A xyloglucan-specific family 12 glycosyl hydrolase from *Aspergillus niger*: recombinant expression, purification and characterization

Emma R. MASTER1, Yun ZHENG, Reginald STORMS, Adrian TSANG and Justin POWLOWSKI2

Centre for Structural and Functional Genomics, Concordia University, 7141 Sherbrooke Street West, Montreal, QC, Canada H4B 1R6

INTRODUCTION

The activities of many microbial CAZymes (carbohydrate-active enzymes) participate in recycling carbon that is photosynthetically fixed as plant cell-wall polymers. Polysaccharides are a primary component of plant cell walls and include cellulose, hemicelluloses and pectin: these can form intermolecular linkages and generate complex biocomposite structures. CAZymes have been identified in bacteria, fungi, plants and animals, and these enzymes have been classified into over 100 distinct GH (glycosyl hydrolase) families ([1]; http://afmb.cnrs-mrs.fr/CAZY/).

In addition to playing an essential role in carbon cycling, CAZymes may be used in many types of industrial processes. For example, forest industries use xylanases and mannanases to boost pulp bleaching, and cellulases to recover recycled fibres [2,3], while food industries employ lipases, pectinases, cellulases, amylases and xylanases to affect the texture and digestibility of their products [4,5]. Emerging technologies for the exploitation of biomass also require CAZymes. Examples include production of biofuels from lignocellulose and fatty waste products using cellulases and xylanases, whose structure has been determined by X-ray crystallography [10,11]. The second was AnXEG12A, a previously uncharacterized GH12 family enzyme with sequence similarity to known XEGs (xyloglucan-specific endo-β-1,4-glucanases (EC 3.2.1.151)). Four xyloglucanases from *Aspergillus* have been characterized previously: a GH74 enzyme from *A. niger* (EglC), a xyloglucanase from *Aspergillus japonicus* and two GH12 enzymes from *Aspergillus aculeatus* (AaXEG12 and AaXEG12-2) [12–15].

Several examples of GHs that are used in industrial processes belong to the GH12 family, and programmes to isolate and characterize family 12 hydrolases systematically for industrial applications have been initiated [16,17]. Enzymes from the GH12 family hydrolyse β-1,4-glucose linkages with a retaining mechanism, and most of those that have been tested act on have been processed by a kexin-type protease, which removed a short prosequence. The substrate specificity was restricted to xyloglucan, with cleavage at unbranched glucose in the backbone. The apparent kinetic parameters were similar to those reported for other xyloglucan-degrading endoglucanases. The pH optimum (5.0) and temperature resulting in highest enzyme activity (50–60°C) were higher than those reported for a GH12 family xyloglucanase from *Aspergillus aculeatus*, but similar to those of cellulose-specific endoglucanases from the GH12 family.

Phylogenetic, sequence and structural comparisons of GH12 family endoglucanases helped to delineate features that appear to be correlated to xyloglucan specificity.

Key words: *Aspergillus*, carbohydrate, endoglucanase, glycosyl hydrolase (GH), xyloglucanase, xyloglucan-specific endo-β-1,4-glucanase (XEG).

In a recent genomics-based enzyme discovery initiative, ESTs (expressed sequence tags) that encode putative secreted enzymes were isolated from a broad range of fungi (https://fungalgenomics.concordia.ca). Fungal species were selected based on having adapted to environmental niches that encourage production of diverse enzymes potentially useful in industrial applications. Many genes that encode putative CAZymes from *Aspergillus niger* were discovered in this analysis [9], including two family 12 GHs. One of these was EglA, a cellulase active on CM-cellulose and β-glucan, whose structure has been determined by X-ray crystallography [10,11]. The second was AnXEG12A, a previously uncharacterized GH12 family enzyme with sequence similarity to known XEGs (xyloglucan-specific endo-β-1,4-glucanases (EC 3.2.1.151)). Four xyloglucanases from *Aspergillus* have been characterized previously: a GH74 enzyme from *A. niger* (EglC), a xyloglucanase from *Aspergillus japonicus* and two GH12 enzymes from *Aspergillus aculeatus* (AaXEG12 and AaXEG12-2) [12–15].

Several examples of GHs that are used in industrial processes belong to the GH12 family, and programmes to isolate and characterize family 12 hydrolases systematically for industrial applications have been initiated [16,17]. Enzymes from the GH12 family hydrolyse β-1,4-glucose linkages with a retaining mechanism, and most of those that have been tested act on...
unbranched polymers, including cellulose and β-glucan [17,18]. Phylogenetic analyses have allowed differentiation of four subfamilies: fungal GH12 enzymes fall into subfamilies 12-1 and 12-2, enzymes from Streptomyces fall into subfamily 12-3, and subfamily 12-4 comprises enzymes from thermophiles [16,19].

A new class of polysaccharide-degrading enzymes is emerging from this family, which specifically hydrolyses xyloglucan, XEG [13,19]. Xyloglucan is found as a storage polymer in dicotyledon seeds and is believed to function in cell-wall expansion in primary cell walls of dicotyledonous plants and non-graminaceous monocots [20] and references therein. Like cellulose, this polymer consists of a linear backbone of β-1,4-glucan linkages, but is distinguished by having up to 75% of β-D-Glcp (β-D-glucopyranose) residues covalently linked to α-D-Xylp (α-D-xylopyranose) at the O-6 position [21]. Depending on the source of xyloglucan, some α-D-Xylp residues are further linked to β-D-Galp (β-D-galactopyranose) or α-L-Araf (α-L-arabinofuranosan), and some galactose residues may be extended by α-L-Fucp (α-L-fucopyranose) [21]. Xyloglucanases have been used to characterize molecular alterations in plant cell-wall mutants [22], engineer the physical properties and biochemical composition of plant fibres [23] and link chemically modified xylo-oligosaccharides to xyloglucan polymers that can be used to impart novel functionality to cellulose microfibrils [7]. There is also growing interest in applying xyloglucanases for food and textile processing and in pharmaceutical applications [24,25].

To date, xyloglucanases have been isolated from bacteria, fungi and plants, and, in addition to representatives from GH12, include examples from GH74 and GH5 [12,14,26,27]. The CAZY database also lists examples of xyloglucanases from GH26 and GH44, which are the subject of patent applications or are unpublished, as well as other enzymes having xyloglucanase activity in addition to other activities. Most of the xyloglucan-specific hydrolases that have been well characterized belong to family GH74, and the crystal structures of two such enzymes, XGH74A from Clostridium thermocellum, and OXG-RCBH from Geotrichum sp. M128, have been reported [28,29]. Structures of one each of GH5 and GH12 bacterial xyloglucan-specific endoglucanases have also been published recently [30]. For the GH12 xyloglucanase, the tertiary structure was very similar to that of the closest non-xyloglucan-degrading GH12 relative, with some amino acid substitutions that appear to favour xyloglucan binding. However, since the abilities to hydrolyse xyloglucan have not been reported for most of the GH12 family endoglucanases, these authors concluded that it was difficult to generalize about substrate specificity determinants from comparison of only two structures [30].

Here, we report the expression, purification and characterization of a family 12 xyloglucanase from A. niger (AnXEG12A) and discuss molecular features that distinguish xyloglucanases from other GH12 enzymes. The identification of structural features related to substrate specificity may allow future modification of the activities of these and other GH12 endoglucanases.

MATERIALS AND METHODS

General

SDS/PAGE analyses were performed using 10 or 12 % resolving gels and prestained protein markers from Bio-Rad. Proteins were visualized by Coomassie Blue staining [31] or silver staining [32]. Protein concentration was determined using the DC Protein Assay (Bio-Rad) or the BCA (bicinchoninic acid) assay (Pierce) with BSA as the standard, according to the manufacturer’s instructions. Spectrophotometric measurements were performed using a Cary 50 UV–visible spectrophotometer (Varian Instruments).

Enzyme substrates

Low-viscosity CM-cellulose, laminarin, lichenan (Icelandic moss), xylan from beechwood, birchwood and oat spelt were purchased from Sigma. CM-cellulose-4M, CM-pachyman, pullan, CM-curdlan, Konjac glucomannan, galactomannan and xyloglucan from tamarind seed (high viscosity) were purchased from Megazyme. Substrates were usually prepared as 1 % stock solutions, except CM-pachyman and CM-curdlan, which were prepared as 0.5 % stock solutions in MilliQ water.

Gene cloning and recombinant expression of AnXEG12A

An EST database for A. niger was previously assembled as part of a gene discovery programme for fungal species adapted to different environmental niches [9]. A gene predicted by BLAST [33] sequence comparisons to encode a xyloglucan-specific endoglucanase was designated AnXEG12A and selected for further study. For recombinant expression, AnXEG12A was amplified from cDNA by PCR using the following forward and reverse primers respectively: 5′-GGCAGCCCATGAAGGTT-TCTCGCTTTTTCCG-3′ and 5′-AGCCTAAGCGCCCTTAC-TCGATGGAAACGGAGTACT-3′.

Gateway LR Clonase™ (Invitrogen) was used to transfer the amplified fragment to ANI7G, a Gateway-compatible integrative expression vector previously constructed from ANI7p [34]. Protoplasts of A. niger strain NS93 glaA::hisG were transformed with ANI7pG_AnXEG12A as described previously [35], and transformants were selected on MM (minimal medium) [36] without uracil and uridine.

A. niger transformants for AnXEG12A expression were cultured without shaking in MMJ, which is similar to MM except that it includes 15 % (w/v) maltose as the sole carbon source and the amounts of nitrogen source, salts and trace elements were increased 4-fold. Triplicate cultures of 16 independent transformants, and cultures produced with pooled conidia from the same 16 transformants, were grown in 200 μl of MM J in 300 μl microtitre plates. The amount of xyloglucanase activity expressed by the 16 independent transformants was compared with the activity expressed by the pooled transformants and the same strain harbouring ANI7pG.

Production and purification of recombinant AnXEG12A

A standard Petri dish containing 25 ml of MM J medium was inoculated with positive transformants to a final concentration of 1 × 10⁶ conidia/ml, and the culture was grown for 5 days at 30°C without agitation. Culture supernatant was filtered through a 0.45 μm membrane filter (Whatman), and the filtrate (17.2 ml) was then concentrated using a 10 kDa cut-off JumboSep centrifugal filter unit (Pall). The concentrated culture medium was replaced by 10 mM sodium acetate buffer including 0.1 M NaCl (pH 5.0) by repeated dilution and concentration (to 2 ml) using the 10 kDa cut-off centrifugal filter unit and then applied to a Superdex™ 75 FPLC gel filtration column (1 cm × 30 cm; Amersham Biosciences) equilibrated with the same buffer. Fractions (1 ml) were collected at a flow rate of 1 ml/min and were assayed for activity on xyloglucan. Active fractions were pooled, concentrated and exchanged to 10 mM Tris/HCl (pH 8.0) using a 10 kDa cut-off centrifugal filter unit, and then applied to a HiTrapQ ion-exchange column (Amersham Biosciences). Fractions were eluted with a linear gradient of
0–1 M NaCl in 10 mM Tris/HCl (pH 8.0) over 20 column volumes. Fractions (0.5 ml) were collected at a flow rate of 1 ml/min and assayed for activity on xyloglucan. Fractions with peak activity were combined, concentrated and exchanged into 10 mM sodium acetate buffer (pH 5.0) using a 10 kDa cut-off Jumbosep centrifugal filter unit. All chromatographic steps were performed using an Akta chromatography system (Amersham Biosciences) at room temperature (22 °C) and fractions were stored at 4 °C. Purified protein was stored at −80 °C until further use.

MS analysis of purified AnXEG12A
The identity of the purified protein was confirmed by SDS/PAGE combined with peptide mass ‘fingerprinting’. Tryptic fragments were generated by incubating a gel fragment containing AnXEG12A, cut from an SDS/12 % PAGE gel, with trypsin (Promega; sequencing grade): the procedure followed was essentially as described by Promega (Technical Bulletin TB-309). After drying the trypsinated gel extracts, samples were redissolved in 0.1 % trifluoroacetic acid (10 μl) and purified using C18 Ziptips (Millipore) according to the manufacturer’s instructions. Purified peptide samples were either mixed with a matrix (α-cyano-4-hydroxycinnamate) and then spotted on to a MALDI (matrix-assisted laser-desorption ionization) plate for analysis using an M@LDI–TOF-LR mass spectrometer (Micromass), or diluted with acetonitrile/water (1:1, v/v) containing 0.05 % trifluoroacetic acid for analysis using a Q-TOF2 ESI (electrospray ionization) mass spectrometer (Micromass). The M@LDI mass spectrometer was operated in reflectron mode and calibrated using six peptides (bradykinin, angiotensin I, substance P, Glu-fibrinopeptide, renin and adrenocorticotropic hormone fragment 18–39) with acetonitrile prior to mixing with an equal volume of matrix. The samples were then spotted on to the MALDI plate and calibrated using Glu-fibrinopeptide.

Enzyme activity with xyloglucan
Standard activity assays were performed at 50 °C in 50 mM citrate buffer (pH 5.0). The concentration of xyloglucan was 0.5 % unless noted otherwise. Reactions were initiated by adding an appropriate dilution of enzyme and incubating for a fixed period of time before withdrawing samples for assay of reducing sugar product. Usually, reducing sugar assays were performed in microtitre plate format using BCA and CuSO4, with standard curves prepared using glucose, as described previously [37,38]. Attenuance values were corrected for background attenuation of reaction mixtures in the absence of enzyme. One unit of enzyme activity is defined as the amount of enzyme required to produce 1 μmol of product/min at 50 °C and pH 5.0. Alternatively, reactions were terminated and analysed using DNSA (3,5-dinitrosalicylic acid), and reducing ends were quantified relative to known concentrations of glucose [39]. All assays were performed at least in triplicate.

Enzyme activity with other polymeric substrates
Substrate specificity was examined in assays carried out in 50 mM citrate buffer (pH 5.0) at 50 °C with 0.23 ng/μl AnXEG12A and 0.5 % polymeric substrate for all substrates except CM-pachyman and CM-curdlan, which were present at 0.25 %. After 18 h, 10 μl of reaction mixture (diluted 20-fold only when xyloglucan was the substrate) was analysed for reducing sugar in the BCA-based assay. Samples were also analysed using the DNSA assay for comparison.

Effect of pH and temperature on AnXEG12A activity
The effects of pH and temperature on the activity of AnXEG12A were determined using xyloglucan (0.5 %) as the substrate. To determine the pH profile of AnXEG12A activity, reactions were performed at 50 °C in 50 mM glucose buffer (pH 2.0–3.5), sodium acetate buffer (pH 4.0–5.0), citrate buffer (pH 5.5–6.5), phosphate buffer (pH 7.0–7.5) or Tris/HCl buffer (pH 8.0–9.0). The effect of temperature on AnXEG12A activity was determined using 50 mM citrate buffer (pH 5.0) and incubating reaction mixtures at 22, 30, 40, 50, 60, 70 and 80 °C. Reactions were initiated by adding enzyme (0.27 ng/μl) and progressed for 15 min. Reactions were terminated and analysed using the BCA-based reducing sugar assay.

Temperature stability
For measurement of temperature stability, reaction mixtures with AnXEG12A (0.25 ng/μl) in 50 mM citrate buffer (pH 5.0) were incubated without substrate for up to 2 h at 22, 30, 40, 50, 60, 70 and 80 °C. Samples were cooled, mixed with 0.5 % xyloglucan and incubated for 30 min at 50 °C before termination and analysis using the DNSA assay described above.

Kinetic measurements
Kinetic parameters were estimated using the standard activity assay with concentrations of xyloglucan ranging from 0.0125 to 0.5 %. Reactions were initiated by adding AnXEG12A (0.25 ng/μl) and incubated for 10–30 min at 50 °C. Reactions were terminated and analysed using the BCA-based assay described above. Apparent kinetic constants were calculated using initial rates with Lineweaver–Burk plots, or using non-linear regression to fit them directly to the Michaelis–Menten equation.

MALDI–TOF (MALDI-time-of-flight) oligosaccharide analysis
Samples of xyloglucan degradation products were analysed by MALDI–TOF MS. Reaction mixtures were as described above for the standard activity assays except that the buffer was 10 mM sodium acetate buffer (pH 5.0). Aliquots were withdrawn from reaction mixtures at different times and frozen immediately at −70 °C until analysis. For analysis, samples were thawed rapidly and diluted 10–50-fold in ice-cold 50 % (v/v) acetonitrile prior to mixing with an equal volume of matrix. The matrix used was 2,5-dihydroxybenzoic acid (10 mg/ml in water) [22]. Samples (1 μl) were then spotted on to the MALDI plate and analysed using a Micromass M@LDI-LR mass spectrometer operated in positive ion mode. The spectrometer was calibrated using a mixture of six peptides, and spectra were corrected using adrenocorticotropic hormone fragment 18–39 as a lockmass, as described above: the matrix used for peptide samples was α-cyano-4-hydroxycinnamate. Spectra were compared with previously published spectra of tamarind xyloglucan degradation products [22,40].

Sequence analysis
The sequence of AnXEG12A was determined as described previously [9] and submitted to GenBank® Nucleotide Sequence Database with accession number DQ486529. CAZyme nomenclature was used to designate this enzyme based on amino acid and structural similarity to previously characterized CAZymes.
Alignment of deduced amino acid sequences of AnXEG12A with other GH12 enzymes was obtained using ClustalW (http://www.ebi.ac.uk/clustalw/) and Boxshade (http://www.ch.embnet.org/software/BOX_form.html) servers. The presence of a signal peptide was predicted using SignalP [42]. For phylogenetic analysis, 100 bootstrapped data sets were generated using SEQBOOT. Protein distance matrices were obtained using the PHYLIP 3.67 Neighbor-joining algorithm PRODIST, and trees were drawn using NEIGHBOR and FigTree version 1.0 ([43] and http://tree.bio.ed.ac.uk/software/figtree/).

Homology modelling

Homology modelling was carried out using the Swiss-Model Automated Comparative Protein Modeling Server [45]. Suitable templates were identified using the Template Identification tool with Gapped BLAST [33]. The seven highest-scoring recommended templates were selected, with E-values ranging from $6 \times 10^{-14}$ to $5 \times 10^{-15}$ over the full length of the processed AnXEG12A sequence. These templates were: Cel12A enzymes from Trichoderma reesei (PDB ID: 1HAV and variants: 1OA2 and 1OLQ), Hypocrean Schweinitzii (PDB ID: 1OA3), Streptomyces sp. 11AG8 (PDB ID: 1A4A), Humicola grisea (PDB ID: 1OLR); and EgIA from A. niger (PDB ID: 1KS5). A multiple sequence alignment with the processed AnXEG12A sequence was generated using T-Coffee [46] and used to start the modelling process in Swiss-Model Alignment mode. Models were then generated based on each of the individual templates.

RESULTS AND DISCUSSION

Isolation of a putative xyloglucan-specific endoglucanase gene from A. niger

A gene, AnXEG12A, encoding a putative xyloglucan-specific endoglucanase was identified in a CDNA library of A. niger as described in the Materials and methods section. The full sequence of the CDNA was submitted to GenBank® with accession number DQ486529. To examine the genomic organization of this gene, AnXEG12A-encoding cDNA was compared with version 1 of A. niger CBS 513.88 (GenBank® accession numbers CAK46524 and CAK38314 respectively) [47]. A fourth A. niger CBS 513.88 GH12 endoglucanase (GenBank® accession number CAK38300) aligns to Contig 14 between nt 558391 and 559125 of the A. niger A.T.C.C. 1015 genome sequence, where there is no predicted gene model. Genes encoding only two of the four enzymes, EgIA and AnXEG12A, were present in our cDNA library, derived from A. niger strain A733 (Fungal Genetics Stock Center, Kansas City, KS, U.S.A.). Of the four GH12 endoglucanases, only EgIA has been characterized previously.

Expression and purification of AnXEG12A

Of the 16 transformants screened for recombinant expression of AnXEG12A, one had a specific activity with xyloglucan that was more than two times higher than the transformant pool. This result is consistent with a previous report comparing the variability of recombinant enzyme activity detected in integrative transformants [34] and suggests that screening is necessary to identify Aspergillus host cells that optimally produce recombinant protein.

Purification of AnXEG12A is summarized in Table 1: the substrate for the activity assays was xyloglucan. Since AnXEG12A was predicted to be significantly smaller than most of the endogenous proteins that are secreted by the A. niger recombinant host, the initial purification step was size-exclusion gel chromatography. Subsequent anion-exchange chromatography resulted in two peaks with xyloglucanase activity. The first activity peak to be eluted contained 5 mg of pure protein (Table 1; Figure 1B), whereas the second peak to be eluted had approximately half as much protein, half of the specific activity, and was also pure as judged by SDS/PAGE (results not shown). Unless noted otherwise, characterization was performed with the enzyme that was eluted in the first peak.

The predicted molecular mass and pI of AnXEG12A are 24 kDa and 3.62 respectively after cleavage of the SignalP-predicted signal peptide (MKVLALSALLSLASA). The purified protein band observed on SDS/PAGE ran at a position corresponding to approx. 32 kDa (Figure 1B). Although this suggests the possibility of glycosylation, a similar discrepancy in molecular mass for an AaXEG12 xyloglucanase with approx. 80% sequence identity was shown not to be due to glycosylation [14]. As can be observed in Figure 1(A), the mass of purified AnXEG12A is $23315 \pm 2$ Da, with two minor peaks indicative of sodium adducts. The data shown are for peak 1 from the ion-exchange column, but essentially identical data were observed for peak 2 (results not shown), so if peak 2 enzyme is modified, the modification is not stable under MS conditions. The observed mass is smaller than predicted from cleavage of the signal peptide and is consistent with additional cleavage after the 333-nt-long 3′-untranslated region of the transcript.

Table 1 Purification of recombinant AnXEG12A from A. niger culture supernatant

<table>
<thead>
<tr>
<th>Sample</th>
<th>Protein concentration (mg/ml)</th>
<th>Total (mg)</th>
<th>Specific activity (units/mg)</th>
<th>Activity yield (%)</th>
<th>Fold purification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desalted culture supernatant</td>
<td>64.5</td>
<td>129</td>
<td>38.9</td>
<td>100</td>
<td>–</td>
</tr>
<tr>
<td>Gel filtration</td>
<td>1.65</td>
<td>22.3</td>
<td>71.2</td>
<td>31.7</td>
<td>1.8</td>
</tr>
<tr>
<td>Anion exchange*</td>
<td>1.33</td>
<td>5.0</td>
<td>83</td>
<td>8.3</td>
<td>2.1</td>
</tr>
</tbody>
</table>

*First activity peak eluted from the gel-filtration column.
mass of 23315 Da. Processing of short (≤20 amino acids) prosequences by KexB of the expression host, A. niger, has been reported to occur at twin basic residues in proteins such as glucoamylase [48]. These results also indicate that a putative N-glycosylation site at Asn-32 of AnXEG12A is not glycosylated.

The identity of the purified protein was further confirmed by MS analysis of peptides obtained from an in-gel trypsin digest. Complete digestion of the protein lacking signal peptide is predicted to yield five peptides, which would have monoisotopic masses of 13310.13, 6377.68, 2658.34, 1012.48 and 533.30 Da. A single major peak with a mass of 2658.43 Da (results not shown). Several minor peaks were not identified.

Properties of AnXEG12A and comparison with other family 12 xylglucanases

The $k_{cat}$ (app) and $K_m$ (app) values for AnXEG12A are similar to those of other xylglucan endoglucanases, but significantly different from those of AaXEG12 (Table 2). However, these values are difficult to compare directly, since the kinetic parameters of AnXEG12A and AaXEG12 were obtained at different optimal assay temperatures of 50 and 37°C respectively and using different assays [14]. The results obtained for the enzyme from A. japonicus are more comparable, since the assay method was the same as in the present study. In any event, it is clear from this comparison that XEG12A from A. niger does not have exceptional catalytic power compared with other known xylglucan-specific endoglucanases. The second peak of activity that was eluted in the final AnXEG12A purification step had a specific activity of 36 units/mg and a $K_m$ (app) of 0.34 mg/ml, both very similar to the values for the main peak.

Hydrolysis of xylglucan by AnXEG12A was optimal at pH 5.0 (Figure 2A) and between 50 and 60°C (Figure 2B): a similar temperature optimum was observed for the second peak of activity from the last purification step. These optima are higher than those reported for AaXEG12 [14], but are similar to those of GH12 family cellulases [17]. The half-life of AnXEG12A was greater than 2 h at temperatures up to 50°C, but the enzyme lost approx. 50% of its activity after 1 h at 60°C, and the half-life was less than 10 min at 70°C and above. Again, this is similar to the thermostability of other family 12 GHs [17], but differs from AaXEG12, which lost 80% of its activity after 2 h at 50°C [14]. Thus despite sharing more than 85% sequence identity, AnXEG12A is more stable at a high temperature than AaXEG12. Interestingly, the predicted amino acid sequences of AnXEG12A and AaXEG12 both include residues correlated with improved thermostability of GH12 endoglucanase from Hypocre a jecorina Cell12A [17], including valine for alanine at position 35 and aspartic acid for asparagine at position 174 (Cell12A numbering).

Activity of AnXEG12A on polymeric substrates

AnXEG12A was tested on a range of polymeric substrates, including xylglucan, CM-cellulose-4M, xylan (from birch, beech or oat spelt), CM-pachyman, CM-curdlan, lichenan, laminarin, pullulan, galactomannan and glucomannan. After 18 h incubation, reducing sugar products were detected with xylglucan from tamarind seed, but not from most of the other polymers. Small (≤2% relative to xylglucan) amounts of colour were reproducibly observed in reaction mixtures containing CM-pachyman and CM-curdlan. Hydrolysis of these β-1,3-glucans by XEG12A seems unlikely, and product formation from these substrates remains to be confirmed using more sensitive and specific methods. An essentially identical substrate profile was observed for the second peak of enzyme activity eluted from the final column in the purification protocol (results not shown).

Table 2  Kinetic properties of AnXEG12A and other xylglucanases with xylglucan from tamarind seed

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Molecular mass (kDa)</th>
<th>GH family</th>
<th>Specific activity (units/mg)</th>
<th>$K_m$ (mg/ml)</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>AnXEG12A</td>
<td>23.3</td>
<td>12</td>
<td>86 ± 16</td>
<td>0.39 ± 0.10</td>
<td>34 ± 6.4</td>
<td>The present study</td>
</tr>
<tr>
<td>AaXEG12</td>
<td>23.6</td>
<td>12</td>
<td>260</td>
<td>3.6</td>
<td>113</td>
<td>[14]</td>
</tr>
<tr>
<td>Maitrebranchia XEG</td>
<td>25</td>
<td>12</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>[54]</td>
</tr>
<tr>
<td>Bacillus licheniformis</td>
<td>26</td>
<td>12</td>
<td>2</td>
<td>NA</td>
<td>0.88</td>
<td>[30]</td>
</tr>
<tr>
<td>Aspergillus japonicus XEG</td>
<td>32</td>
<td>NA</td>
<td>98 ± 5</td>
<td>0.67 ± 0.09</td>
<td>53 ± 2</td>
<td>[13]</td>
</tr>
<tr>
<td>Chrysoaspernum lucknowense XEG</td>
<td>78</td>
<td>NA</td>
<td>75 ± 4</td>
<td>0.31 ± 0.04</td>
<td>99 ± 3</td>
<td>[13]</td>
</tr>
<tr>
<td>Aspergillus niger Eg1C</td>
<td>90.5</td>
<td>74</td>
<td>19</td>
<td>NA</td>
<td>NA</td>
<td>[12]</td>
</tr>
<tr>
<td>Geotrichum XEG</td>
<td>80</td>
<td>74</td>
<td>68</td>
<td>NA</td>
<td>NA</td>
<td>[26]</td>
</tr>
<tr>
<td>Tricholoma reesei XEG</td>
<td>75–105</td>
<td>74</td>
<td>45 ± 5</td>
<td>0.30 ± 0.06</td>
<td>55 ± 4</td>
<td>[13]</td>
</tr>
<tr>
<td>Paenibacillus sp. KM21 XEG74</td>
<td>105</td>
<td>74</td>
<td>44</td>
<td>1.2</td>
<td>77.6</td>
<td>[27]</td>
</tr>
<tr>
<td>Thermobifida fusca</td>
<td>79</td>
<td>74</td>
<td>12.2</td>
<td>0.032</td>
<td>16.1</td>
<td>[49]</td>
</tr>
<tr>
<td>Clostridium thermocellum</td>
<td>NA</td>
<td>74</td>
<td>NA</td>
<td>0.76</td>
<td>0.79</td>
<td>[28]</td>
</tr>
<tr>
<td>Paenibacillus sp. KM21 XEG5</td>
<td>40</td>
<td>5</td>
<td>18.4</td>
<td>2</td>
<td>12.7</td>
<td>[27]</td>
</tr>
<tr>
<td>Paenibacillus tabvl</td>
<td>40</td>
<td>5</td>
<td>218</td>
<td>NA</td>
<td>145</td>
<td>[30]</td>
</tr>
</tbody>
</table>
Sequence and structural comparison of GH12 enzymes

A phylogenetic tree including 102 GH12 sequences was constructed using representatives from archaebal, bacterial and eukaryotic micro-organisms (Figure 4). Subfamilies 12-1 to 12-4 were distinguished, as originally reported by Goedegebuur et al. [16]. The archaebal enzymes clustered in subfamily 12-4, which also contains thermophilic enzymes from bacterial species. There are potentially more subfamilies, as indicated by the presence of some sequences that do not cluster within the four main subfamilies. One of those that clusters close to subfamily 12-1 contains fungal enzymes, including an uncharacterized GH12 endoglucanase from A. niger (GenBank® accession number CAK38300.1), while a second contains bacterial enzymes from plant-associated bacterial species, including Xanthomonas campestris (GenBank® accession number AAM42650.1) and Xanthomonas oryzae (GenBank® accession number AAW74331.1). AnXEG12A clustered with subfamily 12-2 enzymes, and the 18 paralogues of xyloglucanases recently identified in Phytophthora sojae and Phytophthora ramorum [18]. Subfamily 12-2 includes four other proven xyloglucan-specific endoglucanases: one from A. nidulans (GenBank® accession number EAA66551.1, [52]); AaXEG12A (GenBank® accession number O94218, [14]); and two Malbranchea xyloglucanases (GenBank® accession number AAN89225, [53]; and AAO95986, [54]), of which one (AAN89225) was also shown to possess XET [xyloglucan:xyloglucosyl transferase (EC 2.4.1.207)] activity. No data on xyloglucanase activity have been reported for most of the other enzymes in family GH12, with the exception of xyloglucanase from Bacillus licheniformis, BXG12 (GenBank® accession number AAN21381.1), which, uniquely, falls in a cluster close to subfamily 12-3 (Figure 4). However, unlike AnXEG12A, BXG12 is not a xyloglucan-specific enzyme, but also hydrolyses CM-cellulose, glucomannan and β-glucan [30].

GH12 sequences from subfamilies 12-1 and 12-2, as well as the small subcluster adjacent to 12-1 (Figure 4), all from fungi, were compared (Figure 5) to identify residues that could contribute to the substrate preference of GH12 enzymes. To date, fungal enzymes with a strong preference for xyloglucan have been reported only in subfamily 12-2 and not in 12-1. The top eight sequences cluster into subfamily 12-2, the next six into subfamily 12-1 and the last three in the subcluster. The catalytic trio of acidic residues (Asp-100, Glu-115 and Glu-201 in AnXEG12A) is present in all sequences shown in Figure 5 (asterisks), and cysteine residues (Cys-4 and Cys-33) that form a disulfide bridge in other family members [17,55] are also largely preserved. A short deletion and insertion follow AnXEG12A residues Asp-112 and Ile-129 respectively that are conserved among the subfamily 12-1 and the last three in the subcluster. The catalytic trio of acidic residues (Asp-100, Glu-115 and Glu-201 in AnXEG12A) is present in all sequences shown in Figure 5 (asterisks), and cysteine residues (Cys-4 and Cys-33) that form a disulfide bridge in other family members [17,55] are also largely preserved. A short deletion and insertion follow AnXEG12A residues Asp-112 and Ile-129 respectively that are conserved among the subfamily 12-2 sequences but not subfamily 12-1 sequences (Figure 5, horizontal bars). Homology-based modelling was used to examine where these regions are likely to be found in the tertiary structure.

As described in the Materials and methods section, models were derived using structures of GH12 cellulases from T. reesei, H. Schweinitzii and H. grisea, and A. niger. A representative AnXEG12A model based on the structure of the CM-cellulose-prefering T. reesei Cel12A enzyme is shown in Figure 6, where it can be seen that the backbones of the two structures overlap extensively. Further comparison reveals that the xyloglucanase-associated deletion (following Asp-112 in AnXEG12A) would shorten a loop region that in Cel12A (Figure 6, green) construits the substrate-binding cleft somewhat at this location: an even

Reaction products from standard activity assays with tamarind xyloglucan and AnXEG12A were analysed by MALDI–TOF MS following incubation with 0.25 ng/μl enzyme at 50°C for 5, 10, 30 and 60 min with tamarind xyloglucan: the increase in concentration of reducing sugars over this time course was approximately linear (results not shown). Tamarind xyloglucan consists of repeating backbone tetramers of XXXG, XXLG (or XLXG) and XLLG, where G is an unbranched residue that is decorated by α-1,6-D-Xylp and L is a β-D-Glcp residue with a β-D-Galp-(1→2)-α-D-Xylp branch. From 5 to 60 min, the main product peaks detected are consistent with the masses of sodium adducts of XXXG (m/z 1085), XXLG or XLXG (m/z 1247), XLLG (m/z 1409), and peaks corresponding to sodium adducts of four of the possible five octameric subunits (m/z 2291, 2453, 2615 and 2778) [22,40] (Figure 3A). A similar reaction run with a 100-fold higher enzyme concentration for 2 h showed that the octameric subunits were eventually degraded to tetromers (Figure 3B): the spectrum shown in Figure 3(B) also shows similar relative ratios for the three tetrmeric units, as has been reported for other endo-xyloglucanases characterized to date [13,27]. No peaks at m/z 659, 629 or 497, representing sodium adducts of LG, XX or XG respectively [40], were detected for samples shown in Figure 3(A) (60 min) or Figure 3B (results not shown). Furthermore, no cleavage at branched glucose residues was detected in various defined oligosaccharides (E. R. Master and J. Powlowski, unpublished work), unlike published results that showed that Cel74A endoglucanase from H. jecorina [50] and the GH74 oligoxylglucan reducing-end-specific cellobiohydrolase from Geotrichum sp. M128 [51] can cleave at branched glucose residues. Taken together, the results reported here indicate that AnXEG12A cleaves xyloglucan at unbranched glucose residues.

Figure 2 Effects of pH (A) and temperature (B) on hydrolysis of xyloglucan from tamarind seed by AnXEG12A

Results are means ± S.D. (n = 3).
longer deletion in GH11 xylanases contributes to a less constricted substrate-binding cleft compared with Cel12A [55]. It should be noted that the sequence between the active site residues Asp-100 and Glu-115 has few conserved residues compared with other regions in the sequences shown in Figure 5, or in a more comprehensive alignment of 52 GH12 sequences (results not shown). This, together with the preponderance of serine, alanine and glycine residues, suggests that this part of the structure may be quite flexible; however, despite this, the catalytic triad residues align very well in the substrate-binding grooves of the two structures (Figure 6). The insertion SST [AnXEG12A residues 130–132 (see Figure 5)] that is associated with the modelled xyloglucanase (Figure 6, red) is adjacent to the so-called ‘cord’ region identified in the structures of Cel12A and related enzymes. The cord contributes residues (especially the strongly conserved Pro-128 and Ile-129) to the substrate-binding cleft that are likely to be involved in binding the reducing end of the substrate [55]; hence, the presence of the insertion may alter the substrate binding properties of xyloglucanase relative to Cel12A. The AnXEG12A models derived using the other homology modelling templates showed similar structural differences in these two regions (results not shown), so it appears that both are likely to contribute to the substrate preference of AnXEG12A and related xyloglucanases.

The glycosylation site observed in Cel12A (Asn-164) [55] is not conserved in the sequence alignment with AnXEG12A, and, although a putative glycosylation site is present at Asn-32, the MS data reported above suggest that it is not modified.

Uncharacterized A. niger GH12 enzymes

As discussed above, genome sequences of A. niger indicate four predicted GH12 endoglucanases. Now that AnXEG12A and EglA have been characterized, it would be interesting to examine the properties of the other two predicted A. niger GH12 enzymes (GenBank® accession numbers CAK38300 and CAG17892). As noted above, CAK38300 is a member of a distinct small subcluster of the phylogenetic tree: Figure 5 shows that while it contains an insertion at the position of the SST insertion in AnXEG12A, the
sequence is instead GWT, and the distinctive deletion observed in subfamily 12-2 sequences is not present. On the other hand, the protein encoded by CAG17892 clusters in subfamily 12-2 (Figure 4), has a GNS insertion at the SST position, and also no deletion. Thus it will be interesting to see where the substrate specificities of these enzymes lie.
specific hydrolase, AnXEG12A, from A. niger.

In conclusion, the present study reports a family 12 xyloglucan-specific hydrolase, AnXEG12A, from A. niger that was cloned, recombinantly expressed to a high level and characterized as part of a programme to isolate fungal enzymes with potential for scientific and industrial applications. Although two peaks of activity were eluted from the final step of the purification protocol, the only observed difference in the smaller peak was a specific activity one-half that of the main peak. AnXEG12A was more stable at higher pH and temperature than other family 12 endoxylglucanases characterized to date. In addition, phylogenetic, sequence and structural comparisons revealed residues that may be important for controlling substrate specificity in GH12 family enzymes: site-directed mutagenesis studies are being performed to test these hypotheses. The isolation and characterization of AnXEG12A contributes to a growing class of xyloglucan-specific hydrolases and provide data that should prove useful for eventually engineering xyloglucanases with improved characteristics.

We thank Alain Tessier from the Centre for Biological Mass Spectrometry, Concordia University, for help with acquiring the mass spectra. We also thank other members of the Fungal Genomics group who contributed to the gene discovery, bioinformatics and enzyme assay platforms. This work was funded by Genome Canada/Genome Quebec, a Strategic Projects grant (to A. T. and R. S.) and a Discovery Grant (to J. P.) from the NSERC (Natural Sciences and Engineering Research Council) of Canada.

REFERENCES


Figure 6 Comparison of the backbone structures of Cel12A from T. reesei (PDB entry 10A2) (blue) and the AnXEG12 model (yellow) generated using Cel12A as a template for homology modelling.

The view is looking at the substrate-binding groove end-on, with the ‘cord region’ across the back. The subfamily 12-2-associated insertion in the cord region is shown in red (residues 129–132 of AnXEG12A), while a loop deleted in subfamily 12-2 (residues 108–110 (TYS) of 1OA2) is shown in green. The side chains shown belong to the conserved catalytic triad of acid residues (cyan for AnXEG12A and blue for 1OA2).

Figure 5 Alignment of the deduced amino acid sequence of AnXEG12A with other GH12 sequences. The accession number for each sequence is indicated in the alignment. The consensus sequence and similar regions are highlighted in black and grey respectively. Sequence variations that are conserved among the GH12 xyloglucanases are marked with horizontal bars, and conserved catalytic residues are indicated by asterisks (*). For details of accession numbers and organisms, see the legend to Figure 4.

Summary

In conclusion, the present study reports a family 12 xyloglucan-specific hydrolase, AnXEG12A, from A. niger that was cloned, recombinantly expressed to a high level and characterized as part of a programme to isolate fungal enzymes with potential for scientific and industrial applications. Although two peaks of activity were eluted from the final step of the purification protocol, the only observed difference in the smaller peak was a specific activity one-half that of the main peak. AnXEG12A was more stable at higher pH and temperature than other family 12 endoxylglucanases characterized to date. In addition, phylogenetic, sequence and structural comparisons revealed residues that may be important for controlling substrate specificity in GH12 family enzymes: site-directed mutagenesis studies are being performed to test these hypotheses. The isolation and characterization of AnXEG12A contributes to a growing class of xyloglucan-specific hydrolases and provide data that should prove useful for eventually engineering xyloglucanases with improved characteristics.