Plant Biotechnology 27, 251–258 (2010)

Original Paper

Heterologous expression of diverse barley *XTH* genes in the yeast *Pichia pastoris*

Nomchit Kaewthai[1,†], Andrew J. Harvey[2,†], Maria Hrmova[2], Harry Brumer[1,*], Ines Ezcurra[1], Tuula T. Teeri[1], Geoffrey B. Fincher[2]

[1] School of Biotechnology, Royal Institute of Technology, Albanova University Centre, 106 91 Stockholm, Sweden,
[2] School of Agriculture, Food and Wine, and the Australian Centre for Plant Functional Genomics, University of Adelaide, Waite Campus, Glen Osmond, Australia

* E-mail: harry@biotech.kth.se Tel: +46-8-5537-8367 Fax: +46-8-5537-8468

Received August 26, 2009; accepted February 16, 2010 (Edited by K. Igarashi)

Abstract Heterologous expression of plant genes, particularly those encoding carbohydrate-active enzymes such as glycoside hydrolases and glycosyl transferases, continues to be a major hurdle in the functional analysis of plant proteomes. Presently, there are few convenient systems for the production of recombinant plant enzymes in active form and at adequate levels for biochemical and structural characterization. The methylotrophic yeast *Pichia pastoris* is an attractive expression host due to its ease of manipulation and its capacity to perform post-translational protein modifications, such as *N*-glycosylation [Daly and Hearn (2005) *J Mol Recognit* 18: 119–138]. Here, we demonstrate the utility of the *P. pastoris* SMD1168H/pPICZ-alpha C system for the expression of a range of *xyloglucan* endo-*transglycosylase/hydrolase* (*XTH*) cDNAs from barley (*Hordeum vulgare*). Although stable transformants were readily obtained by positive selection for vector-induced antibiotic resistance for all of the nine constructs tested, only five isoforms were secreted as soluble proteins into the culture medium, four in active form. Furthermore, production levels of these five isoforms were found to be variable, depending on the transformant, which further underscores the necessity of screening multiple clones for expression of active enzyme. Failure to express certain *XTH* isoforms in *P. pastoris* could not be correlated with any conserved gene or protein sequence properties, and this precluded using rational sequence engineering to enhance heterologous expression of the cDNAs. Thus, while significant advances are reported here, systems for the heterologous production of plant proteins require further development.

Key words: *Pichia pastoris*, plant protein expression, *xyloglucan* endo-*transglycosylase/hydrolase (XTH)* genes, XET, XEH.

As plant science moves into the "post-genomics era" with the appearance of an increasing amount of gene and transcript sequence data, the need for functional characterization of the corresponding encoded proteins becomes ever more pressing (Hrmova and Fincher 2009). In particular, structure-function studies on the diverse carbohydrate-active enzymes (CAZymes) involved in plant cell wall biosynthesis and remodelling (Davies and Henrissat 2002) are especially topical in the context of increasing yields of agricultural crops, expanding cellulosic fiber utilization, and improving cell wall saccharification to produce sugar feedstocks for biofuel production. A comprehensive census of the *Arabidopsis thaliana* genome predicts that 730 open reading frames (1–2% of the genome) encode glycosyl transferases and glycoside hydrolases (Henrissat et al. 2001), while analysis of the genome of the black cottonwood (*Populus trichocarpa*) indicates a significant expansion CAZymes in this wood-forming species (Geisler-Lee et al. 2006). Comparative analysis of glycosyl transferases in the rice genome likewise indicates large numbers of CAZymes (Cao et al. 2008), and it is anticipated that this diversity will be observed in other grass genomes (Bolot et al. 2009). However, a growing problem is that the accumulation of sequence information is rapidly outpacing functional studies of plant CAZyme family members; only very few of the glycoside hydrolases (Minic 2008) and biosynthetic glycosyl transferases (Liepman et al. 2010) have been biochemically characterized thus far.

The plant *xyloglucan* endo-*transglycosylase/hydrolase* (*XTH*) genes comprise a subfamily of glycoside hydrolase (GH) family 16 (Michel et al. 2001) in the CAZy classification (Cantarel et al. 2009). *XTH* genes encode enzymes capable of catalyzing the cleavage and re-ligation and the irreversible hydrolysis of the xyloglucan polysaccharide (Baumann et al. 2007). Consequently, members of the *XTH* subfamily play intrinsic roles in

† These authors contributed equally to this work.
This article can be found at http://www.jspcmb.jp/

cell wall morphogenesis (Cosgrove 2005) and storage polysaccharide mobilization (de Silva et al. 1993). Additionally, evidence suggests that some members may be able to catalyze *hetero*-transglycosylation reactions, which result in the covalent coupling of xyloglucan with unbranched and unsubstituted β-D-glucans (Baumann et al. 2007; Hrmova et al. 2007; Mohand and Farkaš; 2006). *XTH* genes are numerous in plants, where 30–40 members are found in *A. thaliana* (Yokoyama and Nishitani 2001), *P. trichocarpa* (Baumann et al. 2007; Geisler-Lee et al. 2006), and *O. sativa* (Yokoyama et al. 2004)). The large numbers of *XTH* genes in these species is likely to reflect not only the requirement for tissue- and time-specific expression of the different genes, but it might also reflect functional differences among individual isozymes. Despite a developing knowledge of the spatial localization of these enzymes *in planta* (Becnel et al. 2006), very few *XTH* gene products have been characterized with respect to catalytic specificity and enzyme structure (see Baumann et al. 2007) and references therein). This has been due, in part, to a lack of an efficient, generally applicable system for the heterologous production of individual *XTH* gene products.

For instance, *XTH* genes from fruits have been heterologously expressed in *Escherichia coli,* but subsequent protein solubilization and refolding steps were required to yield low amounts of active enzymes (Arrowsmith and de Silva 1995; Schröder et al. 1998). Indeed, the majority of *XTH* gene products contain a structurally-important *N*-glycosylation site (Johansson et al. 2004; Kallas et al. 2005), which rules out the use of bacterial systems such as *E. coli* as general production hosts. Among eukaryotic hosts capable of post-translational glycosylation of proteins, the *Spodoptera frugiperda* SF9 insect cell system was the first to be successfully used for the expression of *XTH* genes, which include BRU1 from *Glycine max* (Oh et al. 1998) and *AtXTH14*, *AtXTH22*, *AtXTH24* and *AtXTH27* from *Arabidopsis thaliana* (Campbell and Braam 1999). In these examples the enzymes were obtained in soluble, active form, but the levels of expression were low. Furthermore, protein production in animal cells is significantly more technically demanding than in microbial hosts.

The yeast *Pichia pastoris* has become a prime choice for the heterologous expression of eukaryotic genes over the past decade, owing to the yeast's capacity to mediate post-translational modifications and a general ease of use compared with animal cells and other systems. Linearized foreign DNA can be inserted with high efficiency *via* homologous recombination to generate stable cell lines. Moreover, the levels of endogenous, non-recombinant proteins secreted from *P. pastoris* are very low, which facilitates downstream purification (Daly and Hearn 2005). The expression of several *XTH* genes from various plants in *P. pastoris* has now been reported, including individual members from tomato (Catala et al. 2001; Chanliaud et al. 2004; Saladié et al. 2006), cauliflower (Henriksson et al. 2003), nasturtium (Baumann et al. 2007; Chanliaud et al. 2004), hybrid aspen (Johansson et al. 2004; Kallas et al. 2005), silver birch and *Gerbera hybrida* (Toikkanen et al. 2007), *Selaginella kraussiana* (Van Sandt et al. 2006) and barley (Hrmova et al. 2009).

As part of an ongoing effort to catalog the catalytic diversity in the *XTH* gene family, we present here a systematic study of the recombinant expression of nine *XTH* cDNAs from the agriculturally important cereal, barley (*Hordeum vulgare*), which is presently the focus of intense genomics efforts (Bolot et al. 2009; Sreenivasulu et al. 2008). Application of a multi-step screen devised to identify high-expressing clones indicates that *P. pastoris*, while useful for expression of some *HvXTH*s, discriminates against certain isoforms, with an overall success rate of approximately 50%.

Materials and methods

Cloning and expression of HvXTH genes in Pichia pastoris SMD1168H

The GenBank accession numbers of the *HvXTH* genes used in this study are specified in Table 1. The *HvXTH* cDNA molecules were cloned using standard techniques (Hrmova et al. 2009) into pPICZ-alpha C containing the AOX promoter. In each construct, the sequence encoding the predicted native signal peptide was replaced with one encoding the yeast alpha-factor secretion signal peptide. Competent SMD1168H cells were transformed with 5–10 μg of plasmid DNA, which had previously been linearized with *Pme*I as described by the manufacturer (Invitrogen *Pichia* Expression Kit manual). After transformation, the cells were placed on YPDS (yeast extract-peptone-dextrose-sorbitol) agar plates containing zeocin (100 μg ml^{-1}) and incubated at 30ºC for 3–5 days. The colonies from YPDS zeocin plates were further grown on YPD (yeast extract-peptone-dextrose) containing zeocin (100 μg ml^{-1}) for expression studies.

The purified clones were expressed in a small scale to identify positive clones, by innoculating 5 ml of BMGY (buffered glycerol/complex medium) in a 50 ml Falcon tube. The cultures were grown overnight at 30ºC and 180–220 rpm, at which time the media were changed to BMMY (buffered methanol/complex medium). The cultures were induced for 4-5 days at 22ºC, with methanol at a final concentration of 0.5% (v/v) on the first day and 1% (v/v) for the remainder of the cultivation time. On day 4, OD_{600} was measured, the cells were centrifuged at 3400 *g* for 10 min, and the supernatant was used for screening protein expression using the Bradford protein assay and SDS-PAGE (NuPAGE, Invitrogen). For SDS-PAGE analysis, 20 μl of supernatant was added to SDS-PAGE loading buffer; non-expressing clones were verified by precipitation of 700 μl of supernatant with 20% acid, followed by dissolution in

Table 1. GenBank accession numbers of barley *HvXTH1-HvXTH9* genes and properties of the proteins produced in *Pichia pastoris*

Isoform	GenBank Accession numbers	Expected molar mass (M_r)[a]	Predicted number of N-glycosylation sites[b]	Expression detected by Bradford and SDS-PAGE	Clone(s) selected for large-scale production (clone #)	Large-scale expression levels	Protein yield after IMAC purification (mg l^{-1} culture supernatant)	Specific XET activity of purified protein (pkat mg^{-1})
HvXTH1	X91660	32999	2	No	1–10 & 1–13	None		—
HvXTH2	X93173	32085	1	Yes	2–2	Excellent	—	n.d[c]
HvXTH3	X93174	33751	1	Yes	3–1	Excellent	12	5296[f]
HvXTH4	X93175	32990	1	Yes	4–10	Modest	2	1804[f]
HvXTH5d	X91659	34781	1	No	5–11 & 5–17	None		—
HvXTH6	EU247793	32161	1	Yes	6–2	Excellent	4	5630[e,f]
HvXTH7	FJ917200	33853	0	No	7–12 & 7–17	None		—
HvXTH8	FJ917201	36780	1	Yes	8-6	Modest	1	1311[f]
HvXTH9	FJ917202	32313	1	No	9–1 & 9–4	None		—

[a] Including the N-terminal EAEA alpha-factor cleavage site, Cla I and Xba I cloning overhangs, c-myc epitope and 6xHis tags. With *HvXTH6* we determined that EAEA is retained at the N-terminus (Hrmova et al. 2009), thus this site is included in mass calculations for all protein sequences.

[b] Predicted by NetNGlyc 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc/).

[c] n.d., not detected.

[d] The *HvXTH5* gene product was previously extracted from the native source. The specific XET activity (EC 2.4.1.207) was 3,630 pkat mg^{-1} and the enzyme was devoid of xyloglucan *endo*-hydrolase activity (EC 3.2.1.151) (Hrmova et al. 2007; Hrmova et al. 2009).

[e] The specific activity of the native, partially purified *HvXTH6* gene product is 518 pkat mg^{-1} (Hrmova et al. 2009).

[f] These recombinant enzymes had no detectable xyloglucan *endo*-hydrolase activity (EC 3.2.1.151) using the Somogyi-Nelson reducing sugar assay as described in (Hrmova et al. 2009); all *HvXTH* cDNAs thus encode XET enzymes.

20 μl water and addition of loading buffer prior to SDS-PAGE. *Pichia pastoris* clones expressing *Ptt*XET16-34 (Kallas et al. 2005) and pPICZ-alpha C, which contained no insert, were used as positive and negative controls, respectively.

Western blot analysis

Western blot analyses were performed using 0.45 µm nitrocellulose blotting membranes (Millipore, cat. #HATF00010), a mouse monoclonal anti-polyHistidine-alkaline phosphatase IgG2a isotype antibody (Sigma, cat. #A5588) and the BCIP/NBT-purple liquid reagent for membranes (Sigma, cat. #B3679)

Protein N-deglycosylation

Protein N-deglycosylation was performed using the N-glycosidase F Deglycosylation Kit as described by the manufacturer (Roche). Following N-deglycosylation, samples were analyzed on SDS-PAGE gels stained with Coomassie Brilliant blue R-250.

Large-scale production and purification of HvXTH

Selected clones (Table 1) were grown in 400 ml cultures in 2 l shaking flasks. Secreted *HvXTH* gene products were purified by immobilized metal ion affinity chromatography (IMAC) on Co^{2+}-charged Talon resin (Clontech, Mountain View, CA, USA). All purification steps were conducted at 4ºC. The active fractions were concentrated, followed by dialysis in 50 mM ammonium acetate buffer at pH 6.0. The purified proteins were stored at 4ºC and were found to be stable for at least for 6 months (data not shown).

Enzyme assays

Xyloglucan *endo*-transglycosylase (XET, xyloglucan: xyloglucosyl transferase, EC 2.4.1.207) activity was measured using a modification (Hrmova et al. 2007) of the radiometric method devised by Fry, which quantifies the incorporation of radioactive [1-^{3}H]-labelled xyloglucan-derived saccharide heptaitol XXXGol (where X=[α-D-Xyl*p*(1→6)]-β-D-Glc*p*(1→4), and Gol=D-glucitol) into tamarind seed xyloglucan polysaccharide (Fry et al. 1992). Enzyme activity is expressed in katals, where 1 katal represents 1 mol of product formed per s; specific activity is expressed in pkatals · mg^{-1} protein.

Bioinformatics

Protein *N*-glycosylation sites were predicted using NetNGlyc 1.0, available at URL http://www.cbs.dtu.dk/services/NetNGlyc/. Protein isoelectric points (p*I* values) were calculated using the *Compute pI/MW Tool* at the ExPASy Proteomics Server, URL http://www.expasy.ch/. Analysis for AT-rich regions of *HvXTH* cDNA sequences was performed using the software program *freak* (part of the EMBOSS suite, accessed via the Internet at http://vm-bioinfo.toulouse.inra.fr/emboss/) with identical parameters to those used by (Boettner et al. 2007): stepping value=1, window size=30. (accessed *via* the Mobyle portal at http://mobyle.pasteur.fr/). The codon adaptation index (CAI), effective number of codons (Nc), GC content, protein hydrophobicity (GRAVY score), and frequency of aromatic protein residues in *HvXTH* gene products were analysed using *CodonW via* the Mobyle portal at URL http://mobyle.pasteur.fr/. Potential PEST protein degradation sequence motifs were analysed using *epestfind via* the Mobyle portal.

Results and discussion

Expression screening of P. pastoris clones

Nine *xyloglucan endotransglycosylases/endohydrolase* (*XTH*) native cDNA sequences from barley (*Hordeum vulgare*) were cloned into the pPICZ alpha C vector for expression and secretion by *Pichia pastoris* SMD1168H (Figure 1). In each case, stable transformants were obtained by homologous recombination into the *P. pastoris* genome and selection for resistance to the antibiotic zeocin (Daly and Hearn 2005). During the course of earlier work on the recombinant expression of *XTH* cDNA from dicots, it was routinely observed that *P. pastoris* clones could be positive both in the selection step and in a PCR screen for genomic integration of the full-length vector insert, but did not produce soluble protein (Baumann et al. 2007; Henriksson et al. 2003; Kallas et al. 2005; and N.K. unpublished results). However, it was also observed that a proportion of positive clones did indeed secrete soluble, recombinant protein and that these could be detected only if individual clones were screened using antibodies against the expressed enzyme. Previously, we have used colony blotting with anti-protein antibodies ("Yeastern" blotting; Holmquist et al. 1997) for this purpose (Kallas et al. 2005). Here, we developed a simplified screening method based on the Bradford protein assay (Bradford 1976), which omits both the PCR screen and the sometimes ambiguous colony blot.

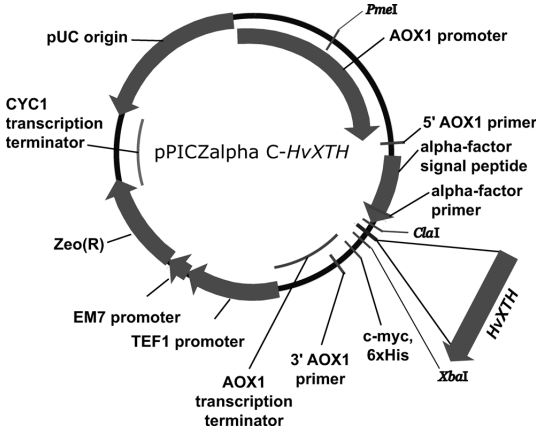

Figure 1. pPICZ-alpha C vector map indicating the position of the *HvXTH* cDNA insert.

After selection on zeocin-containing plates, colonies were individually picked and used to inoculate 5 ml liquid cultures. Following induction by methanol over 4–5 days, protein expression driven by the alcohol oxidase (AOX) promoter was induced, whereafter the culture supernatant was assayed directly by the Bradford method to measure total soluble protein. Comparison of the protein amount with respect to the culture cell density (measured as OD_{600}) gave a good indication of the relative amount of *XTH* product secreted by each clone (Figure 2). A value of ca. 0.001 mg ml^{-1} $(OD_{600})^{-1}$, observed for the empty-vector negative control and several non-productive clones, was taken as an average cut-off value arising from background *P. pastoris* proteins (Figure 2). Examination of Bradford-positive clones by SDS-PAGE indicated a predominant band at the expected molar mass for five of the nine *HvXTH* targets, with minimal amounts of contaminating proteins (Figure 3). Western blotting using an antibody specific for the c-myc epitope tag (Figure 1) confirmed that the bands observed in SDS-PAGE corresponded to a single product of the *XTH* expression construct in each case (Figure 4), with the exception of *HvXTH2*, which showed two bands (discussed below). In contrast, SDS-PAGE analysis of supernatants from Bradford-negative clones showed multiple protein bands that originated from endogenous *P. pastoris* proteins (Figure 3, *HvXTH7*), while Western blot analysis was negative in all cases (*HvXTH1, 5, 7, 9*; Figure 4). The data reflect an intriguing gene target-specific dependence of protein production in *P. pastoris*, which is nonetheless independent of selection for the antibiotic resistance

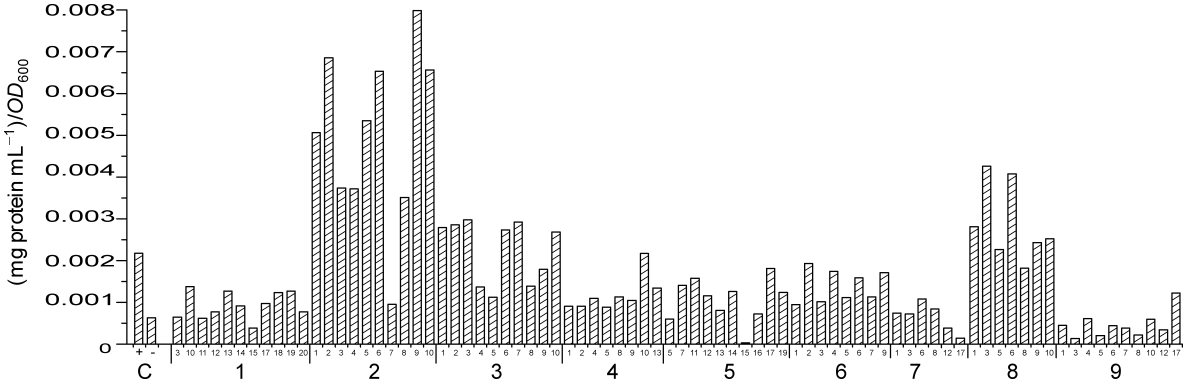

Figure 2. Preliminary screening of high expression clones of *HvXTH* genes by the Bradford protein assay versus cell density. Large numbers on x-axis denote *HvXTH* construct (*HvXTH1-HvXTH9*), small numbers denote unique clone number. C, control samples: +, *Ptt*XET16-34 (Kallas et al. 2005); −, pPICZ-alpha C empty vector.

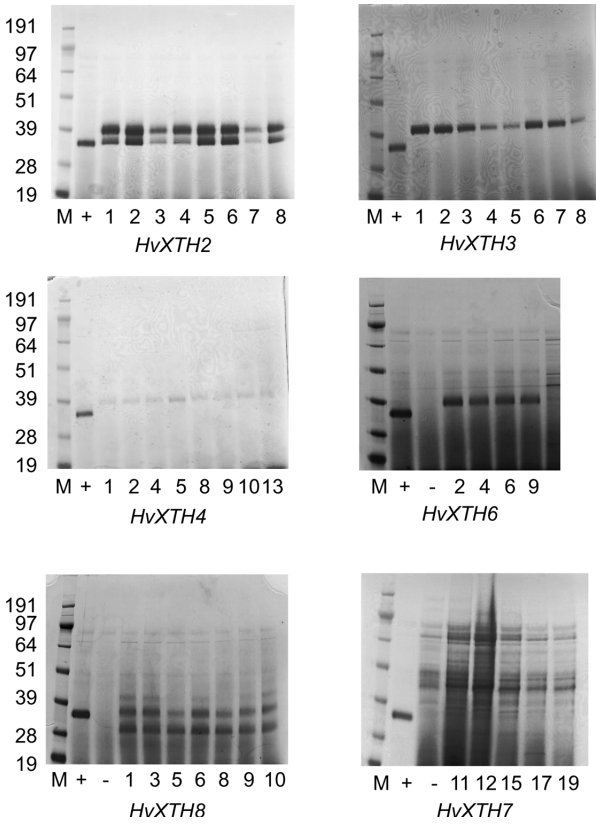

Figure 3. SDS-PAGE of *HvXTH* products in the culture supernatant after four days cultivation. Clone numbers are shown below each lane. M, molecular mass marker (kDa); +, positive control (*Ptt*XET16-34); −, empty vector negative control. *HvXTH2, 3, 4, 6,* and *8* were loaded with 20 µl supernatant per lane. *HvXTH7* was loaded with concentrated samples from trichloroacetic acid precipitation from 700 µl supernatant. Gels for *HvXTH1, 5,* and *9* (not shown) were essentially identical to that of *HvXTH7*.

marker: for those expression targets which do express (*HvXTH2, 3, 4, 6, 8*), secreted protein is observed in the majority of the clones, whereas protein production is negative in *all* clones for other *XTH* variants (*HvXTH1, 5, 7, 9*; Figures 3 and 4). The repeated success in obtaining secreting clones with certain sequences mirrors the results of Toikkanen et al. (2007), where about 90% of *Gerbera hybrida* XET transformants had correct integration of the expression cassette and secreted XET in active form.

Scale-up of recombinant enzyme production

In general, screening of small-scale (5 ml) cultures reliably predicted success during scale-up and purification for biochemical characterization of *XTH* protein products. As indicated in Table 1, one or two clones of Bradford/PAGE-positive or -negative transformants, respectively for each *HvXTH* gene target were selected for protein production in 400 ml cultures. As expected from the Bradford/PAGE screen, no protein was produced by at least two independent clones containing

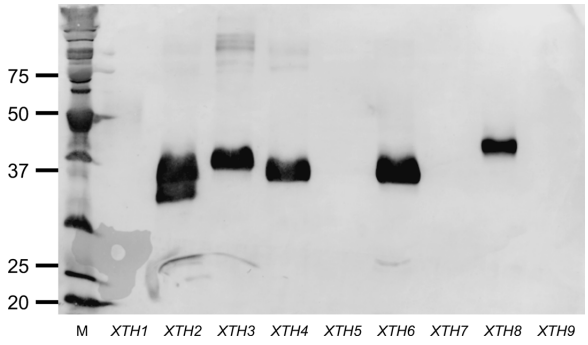

Figure 4. Western blot analysis of *HvXTH* gene products using anti-(His_6) antibody. M, relative position of molecular mass marker (kDa).

HvXTH1, 5, 7, and *9*. In contrast, clones containing *HvXTH3, 4, 6,* and *8* vectors all produced significant amounts of soluble, purifiable protein, all of which were shown to have xyloglucan *endo*-transglycosylase activity (Table 1). In the case of *HvXTH6*, the product of which could only be partially purified from barley tissues, the heterologously produced enzyme from *P. pastoris* possessed an approximately 10-fold higher specific activity (Table 1). Notably, the ability to obtain the enzyme via heterologous expression in *P. pastoris* greatly facilitated the detailed kinetic analysis of this isoform (Hrmova et al. 2009). Likewise, expression in *P. pastoris* has previously been essential for the production of the large quantities of enzyme necessary for crystallography of XET and XEH three-dimensional structures (Baumann et al. 2007; Johansson et al. 2004), as well as other biochemical studies (see Introduction).

Effects of protein N-glycosylation on expression

Although the expression of *HvXTH2* resulted in the secretion of the highest levels of soluble protein of all *HvXTH* targets, in both small and large scale cultures (Table 1, Figures 2, 3), this protein could not be purified and did not possess measurable enzyme activity. As shown in Figure 3, the crude culture supernatant of eight independent clones contained two predominant proteins, as revealed by SDS-PAGE. Treatment of the supernatant with protein *N*-glycosidase F (PNGase F), which removes N-linked glycans from proteins by cleavage of the Asn-GlcNAc bond, resulted in the disappearance of the two bands and their replacement with a protein of apparent lower molar mass (Figure 5). In comparison, treatment of the single *HvXTH3* product with protein *N*-glycosidase F resulted in a similar mobility shift, corresponding to less than 2 kDa reduction in molecular mass (Figure 5). The same phenomenon was observed for *HvXTH4* and *HvXTH6* products (data not shown). All of the *HvXTH* cDNA sequences in this study encode proteins containing at least one potential *N*-glycosylation site (Table 1), with the exception of the *HvXTH1* (two sites) and *HvXTH7* (zero sites) products.

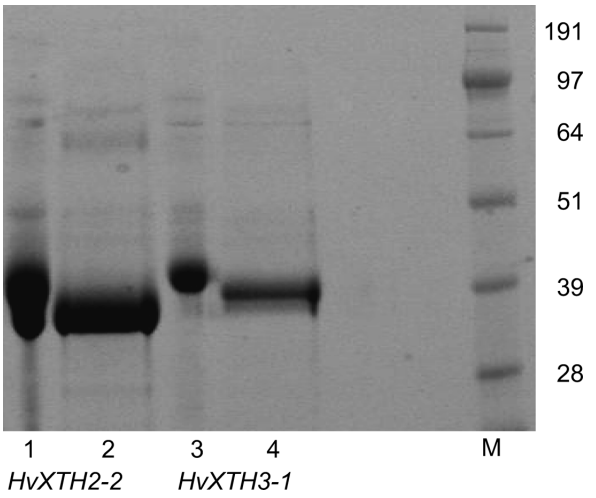

Figure 5. Enzymatic *N*-deglycosylation of *HvXTH* gene products analyzed by SDS-PAGE. *HvXTH2* (clone 2) and *HvXTH3* (clone 1) culture supernatants before (Lanes 1 & 3, respectively) and after (Lanes 2 & 4, respectively) N-glycosidase F treatment. M, molecular mass markers (kDa).

The observation of soluble, but inactive protein glycoforms in the case of *HvXTH2* suggests an error in N-glycan processing, which may affect protein folding, trafficking, stability, and/or enzyme activity directly. Analysis of the original *Ptt*XET16-34 3-D structure (Johansson et al. 2004) has demonstrated that the conserved *N*-glycosylation site in *XTH* gene products, which are also present in the *HvXTH* products studied here (Hrmova et al. 2009 and M.H. unpublished data), is located near one end of the active site cleft, with the di-GlcNAc core of the N-glycan involved in key structural interactions with the polypeptide (Kallas et al. 2005). Further, removal of this N-glycan in *Ptt*XET16-34 by site directed mutagenesis results in greatly reduced secretion levels and intracellular accumulation of protein in *P. pastoris* (Kallas et al. 2005), while enzymatic cleavage of N-glycans greatly reduces the stability of recombinant hybrid aspen and cauliflower XETs (Henriksson et al. 2003; Kallas et al. 2005). It is therefore tempting to speculate that the inability of *P. pastoris* to properly translate certain plant *XTH* genes may arise from errors in N-glycan attachment or in variations of sugar composition in N-glycans, which would affect retention and processing in the endoplasmic reticulum. In particular, the lack of an *N*-glycosylation site on *HvXTH7* may not allow retention by the calnexin/calreticulin system, while the two *N*-glycosylation sites on *HvXTH1* may lead to over-glycosylation and incorrect folding.

Gene and protein sequence analysis

While *N*-glycan processing-related issues may be part of the difficulty in expressing plant *XTH* genes in *P. pastoris*, any number of the myriad steps in gene expression and protein export may present bottlenecks. In an attempt to uncover general sequence-specific traits, which might affect the success or failure of *XTH* expression in *P. pastoris*, we employed a range of bioinformatics tools. Primary sequence analysis indicated that no obvious relationships could be delineated between successful and unsuccessful *HvXTH* heterologous expression products. Successful and unsuccessful expression targets did not separate into defined phylogenetic clades based on alignments with *XTH*s from genomic sequences (see Baumann et al. (2007) for one such phylogeny). In a recent comparative analysis of recombinant human protein production in *P. pastoris*, using a sample of 79 genes, Boettner et al. concluded that a high abundance of AT-rich regions in target cDNA and a high isoelectric point (pI) of the encoded protein were the most significant factors detrimental to expression of heterologous proteins without a yeast homolog (Boettner et al. 2007). Application of the same analysis method to *HvXTH* cDNA targets did not reveal a correlation between AT-rich regions and protein secretion. Rather, with the exception of short regions at the 3′ end of certain targets, all *HvXTH* sequences were below the 0.6 frequency cut-off value defined by (Boettner et al. 2007) as delineating AT-rich sequences (data not shown). Prediction of the p*I* values of *HvXTH* products indicated that all had pI values less than 6.7, except for the *HvXTH4* gene product (pI 8.8), which was in fact successfully expressed and secreted by *P. pastoris* (Table 1). Further analysis of *HvXTH* cDNA sequences and predicted protein properties using *CodonW* showed no correlation between expression and the codon adaptation index (CAI), effective number of codons (Nc), GC content, protein hydrophobicity (GRAVY score), or the frequency of aromatic protein residues (data not shown). Analyses for PEST protein degradation sequence motifs (Boettner et al. 2007) were also negative for all *HvXTH* products.

Conclusion

A lack of correlation between successful protein production and gene/protein sequence characteristics indicates that there are as yet unknown sequence features that complicate the heterologous expression of plant genes in *P. pastoris*. The 50% success rate observed with barley *XTH* genes seems to reflect that observed for human genes (44 of 79 showing no expression) (Boettner et al. 2007), and supports anecdotal evidence from our own laboratory obtained with the heterologous expression of dicot *XTH* genes in *P. pastoris*. Clearly, further in-depth analysis correlating sequence or structural features with failed heterologous expression in the *P. pastoris* system will be required to improve the success rate through sequence engineering. In the future,

the use of plant systems for heterologous expression may prove to be effective as a route to obtain plant proteins/enzymes for structure-function studies (Farrokhi et al. 2008; Gleba et al. 2007; Plesha et al. 2009; Streatfield 2007). However, recent inspection of the Protein Data Bank (http://www.rcsb.org) indicates that plant expression systems have not yet been harnessed for tertiary structure studies. In this context, we have cloned the cDNAs of *HvXTH4* and *HvXTH5* with and without apoplast-targetting signal peptides into pCambia vectors for expression in tobacco. Expression of functional *HvXTH* products occurred at very low levels (G.B.F and M.H. unpublished data), and thus this approach was unsuitable for subsequent structural or any other studies. Consequently, present data suggest that *Pichia pastoris*, despite its modest success rate, still remains the most practically useful system for the heterologous expression of plant *XTH* genes, if not plant genes in general.

Acknowledgements

The authors would like to acknowledge The Ministry of Science and Technology, Thailand; The Swedish Science Council; Formas; and BIOMIME, The Swedish Center for Biomimetic Fiber Engineering, for financial support. This work was also supported by grants from the Australian Research Council and the South Australian State Government (to G.B.F. and M.H.). Dr. Martin Baumann is thanked for fruitful discussions regarding transformant screening.

References

Arrowsmith DA, de Silva J (1995) Characterization of 2 Tomato Fruit-Expressed cDNAs Encoding Xyloglucan Endo-Transglycosylase. *Plant Mol Biol* 28: 391–403

Baumann MJ, Eklöf JM, Michel G, Kallas ÅM, Teeri TT, Czjzek M, 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

Becnel J, Natarajan M, Kipp A, Braam J (2006) Developmental expression patterns of *Arabidopsis XTH* genes reported by transgenes and Genevestigator. *Plant Mol Biol* 61: 451–467

Boettner M, Steffens C, von Mering C, Bork P, Stahl U, Lang C (2007) Sequence-based factors influencing the expression of heterologous genes in the yeast *Pichia pastoris*-A comparative view on 79 human genes. *J Biotechnol* 130: 1–10

Bolot S, Abrouk M, Masood-Quraishi U, Stein N, Messing J, Feuillet C, Salse J (2009) The 'inner circle' of the cereal genomes. *Curr Opin Plant Biol* 12: 119–125

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72: 248–254

Campbell P, Braam J (1999) *In vitro* activities of four xyloglucan endotransglycosylases from *Arabidopsis*. *Plant J* 18: 371–382

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res* 37: D233–D238

Cao PJ, Bartley LE, Jung KH, Ronald PC (2008) Construction of a rice glycosyltransferase phylogenomic database and identification of rice-diverged glycosyltransferases. *Mol Plant* 1: 858–877

Catala C, Rose JKC, York WS, Albersheim P, Darvill AG, Bennett AB (2001) Characterization of a tomato xyloglucan endotransglycosylase gene that is down-regulated by auxin in etiolated hypocotyls. *Plant Physiol* 127: 1180–1192

Chanliaud E, de Silva J, Strongitharm B, Jeronimidis G, Gidley MJ (2004) Mechanical effects of plant cell wall enzymes on cellulose/xyloglucan composites. *Plant J.* 38: 27–37

Cosgrove DJ (2005) Growth of the plant cell wall. *Nat Rev Mol Cell Biol* 6: 850–861

Daly R, Hearn MTW (2005) Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. *J Mol Recognit* 18: 119–138

Davies GJ, Henrissat B (2002) Structural enzymology of carbohydrate-active enzymes: implications for the post-genomic era. *Biochem. Soc Trans* 30: 291–297

de Silva J, Jarman CD, Arrowsmith DA, Stronach MS, Chengappa S, Sidebottom C, Reid JSG (1993) Molecular characterization of a xyloglucan-specific endo-(1→4)-beta-D-glucanase (xyloglucan endotransglycosylase) from nasturtium seeds. *Plant J* 3: 701–711

Farrokhi N, Hrmova M, Burton RA, Fincher GB (2008) Heterologous and Cell Free Protein Expression Systems. *Meth Mol Biol* 513: 175–198

Fry SC, Smith RC, Renwick KF, Martin DJ, Hodge SK, Matthews KJ (1992) Xyloglucan endotransglycosylase, a new wall-loosening enzyme activity from plants. *Biochem J* 282: 821–828

Geisler-Lee J, Geisler M, Coutinho PM, Segerman B, Nishikubo N, Takahashi J, Aspeborg H, Djerbi S, Master E, Andersson-Gunneras S, Sundberg B, Karpinski S, Teeri TT, Kleczkowski LA, Henrissat B, Mellerowicz EJ (2006) Poplar carbohydrate-active enzymes. Gene identification and expression analyses. *Plant Physiol* 140: 946–962

Gleba Y, Klimyuk V, Marillonnet S (2007) Viral vectors for the expression of proteins in plants. *Curr Opin Biotechnol* 18: 134–141

Henriksson H, Denman SE, Campuzano IDG, Ademark P, Master ER, Teeri TT, Brumer H (2003) N-linked glycosylation of native and recombinant cauliflower xyloglucan endotransglycosylase 16A. *Biochem J* 375: 61–73

Henrissat B, Coutinho PM, Davies GJ (2001) A census of carbohydrate-active enzymes in the genome of *Arabidopsis thaliana*. *Plant Mol Biol* 47: 55–72

Holmquist M, Tessier DC, Cygler M (1997) High-level production of recombinant *Geotrichum candidum* lipases in yeast *Pichia pastoris*. *Protein Expr Purif* 11: 35–40

Hrmova M, Farkaš V, Harvey AJ, Lahnstein J, Wischmann B, Kaewthai N, Ezcurra I, Teeri TT, Fincher GB (2009) Substrate specificity and catalytic mechanism of a xyloglucan xyloglucosyl transferase HvXET6 from barley (*Hordeum vulgare* L.). *FEBS J* 276: 437–456

Hrmova M, Farkaš V, Lahnstein J, Fincher GB (2007) A barley xyloglucan xyloglucosyl transferase covalently links xyloglucan, cellulosic substrates, and (1,3;1,4)-beta-D-glucans. *J Biol Chem* 282: 12951–12962

Hrmova M, Fincher GB (2009) Functional genomics and structural biology in the definition of gene function. *Meth Mol Biol* 513:

199–227

Johansson P, Brumer H, Baumann MJ, Kallas ÅM, Henriksson H, Denman SE, Teeri TT, Jones TA (2004) Crystal structures of a poplar xyloglucan endotransglycosylase reveal details of transglycosylation acceptor binding. *Plant Cell* 16: 874–886

Kallas ÅM, Piens K, Denman SE, Henriksson H, Fäldt J, Johansson P, Brumer H, Teeri TT (2005) Enzymatic properties of native and deglycosylated hybrid aspen (*Populus tremula x tremuloides*) xyloglucan endotransglycosylase 16A expressed in *Pichia pastoris*. *Biochem J* 390: 105–113

Liepman AH, Wightman R, Geshi N, Turner SR, Scheller HV (2010) Arabidopsis– a powerful model system for plant cell wall research. *Plant J* 61: 1107–1121

Michel G, Chantalat L, Duee E, Barbeyron T, Henrissat B, Kloareg B, Dideberg O (2001) The kappa-carrageenase of *P. carrageenovora* features a tunnel-shaped active site: A novel insight in the evolution of clan-B glycoside hydrolases. *Structure* 9: 513–525

Minic Z (2008) Physiological roles of plant glycoside hydrolases. *Planta* 227: 723–740

Mohand FA, Farkaš V (2006) Screening for hetero-trans-glycosylating activities in extracts from nasturtium (*Tropaeolum majus*). *Carbohyd Res* 341: 577–581

Oh MH, Romanow WG, Smith RC, Zamski E, Sasse J, Clouse SD (1998) Soybean BRU1 encodes a functional xyloglucan endotransglycosylase that is highly expressed in inner epicotyl tissues during brassinosteroid-promoted elongation. *Plant Cell Physiol* 39: 124–130

Plesha MA, Huang TK, Dandekar AM, Falk BW, McDonald KA (2009) Optimization of the bioprocessing conditions for scale-up of transient production of a heterologous protein in plants using a chemically inducible viral amplicon expression system. *Biotechnol Prog* 25: 722–734

Saladié M, Rose JKC, Cosgrove DJ, Catala C (2006) Characterization of a new xyloglucan endotransglucosylase/hydrolase (XTH) from ripening tomato fruit and implications for the diverse modes of enzymic action. *Plant J* 47: 282–295

Schröder R, Atkinson RG, Langenkamper G, Redgwell RJ (1998) Biochemical and molecular characterisation of xyloglucan endotransglycosylase from ripe kiwifruit. *Planta* 204: 242–251

Sreenivasulu N, Graner A, Wobus U (2008) Barley Genomics: An Overview. *International Journal of Plant Genomics* Article ID 486258: 13 pages. doi:10.1155/2008/486258

Streatfield SJ (2007) Approaches to achieve high-level heterologous protein production in plants. *Plant Biotechnol J* 5: 2–15

Toikkanen JH, Niku-Paavola ML, Bailey M, Immanen J, Rintala E, Elomaa P, Helariutta Y, Teeri TH, Fagerström R (2007) Expression of xyloglucan endotransglycosylases of *Gerbera hybrida* and *Betula pendula* in *Pichia pastoris*. *J Biotechnol* 130: 161–170

Van Sandt VST, Guisez Y, Verbelen JP, Vissenberg K (2006) Analysis of a xyloglucan endotransglycosylase/hydrolase (XTH) from the lycopodiophyte *Selaginella kraussiana* suggests that XTH sequence characteristics and function are highly conserved during the evolution of vascular plants. *J Exp Bot* 57: 2909–2922

Yokoyama R, Nishitani K (2001) A comprehensive expression analysis of all members of a gene family encoding cell-wall enzymes allowed us to predict cis-regulatory regions involved in cell-wall construction in specific organs of arabidopsis. *Plant Cell Physiol* 42: 1025–1033

Yokoyama R, Rose JKC, Nishitani K (2004) A surprising diversity and abundance of xyloglucan endotransglucosylase/hydrolases in rice. Classification and expression analysis. *Plant Physiol* 134: 1088–1099