXET6 had a low rate with XG and $[\alpha(1-4)\text{GalAp}]_5$, and almost no activity with $\{[\beta(1-4)\text{GalP}]_4[\text{Araf}]_5\}_n$, $[\beta(1-4)\text{ManP}]_6$ and $[\beta(1-4)\text{ManP}]_6[\text{GalP}]_2$. Remarkably, the reaction rates of HvXET6 with the XG/XLLG and XG/XXLG substrate pairs occurred at faster rates than with XG/XXXG (Supplemental Figure 5a; inset) or with XG/ $[\beta(1-4,1-3,1-4)\text{GlcP}]_4$ C compared to the XG/ $[\beta(1-3,1-4,1-4)\text{GlcP}]_4$ A or XG/ $[\beta(1-4,1-4,1-3)\text{GlcP}]_4$ B pairs (Supplemental Figure 5d; inset). Again, the position of the (1,3)- β -linkage in acceptors was the key, such that these substrates could be accommodated in the enzyme active sites.

3D protein modelling of HvXET3 and HvXET4, and engineering of TmXET6.3

To understand the spatial disposition of amino-acid residues that play key roles in the substrate specificity of the HvXET3 and HvXET4 isoforms, we constructed 3D homology models (Figure 4), using the crystal structure of poplar PttXET16A as the template (PDB accession 1UN1). When the sequences of poplar PtXET16A (Q8GZD5), nasturtium TmXET6.3 (Q07524), barley HvXET3 to HvXET6 (P93671, P93672, P93668, B1P1S7) were aligned using PROMALS3D (Pei *et al.*,

2008), we observed the clustered segregation of entries into three groups pointing at their evolutionary relationships. While HvXET3, HvXET4, and HvXET6 formed a cluster of three entries, nasturtium XET and barley HvXET5 with the poplar enzyme formed two other stand-alone clusters (Figure 4a). Notably, the secondary structure (ss) element dispositions of sheets (showed in blue fonts in Figure 4a) and loops showed similar distributions when examined in the sequence-aligned (Figure 4a) or 3D (Figures 4b, 4c) arrangements. Absolute conservation and similarity of amino-acid (aa) residues (on the scale 9-6; *cf.* Figure 4a) indicated that the positions of sheets in all structures were highly conserved in agreement with their β -jellyroll fold that forms the distinctive architecture of XET enzymes (Johansson *et al.*, 2004; Vaaje-Kolstad *et al.*, 2010a). The root-mean-square-deviation (RMSD) values between superposed poplar PtXET16A (272 residues) and 3D models of respective HvXET3 (263 residues) or HvXET4 (255 residues) were 0.39 Å and 0.70 Å, and the RMSD values between superposed TmXET6.3 (250 residues) and respective HvXET3 or HvXET4 were 0.52 Å and 0.70 Å (Figure 4). This substantiated the conclusion that proteins were structurally highly similar, as expected.

When the XXXG-donor and $[\alpha(1-4)\text{GalAp}]_5$ -acceptor substrate pair was docked in the enzyme active sites of HvXET3 and HvXET4 using the Induced Fit Docking protocol (Schrödinger release 2015-2) the XXXG donor was positioned in the -4 to -1 subsites and the $[\alpha(1-4)\text{GalAp}]_5$ acceptor in the +1 to +5 subsites (Figures 4 b and 4c) (the nomenclature of subsites according to Biely *et al.*, 1991 and Davies *et al.*, 1997). We could identify several separations at 2.6 Å to 3.5 Å between Y67, W166, and XXXG in both isoforms while in HvXET4 also S164 contributed; $[\alpha(1-4)\text{GalAp}]_5$ was bound by N96 and E106, and in HvXET3 also by H94. Additional positively charged residues contributed to the binding of negatively charged $[\alpha(1-4)\text{GalAp}]_5$, such as C-terminal R261 in HvXET34 and K260 in HvXET4. The latter residues are absent in HvXET5 in equivalent positions, which is catalytically incompetent with the XG/ $[\alpha(1-4)\text{GalAp}]_5$ substrate pair, but are present in TmXET6.3 that did not catalyse this reaction. We concluded that these C-terminal positively charged residues may contribute, but are not required for the hetero-transglycosylation with the XG/ $[\alpha(1-4)\text{GalAp}]_5$ substrate pair. In both isoforms, the overlapped positions of W75 and Y110 in TmXET6.3 (magenta) and H75 and R110 in HvXET3 or HvXET4 (navy blue) (Figures 4b and 4c) co-located at the edges of β -jellyroll scaffolds, where the β -sheet secondary structures were likely to be confined.

The covalent bond during transglycosylation reactions in HvXET3 and HvXET4 would be mediated by E77 (catalytic nucleophile) and E81 (catalytic acid/base) situated at the -1/+1 subsites

(Figures 4 b and 4c). These residues (at the C α carbons) were at 6.4-6.7 Å and 11.8-12.0 Å separations from W75 and Y110 in HvXET3 or HvXET4, respectively. However, in both HvXET3 or HvXET4 the catalytic acid/base E81 interacted tightly at 2.9-3.4 Å with the COOH-group of [α (1-4)GalAp]₅ and E77 made close contacts of 2.8-2.9 Å with the reducing end glucose at the C2 of XXXG. Initial insights in the structure of the XXXG~[α (1-4)GalAp]₅ covalent complex formed in both HvXET isoforms suggested that the two substrates could be linked through either (1,4)- or (1,3)- β -D-glycosidic linkages (data not shown).

The emergence of the hetero-transglycosylation activity with the XG/[α (1-4)GalAp]₅ substrate pair by mutant TmXET6.3

The findings with the barley HvXET3 and HvXET4 isoforms (Figure 4) led us to create single and double mutants of TmXET6.3 that could catalyse the hetero-transglycosylation reaction using the XG/[α (1-4)GalAp]₅ substrate pair (Figure 5). Through site-directed mutagenesis, we prepared single mutants W75H and Y110R, in addition to the W75H/Y110R double mutant of TmXET6.3 (Supplemental Tables S1, S2). In all instances, these non-conservative variations led to the emergence of the hetero-transglycosylation activity with the XG/[α (1-4)GalAp]₅ substrate pair in TmXET6.3, whereas the WT enzyme could not catalyse this reaction (Figure 5). It was rewarding to see that the activity of the double mutant W75H/Y110R showed a nearly two-fold increase, compared to the activity of the single W75H or Y110R mutants of TmXET6.3. It was noteworthy, that amongst all investigated barley isoforms, HvXET5 – a well-characterised barley XET enzyme (Hrmova *et al.*, 2007) – was unable to catalyse the hetero-transglycosylation reaction with the XG/[α (1-4)GalAp]₅ substrate pair (Figure 5), while HvXET3 and HvXET4, and HvXET6 (Hrmova *et al.*, 2009) could mediate this reaction to various extents (Figure 5). In barley isoforms, these data pointed to the requirement of a characteristic residue signature that is needed for the hetero-transglycosylation activity with the XG/[α (1-4)GalAp]₅ substrate pair.

Comparative transcriptomics of the *HvXET* gene family based on RT-qPCR

RT-qPCR-derived normalised expression levels of selected *HvXET* gene transcripts, using gene-specific primers listed in Supplemental Table S3 in a variety of barley tissues, indicating that the expression levels of *HvXET3*, *HvXET4*, and *HvXET6* genes diversified (Figure 6). The highest levels of *HvXET3* (given in arbitrary units – AU; Figure 6) were observed in the first leaf base; this is a relatively high expression level, *e.g.* similar to that of *HvCsIF6* implicated in the

biosynthesis of (1,3;1,4)- β -D-glucans (Burton *et al.*, 2008). This was followed by a dramatic decline of the expression levels in all other examined tissues, for example in flowers at anthesis (Figure 6a), while in roots, grain, flowers (after various days of pollination – DAP) and coleoptiles (after various days of germination – DAG), the *HvXET3* expression levels were low (Figure 6a). Conversely, the *HvXET4* transcripts showed comparable levels in at least four tissues of spikes, anther at pre-anthesis, and in the first leaf base and root tips (Figure 6a). Since the expression levels of *HvXET6* were in general even lower than those of *HvXET3* and *HvXET4*, these data were plotted on a different scale, allowing us to compare the expression levels between all examined *HvXET* genes (Figures 6a, 6b). The highest expression levels of *HvXET6* were identified in flowers at anthesis, followed by those in the peduncle and the first leaf base, while other organs/tissues showed declining expression levels (Figure 6b).

Members of the *HvXET* gene family form co-expression clusters

Based on protein sequences analyses of the barley genome (IBSC version 2; sourced by the International Barley Genome Sequencing Consortium) (Mayer *et al.*, 2012) we identified a total of 45 sequences. However, nine of those contained only the PF00722 (Glyco_hydro_16) domain, while other sequences were incomplete or of poor quality and were not considered. Thus, we identified 36 *XTH* genes that represented 36 barley XTH enzymes classified in the GH16 family of CAZy (Lombard *et al.*, 2014), containing the PF00722 and PF06955 (XET_C) functional domains listed in the Pfam-A v29 database (El-Gebali *et al.*, 2019) and the highly conserved catalytic motif DE/DhE (*l*-aliphatic, *h*-hydrophobic) required for the XTH activity.

The gene expression of these 36 XTH barley enzymes was examined from eight tissue RNA-sequence data sets. The resultant heatmap utilising MeV 4.9.0 (Saeed *et al.*, 2003) showed that the *HvXET3* and *HvXET6* transcripts formed a co-expressing cluster ($r^2 \Rightarrow 0.98$), while *HvXET4* and *HvXET5* clustered independently (Supplemental Figures S6a, S6b; Supplemental Data Set S1). We found that in the internodes the *HvXET3* and *HvXET6* co-expression coincided with two other *XET* genes (blue and orange lines in Supplemental Figure 6b). The quantitative data of the *HvXET3/HvXET6* co-expression cluster and the two other *HvXET* genes are detailed in Supplemental Data Set S1 (red types highlighted in grey).

Large-scale molecular phylogeny analyses of XTH enzymes

To reconstruct the evolutionary history of the barley XTH enzymes within the GH16 family, we generated a best-known maximum likelihood tree using RAxML (Stamatakis, 2014) consisting of 419 GH16 protein sequences across six monocots, four eudicots and a basal Angiosperm *Amborella trichopoda*, including 36 barley XTH sequences (Supplemental Data Set S2). Non-specific HvXET3 (UniProt accession P93671), HvXET4 (P93671), HvXET5 (P93668) and HvXET6 (B1P1S7 or F2DM52) and TmXET6.3 (V5ZEF7) were also included along with all experimental XET/XEH crystal structures (yellow nodes in Figures 7, S7) of the poplar PttXET16 (PDB accession 1UN1; Johansson *et al.*, 2004) and the apo-structure of TmNXG1 endo-xyloglucanase 1 (2UWA; Baumann *et al.*, 2007) and the TmNXG1-DYNIIG mutant in complex with XLLG (2VH9; Mark *et al.*, 2008). The data indicated that there was a significant gene duplication either before the emergence of angiosperms or earlier in their evolutionary history. The RAxML tree resolves HvXET3 and HvXET4 isoforms (both in orange nodes) that exhibited the relatively high hetero-transglycosylation activity with the $[\alpha(1-4)\text{GalAp}]_5$ acceptor to be monophyletic (Figures 5, 7, S7). On the other hand, TmXET6.3 (brown node) and HvXET5 (orange node), the latter co-located on the tree with PttXET16 (yellow node), clustered independently as expected (Figures 7, S7) – these enzymes did not catalyse the hetero-transglycosylation reactions with the XG/ $[\alpha(1-4)\text{GalAp}]_5$ substrate pair. While the HvXET6, isoform (orange node) that showed a low activity with the XG/ $[\alpha(1-4)\text{GalAp}]_5$ pair, clustered in an independent position on the tree (Figures 7, S7), the XEH enzymes 2UWA and 2VH9 (overlapping yellow nodes), did not show any evolutionary links to the barley, nasturtium and poplar XET enzymes (Figures 7, S7). To indicate the evolutionary divergence between XTH and broad-specific endo-(xylo)glucanases, in the tree we highlighted the position of *Populus trichocarpa* Potri.002G153200 (lacking the functional XET_C domain but retaining the catalytic DE/DhE motif) that was identified in CAZy as a broad-specific endo-(xylo)glucanase (EC 3.2.1.151 and EC 3.2.1.73) (Supplemental Figure S7; cyan node).

DISCUSSION

The HvXET5 and HvXET6 isoforms isolated from germinated barley seedlings were the first identified XET enzymes with homo- and hetero-transglycosylation activities (Hrmova *et al.*, 2007; Hrmova *et al.*, 2009), although their *in-planta* localisation remained unclear. In this work, we focused on HvXET3 (UniProt accession P93671) and HvXET4 (P93671) (Kaewthai *et al.*, 2010; Vaaje-Kolstad *et al.*, 2010a and 2010b) isoforms annotated by Schünmann *et al.* (1997), who

observed their expression in elongating leaves. Before substrate specificities were explored, we examined the physicochemical parameters of HvXET3 and HvXET4, which were similar except their pH tolerance, which was significantly higher for HvXET4. These differences did not reflect the occupancy of their single N-glycosylation sites at N85 (Figure 4a; green box), positioned on loops inter-connecting β -sheets (Figure 4a), where we identified as the most prevalent Hex₈HexNAc₂, Hex₉HexNAc₂ and Hex₁₀HexNAc₂ N-glycans (Supplemental Figure S2). To our best knowledge, this is the first report of N-glycan site identifications in recombinant XET enzymes produced in *Pichia*. This compared to two N-acetyl β -D-glucosamine and up to five hexosyl and pentosyl residues resolved in the native barley exo-hydrolase by X-ray crystallography (Varghese *et al.*, 1999). HvXET3 and HvXET4 also lacked predicted O-glycosylation sites for GlcNAc at 35 Ser or Thr residues, evaluated by DictyOGlyc 1.1 and the 0.4 thresholds (Gupta *et al.*, 1999). We suggest that the higher pH and freeze/thaw cycle stability of HvXET4 compared to HvXET3 is not due to the occupancies of glycosylation sites but is related to a more compact protein fold.

We identified hetero-transglycosylation activities in all investigated barley isoforms with the XG-donor/ α (1-4)GalAp]₅-acceptor substrate pair (Figures 3, S4), except previously characterised HvXET5 (Hrmova *et al.*, 2007) and nasturtium TmXET6.3 (Stratilová *et al.*, 2019) (Figure 5). To find out if this hetero-transglycosylation activity could be localised to barley tissues, we fine-sectioned young barley roots and localised it to the cell walls, using the Alexa Fluor 488-labelled [α (1-4)GalAp]₇/ α (1-4)GalAp]₈ mixture (Figure 1); this validated the RT-qPCR analyses that *HvXET4* and *HvXET3* transcripts occurred in root tips (Figure 6).

To reconcile the differences in substrate specificity between the two groups of enzymes: HvXET3-HvXET4-HvXET6 *versus* HvXET5-TmXET6.3 (Figure 5), we built 3D models and docked the XXXG-donor/ α (1-4)GalAp]₅-acceptor substrate pair in the active sites to define dispositions of the residues responsible for the hetero-transglycosylation activity (Figure 5); these models were compared with those of HvXET6 (Hrmova *et al.*, 2009) and TmXET6.3 (Stratilová *et al.*, 2019). The benefits of conformational characteristics of carbohydrate substrate pairs bound in the active sites of examined barley XET enzymes could now be envisaged as crystal structures are unavailable. This comparison revealed that at least two residues H75 (localised near XXXG) and R110 (localised near [α (1-4)GalAp]₅) confined to the edges of the β -jellyroll fold, could be critical to interactions with this substrate pair situated between the -4 to +5 subsites. In addition, the W166, S164, Y67 residues at the XXXG site (-4 to -1 subsites), and H94, N96, E106 (and C-

terminal K260 and R261) at the ($[\alpha(1-4)\text{GalAp}]_5$) site (+1 to +5 subsites) made closed contacts with substrates. In both enzymes, the catalytic E77 and E81, together with D79 located at the -1/+1 subsites and near H75 (Figures 4b, 4c) interacted with XXXG/ $[\alpha(1-4)\text{GalAp}]_5$. Energetic contributions to each subsite occupied by XXXG/ $[\alpha(1-4)\text{GalAp}]_5$ need to be yet assessed, although the 3D models indicated that the major inputs would arrive from the -1/+1 subsites and acceptor sites. In both barley isoforms, we observed tighter binding in the positive subsites, where the incoming $[\alpha(1-4)\text{GalAp}]_5$ binds, compared to the negative subsites where the XXXG donor is retained (Figures 4b, 4c) as previously observed by Mark *et al.* (2008) with TmNXG1 and PttXET16.

We compared residues of plant XET enzymes in positions 75 and 110 (numbering omits signal peptides) that underlined the transfers of XG fragments to $[\alpha(1-4)\text{GalAp}]_5$ (Figure 4), and found that in HvXET isoforms, position 75 was equivalent to the His residue (Figure 4; Supplemental Table S5). However, HvXET5 with H83 (H75 is the equivalent position in HvXET3, HvXET4) could not catalyse the hetero-transglycosylation reaction, nor could TmXET6.3 that had equivalent W75. On the contrary, position 110 was equivalent to the Arg residue in HvXET3, HvXET4, and HvXET6 that catalysed the reaction with the XG/ $[\alpha(1-4)\text{GalAp}]_5$ substrate pair, while in HvXET5 and TmXET6.3, position 110 corresponded to Tyr. Here, we also analysed position 108, as this Q108 residue was shown to have the key importance for the hetero-transglycosylation activity of TmXET6;3 (Stratilová *et al.*, 2019). The Gln residue was present in XET enzymes except for HvXET5 and PttXET16A that had Arg in this position. Based on this comparison we predict that PttXET16A similarly to HvXET5 may not mediate the reaction with the XG/ $[\alpha(1-4)\text{GalAp}]_5$ pair. We also evaluated the relative representation of XET enzymes with the potential to utilise the XG/ $[\alpha(1-4)\text{GalAp}]_5$ pair and analysed 3,394 entries from UniProtKB (Supplemental Figure S8). Only around 29% (988 entries) had His and Gln in positions 75 and 108, while the most entries (1,886 or 55.6%) had other equivalent residues. The R110 and Q108 incidence rate was 8% (271 entries), while the H75, Q108, and R110 residue frequency was 7.3% (249 entries).

Structural and sequence comparisons (Figure 4; Supplemental Table S5) proposed replacing the non-conservative W75 and Y110 in TmXET6.3 to generate W75H or Y110R single and W75H/Y110R double mutants in the expectation to be able to catalyse the reaction with the XG/ $[\alpha(1-4)\text{GalAp}]_5$ pair (Figure 5). It turned out that the single and double mutants could catalyse this reaction, and that the double mutant showed a nearly two-fold activity increase compared to single mutants (Figure 5). This broadened the substrate specificity in TmXET6.3 mutants;

although, it could not be excluded that the XG/[α (1-4)GalAp]₅ activity in mutant TmXET6.3 could be affected by other residues.

Clustering of *HvXET* transcripts disclosed the heatmap of the *HvXET* gene family, where *HvXET3* and *HvXET6* showed co-expression, and *HvXET4* and *HvXET5* clustered singly (Supplemental Figure 6b, Supplemental Table S4, Supplemental Data Set S2). Amongst XG-active enzymes in the GH16 family of nearly 23,000 sequences, plant XET enzymes are assigned to the GH16_20 subfamily (Viborg *et al.*, 2019). Here, XTH enzymes (*E*-value threshold of 10⁻³⁵ to 10⁻⁶⁵) exhibit either XG endo-transglycosylase or XG endo-hydrolases activity and differ by the presence of the C-terminal extension XET_C (Behar *et al.*, 2018) allowing the former group catalysing transglycosylation reactions. This extension represents an archetypal addition to a bacterial or an early-lineage eukaryotic GH16 sequence, which resulted from duplications typical for XG endo-transglycosylases contrary to XG endo-hydrolases (Behar *et al.*, 2018). These authors also assigned the barley GH16 enzyme lacking XET_C (NCBI: AK356496.1; genomic locus: HORVU.MOREX.r2.4HG0344700.1) that we confirm exists in the IBSC barley genome – such sequences (termed XET/XTH-like) could be also found in the *Aegilops tauschii* (NCBI: XP_020193613.1) and *Oryza sativa* (NCBI: XP_025879674.1; genomic locus: LOC_Os03g02610.1) genomes. It would be worthy to look more deeply at the phylogenetic origin of these entries and how they relate to XTH enzymes.

In the barley IBSC genome, we identified 36 XTH sequences that are classified in the GH16 family of CAZy. This database lists entries based on structural folds and sequence homology (Lombart *et al.*, 2014), and assigns catalytic activity, although, the evidence for it is often circumstantial or absent. Hence, the substrate specificity definition of any enzyme is fundamental as it could point to yet an unidentified role. To resolve phylogenetic relationships of investigated barley XET isoforms, we built the Randomised Accelerated Maximum Likelihood (RAxML) (Stamatakis, 2014) tree of 419 GH16 protein sequences across six monocots and the basal Angiosperm *Amborella trichopoda*, although the specificity of these enzymes remains often unassigned. Our analyses suggested that the GH16 family experienced multiple gene duplication events before the emergence of Angiosperms (Figures 7, S6, Supplemental Data Set S2). Here, *HvXET3* with *HvXET4*, and *HvXET5* were separated by one of these splits – meaning they diversified at least 200 million years ago. As the *HvXET3/HvXET4* group remained deeply unresolved and contained the majority of diversifications that the GH16 family underwent, there were also six to seven uncharacterised and diverged clades that existed before the emergence of

Angiosperms, for example in *Physcomitrella* and *Pinus taeda*. There are no 3D structures available in the HvXET3/HvXET4 group – the largest group in the examined plant GH16 enzymes, though those of the two most conserved clades were resolved (Figures 7, S7).

As discussed above, the HvXET3 and HvXET4 isoforms (orange nodes; relatively high hetero-transglycosylation activity with XG/[$\alpha(1-4)$ GalAp]₅) clustered together, but also clustered with TmXET6.3 (brown node; not active with XG/[$\alpha(1-4)$ GalAp]) (Figures 7, S7). Conversely, HvXET5 (orange node; not active with XG/[$\alpha(1-4)$ GalAp]₅) co-located in the tree with PttXET16 (yellow node), while HvXET6 (orange node; low hetero-transglycosylation activity with XG/[$\alpha(1-4)$ GalAp]₅) branched early in the clade of HvXET3 and HvXET4, although this needs to be further examined. Thus, the subtle differences in the configuration of residues in the active sites together with loop dispositions (Figure 4) are likely to change binding patterns at acceptor binding sites and favour hetero-transglycosylation with the XG/[$\alpha(1-4)$ GalAp]₅ pair. Similar assumptions concerning the intricacy between strictly hydrolytic and transglycosylating enzymes were drawn for poplar and nasturtium XG-active enzymes (Mark *et al.*, 2008).

We suggest that the significance of hetero-transglycosylations using XG and de-esterified pectin [$\alpha(1-4)$ GalAp]₅ fragments in plant cell walls by non-specific XET enzymes may be related to structural roles of pectins in primary and secondary cell walls. Although walls are rigid exoskeletons, they also must be versatile with dynamic properties to fulfill multiple roles during plant cell growth and development. This flexibility could result from the complex interactions between major cell wall polysaccharides such as cellulose, (glucurono-arabino) xylans, (gluco)mannans, mixed linkage glucans, XGs and pectins, and structural proteins including expansins, polyphenolic compounds, and inorganic molecules. It was shown that the major hemicellulose in type I cell walls of dicots and noncommelinid monocots are XGs and that these walls contain homogalacturonan and rhamnogalacturonan-I pectic polysaccharides (Caffall and Mohnen, 2009). On the contrary, type II cell walls of grasses and commelinid monocots contain (glucurono-arabino) xylans and pectins in lower amounts (Carpita and Gibeaut, 1993; Vogel, 2008; Scheller and Ulvskov, 2010).

It was shown that XGs and pectins in plant tissues occur in the form of covalent (Cumming *et al.*, 2005; Cornuault *et al.*, 2014) or non-covalent (Zykwinska *et al.*, 2008; Nonogaki, 2019) XG-pectin heteropolymers, and that XG-pectin heteropolymers represent rather minor components in walls. However, these heteropolymers may offer alternative mechanisms for modulation of the wall structure through *e.g.* the activities of broad-specific XET enzymes (Park and Cosgrove,

2015), and as shown in this work. Nonetheless, the significance of hetero-transglycosylation reactions for the cell wall structure needs to be further tested, *e.g.* by mutant knock-out plants lacking this enzyme activity. The fundamental complexities between the major structural polysaccharides in monocot and dicot plants have been outlined (Carpita and McCann, 2000; Thompson, 2005; Cosgrove, 2018; Bulone *et al.*, 2019; Penning *et al.*, 2019), yet there is room for investigations of specificities of enzymes that mediate wall polysaccharide synthesis and remodelling. To this end, it is of importance to examine substrate specificity/poly- or multi-specificity of XET enzymes that catalyse homo- and hetero-transglycosylation reactions described here. This knowledge is prone to offer new insights into the complexity of plant cell walls and its mechanics. Our work suggests that there is yet ‘another building block (brick) in the plant cell wall’ that has the potential to have a role in the cell wall structure and underlying bio-synthetic mechanisms.

METHODS

***In situ* microscopic analyses of young barley roots incubated with Alexa Fluor 488-labelled oligosaccharides**

The mixture of hepta-galacturonic acid ($[\alpha(1-4)\text{GalAp}]_7$) and octa-galacturonic acid ($[\alpha(1-4)\text{GalAp}]_8$) oligosaccharides (catalogue number GAT112) (Elicityl, Crolles, France), XXFG (GLU1150) (Elicityl) and β -D-glucosamine (GlcN) (Sigma-Aldrich, St. Louis, MO, USA) were conjugated with Alexa Fluor™ 488 hydroxylamine (A30629) (Thermo Fisher Scientific, Waltham, MA, USA) as described (Mravec *et al.*, 2014). Barley seeds (*Hordeum vulgare* L.) were sterilised for 10 min in a bleach solution (household 3-8% w/v sodium hypochlorite), washed twice in water, let soaking in water for 1 hour and germinated on a sterile wetted Whatman filter paper in a plastic chamber at ambient (approximately 20 °C) temperature for four days. Around 3 mm-long young root apical segments were dissected and incubated in 0.025 M MES-KCl buffer, pH 5.7 with fluorescent (Alexa Fluor™ 488-labelled) probes at a 1:300 (v/v) dilution from stocks of 1 mg/ml. Roots were pre-treated with Proteinase K (Invitrogen, Carlsbad, CA, USA) at the concentration of 0.2 mg/ml for 30 min at ambient temperature of approximately 20 °C in the phosphate-buffered solution, pH 7.4 (0.0018 M KH_2PO_4 , 0.01 M Na_2HPO_4 , 0.0027 M KCl, 0.137 M NaCl). After 30 min incubations with labelled oligosaccharides, the whole roots with root hairs were washed twice in 0.025 M MES-KCl buffer, pH 5.7, and mounted in the water on glass slides containing a cavity. Zones with elongated root hairs (approximately two mm from the apex) were

scanned using the LSCM Leica SP5 microscope (Leica, Wetzlar, Germany) with a UV diode (405 nm) and an argon laser (488 nm) using the same settings for all samples. Images were processed using GIMP II (GNU Image Manipulation Program) software for contrast improvement.

Substrates for enzyme reactions

The source of polysaccharide substrates is described by Stratilová *et al.* (2019). Oligosaccharides derived from arabinogalactan, arabinoxylan, galactomannan, laminarin, pullulan, xylan polysaccharides were prepared as described (Kosík *et al.*, 2010). Unlabeled cello-, manno-, xyloglucan- and (1,3;1,4)- β -D-glucan-derived oligosaccharides were purchased from Megazyme. Pentagalacturonic acid [α (1-4)GalAp]₅ was a gift from Anna Malovíková (Institute of Chemistry of Slovak Academy of Sciences, Slovakia). All SR-labelled acceptor substrates were prepared as described (Kosík and Farkaš, 2008; Stratilová *et al.*, 2019). Abbreviations, descriptions, and chemical structures of commercial and purified acceptor oligosaccharides are defined in Figure 2 and Supplemental Scheme S1. Concentrations of labelled oligosaccharides were calculated based on the molar absorption coefficient for Lissamine rhodamine B sulfonyl chloride, $\epsilon_{566} = 85000 \text{ M}^{-1} \text{ cm}^{-1}$ (Anaspec, CA, USA).

Enzyme activity assays

Heterologously expressed HvXET enzymes were prepared as described (Hrmova *et al.*, 2010a; Kaewthai *et al.*, 2010). Here, secreted enzymes from *Pichia pastoris* cultures were dialysed and purified by Immobilized Metal Affinity Chromatography using imidazole gradients. Fractions with XET activities were desalted, concentrated, and stored in aliquots at -20 °C; enzymes were stable for at least for 12 months (Hrmova *et al.*, 2010a). Heterologously expressed WT and mutant TmXET6.3 were dialysed into 0.1 M ammonium acetate buffer, pH 5.5, and used as enzyme sources. Enzyme activity assays were performed using the high-performance liquid chromatography (HPLC) method with fluorimetric detection. XG or HEC donor substrates, SR-labelled oligosaccharide acceptor substrates, and enzyme preparations were suspended in 0.1 M ammonium acetate buffer, pH 5.5, in the volumetric ratio of 5:1:4; the final concentrations of donor and SR-labelled oligosaccharides acceptor substrates were 1.5% (w/v) and 25 μM , respectively. Mixtures were incubated at 25 °C for various time intervals placed directly in vials in an autosampler of the Dionex UltiMate™ 3000 HPLC device (Thermo Fischer Scientific, Waltham, MA, USA) equipped with the Dionex UltiMate™ FLD-3100 fluorescence (Thermo

Fischer Scientific) and the RefractoMax 521 refractive index (Thermo Fischer Scientific) detectors. Analyses were performed through isocratic size-exclusion chromatography on the 7.8 mm x 300 mm TSKgel G3000 SWXL column (TosoHaas, Tokyo, Japan). Elution proceeded with 0.1 M ammonium acetate, pH 5.7 containing 20% (v/v) aqueous acetonitrile ($\phi_r = 0.2$) at a flow rate 0.5 ml/min. The fluorescence detector was programmed using respective 530 nm and 575 nm excitation and emission wavelengths. Chromeleon 6.80 software (Thermo Fischer Scientific) was used for device control and data acquisition. All assays were performed in technical triplicates. Standard deviations were calculated *via* Excel in Microsoft Office Professional 2016.

The modified quantitative ‘dot-blot’ method (Fry, 1997) was used for pH optima and stability, and thermal optima and stability of enzymes, as follows. Reaction mixtures containing XG, XXLG-SR, and enzyme preparations (in the same volumetric and concentration ratios as those described for the HPLC method) were terminated by the addition of an equal volume of 70% (v/v) formic acid. The volumes of 5 μ l aliquots of reaction mixtures were spotted onto Whatman 3MM paper pieces corresponding in size and shape to the 96-well microtitration ELISA plate and dried in hot air. The unbound XXLG-SR was removed by washing in 60% (v/v) ethanol containing 5% (v/v) formic acid for 12 h, followed by repeated washing in 60% (v/v) ethanol. After drying of plates, the intensity of fluorescence on the paper was quantified using a microplate reader equipped with a fluorescence detector using the Mithras² LB 943 multimode reader (Berthold Technologies, Bad Wildbad, Germany) with a filter for SR. All assays were performed in technical triplicates with standard errors not exceeding 10%, calculated in Excel in Microsoft Office Professional 2016.

N-Deglycosylation of HvXET3, HvXET4, and HvXET6 isoforms

Purified XET enzyme preparations containing approximately 150 μ g of proteins were mixed with 40 μ l of 10 mM Tris-HCl buffer, pH 7.5 containing 0.1% (w/v) SDS and incubated with 10 mM dithiothreitol and 25 mM iodoacetamide for 45 min, following the standard protein reduction/alkylation protocol (Shevchenko *et al.*, 2006). To release N-glycans, samples were incubated with 1 unit of the peptide-N-glycosidase F (Roche, Basel, Switzerland) at 37 °C for 16 h. N-glycans were isolated by non-porous graphitised carbon solid-phase extraction (SPE) (Supelclean ENVI-Carb SPE) and desalted by reverse phase SPE (LiChroprep RP18, 25-40 μ m) (Hykollari *et al.*, 2017). To increase signal intensities, free N-glycans were permethylated (Palmigiano *et al.*, 2018) and analysed by the UltrafleXtreme MALDI ToF/ToF mass spectrometer

(Bruker Daltonics, MA, USA) in the reflectron positive ion mode using as a matrix solution 2% (w/v) 2,3-dihydroxybenzoic acid in 30% (v/v) acetonitrile with the addition of 10 mM NaOH to unify the adduct formation (Pazitna *et al.*, 2020). Raw data were processed by FlexAnalysis (Bruker Daltonics, MA, USA) and GlycoWork Bench software (Ceroni *et al.*, 2008).

pH and temperature optima, and stabilities of HvXET3 and HvXET4

The fluorimetric method was used for all analyses using the modified 'dot blot' method as described above, including donor and acceptor substrate concentrations. For pH optima and stability, 0.1 M citrate-phosphate buffers, pH 4-7, and 0.1 M phosphate buffers, pH 7-8 were used for donor substrate preparations. Enzyme sources were filtered cultivation media with expressed enzymes (Hrmova *et al.*, 2010a; Stratilová *et al.*, 2019) that were diluted in a 1:1 ratio (by volume) in specified buffers. Enzyme activities during all incubations were determined at 25 °C for 30 min in 24 h intervals. Values for pH optima were expressed in % relative to 100% activity at pH 6.0. Values for pH stability were related to the activity determined on the first day of incubations. For thermal optima and stability, filtered cultivation media with expressed enzymes were diluted in a 1:1 ratio (by volume) in 0.1 M ammonium acetate buffer, pH 5.5. These preparations were pre-incubated (0 h, 2 h, 4 h, 6 h) at the 1-40 °C temperature range (1 °C, 5 °C, 10 °C, 15 °C, 20 °C, 25 °C, 30 °C, 35 °C, 40 °C). Activities were assayed as described above. Values for thermal optima and stability were related to the activities determined at 0 h at a given thermal optimum. Freeze/thaw stability profiles were determined after five freezing at -20 °C (1 h)/thaw (10 min at 20 °C) cycles. Activities were expressed as relative activities to 100%. Assays were performed in three technical triplicates. Standard deviation values were calculated in Excel in Microsoft Office Professional 2016.

SDS-PAGE and protein concentration determination

The purity of barley XET enzymes, and WT and single and double mutants of TmXET6.3 were determined on 10% (w/v) SDS-PAGE gels performed as described (Hrmova *et al.*, 2007); low Molecular Mass Standards (Sigma) were used as standards. Protein concentrations were estimated using the Qubit™ protein assay (Life Technologies, Carlsbad, CA, USA) and the Qubit fluorometer 2.0 (Invitrogen).

Generation of TmXET6.3 mutants

To generate single W75H and Y110R, and double W75H/Y110R mutants, point mutations were introduced by the QuikChange site-directed mutagenesis kit, using the manufacturer protocol and pJET-C'XET specific primers (Supplemental Tables S1 and S2). All other cloning procedures used were as described by Stratilová *et al.* (2019). All sequences obtained by QuikChange site-directed mutagenesis were confirmed by automated Sanger sequencing in both directions. Plasmids were isolated using the Plasmid Midi Kit (Qiagen) and linearised by *SacI* before transformation into *P. pastoris*.

Expression of nasturtium and barley XET proteins and TmXET6.3 mutants

P. pastoris strain GS115 was used as the host to express barley XET isozymes (Hrmova *et al.*, 2010a; Kaewthai *et al.*, 2010), and WT and mutant TmXET6.3 (Stratilová *et al.*, 2019).

3D protein modelling and docking of donor and acceptor substrates in HvXET3 and HvXET4

3D models of HvXET3 and HvXET4 were constructed using the SWISS-MODEL server that uses the ProMod3 modelling engine (Waterhouse *et al.*, 2018) and the poplar PttXET16A structural template (PDB accession 1UN1; Johansson *et al.*, 2004), rather than the TmNXG1 endo-xyloglucanase 1 (PDB accession 2UWA; Baumann *et al.*, 2007). PttXET16A has the XET mode of action, while TmNXG1 is the XEH enzyme. Sequence identities, determined with the EMBOSS-Needle Global alignment algorithm (Needleman and Wunsch, 1970) of barley XETs to PttXET16A were between 51-52%, while those to TmNXG1 were between 37-38%. The coordinates of the XXXG donor were taken from the crystal structure of the TmNXG1-DYNIIG mutant in complex with XLLG (2VH9; Mark *et al.*, 2008) after removing two-terminal Gal residues. The coordinates of [α (1-4)GalAp]₅ were built using the GLYCAM-web server tools (<https://dev.glycam.org/>). The LigPrep tool (Schrödinger Release 2015-2: Protein Preparation Wizard) was used for the preparation of HvXET3 and HvXET4, and donor and acceptor substrates for docking into the enzyme active sites that proceeded with Glide (Schrödinger Release 2015-2) and the Extra Precision protocol (Friesner *et al.*, 2006). Substrates were docked in the active sites using core constraints of ligands positioned in the 1UMZ protein-ligand complex (Johansson *et al.*, 2004). Donor and acceptor substrates were re-fitted using Induced Fit Docking (Schrödinger Release 2015-2). The flexibility of residues in HvXET3 and HvXET4 was set to be within a maximum distance of 7 Å from substrates. The most favourable enzyme-donor-acceptor complex

poses were chosen based on Induced Fit docking scores (Schrödinger Release 2015-2). Graphical images of protein-ligand complexes were created in PyMol GSLS version 4.50 (Schrödinger, USA), and RMSD values were calculated using SSM (secondary-structure matching) in CCP4mg that performs a locally optimised superposition after a global superposition (Krissinel and Hendrick, 2004).

RT-qPCR analyses

RT-qPCR analyses were carried with gene-specific primers listed in Supplemental Table S3, according to Burton *et al.* (2008).

Identification of GH16 sequences in the barley genome

The barley IBSC genome (version 2) was downloaded from the Ensembl Plants website (<http://plants.ensembl.org/index.html>) and uploaded onto the CLC Genomics Workbench ver9.5.2 (Qiagen, Aarhus A/S www.clcbio.com). The genome was appended with genes, mRNA entries, and coding sequences annotations. Protein sequences collected from the IBSC genome were analysed (Finn *et al.*, 2014) using the Pfam-A v29 database in the CLC genomics workbench to generate the list of proteins and associated Pfam entries. The GH16 CAZY proteins contained the central PF00722 (Glyco_hydro_16) and the shorter C-terminal PF06955 (XET_C) Pfam domains. All sequences were curated using the FGENESH+ gene prediction tool (Solovyev, 2007). A total of 45 sequences was identified in the IBSC genome, however, nine of those contained only the PF00722 domain. Thus, we considered 36 *XTH* genes.

Barley RNA-sequence analyses in tissues

Reads for the RNA-sequence study ERA197495 (Mayer *et al.*, 2012) were downloaded using the CLC genomics Workbench from the Sequence Read Archive (<https://submit.ncbi.nlm.nih.gov/subs/sra/>). Reads were quality trimmed and assembled to the barley IBSC genome. The expression for each tissue replicate was calculated in transcripts *per* million (TPM). The experiment was normalised by scaling and the normalised means of tissue replicates were used in the analyses.

The clustered analysis of the expression profiles of the GH16 family

The clustered analysis of the expression profiles of 36 *XTH* genes identified as described above, was carried out using the Multiexperiment Viewer (Saeed *et al.*, 2003). Hierarchical clustering was carried out on the expression data with Pearson Correlation using the absolute distance and average linkage clustering.

Large-scale molecular phylogeny analyses of selected GH16 family enzymes: reconstruction of the evolutionary origin of barley XET/XEH enzymes

Representative sequences of the GH16 family from six monocot genomes: *Hordeum vulgare*, *Oryza sativa*, *Brachypodium distachyon*, *Setaria italica*, *Sorghum bicolor* and *Ananas comosus*; four eudicot genomes: *Arabidopsis thaliana*, *Solanum lycopersicum*, *Populus trichopoda* and *Glycine max*, and the basal Angiosperm *Amborella trichopoda*, were obtained from Phytozome v12 (Goodstein *et al.*, 2012) according to the presence of the PF00722 (Glyco_hydro_16) sequence Pfam motif. Pfam is a widely used and comprehensive conserved protein family database derived from UniProt Reference Proteomes (El-Gebali *et al.*, 2019). Multiple sequence alignment was built using GUIDANCE2 (Sela *et al.*, 2015; Landan and Graur, 2008; Penn *et al.*, 2010) and highly gapped sites were manually filtered, but otherwise, no column removal was performed. Phylogenetic trees were built using the Randomised Axelerated Maximum Likelihood version rapid bootstrapping algorithm 8 (RAxML) (Stamatakis, 2014) with the WAG protein substitution model, GAMMA rate heterogeneity, and the autoMRE bootstrap criterion. In total, five replicate RAxML analyses were performed and the tree with the highest GAMMA-based likelihood was selected as final.

DATA AVAILABILITY

All data can be found within the manuscript and its Supporting Information. The nucleotide sequences of XET enzymes are available in GenBank under the accessions: TmXET6.3 (HF968473), HvXET3 (X93174), HvXET4 (X93175), HvXET5 (X91659) and HvXET6 (EU247793). HvXET3 and HvXET4 structural models in complex with the XXXG/[α (1-4)GalAp]₅ substrate pair are available in the Protein Model DataBase (PMDb) under the accessions PM0082622 and PM0082623, respectively.

ACKNOWLEDGEMENTS

This work was supported by the 2/0058/16 and 2/0137/20 grants from VEGA (Slovakia) to ES and SK, the APP0075 DoktoGrant (Slovakia) to BS, the Villum Foundation 00017489 grant to JM and the DP120100900 grant from the Australian Research Council (Australia) to MH. We acknowledge the funding from the University of Adelaide (Australia) to NS, JGS and MH, and the Huaiyin Normal University (China) to MH. We thank IBH Wilson from the Universität für Bodenkultur (Austria) for providing the pPICZ α -His/FLAG plasmid and AJ Harvey (Yeditepe University, Turkey) for assistance. The authors thank the anonymous reviewer(s) for their contribution to the peer review of this work.

AUTHOR CONTRIBUTIONS

JM, BS, and DK performed in planta experiments. SŠ constructed plasmids and selected production clones. ES, SG, and MH purified proteins and performed biochemical characterisations. BS, ES, and KV quantified enzyme activities, and ZP completed N-glycan analyses. BS and SK built 3D molecular models and docked ligands, and MH generated structural graphics. BS and ES conducted bioinformatics analyses, and JGS performed large-scale phylogeny explorations. NS finalised RT-qPCR, RNA-sequence analyses, and identified GH16 entries in the barley genome. ES and MH conceived and designed experiments. MH wrote the manuscript. All co-authors discussed the data and contributed to writing.

CONFLICT OF INTEREST

The authors declare they do not have any conflicts of interest.

SUPPORTING INFORMATION

Supporting Information may be found in the online version of this article.

Supplemental Figure S1. SDS-PAGE profiles of HvXET3, HvXET4, and HvXET6 purified by Immobilized Metal Affinity Chromatography. Protein loadings in lane 2 (15 μ g), lane 3 (20 μ g), and lane 4 (2 μ g) were quantified by the QubitTM assay and visualised by Coomassie Brilliant Blue. Molecular masses of standards (St.) are specified in kilodalton values.

Supplemental Figure S2. MALDI ToF/ToF spectra of N-glycans released by the peptide-N-glycosidase F from HvXET3, HvXET4, and HvXET6 isoforms (top three respective panels)

expressed in *Pichia*, and secreted to the media. Three major Hex₈HexNAc₂ and Hex₉HexNAc₂ Hex₁₀HexNAc₂ N-glycans, and other poly-hexosylated HexNAc₂ species (purple font) were identified in HvXET isoforms. The bottom panel refers to the spectrum of poly-hexosylated species (red font) secreted from *Pichia* cells with an empty vector.

Supplemental Figure S3. The pH (a) and temperature (b) dependences of HvXET3, and HvXET4, that were determined fluorometrically with the XG/XLLG-SR pair. Values are expressed in % relative to 100% activity at 25 °C or pH 6.0. Error bars represent standard deviation values for three replicates, calculated *via* Excel in Microsoft Office Professional 2016.

Supplemental Figure S4. Thermal stability (a, b), pH-dependent activity tolerance (c, d), and stability during freeze/thaw cycles (e) in HvXET3 and HvXET4. Thermal stability was tested in the 1-40 °C temperature range, activity tolerance in the 4-8 pH range, and stability during five freeze/thaw cycles. Stability in the 4-4.5 and 6-8 pH ranges is plotted separately for clarity. Error bars represent standard deviation values for three replicates, calculated *via* Excel in Microsoft Office Professional 2016.

Supplemental Figure S5. Homo- and hetero-transglycosylation activities catalysed by HvXET3 and HvXET4 with different substrate pairs. (a), Enzyme activities determined with XG (a) and HEC (b), as donors and a variety of XG-derived acceptor oligosaccharides. Inset in panel (a) indicates activities with HvXET6. (c), Enzyme activities determined with XG as the donor and [β(1-4)Glc_p]₃₋₅. (d), Enzyme activities were determined with XG as the donor and [β(1-3,1-4,1-4)Glc_p]₄ A, [β(1-4,1-4,1-3)Glc_p]₄ B, [β(1-4,1-3,1-4)Glc_p]₄ C as the acceptors. Insets in panels (a) and (d) indicate activities with HvXET6. Abbreviations, descriptions, and chemical structures of acceptors are defined in Figure 2 and Supplemental Scheme S1. The average of three technical replicates (n = 3) is plotted with standard deviation values, calculated *via* Excel in Microsoft Office Professional 2016.

Supplemental Figure S6. A clustered heatmap of RNA sequences of the *HvXET* gene family in eight barley tissues as described in Mayer *et al.* (2012) (a), and (b) co-expression of the *HvXET3* and *HvXET6* cluster (light and dark blue lines; $r^2 = >0.95$) with two other *XET* genes (lines coloured in shades of orange) mainly in internodes (four days after germination). The expression

is calculated as % of maximal transcripts *per* million (TPM) from all tissues. Abbreviations: Leaf = Shoots from seedlings (10 cm shoot stage), Root = (10 cm shoot stage), Internode = Developing tillers, third internode (42 days after pollination), Inflorescence 5 mm = Young developing inflorescences (5 mm), Inflorescence 15 mm = Developing inflorescences (10-15 mm), Caryopsis 5 DAP = Developing grain (five days after pollination), Caryopsis 5 DAP = Developing grain (15 days after pollination), Embryo 4 DAG = embryo (four days after germination).

Supplemental Figure S7. Randomised Accelerated Maximum Likelihood (RAxML) (Stamatakis, 2014) tree of 419 GH16 protein sequences across six monocots (green nodes), four eudicots (blue nodes) and the basal Angiosperm *Amborella trichopoda* (magenta nodes). HvXET3 (UniProt accession P93671), HvXET4 (P93671), HvXET5 (P93668), and HvXET6 (B1P1S7 or F2DM52) are in orange nodes and TmXET6.3 (V5ZEF7) in a brown node. The *Populus trichocarpa* broad-specific endo-(xylo)glucanase Potri.002G153200 (EC 3.2.1.151 and EC 3.2.1.73) listed in CAZy is represented in a cyan node. Experimentally determined PttXET16 XET (PDB 1UN1), TmNXG1 XEH apo (2UWA), and TmNXG1-DYNIIG XEH in complex with XLLG (2VH9) are in yellow nodes. Bootstrap support values for deep nodes are indicated at bipartitions. The list of sequences and the FASTA alignment files used to generate the RAxML tree are included in Supplemental Data Set S2.

Supplemental Figure S8. The relative representation of XET enzymes with the potential to utilise [$\alpha(1-4)\text{GalAp}$]₅ as the acceptor substrate. Large-scale bioinformatics analyses of amino-acid residue positions of available UniProtKB XET entries (3,394) indicated the presence of both H75 and Q108 in 988 (29.1%), both R110 and Q108 in 271 (8.0%) and of all three residues H75, Q108 and R110 in 249 (7.3%) respective cases. The occurrence of XET enzymes lacking these residues was found in 1,886 cases (55.6%). Residue numbering corresponds to those of HvXET3, HvXET4, and TmXET6.3, without signal peptides.

Supplemental Table S1. Plasmids used for the generation of TmXET6.3 mutants *via* site-directed mutagenesis.

Supplemental Table S2. Oligonucleotides used for the generation of mutants using site-directed mutagenesis. Mutagenic base substitutions used for site-directed mutagenesis are in bold.

Supplemental Table S3. Barley IBSC (version 2) genome locus and MorexV2 locus of control and *HvXET* genes investigated by RT-qPCR using gene-specific primers. PCR product sizes in base pairs and optimal acquisition temperatures of analysed genes are also indicated.

Supplemental Table S4. Expression levels of the *HvXET* gene family (36 members) in seven barley tissues, including those of *HvXET3*, *HvXET4*, *HvXET5*, and *HvXET6* investigated in this work.

Supplemental Table S5. Amino-acid residues in positions 75, 108, and 110, influencing the ability of XET enzymes to catalyse the transfer of XG fragments onto [$\alpha(1-4)$ GalAp]₅.

Supplemental Scheme S1. Abbreviations, descriptions, and chemical structures of commercial and isolated acceptors used in this work.

Supplemental Data Set S1. Co-expression matrix of the *HvXET* gene family (36 members) in seven barley tissues, including *HvXET3*, *HvXET4*, *HvXET5*, and *HvXET6* (highlighted in bold red fonts) investigated in this work. Co-expressing *HvXET* genes to the *HvXET3* and *HvXET6* gene cluster (light and dark blue lines; $r^2 \Rightarrow 0.95$) are shaded in grey.

Supplemental Data Set S2. The list of sequences of the GH16 family (sheet 1) and the FASTA alignment file (sheet 2) used for the generation of the RAxML phylogenetic tree.

REFERENCES

- Abasolo, W., Eder, M., Yamauchi, K., Obel, N., Reinecke, A., Neumetzler, L., Dunlop, J.W., Mouille, G., Pauly, M., Höfte, H. and Burgert, I.** (2009) Pectin may hinder the unfolding of xyloglucan chains during cell deformation: implications of the mechanical performance of *Arabidopsis* hypocotyls with pectin alterations. *Mol. Plant* **2**, 990-999.
- Ait Mohand, F. and Farkaš, V.** (2006) Screening for hetero-transglycosylating activities in extracts from nasturtium (*Tropaeolum majus*). *Carbohydr. Res.* **341**, 577-581.
- Baumann, M.J., Eklöf, J.M., Michel, G., Kallas, Å.M., Teeri, T.T., Czjzek, M. and Brumer, H.** (2007) Structural evidence for the evolution of xyloglucanase activity from xyloglucan

endo-transglycosylases: biological implications for cell wall metabolism. *Plant Cell* **19**, 1947-1963.

Behar, H., Graham, S.W. and Brumer, H. (2018) Comprehensive cross-genome survey and phylogeny of glycoside hydrolase family 16 members reveals the evolutionary origin of EG16 and XTH proteins in plant lineages. *Plant J.* **95**, 1114-1128.

Biely, P., Vrřanská, M. and Claeysens, M. The *endo*-1,4- β -glucanase I from *Trichoderma reesei*. Action on β -1,4-oligomers and polymers derived from D-glucose and D-xylose. (1991) *Eur. J. Biochem.* **200**, 157-163.

Bourquin, V., Nishikubo, N., Abe, H., Brumer, H., Denman, S., Eklund, M., Christiernin, M., Teeri, T.T., Sundberg, B. and Mellerowicz, E.J. (2002) Xyloglucan *endo*transglycosylases have a function during the formation of secondary cell walls of vascular tissues. *Plant Cell* **14**, 3073–3088.

Bulone, V., Schwerdt, J.G. and Fincher, G.B. (2019) Co-evolution of enzymes involved in plant cell wall metabolism in the grasses. *Front. Plant Sci.* **10**, 1009.

Burton, R.A., Jobling, S.A., Harvey, A.J., Shirley, N.J., Mather, D.E., Bacic, A. and Fincher, G.B. (2008) The genetics and transcriptional profiles of the cellulose synthase-like *HvCs1F* gene family in barley. *Plant Phys.* **146**, 1821-1833.

Caffall, K.H. and Mohnen, D. (2009) The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. *Carbohydr. Res.* **344**, 1879-1900.

Carpita, N.C. and Gibeaut, D.M. (1993) Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J.* **3**, 1–30.

Carpita, N. and McCann, M. (2000) The cell wall. In *Biochemistry and Molecular Biology of Plants* (Buchanan, B.B., Wilhelm, G. and Jones, R.L. eds). Rockville, IL: American Society of Plant Physiologists, pp 52-108.

Ceroni, A., Maass, K., Geyer, H., Dell, A. and Haslam, S.M. (2008) GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans. *J. Proteome Res.* **7**, 1650-1569.

Cosgrove, D.J. (2005) Growth of the plant cell wall. *Nat. Rev. Mol. Cell Biol.* **6**, 850-861.

Cosgrove, D.J. (2018) Diffuse growth of plant cell walls. *Plant Physiol.* **176**, 16-27.

- Cornuault, V., Manfield, I.W., Ralet, M.C. and Knox, J.P.** (2014) Epitope detection chromatography: a method to dissect the structural heterogeneity and inter-connections of plant cell-wall matrix glycans. *Plant J.* **78**, 715-722.
- Cumming, C.M., Rizkallah, H.D., McKendrick, K.A., Abdel-Massih, R.M., Baydoun, E.A. and Brett, C.T.** (2005) Biosynthesis and cell-wall deposition of a pectin–xyloglucan complex in pea. *Planta* **222**, 546-555.
- Davies, G.J., Wilson, K.S. and Henrissat B.** (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem J.* **321**, 557-559.
- El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., Qureshi, L.J., Richardson, L.J., Salazar, G.A., Smart, A., Sonnhammer, E.L.L., Hirsh, L., Paladin, L., Piovesan, D., Tosatto, S.C.E. and Finn, R.D.** (2019) The Pfam protein families database in 2019. *Nucleic Acids Res.* **38**, D211-D222.
- Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E.L., Tate, J. and Punta, M.** (2014) Pfam: the protein families database. *Nucleic Acids Res.* **42**, D222-D230.
- Friesner, R.A., Murphy, R.B., Repasky, M.P., Frye, L.L., Greenwood, J.R., Halgren, T.A., Sanschagrin, P.C. and Mainz, D.T.** (2006) Extra Precision Glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **49**, 6177-6196.
- Fry, S.C., Smith, R.C., Renwick, K.F., Martin, D.J., Hodge, S.K., Matthews, K.J.** (1992) Xyloglucan endotransglycosylase, a new wall-loosening enzyme activity from plants. *Biochem J.* **282**, 821-828.
- Fry, S.C.** (1997) Novel “dot-blot” assays for glycosyltransferases and glycosylhydrolases: optimisation for xyloglucan endotransglycosylase (XET) activity. *Plant J.* **11**, 1141-1150.
- Gibeaut, D.M., Pauly, M., Bacic, A. and Fincher, G.B.** (2005) Changes in cell wall polysaccharides in developing barley (*Hordeum vulgare*) coleoptiles. *Planta* **221**, 729-738.
- Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N. and Rokhsar, D.S.** (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* **40**, D1178-D1186.
- Gerttula, S., Zinkgraf, M., Muday, G.K., Lewis, D.R., Ibatullin, F.M., Brumer, H., Hart, F., Mansfield, S.D., Filkov, V. and Groover, A.** (2015) Transcriptional and hormonal regulation of gravitropism of woody stems in *Populus*. *Plant Cell* **27**, 2800-2813.

- Gupta, R., Jung, E., Gooley, A.A., Williams, K.L., Brunak, S. and Hansen, J.** (1999) Scanning the available *Dictyostelium discoideum* proteome for O-linked GlcNAc glycosylation sites using neural networks. *Glycobiology* **9**, 1009-1022.
- Hayashi, T.** (1989) Xyloglucans in the primary cell wall. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* **40**, 139-168.
- Hrmova, M., Farkaš, V., Lahnstein, J. and Fincher, G.B.** (2007) A barley xyloglucan xyloglucosyl transferase covalently links xyloglucan, cellulosic substrates, and (1,3;1,4)- β -D-glucans. *J. Biol. Chem.* **282**, 12951-12962.
- Hrmova, M., Farkaš, V., Harvey, A.J., Lahnstein, J., Wischmann, B., Kaewthai, N., Ezcurra, I., Teeri, T.T. and Fincher, G.B.** (2009) Substrate specificity and catalytic mechanism of a xyloglucan xyloglucosyl transferase HvXET6 from barley (*Hordeum vulgare* L.). *FEBS J.* **276**, 437-456.
- Hsieh Y.S.Y and Harris, P.J.** (2009) Xyloglucans of monocotyledons have diverse structures. *Mol. Plant* **2**, 943-965.
- Hykollari, A., Paschinger, K., Eckmair, B. and Wilson, B.H.** (2017) Analysis of invertebrate and protist N-glycans. *Meth. Mol. Biol.* **1503**, 167-184.
- Johansson, P., Brumer, H. 3rd, Baumann, M.J., Kallas, A.M., Henriksson, H., Denman, S.E., Teeri, T.T. and Jones, T.A.** (2004) Crystal structures of a poplar xyloglucan endotransglycosylase reveal details of transglycosylation acceptor binding. *Plant Cell* **16**, 874-886.
- Kaewthai, N., Harvey, A.J., Hrmova, M., Brumer, H., Ezcurra, I., Teeri, T.T. and Fincher, G.B.** (2010) Recombinant expression of a diversity of barley *XTH* genes in the yeast *Pichia pastoris*. *Plant Biotech.* **27**, 251-258.
- Keegstra, K., Talmadge, K.W., Bauer, W.D. and Albersheim, P.** (1973) The Structure of plant cell walls: III. A model of the walls of suspension-cultured sycamore cells based on the interconnections of the macromolecular components. *Plant Physiol.* **51**, 188-197.
- Kiefer, L.L., York, W.S., Darvill, A.G. and Albersheim, P.** (1989) Xyloglucan isolated from suspension-cultured sycamore cell walls is O-acetylated. *Phytochemistry* **28**, 2105-2107.
- Kooiman, P.** (1957) Amyloids of plant seeds. *Nature* **179**, 107-109.
- Kosík, O., Auburn, R.P., Russel, S., Stratilová, E., Garajová, S., Hrmova, M. and Farkaš, V.** (2010) Polysaccharide microarrays for high-throughput screening of transglycosylase activities in plant extracts. *Glycoconj. J.* **27**, 79-87.

- Krissinel, K. and Hendrick, K.** (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr.* **D60**, 2256-2268.
- Landan, G. and Graur, D.** (2008) Local reliability measures from sets of co-optimal multiple sequence alignments. *Pac. Symp. Biocomput.* **13**, 15-24.
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M. and Henrissat, B.** (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, D490-D495.
- Mark, P., Baumann, M.J., Eklöf, J.M., Gullfot, F., Michel, G., Kallas, Å.M., Teeri, T.T., Brumer, H. and Czjzek, M.** (2008) Analysis of nasturtium *TmNXG1* complexes by crystallography and molecular dynamics provides detailed insight into substrate recognition by family GH16 xyloglucan *endo*-transglycosylases and *endo*-hydrolases. *Proteins* **75**, 820-836.
- Mayer, K.F., Waugh, R., Brown, J.W. et al.** International Barley Genome Sequencing Consortium (2012) A physical, genetic and functional sequence assembly of the barley genome. *Nature* **491**, 711-716.
- Mohler, K.E., Simmons, T.J. and Fry, S.C.** (2013) Mixed-linkage glucan:xyloglucan endotransglucosylase (MXE) re-models hemicelluloses in *Equisetum* shoots but not in barley shoots or *Equisetum* callus. *New Phytol.* **197**, 111-122.
- Mravec, J., Kračun, S.K., Rydahl, M.G., Westereng, B., Miart, F., Clausen, M.H., Fangel, J.U., Daugaard, M., Van Cutsem, P., De Fine Licht, H.H., Höfte, H., Malinovsky, F.G., Domozych, D.S. and Willats, W.G.** (2014) Tracking developmentally regulated post-synthetic processing of homogalacturonan and chitin using reciprocal oligosaccharide probes. *Development* **141**, 4841-4850.
- Needleman, S.B. and Wunsch, C.D.** (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* **48**, 443-453.
- Nishikubo, N., Awano, T., Banasiak, A., Bourquin, V., Ibatullin, F., Funada, R., Brumer, H., Teeri, T.T., Hayashi, T., Sundberg, B. and Mellerowicz, E.J.** (2007). Xyloglucan endo-transglycosylase (XET) functions in gelatinous layers of tension wood fibers in poplar—a glimpse into the mechanism of the balancing act of trees. *Plant Cell Physiol.* **48**, 843–855.
- Nishitani, K. and Tominaga, R.** (1992) Endo-xyloglucan transferase, a novel class of glycosyltransferase that catalyzes transfer of a segment of xyloglucan molecule to another xyloglucan molecule. *J. Biol. Chem.* **267**, 21058-21064.

- Nonogaki, H.** (2019) Seed germination and dormancy: The classic story, new puzzles, and evolution. *J. Integr. Plant Biol.* **61**, 541-563.
- Okazawa, K., Sato, Y., Nakagawa, T., Asada, K., Kato, I., Tomita, E. and Nishitani, K.** (1993) Molecular cloning and cDNA sequencing of endoxyloglucan transferase, a novel class of glycosyltransferase that mediates molecular grafting between matrix polysaccharides in plant cell walls. *J. Biol. Chem.* **268**, 25364-25638.
- Palmigiano, A., Messina, A., Bua, R.O., Barone, R., Sturiale, L., Zappia, M., Garozzo D.** (2018) CSF N-glycomics using MALDI MS techniques in Alzheimer's Disease. *Meth. Mol. Biol.* **1750**, 75-91.
- Pauly, M. and Keegstra, K.** (2016) Biosynthesis of the plant cell wall matrix polysaccharide xyloglucan. *Annu. Rev. Plant Biol.* **67**, 235-259.
- Park, Y.B. and Cosgrove, D.J.** (2015) Xyloglucan and its interactions with other components of the growing cell wall. *Plant Cell Physiol.* **56**, 180-94.
- Pažitná, L., Nemčovič, M., Pakanová, Z., Baráth, P., Aliev, T., Dolgikh, D., Argentova, V. and Katrlík, J.** (2020) Influence of media composition on recombinant monoclonal IgA1 glycosylation analysed by lectin-based protein microarray and MALDI-MS. *J. Biotech.* **314-315**, 34-40.
- Pei, J., Kim, B.-H. and Grishin, N.V.** (2008) PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res.* **36**, 2295-2300.
- Peña, M.J., Kong, Y., York, W.S. and O'Neill, M.A.** (2012) A galacturonic acid-containing xyloglucan is involved in *Arabidopsis* root hair tip growth. *Plant Cell* **24**, 4511-4524.
- Penn, O., Privman, E., Ashkenazy, H., Landan, G., Graur, D. and Pupko, T.** (2010) GUIDANCE: a web server for assessing alignment confidence scores. *Nucleic Acids Res.* **38**, W23-W28.
- Penning, B.W., McCann, M.C. and Carpita, N.C.** (2019) Evolution of the cell wall gene families of grasses. *Front. Plant Sci.* **10**, 1205.
- Popper, Z.A. and Fry, S.C.** (2008) Xyloglucan – pectin linkages are formed intraprotoplasmically, contribute to wall-assembly, and remain stable in the cell wall. *Planta* **227**, 781-794.
- Saeed, A.I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M., Sturn, A., Snuffin, M., Rezantsev, A., Popov, D., Ryltsov, A., Kostukovich, E., Borisovsky, I., Liu, Z., Vinsavich, A., Trush, V. and Quackenbush,**

- J.** (2003) TM4: A free, open-source system for microarray data management and analysis. *Biotechniques* **34**, 374-378.
- Shevchenko, A., Tomas, H., Havli, J., Olsen, J.V. and Mann, M.** (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Prot.* **1**, 2856-2860.
- Scheller, H.V. and Ulvskov, P.** (2010) Hemicelluloses. *Annu. Rev. Plant Biol.* **61**, 263-289.
- Schrödinger Release 2015-2: Induced Fit Docking protocol; Glide, Schrödinger, LLC, New York, NY, 2015.
- Schrödinger Release 2015-2: LigPrep, Schrödinger, LLC, New York, NY, 2015.
- Schrödinger Release 2015-2: Protein Preparation Wizard; Epik, Schrödinger, LLC, New York, NY, 2015.
- Shinohara, N., Sunagawa, N., Tamura, S., Yokoyama, R., Ueda, M., Igarashi, K. and Nishitani, K.** (2017) The plant cell-wall enzyme AtXTH3 catalyses covalent cross-linking between cellulose and cello-oligosaccharide. *Sci Rep.* **7**, 46099.
- Schünmann, P.H.D., Smith, R.C., Lang, V., Matthews, R. and Chandler, P.M.** (1997) Expression of XET-related genes and its relation to elongation in leaves of barley (*Hordeum vulgare* L.) *Plant Cell Environ.* **20**, 1439-1450.
- Sela, I., Ashkenazy, H., Katoh, K. and Pupko, T.** (2015) GUIDANCE2: accurate detection of unreliable alignment regions accounting for the uncertainty of multiple parameters. *Nucleic Acids Res.* **43**, W7-W14.
- Simmons, T.J., Mohler, K.E., Holland, C., Goubet, F., Franková, L., Houston, D.R., Hudson, A.D., Meulewaeter, F. and Fry, S.C.** (2015) Hetero-trans- β -glucanase, an enzyme unique to *Equisetum* plants, functionalizes cellulose. *Plant J.* **83**, 753-769.
- Solovyev V.V.** (2007) Statistical approaches in Eukaryotic gene prediction. In *Handbook of Statistical Genetics* (Balding D., Cannings C., Bishop M., eds), Wiley-Interscience; 3rd Edition, 1616 pp.
- Stamatakis, A.** (2014) RAxML Version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313.
- Stratilová, B., Firáková, Z., Kludiny, J., Šesták, S., Kozmon, S., Strouhalová, D., Garajová, S., Ait-Mohand, F., Horváthová, Á., Farkaš, V., Stratilová, E. and Hrmova, M.** (2019) Engineering the acceptor substrate specificity in the xyloglucan endotransglycosylase TmXET6.3 from nasturtium seeds (*Tropaeolum majus* L.). *Plant Mol. Biol.* **100**, 181-197.
- Thompson, D.S.** (2005) How do cell walls regulate plant growth? *J. Exp. Bot.* **56**, 2275-2285.

Vogel, J. (2008) Unique aspects of the grass cell wall. *Curr. Opin. Plant Biol.* **11**, 301-317.

Vaaje-Kolstad, G., Farkaš, V., Hrmova, M. and Fincher, G.B. (2010a) Xyloglucan xyloglucosyl transferases from barley (*Hordeum vulgare* L.) bind oligomeric and polymeric xyloglucan molecules in their acceptor binding sites. *Biochim. Biophys. Acta - General Subjects* **1800**, 674-684.

Vaaje-Kolstad, G., Farkaš, V., Fincher, G.B. and Hrmova, M. (2010b) Barley xyloglucan xyloglucosyl transferases bind xyloglucan-derived oligosaccharides in their acceptor binding regions in multiple conformational states. *Arch. Biochem. Biophys.* **496**, 61-68.

Varghese, J. N., Hrmova, M. and Fincher, G. B. (1999) Three-dimensional structure of a barley β -D-glucan exohydrolase, a family 3 glycosyl hydrolase. *Structure* **7**, 179-190.

Viborg, A.H., Terrapon, N., Lombard, V., Michel, G., Czjzek, M., Henrissat, B. and Brumer, H. (2019) A subfamily roadmap of the evolutionarily diverse glycoside hydrolase family 16 (GH16). *J. Biol. Chem.* **294**, 15973-15986.

Vogel, J. (2008) Unique aspects of the grass cell wall. *Curr. Opin. Plant Biol.* **11**, 301-307.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R. and Schwede, T. (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296-W303.

York, W.S., van Halbeek, H., Darvill, A.G. and Albersheim, P. (1990) Structural analysis of xyloglucan oligosaccharides by ^1H -n.m.r. spectroscopy and fast-atom-bombardment mass spectrometry. *Carbohydr. Res.* **200**, 9-31.

Zykwinska, A., Thibault, J.F. and Ralet, M.C. (2008) Competitive binding of pectin and xyloglucan with primary cell wall cellulose. *Carbohydr. Polym.* **74**, 957-961.

FIGURE LEGENDS

Figure 1. Incorporation of Alexa Fluor™ 488-labelled oligosaccharides in barley root hair cell walls. (a-e) Imaging by confocal microscopy proceeded 30 min after incubation with individual probes on the surface of 4-day old barley roots at approximately 2 mm from the apex. Upper panels: Alexa Fluor 488 (A488) channel; middle panels: UV auto-fluorescence (excited at 405 nm, recorded at 430-470 nm); bottom panels: Differential Interference Contrast (DIC) microscopy scans of samples with buffer only (a), GlcN-A488 (b), XXFG-A488 (c) and $\alpha(1-4)\text{GalAp}]_7\text{-A488}/[\alpha(1-4)\text{GalAp}]_8\text{-A488}$ (d). Notice the incorporation of XXFG-A488 and $\alpha(1-4)\text{GalAp}]_7\text{-A488}/[\alpha(1-4)\text{GalAp}]_8\text{-A488}$.

A488/[$\alpha(1-4)$ GalAp]₈-A488 in cell walls, but the incorporation was compromised in roots pre-treated with Proteinase K (e). Close-up micrographs and overlay with the auto-fluorescence channel suggested the cell wall localisation of $\alpha(1-4)$ GalAp]₇-A488/[$\alpha(1-4)$ GalAp]₈-A488. (f). Scale bars in images correspond to 200 μ m (a-e) or 100 μ m (f).

Figure 2. Abbreviations, descriptions, and chemical structures of commercial and purified acceptor oligosaccharides used in this work. Chemical formulas are defined in Supplemental Scheme S1.

Figure 3. Hetero-transglycosylation activities catalysed by HvXET3 and HvXET4 with donor XG (a) and HEC (b), and chemically distinct SR-labelled acceptors. Inset in panel (a) represents activities of HvXET6 with XG and SR-labelled acceptors. Abbreviations, descriptions, and chemical structures of acceptors are defined in Figure 2 and Supplemental Scheme S1. Transglycosylation activities are expressed in fluorescence values/min. The average of three technical replicates (n=3) is plotted with standard deviation values, calculated *via* Excel in Microsoft Office Professional 2016.

Figure 4. Sequence alignment and β -sandwich architectures with the β -jellyroll topology of HvXET3 and HvXET4 in complex with XXXG (donor) and [$\alpha(1-4)$ GalAp]₅ (acceptor) substrates, superposed with the TmXET6.3 model. (a), Protein sequences of poplar PtXET16A (Q8GZD5), nasturtium TmXET6.3 (Q07524), barley HvXET3-HvXET6 (P93671, P93672, P93668, B1P1S7) were aligned with PROMALS3D (Pei *et al.*, 2008) that cluster in three groups (magenta fonts). H75/W75 and R110/Y110 (black boxes) residues were investigated in this work (numbering omits signal peptides). Predicted and consensus β -sheets secondary structures (ss) are in blue and assigned with the letter 'e', and loops and turns are in black. Absolute conservation and similarity of amino-acid (aa) residues on the scale 9-6 are shown at the top of the diagram (brown and black), respectively. Single N-glycosylation sites are marked in a green box. (b, c), Left panels: 3D models of HvXET3 (green) or HvXET4 (cyan) in complex with the XXXG donor (yellow cpk sticks)/[$\alpha(1-4)$ GalAp]₅ acceptor (cpk sticks) substrate pair are superposed with TmXET6.3 (pink). Residues in HvXET enzymes (cpk blue sticks with dots indicating van der Waals radii) that correspond to those in TmXET6.3 (cpk magenta sticks with dots) are indicated. (b, c), Right panels: Details of binding of the XXXG/[$\alpha(1-4)$ GalAp]₅ pair by HvXET3 (green) or HvXET4

(cyan). Separations of 2.6 Å and 3.5 Å between W166, S164, Y67, and XXXG (-4 to -1 subsites), and H94, N96, E106, K260 and R261 and [$\alpha(1-4)\text{GalAp}$]₅ (+1 to +5 subsites) are shown. Catalytic E77 and E81 residues (red sticks) with conserved D79 (pink) interact with substrates. The -4 to -1 subsites with XXXG and the +1 to +5 subsites with [$\alpha(1-4)\text{GalAp}$]₅ in enzyme active sites are indicated at the bottom of panels (b, c). RMSD values between superposed TmXET6.3 (250 residues) and HvXET3 (263 residues) or HvXET4 (255 residues) are 0.52 Å and 0.70 Å, respectively. Structures in panels (b, c) are clipped for convenience to view bound XXXG and [$\alpha(1-4)\text{GalAp}$]₅.

Figure 5. Hetero-transglycosylation activities with the XG donor and [$\alpha(1-4)\text{GalAp}$]₅-SR acceptor substrates catalysed by barley HvXET3, HvXET4, HvXET5 and HvXET6, and nasturtium WT and single W75H and Y110R, and double W75H/Y110R TmXET6.3 mutants are indicated. The chemical structure of [$\alpha(1-4)\text{GalAp}$]₅ is represented at the top of the figure. Hetero-transglycosylation activities were normalised relative to 100% activity of the XG-XGO8 substrate pair. The average of three technical replicates (n = 3) is plotted with standard deviation values, calculated *via* Excel in Microsoft Office Professional 2016.

Figure 6. RT-qPCR-derived normalised expression levels of selected *HvXET* transcripts in various tissues. (a), Transcript levels (given in arbitrary units AU) of *HvXET3* and *HvXET4*; (b), Levels of *HvXET6* in various tissues are plotted on a different scale, compared to the panel (a). Error bars are standard deviation values of three PCR repeats for each mRNA. Abbreviations: Root m/zone = root maturation zone, Grain 3-5 DAP = grain 3-5 days after pollination, Grain 8-10 DAP = grain 8-10 days after pollination, Coleoptile 3 DAG = coleoptile 3 days after germination.

Figure 7. Randomised Accelerated Maximum Likelihood (RAxML) (Stamatakis, 2014) tree of 419 GH16 protein sequences across six monocots (green nodes), four eudicots (blue nodes) and the basal Angiosperm *Amborella trichopoda* (magenta nodes). HvXET3 (UniProt accession P93671), HvXET4 (P93671), HvXET5 (P93668) and HvXET6 (B1P1S7 or F2DM52) are in orange nodes and TmXET6.3 (V5ZEF7) in a brown node. Experimentally determined PttXET16 XET (PDB 1UN1), TmNXG1 XEH apo (2UWA), and TmNXG1-DYNIIG XEH in complex with XLLG (2VH9) are in yellow nodes. Bootstrap support values for deep nodes are indicated at bipartitions. To improve figure readability, a small clade separated by a long molecular branch was truncated

(black square). This clade contains the *Populus trichocarpa* Potri.002G153200 locus identified in CAZy as a broad-specific endo-(xylo)glucanase (EC 3.2.1.151 and EC 3.2.1.73). The list of sequences and the FASTA alignment file used to generate the RAxML tree are included in Supplemental Data Set S2. This version of the tree including the small clade is available as Supplemental Figure S7.

icle

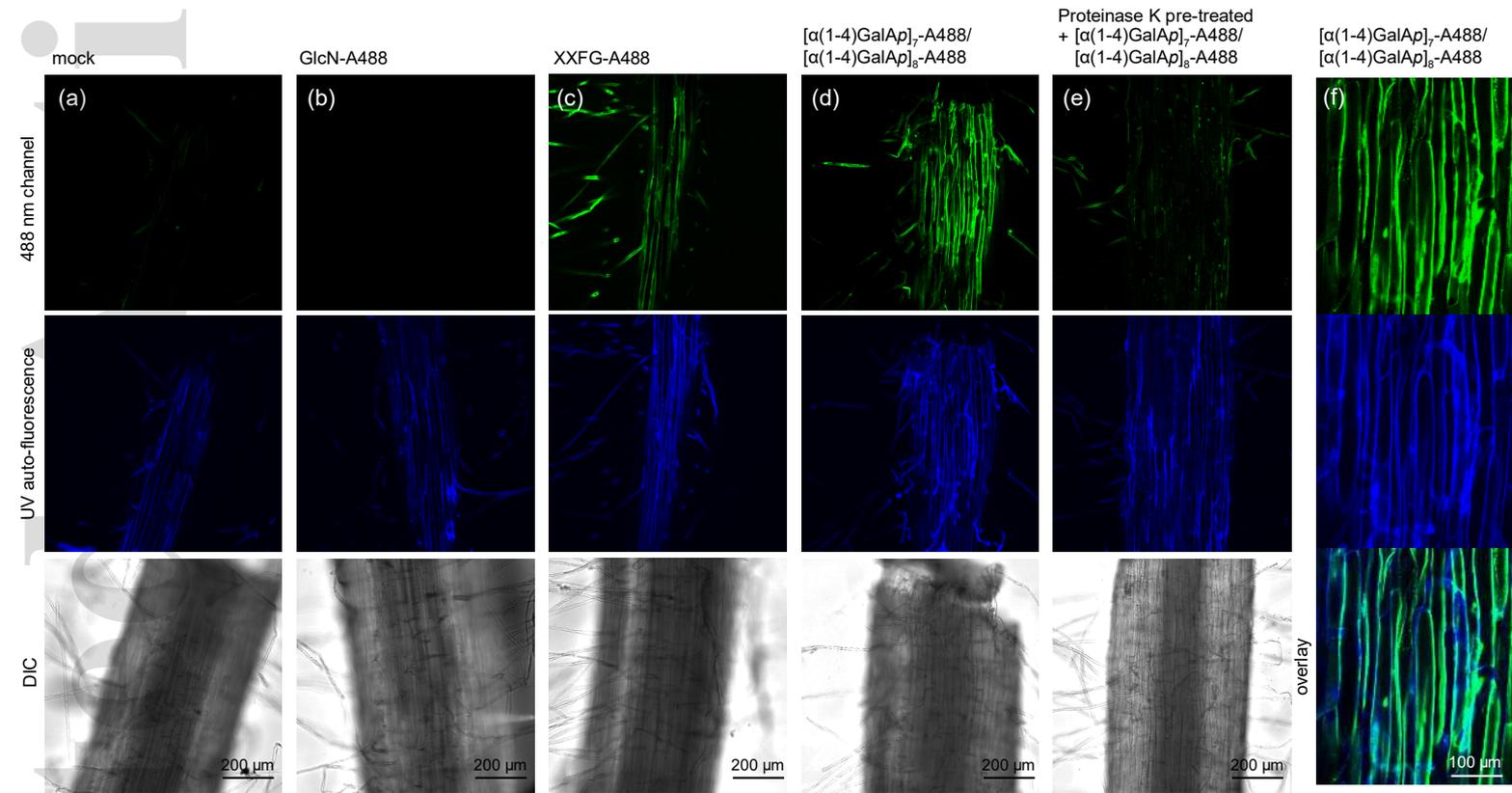
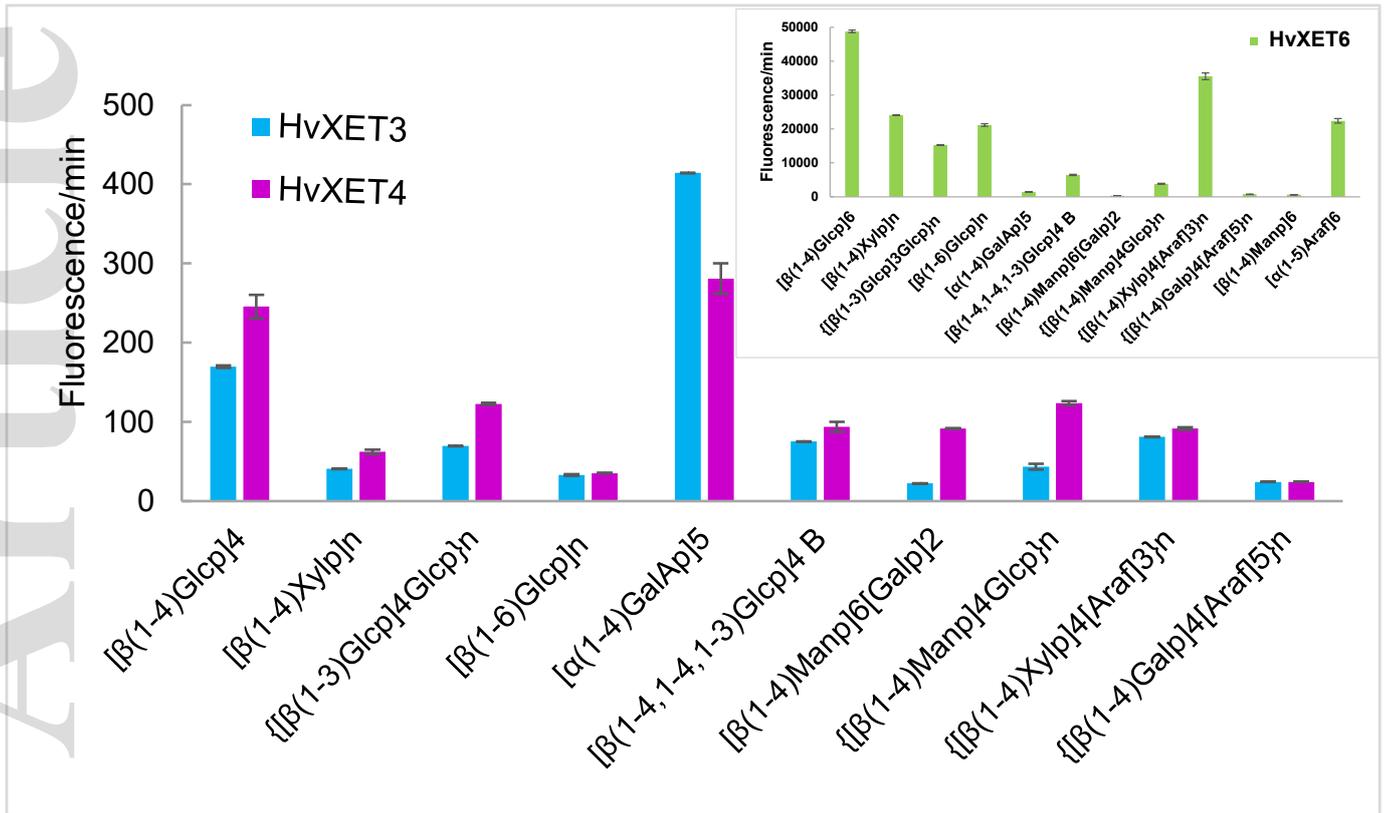


Figure 1

Accep

(a)



(b)

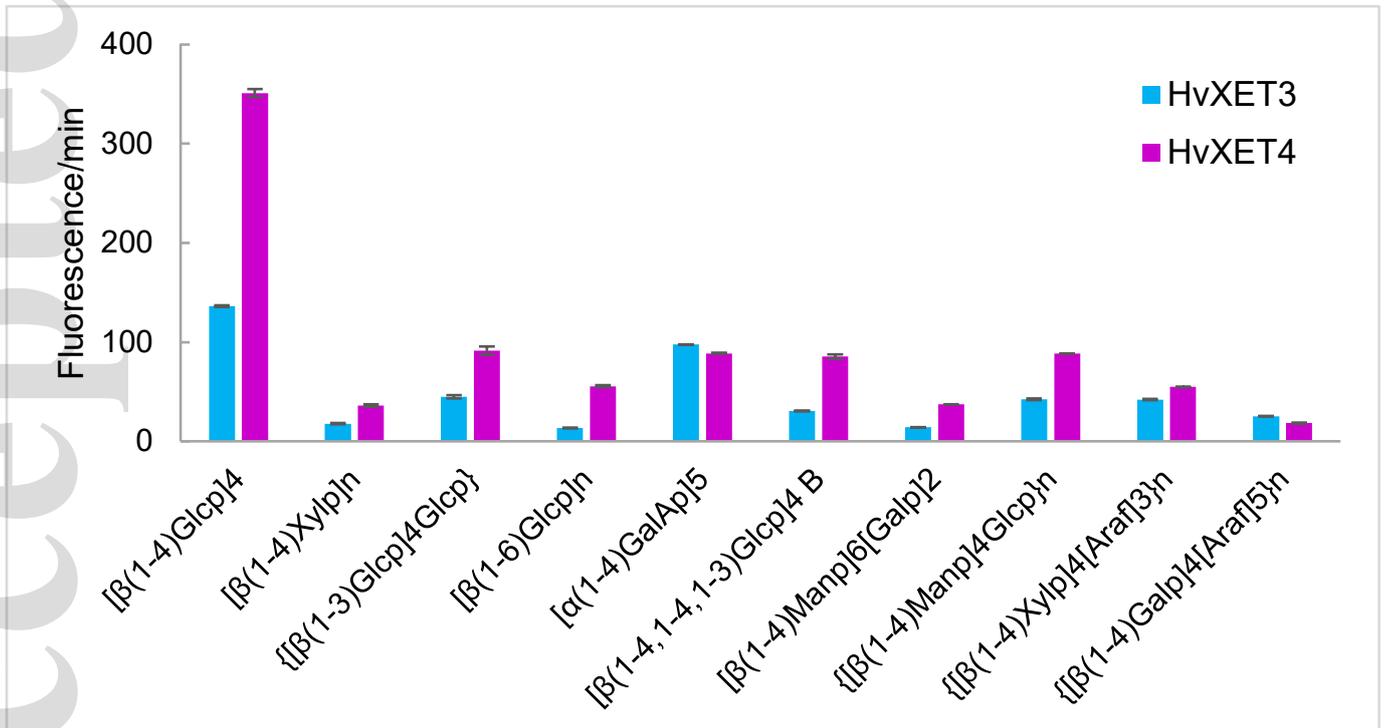


Figure 3

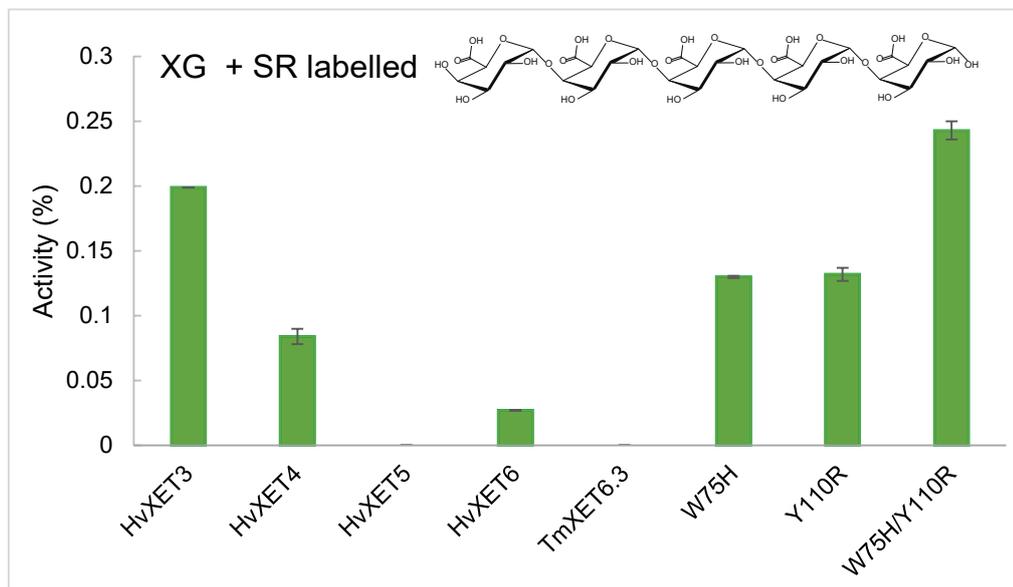


Figure 5

Figure 6

