A sub-proteome of Arabidopsis thaliana mature stems trapped on Concanavalin A is enriched in cell wall glycoside hydrolases.

Zoran Minic, Elisabeth Jamet, Luc Négroni, P. Arsene Der Garabedian, Michel Zivy, Lise Jouanin

To cite this version:
Zoran Minic, Elisabeth Jamet, Luc Négroni, P. Arsene Der Garabedian, Michel Zivy, et al.. A sub-proteome of Arabidopsis thaliana mature stems trapped on Concanavalin A is enriched in cell wall glycoside hydrolases.. Journal of Experimental Botany, Oxford University Press (OUP), 2007, pp.1. <10.1093/jxb/erm082>. <hal-00154429>

HAL Id: hal-00154429
https://hal.archives-ouvertes.fr/hal-00154429
Submitted on 13 Jun 2007

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

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Published in Journal of Experimental Botany (2007):
doi:10.1093/jxb/erm082

A sub-proteome of Arabidopsis thaliana mature stems trapped on Concanavalin A is enriched in cell wall glycoside hydrolases

Zoran Minic¹, Elisabeth Jamet², Luc Négroni³, P. Arsene der Garabedian⁴, Michel Zivy³ and Lise Jouanin¹*

1) Laboratoire de Biologie Cellulaire - Institut National de la Recherche Agronomique, Route de Saint-Cyr, 78026 Versailles Cedex, France.
2) Surfaces Cellulaires et Signalisation chez les Végétaux, UMR 5546 CNRS-Université Paul Sabatier-Toulouse III, 24 Chemin de Borde Rouge, BP42617, 31326-Castanet-Tolosan, France.
3) Institut National de la Recherche Agronomique, Centre National de la Recherche Scientifique, Université Paris XI, Institut National Agronomique Paris-Grignon, la Ferme du Moulon, 91190 Gif-sur-Yvette, France.
4) Laboratoire d’Enzymologie des Acides Nucléiques, Université Pierre et Marie Curie, 96 Boulevard Raspail, 75006 Paris, France

*Corresponding author:
Lise JOUANIN, Institut National de la Recherche Agronomique, Route de Saint-Cyr, 78026 Versailles Cedex, France.
Tel: 33 1 30 83 30 63
Fax: 33 1 30 83 30 99
Email: jouanin@versailles.inra.fr

Abstract
N-glycosylated proteins were isolated from Arabidopsis thaliana mature stems using affinity chromatography on Concanavalin A Sepharose, separated by 2D-electrophoresis and identified using nanoHPLC-MS/MS and MALDI-TOF MS. 102 glycoproteins were identified. 94% of these proteins were predicted by bioinformatics to be targeted to the secretory pathway and 87% of them were predicted to be localized in the cell wall or at the plasma membrane. 30% of these proteins belong to glycoside hydrolases (GHs) families with some of them possibly involved in the hydrolysis of cell wall polysaccharides. The second major class of identified proteins comprises aspartyl and serine proteases. Other proteins are predicted to be oxido-reductases, contain interacting domains, are potentially involved in signalling or have an unknown function. This is to our knowledge the first survey of plant cell wall N-glycosylated proteins.

Key words: Arabidopsis thaliana, cell wall, concanavalin A, glycoside hydrolase, glycosylation, plant stem.
Introduction

Plant stems play different roles in growth and development, transport of nutrients and water and constitute a physical support for the plant. Major components of stem cell walls are polysaccharides (Carpita and Gibeaut, 1993; Roberts, 1994). Polysaccharides represent up to 90% of plant cell walls and constitute three different kinds of polymers: cellulose, hemicelluloses and pectins. Plant cell wall polysaccharide composition and structure change during plant development and are different from one plant species to another (Cosgrove, 1997; Popper and Fry, 2003). Cell wall proteins (CWPs) contribute to wall architecture or are involved in the regulation of growth and development, or defence against biotic or abiotic stresses (Lee et al., 2004; Jamet et al., 2006). Cell wall modifying proteins such as glycoside hydrolases (GHs), esterases, transglycosylases, lyases and peroxidases are involved in the construction, remodelling or turnover of cell wall components (Hernández et al., 1995; Cosgrove, 1997; Fry, 2004; Stolle-Smits et al., 1999; Obel et al., 2002; Reiter, 2006).

The enzymes of GH and transglycosylase superfamilies are particularly important for the reorganization of cell wall polysaccharides after their deposition (Fry, 2004; Minic and Jouanin, 2006). They fall into several families whose distinction is based on amino acid sequence similarities (Henrissat, 1991; 1998). Exo-glycanases attack polysaccharides progressively from the non-reducing end or substituted side groups, thus releasing monosaccharides. Endo-glycanases attack polysaccharide backbones in an endo-fashion. They have a large impact on the molecular mass of polysaccharides. A third group of hydrolases can break some substituted non-carbohydrate groups linked to wall polysaccharides such as O-acetyl, O-methyl and O-feruloyl groups (Fry, 2004). Xyloglucan transglycosylase hydrolases (XTHs) can exhibit both endo-glycanase and transglycosylase activities (Fry, 2004).

Significant progress has been made in proteomic analysis of plant cell walls (Jamet et al., 2006). Interesting results were obtained using cell cultures, culture medium of seedlings, leaves, etiolated hypocotyls, protoplasts and roots (Chivas et al., 2002; Borderies et al., 2003; Boudart et al., 2005; Charmont et al., 2005; Kwon et al., 2005; Jamet et al., 2006; Zhu et al., 2006). One cell wall proteome of mature stems was described in *Medicago sativa* L. (Watson et al., 2004). All these studies were based either on elution of cell wall proteins from living cells or on extraction of proteins from purified cell walls with salt solutions. However, since all CWPs are secreted proteins, they can be N-glycosylated with sugars such as D-glucose and D-mannose during their passage through endoplasmic reticulum and Golgi (Lerouge et al., 1998). It should be possible to trap them on Concanavalin A (Con A) which is a lectin extracted from *Canavalia ensiformis* L. able to bind molecules containing α-D-mannopyranosyl, β-D-glucopyranosyl or sterically-related residues (Carlsson et al., 1998). Recently, the N-glycoproteomes of human urine and human bile were analysed using Con A Sepharose affinity chromatography followed by 2D-electrophoresis and mass spectrometry (Kristiansen et al., 2004; Wang et al., 2006). The majority of the proteins identified were predicted to be extracellular or membrane components. Con A affinity chromatography was also used for the characterisation of N-linked glycoproteins of *Ceanorhabditis elegans* (Kaji et al., 2003) and of GHs from various plant organs (Sheldon et al., 1998; Wilson and Altmann, 1998; Minic et al., 2004; Li and Kushad, 2005; Minic et al., 2006; Van Riet et al., 2006). In this work, we have developed a new proteomic approach starting from a crude protein extract and using Con A Sepharose affinity chromatography to identify soluble cell wall N-linked glycoproteins. This glycoproteome is significantly enriched for putative cell wall GHs compared to previous cell wall proteomes.
Materials and methods

Plant material
Wild-type Arabidopsis thaliana, Wassilewskija ecotype, was grown in the greenhouse at 20°C to 22°C with a 16 h-photoperiod at 150 μE.m⁻².s⁻¹. Inflorescence stems of plants at mature stage (18–22 cm) were used for analysis.

Preparation of protein extracts from stems of A. thaliana
Mature stems of A. thaliana measuring 18-22 cm in length at the late flowering stage were used for analysis. Approximately 10 g of stem tissues were suspended in 12 mL of ice-cold extraction buffer and grinded in a mortar with a pestle for 5 min. The extraction buffer consisted in 25 mM BisTris pH 7.0 (HCl), 200 mM CaCl₂, 10% (v/v) glycerol, 4 μM Nacacodylate, 1/200 (v/v) protease inhibitor cocktail (P-9599, Sigma Chemical, St Louis, MO, USA). The ground material was centrifuged twice at 4°C for 3 min at 10,000 g, and the supernatant was further centrifuged for 15 min at 17,000 g. The resulting supernatant was used for chromatographic analyses.

Con A Sepharose affinity chromatography
A 1 x 6-cm column was filled with 3 mL of Con A Sepharose (Sigma Chemical, St Louis, MO, USA) and washed with 6 mL of 20 mM Tris pH 7.4 (HCl), 1 mM CaCl₂/MgCl₂/MnCl₂ and 0.5 M NaCl buffer. The soluble protein extract (10 mL) was added and then washed with 15 mL of this buffer at a flow rate of 5 mL.h⁻¹. Proteins were eluted with 0.2 M methyl-α-glucopyranoside in the same buffer. The eluate was collected, concentrated by "Ultrafree-CL" (10 kDa) (Sigma Chemical, St Louis, MO, USA) to 300 μL and dialysed against 7 M urea, 5 mM K₂CO₃, 0.125% SDS, 0.6% Triton X-100, 1 mM DTT, 2% carrier ampholytes 3-10 (GE Healthcare Europe GmbH, Orsay, France).

Glycoside hydrolase activities
The reaction mixture contained 2 mM pNP-glycosides (Sigma Chemical, St Louis, MO, USA), 0.1 M acetate buffer (pH 5.0), 2 mM sodium azide, and 50 μL of protein extract in a total volume of 0.5 mL. The reaction was carried out at 37°C for 5 to 60 min (depending on activity) and stopped by the addition of 0.5 mL 0.4 M sodium chloride. Controls were stopped at time 0. Concentration of the resulting pNP was determined spectrophotometrically at 405 nm by comparison to a calibration curve. Standard deviations values for 3 replicate assays were less than 5%.

2D-electrophoresis
Isoelectric focusing (IEF) was performed using 24 cm-immobilized pH gradient (IPG) strips (GE Healthcare Europe GmbH, Orsay, France) with a linear pH gradient from 4 to 7 and 250 μg of protein were applied on an IPG strip for in-gel rehydration in 7 M urea, 2 M thiourea, 2% CHAPS, 10 mM DTT, 2% IPG buffer pH 4-7 (Méchin et al., 2004). Focusing was achieved using a Protean IEF Cell (Bio-Rad, Hercules, CA, USA). An active rehydration was performed at 22°C during 12 h at 50 V prior to focusing. To improve sample entry, the voltage was increased step by step from 50 to 10,000 V (0.5 h at 200 V, 0.5 h at 500 V, 1 h at 1000 V then 10,000 V for a total of 94,000 V h). After IEF, IPG strips were successively incubated in 50 mM Tris pH 8.8 (HCl), 6 M urea, 30% glycerol, 2 % SDS, 1% DTT for 15 min, and in 50 mM Tris pH 8.8 (HCl), 6 M urea, 30% glycerol, 2 % SDS, 2.5% iodoacetamide for 15 min (Görg et al., 1987). Strips were further sealed on top of the 1 mm-thick second dimensional gel (24 x 24 cm) with the help of 1% low-melting agarose in SDS-electrophoresis buffer (25 mM Tris, 0.2 M glycine, 0.1% SDS). Continuous gels (11% T,
2.67% C gels with PDA as cross-linking agent) were used. Separation was carried out at 20 V for 1 h and subsequently at a maximum of 30 mA/gel, 120 V overnight, until the bromophenol blue front had reached the end of the gel.

**Protein staining**
Following 2D-electrophoresis, gels were stained with colloidal Coomassie blue G250 according to Mechlin et al. (2004).

**Identification of proteins by mass spectrometry**
The individual protein spots obtained after 2D-electrophoresis were excised and in-gel digested with trypsin according to a standard protocol (Santoni et al., 2003). Tryptic peptides from each protein were analyzed by nanoHPLC-MS/MS or MALDI-TOF MS as previously described (Minic et al., 2004; Mechlin et al., 2004). Proteins analyzed by MALDI-TOF MS were identified via automated NCBI non redundant protein database (http://www.ncbi.nlm.nih.gov/) searching using the MASCOT programme (http://www.matrixscience.com/search_form_select.html). Only mowse scores exceeding threshold (p<0.5) were considered as positive results. Identification of proteins with nanoHPLC-MS/MS (ion trap) was performed with Biowoks™ (Thermo scientific, San Jose, USA). The main search parameters were methionine oxidation as differential modification and trypsin as enzyme. One miss cleavage was allowed. The *A. thaliana* protein database was downloaded from the mips website (http://mips.gsf.de/projects/plants). Identification was considered significant when the proteins were identified with at least 2 different tryptic peptides as first candidate, Xcorr > 1.7, 2.2 and 3.3 for respectively mono-, di- and tri-charged peptides and delta Cn >0.1.

**Bioinformatics analyses**
Sub-cellular localization and length of signal peptides were predicted using PSORT (http://psort.nibb.ac.jp/) and TargetP (http://www.cbs.dtu.dk/services/TargetP/) (Nielsen et al., 1997; Emanuelsson et al., 2000). Prediction of transmembrane domains was done with Aramemnon (http://aramemnon.botanik.uni-koeln.de/) (Schwacke et al, 2003). Molecular masses and pI values were calculated using the aBi program (http://www.up.univ-mrs.fr/~wabim/d_abim/compo-p.html). Homologies to other proteins were searched for using BLAST programs (http://www.ncbi.nlm.nih.gov/BLAST/) (Altschul et al., 1990). Identification of protein families and functional domains was performed using MyHits (http://myhits.isb-sib.ch/cgi-bin/motif_scan) and InterProScan (http://www.ebi.ac.uk/InterProScan/) (Quevillon et al., 2005). GHs and CEs were classified according to the CAZY database (http://www.cazy.org/CAZY/) (Coutinho et al., 1999). Peroxidases were named as in the PeroxiBase (http://peroxidase.isb-sib.ch/index.php) (Bakalovic et al., 2006).

**Protein measurements**
Protein concentration was determined by the method of Bradford (1996) using bovine serum albumin dissolved in extraction buffer as the standard.
Results and discussion

*Extraction of glycoside hydrolases from stem tissues of *A. thaliana*

In a first attempt to study GHs from stem tissues of *A. thaliana* by using a proteomic approach it was necessary to establish a protocol for the extraction of these enzymes. Based on published experimental data on the purification of GHs from various plant organs (Sheldon et al., 1998; Wilson and Altmann, 1998; Minic et al., 2004; Li and Kushad, 2005; Van Riet et al., 2006), we developed a 2-step extraction procedure. Stems were ground in a buffer containing 200 mM CaCl$_2$, followed by Con A Sepharose affinity chromatography. CaCl$_2$ was chosen as the most efficient salt for CWP extraction (Boudart et al., 2005). This protocol is different from those used in previous cell wall proteomic studies (Feiz et al., 2006): (i) the initial step is a grinding in a buffer containing 200 mM CaCl$_2$ to release CWPs instead of a low ionic strength buffer usually used to prevent CWP elution; (ii) there is no step of cell wall isolation to avoid loosing CWPs weakly-bound to cell walls during the centrifugation steps required for cell wall isolation; (iii) the last step is an affinity chromatography to trap N-glycosylated proteins. Results show that the affinity chromatography step resulted in a significant increase in specific activities of several exo-GHs using artificial substrates compared to what was measured in the dialysed crude protein extract. These increases varied from 2.0 for β-D-xylosidase to 6.1 for β-D glucosidase (Table 1). On the basis of this observation, this protocol was adapted to analyze the N-glycoproteome of mature stems.

**Table 1.** Specific activities of several glycoside hydrolases after Con A Sepharose affinity chromatography.

All enzyme activities were measured *in vitro* at 37 °C, using 50 μL of protein and pNP-glycosides as substrates.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Specific activities (nmol/min/mg protein)</th>
<th>Recovery (%)</th>
<th>Ratio of specific activities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Crude protein extract</td>
<td>After Con A Sepharose</td>
<td>Con A Sepharose/ Crude protein extract</td>
</tr>
<tr>
<td>β-D-xylosidase</td>
<td>32</td>
<td>89</td>
<td>43</td>
</tr>
<tr>
<td>α-L-arabinofuranosidase</td>
<td>103</td>
<td>442</td>
<td>66</td>
</tr>
<tr>
<td>β-D-glucuronidase</td>
<td>10</td>
<td>22</td>
<td>34</td>
</tr>
<tr>
<td>β-D-mannosidase</td>
<td>26</td>
<td>108</td>
<td>64</td>
</tr>
<tr>
<td>α-D-mannosidase</td>
<td>47</td>
<td>172</td>
<td>56</td>
</tr>
<tr>
<td>β-D-glucosidase</td>
<td>82</td>
<td>497</td>
<td>93</td>
</tr>
<tr>
<td>α-D-galactosidase</td>
<td>85</td>
<td>468</td>
<td>84</td>
</tr>
<tr>
<td>β-D-galactosidase</td>
<td>393</td>
<td>825</td>
<td>32</td>
</tr>
<tr>
<td>α-D-glucosidase</td>
<td>8</td>
<td>31</td>
<td>59</td>
</tr>
</tbody>
</table>
Proteomic analysis after enrichment of the soluble protein extract in glycoside hydrolases by lectin affinity chromatography

The proteomic analysis was performed using protein extracts from 18-22 cm mature stems at the late flowering stage. About 10 g of stem tissues were used for the extraction of proteins. After grinding and centrifugation, the crude protein extract contained 2.5 mg of protein as determined by the Bradford method (1996). A fraction of this protein extract was subjected to Con A Sepharose affinity chromatography. Eluted proteins were concentrated and dialysed, resulting in a fraction of 400 µL containing about 300 µg of protein. A sample containing 250 µg of protein was subjected to 2D-electrophoresis. Proteins were detected by colloidal Coomassie blue staining (Fig. 1). The number of resolved spots was about 200. Fifty-seven spots resolved by 2D-electrophoresis were analyzed using MALDI-TOF. Spots corresponding to proteins having molecular mass smaller than 20 kDa were not analyzed since they are not expected to contain GHs on the basis of calculations made from genes predicted to encode such proteins (Minic and Jouanin, 2006). The other proteins visible on the gel were also analyzed, but due to small quantity or mixture with other proteins their scores were not significant. Fifteen spots localized at the basic side of the gel were subjected to tryptic digestion and proteins were identified using nanoHPLC-MS/MS. Each of them was expected to contain more than one protein since previous studies showed that most CWP are basic (Jamet et al., 2006). A total number of 102 different proteins was identified (Table 2; Tables S1 and S2 at JXB online). Many of these proteins were present in several of the spots resolved by 2D-electrophoresis suggesting post-translational modifications such as glycosylations. On the basis on these results, those spots were collected into thirty-five groups as shown in Fig 1. Conversely, as expected, most of the 15 spots at the basic side of the gel contained more than one protein.

![Image](image_url)

**Fig 1.** Analysis of *A. thaliana* proteins by 2D-electrophoresis. The 2D-gel was loaded with 250 µg of the fraction obtained after Con A Sepharose affinity chromatography from stem tissues of *A. thaliana*. The gel was stained with colloidal Coomassie blue. Fifteen spots were picked for nanoHPLC-MS/MS (1 to 15) and fifty-seven for MALDI-TOF analyses. In the larger case, numbering refers to groups of spots containing the same protein. Arrows in the circles represent same identified proteins. Molecular mass markers are indicated on the right.
Bioinformatic prediction of sub-cellular localization and N-glycosylation of identified proteins

PSORT, TargetP and Aramemnon programmes were used to predict the sub-cellular localization of the identified proteins. Seventy-seven out of the 102 identified proteins (77%) were predicted to be localized in the cell wall matrix, 13 at the plasma membrane, 6 into the endoplasmic reticulum, 2 in the cytoplasm and 3 in the chloroplast. Six proteins were predicted to be either targeted to vacuoles or to the cell wall. However, vacuolar targeting is not well-established in plants and the predictions are not yet very reliable (Hadlington and Denecke, 2000). Altogether, about 90% of the proteins have a predicted N-terminal signal peptide, which means that all of these proteins are targeted to the secretory pathway. Two proteins could not be assigned to any sub-cellular compartment due to discrepancies between predictions with PSORT and TargetP. Seven proteins were predicted to harbour a glycosyl phosphatidyl inositol (GPI) anchor. As expected, all of the identified proteins contained N-glycosylation sites as predicted by the MyHits programme (see Supplementary Table S1 at JXB).

These results show that the proposed protocol allowed the isolation of a protein fraction essentially composed of N-linked glycoproteins targeted to either the cell wall or to the plasma membrane. Despite the absence of a cell wall purification step, it should be noted that the proportion of proteins with a predicted intracellular localization was very low (12%). This protocol thus appears as an efficient alternative to previously described protocols used for A. thaliana cell wall proteomic analyses. Previous protocols include: (i) non-destructive methods such as analysis of culture media (Borderies et al., 2003; Charmont et al. 2005), washing of cells cultured in liquid medium with salt solutions (Borderies et al., 2003; Kwon et al. 2005) and vacuum infiltration of leaves (Boudart et al., 2005), (ii) destructive methods, i.e., cell wall purification, prior to CWP extraction with various buffers (Chivasa et al. 2002; Bayer et al., 2006; Feiz et al., 2006). The choice of a particular protocol will depend on the aim of the study and on the plant organ of interest.

Identification and functional classification of proteins

Identification of protein families and functional domains were performed using several bioinformatic programmes. Proteins were classified according to the predicted functional classes of CWPs proposed by Jamet et al. (2006) (Tables 2, 3). Proteins belonging to seven functional classes were found according to the presence of functional domains predicted as described in the Material and methods: (i) proteins acting on polysaccharides include GH and esterases; (ii) oxido-reductases mainly include peroxidases and multicopper oxidases; (iii) proteins with interacting domains include proteins with lectin or LRR (leucine rich repeat) domains as well as enzyme inhibitors; (iv) proteins involved in signaling processes include fasciclin AGPs (arabinogalactan proteins); (v) proteases; (vi) proteins of yet unknown function; (vii) miscellaneous. However, this classification is provisional since the biological role of many of these proteins remains to be determined (Tatosov et al., 1997). For example, an enzyme of the GH 3 family (XYL3) shows amino acid homology with β-D-xylosidase. However, it was identified as an enzyme that efficiently hydrolyzed arabinosyl residues from arabinans, suggesting that it works as an α-L-arabinofuranosidase (Minic et al., 2006).
Table 2. Proteins identified through proteomic analysis of A. thaliana mature stems

<table>
<thead>
<tr>
<th>Predicted or known functions (^a)</th>
<th>Accession</th>
<th>Spot number (^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Proteins acting on polysaccharides</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH family 1</td>
<td>At2g44450</td>
<td>7</td>
</tr>
<tr>
<td>GH family 1</td>
<td>At3g18080</td>
<td>13, 14, 15</td>
</tr>
<tr>
<td>GH family 1</td>
<td>At5g54570</td>
<td>47</td>
</tr>
<tr>
<td>GH family 1 (thioglucoside hydrolase 1) (TGG1)</td>
<td>At5g26000</td>
<td>41</td>
</tr>
<tr>
<td>GH family 3 (BXL1) (beta-xylosidase 1)</td>
<td>At5g49360</td>
<td>5, 9, 11, 13</td>
</tr>
<tr>
<td>GH family 3 (BXL2) (beta-xylosidase 2)</td>
<td>At1g02640</td>
<td>13, 15</td>
</tr>
<tr>
<td>GH family 3</td>
<td>At5g20950</td>
<td>5, 6, 8</td>
</tr>
<tr>
<td>GH family 16 (xyloglucan endotransferase) (EXGT-A1) (At-XTH4)</td>
<td>At2g06850</td>
<td>13, 14, 15</td>
</tr>
<tr>
<td>GH family 16 (xyloglucan endotransferase) (Meri5) (At-XTH24)</td>
<td>At4g30270</td>
<td>15</td>
</tr>
<tr>
<td>GH family 17</td>
<td>At3g07320</td>
<td>6, 7</td>
</tr>
<tr>
<td>GH family 17</td>
<td>At5g56590</td>
<td>11, 13</td>
</tr>
<tr>
<td>GH family 17</td>
<td>At4g31140</td>
<td>33</td>
</tr>
<tr>
<td>GH family 19 (chitinase), carbohydrate-binding module family 18 (chitin binding function)</td>
<td>At2g43570</td>
<td>31</td>
</tr>
<tr>
<td>GH family 20 (N-acetyl-beta-glucosaminidase)</td>
<td>At1g65590</td>
<td>7, 10</td>
</tr>
<tr>
<td>GH family 20 (N-acetyl-beta-glucosaminidase)</td>
<td>At3g55260</td>
<td>22</td>
</tr>
<tr>
<td>GH family 28 (polygalacturonase)</td>
<td>At1g80170</td>
<td>8, 9, 10</td>
</tr>
<tr>
<td>GH family 28 (polygalacturonase)</td>
<td>At3g07830</td>
<td>24</td>
</tr>
<tr>
<td>GH family 28 (polygalacturonase)</td>
<td>At3g16850</td>
<td>29</td>
</tr>
<tr>
<td>GH family 31 (alpha-glucosidase) (AtGLU1)</td>
<td>At5g11720</td>
<td>9, 13</td>
</tr>
<tr>
<td>GH family 31 (alpha-xylosidase) (XYL1)</td>
<td>At1g68560</td>
<td>17</td>
</tr>
<tr>
<td>GH family 32 (beta-fructofuranosidase) (ATBETAFRINGT4)</td>
<td>At1g12240</td>
<td>28</td>
</tr>
<tr>
<td>GH family 35 (beta-galactosidase) (BGAL1)</td>
<td>At3g13750</td>
<td>8</td>
</tr>
<tr>
<td>GH family 35 (beta-galactosidase) (BGAL9)</td>
<td>At2g32810</td>
<td>10</td>
</tr>
<tr>
<td>GH family 35 (beta-galactosidase) (BGAL10)</td>
<td>At5g63810</td>
<td>7, 13</td>
</tr>
<tr>
<td>GH family 38 (alpha-mannosidase)</td>
<td>At5g13980</td>
<td>13</td>
</tr>
<tr>
<td>GH family 38 (alpha-mannosidase)</td>
<td>At3g26720</td>
<td>42</td>
</tr>
<tr>
<td>GH family 51 (alpha-L-arabinofuranosidase)</td>
<td>At3g10740</td>
<td>27</td>
</tr>
<tr>
<td>GH family 79 (endo-beta-glucuronidase / heparanase)</td>
<td>At5g07830</td>
<td>10</td>
</tr>
<tr>
<td>GH family 79 (endo-beta-glucuronidase / heparanase)</td>
<td>At5g34940</td>
<td>7</td>
</tr>
<tr>
<td>CE family 8 (pectin methylesterase)</td>
<td>At3g43270</td>
<td>10</td>
</tr>
<tr>
<td>CE family 13 (pectin acylesterase)</td>
<td>At5g23870</td>
<td>9</td>
</tr>
</tbody>
</table>

**Oxido-reductases**

<p>| homolog to SKU5 (SKS4) (multicopper oxidase domain) | At4g22010 | 6 |
| homolog to SKU5 (SKS5) (multicopper oxidase domain) | At1g76160 | 7, 8 |
| homolog to SKU5 (SKS6) (multicopper oxidase domain) | At1g41830 | 7, 8, 12 |</p>
<table>
<thead>
<tr>
<th>Homolog Name</th>
<th>Accession Number</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>homolog to SKU5 (SKS7) (multicopper oxidase domain)</td>
<td>At1g21860</td>
<td>7, 12, 13</td>
</tr>
<tr>
<td>homolog to SKU5 (SKS17) (multicopper oxidase domain)</td>
<td>At5g66920</td>
<td>10</td>
</tr>
<tr>
<td>SKU5 (multicopper oxidase domain)</td>
<td>At4g12420</td>
<td>3, 6, 7, 8, 9</td>
</tr>
<tr>
<td>homolog to peroxidase (AtPrx22)</td>
<td>At2g38380</td>
<td>11</td>
</tr>
<tr>
<td>homolog to peroxidase (AtPrx28)</td>
<td>At3g03670</td>
<td>49</td>
</tr>
<tr>
<td>homolog to peroxidase (AtPrx33)</td>
<td>At3g49110</td>
<td>11, 13</td>
</tr>
<tr>
<td>homolog to peroxidase (AtPrx34)</td>
<td>At3g49120</td>
<td>23</td>
</tr>
<tr>
<td>homolog to <em>C. pepo</em> ascorbate oxidase (P37064)</td>
<td>At5g21100</td>
<td>11</td>
</tr>
<tr>
<td>homolog to berberine bridge enzyme (S)-reticulin:oxygen oxido-reductase</td>
<td>At4g20830</td>
<td>13</td>
</tr>
<tr>
<td>homolog to germin (AtGER3)</td>
<td>At5g20630</td>
<td>32</td>
</tr>
</tbody>
</table>

**Proteins with interacting domains**

<table>
<thead>
<tr>
<th>Homolog Name</th>
<th>Accession Number</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>homolog to carrot EDGP and tomato XEGIP</td>
<td>At1g03230</td>
<td>9, 10</td>
</tr>
<tr>
<td>homolog to carrot EDGP and tomato XEGIP (expressed protein (LRR repeats))</td>
<td>At1g03220</td>
<td>39</td>
</tr>
<tr>
<td>expressed protein (LRR repeats)</td>
<td>At1g33590</td>
<td>26</td>
</tr>
<tr>
<td>expressed protein (LRR repeats)</td>
<td>At1g33600</td>
<td>8, 9</td>
</tr>
<tr>
<td>expressed protein (LRR repeats)</td>
<td>At2g42800</td>
<td>14</td>
</tr>
<tr>
<td>polygalacturonase inhibiting protein (PGIP2)</td>
<td>At5g06870</td>
<td>13, 15</td>
</tr>
<tr>
<td>homolog to lectin (legume lectin beta domain)</td>
<td>At1g53070</td>
<td>13</td>
</tr>
<tr>
<td>homolog to lectin (legume lectin beta domain)</td>
<td>At1g53080</td>
<td>13, 14, 15</td>
</tr>
<tr>
<td>Homolog to lectin (legume lectin beta domain)</td>
<td>At3g16530</td>
<td>1, 13, 14, 15</td>
</tr>
<tr>
<td>Homolog to lectin (curculin-like)</td>
<td>At1g78850</td>
<td>7, 8</td>
</tr>
<tr>
<td>Homolog to lectin (curculin-like)</td>
<td>At1g78860</td>
<td>7, 8</td>
</tr>
<tr>
<td>Homolog to lectin (curculin-like)</td>
<td>At5g18470</td>
<td>7, 8, 10</td>
</tr>
</tbody>
</table>

**Signalling**

| Fasciclin-like arabinogalactan protein (AtFLA8) | At2g45470 | 1, 13, 14, 15 |
| Fasciclin-like arabinogalactan protein (AtFLA13) | At5g44130 | 30 |

**Proteases**

<p>| Homolog to aspartyl protease | At1g09750 | 11, 12, 13 |
| Homolog to aspartyl protease | At3g52500 | 8, 9           |
| Homolog to aspartyl protease | At5g07030 | 10, 11, 12, 13 |
| Homolog to aspartyl protease | At3g54400 | 21            |
| Homolog to serine protease   | At1g30600 | 7             |
| Homolog to serine protease   | At2g39850 | 50            |
| Homolog to serine protease   | At4g21630 | 12, 13, 14, 15 |
| Homolog to serine protease   | At4g21650 | 10, 13, 14, 15 |
| Homolog to serine protease   | At5g67360 | 16            |
| Homolog to serine protease   | At3g14067 | 18            |
| Xylem serine peptidase 1 (XSP1) | At4g00230 | 9             |</p>
<table>
<thead>
<tr>
<th>Homolog</th>
<th>Accession</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>homolog to serine carboxypeptidase (SCPL11)</td>
<td>At2g22970</td>
<td>14, 15</td>
</tr>
<tr>
<td>homolog to serine carboxypeptidase (SCPL12)</td>
<td>At2g22920</td>
<td>15</td>
</tr>
<tr>
<td>homolog to serine carboxypeptidase (SCPL34)</td>
<td>At5g23210</td>
<td>12, 13</td>
</tr>
</tbody>
</table>

**Miscellaneous**

<table>
<thead>
<tr>
<th>Homolog</th>
<th>Accession</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>homolog to anther specific proline-rich protein APG (lipase acylhydrolase domain)</td>
<td>At1g54000</td>
<td>13</td>
</tr>
<tr>
<td>homolog to anther specific proline-rich protein APG (lipase acylhydrolase domain)</td>
<td>At1g67830</td>
<td>12</td>
</tr>
<tr>
<td>homolog to purple acid phosphatase</td>
<td>At2g16430</td>
<td>7</td>
</tr>
<tr>
<td>homolog to purple acid phosphatase</td>
<td>At2g27190</td>
<td>13</td>
</tr>
<tr>
<td>homolog to purple acid phosphatase</td>
<td>At5g34850</td>
<td>19</td>
</tr>
<tr>
<td>homolog to <em>Chlamydomonas reinhardtii</em> apospory associated protein (aldose-1-epimerase domain)</td>
<td>At4g25900</td>
<td>13</td>
</tr>
<tr>
<td>homolog to glycerophosphoryl diester phosphodiesterase</td>
<td>At4g26690</td>
<td>35</td>
</tr>
</tbody>
</table>

**Unknown function**

<table>
<thead>
<tr>
<th>Homolog</th>
<th>Accession</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>expressed protein</td>
<td>At1g21680</td>
<td>5, 6</td>
</tr>
<tr>
<td>expressed protein</td>
<td>At1g21670</td>
<td>2, 3</td>
</tr>
<tr>
<td>expressed protein</td>
<td>At3g14920</td>
<td>14, 15</td>
</tr>
<tr>
<td>expressed protein (DUF642)</td>
<td>At4g32460</td>
<td>44, 13</td>
</tr>
<tr>
<td>expressed protein (DUF642)</td>
<td>At5g11420</td>
<td>10, 11, 12</td>
</tr>
<tr>
<td>expressed protein (DUF642)</td>
<td>At5g25460</td>
<td>12</td>
</tr>
<tr>
<td>expressed protein (DUF1005)</td>
<td>At4g29310</td>
<td>13</td>
</tr>
<tr>
<td>expressed protein (DUF1184)</td>
<td>At4g18080</td>
<td>46</td>
</tr>
<tr>
<td>expressed protein (DUF1680)</td>
<td>At5g12950</td>
<td>5</td>
</tr>
<tr>
<td>expressed protein (cupin domains)</td>
<td>At3g22640</td>
<td>34</td>
</tr>
</tbody>
</table>

**Intracellular proteins**

| homolog to thioredoxin | At1g21750 | 45 |
| GH family 1 | At2g25630 | 7 |
| homolog to serine protease | At2g05920 | 3, 9 |
| homolog to serine carboxypeptidase | At4g36195 | 43 |
| homolog to phosphorylase | At4g24350 | 12 |
| homolog to gamma-glutamyltranspeptidase | At4g39640 | 12 |
| homolog to aldo-keto reductase | At2g27680 | 37 |
| homolog to peroxiredoxin | At3g52960 | 25 |
| homolog to copper amine oxidase | At4g12990 | 21 |
| expressed protein (PH domain) | At2g30880 | 36 |
| expressed protein (RNA recognition motif) | At3g52980 | 40 |
| expressed protein | At2g41950 | 48 |
| expressed protein (CBS domains) | At3g48530 | 38 |

* Functions are predicted as described in Experimental
* Spot number refers to Fig. 1.
Table 3. *Predicted functional classes of proteins in the A. thaliana stem sub-proteome.* Functional classes have been defined according to Jamet et al. (2006). A list of all proteins identified in this study is provided in Table 2. The detailed bioinformatics functional analysis is given in Supplementary Table S1 at *JXB* online.

<table>
<thead>
<tr>
<th>Functional classes</th>
<th>Number of proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proteins acting on polysaccharides</td>
<td>31</td>
</tr>
<tr>
<td>Glycoside hydrolases</td>
<td>29</td>
</tr>
<tr>
<td>Esterases</td>
<td>2</td>
</tr>
<tr>
<td><strong>Oxido-reductases</strong></td>
<td>13</td>
</tr>
<tr>
<td>Peroxidases</td>
<td>4</td>
</tr>
<tr>
<td>Multicopper oxidases</td>
<td>6</td>
</tr>
<tr>
<td>Others</td>
<td>3</td>
</tr>
<tr>
<td><strong>Proteins with interacting domains</strong></td>
<td>12</td>
</tr>
<tr>
<td>Lectin domains</td>
<td>6</td>
</tr>
<tr>
<td>LRR domains</td>
<td>3</td>
</tr>
<tr>
<td>Others</td>
<td>3</td>
</tr>
<tr>
<td><strong>Signalling</strong></td>
<td>2</td>
</tr>
<tr>
<td><strong>Proteases</strong></td>
<td>14</td>
</tr>
<tr>
<td>Serine proteases</td>
<td>10</td>
</tr>
<tr>
<td>Aspartyl proteases</td>
<td>4</td>
</tr>
<tr>
<td><strong>Miscellaneous</strong></td>
<td>7</td>
</tr>
<tr>
<td>Homologs to phosphatase</td>
<td>3</td>
</tr>
<tr>
<td>Homologs to proline-rich protein (lipase acid hydrolase domain)</td>
<td>2</td>
</tr>
<tr>
<td>Others</td>
<td>2</td>
</tr>
<tr>
<td><strong>Unknown function</strong></td>
<td>10</td>
</tr>
<tr>
<td><strong>Intracellular proteins</strong></td>
<td>13</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>102</td>
</tr>
</tbody>
</table>

Thirty-three proteins were expected to act on polysaccharides (Table 2). Furthermore, 30 proteins (29%) belong to the superfamily of GHs, 29 of which were predicted to be extracellular or plasma membrane-associated. The second largest group comprises 16 proteases, 14 of which were predicted to be localized in the extracellular matrix. Together GHs and proteases represent 47% of identified proteins. Among other proteins, oxido-reductases, proteins with interacting domains, miscellaneous proteins, proteins of unknown function, signalling and intracellular proteins were identified.
Proteins from the same functional classes as in previous cell wall proteomic studies were found, but 37 proteins were not identified before (Chivasa et al. 2002; Borderies et al. 2003; Borner et al., 2003; Schultz et al., 2004; Boudart et al., 2005; Charmont et al. 2005; Kwon et al. 2005; Bayer et al., 2006; Feiz et al. 2006). This stem N-glycoproteome thus appears to be very specific. Among the 90 proteins predicted to be at the plasma membrane or in the cell wall, only 5 also have been found in previously described proteomes: cell suspension cultures, rosette leaves and etiolated hypocotyls (Jamet et al., 2006). They encode a β-xylanase that belongs to the GH 31 family (Atlg68560), a multicopper oxidase (Atlg76160), two lectins (Atlg78850, Atlg78860) and a protein homologous to the carrot extracellular dermal glycoprotein (EDGP) and to the tomato xylanase specific endoglucanase inhibitor protein (XEGIP) (Atlg03220) (Qin et al., 2003). However, some protein families are missing. Since proteins with molecular masses lower than 20 kDa were not analyzed, it was not possible to identify homologs to protease or pectin methylesterase inhibitors, non-specific lipid transfer proteins, and blue copper binding proteins. Only one protein homolog to germins was identified. Although several expansins, which molecular masses are between 25 and 30 kDa, were previously identified in cell wall proteomes (Jamet et al., 2006), none was found in this study. This might be explained either by their low abundance or their low level of N-glycosylation. Finally, no structural protein could be identified either because of their strong binding to the extracellular matrix, or the absence of N-glycans.

**Possible roles of proteins identified in stem tissues of A. thaliana**

Proteins acting on polysaccharides constitute the major functional class. According to Coutinho et al. (1999), they belong to 12 GH families and to 2 carbohydrate esterase (CE) families (Fig 2, Table 2). These enzyme families have diverse biological functions in defence, signalling, hydrolysis of starch, and cell wall modifications (Minic and Jouanin, 2006). A total of 18 GHs belonging to 7 different GH families were found. Three of them, α-L-arabinofuranosidase (At3g10740), α-L-arabinofuranosidase/β-D-xylosidase (At5g49360) and β-glucosidase AtGLU1 (At5g11720), were recently purified and characterized (Minic et al., 2004; Monroe et al., 1999). The XYL1 β-xylosidase and two XTHs (Meri5/At4g30270, EXGT-A1/At2g06850) have been studied previously using both biochemical and genetic approaches (Akamatsu et al., 1999; Sampedro et al., 2001; Rose et al., 2002). A pectin methyl- and a pectin acyl-esterase were also identified. Previous studies have shown that pectin methylesterase activity is inversely correlated to the growth rate of expanding tissues, suggesting its possible involvement in wall rigidification (McQueen-Mason and Cosgrove, 1995).

Possible substrates *in muro* of the majority of these enzymes are xylglucans and pectins. Most of them act as exo-enzymes whereas only enzymes belonging to GH families 16 and 28 act as endo-GHs on xylglucans and homogalacturonans, respectively. Other cell wall GHs can hydrolyse β1,4 glucan, arabinoxylan and xylan. These results suggest that xylglucan and pectins, that are composed of homogalacturonans (HG), arabinans and galactans (RG-I), undergo structural changes in mature stem. However, some GH families can act on several natural polysaccharides showing broad substrate specificity. This low specificity has been reported in the case of several purified cell wall GHs (Leach et al., 1995; Kim and et al., 2000; Sampedro et al., 2001; Steele et al., 2001; Rose et al., 2002; Lee et al., 2003; Minic et al., 2004; Minic et al., 2006). It has been hypothesized that it allows efficient modification of complex cell wall polysaccharides without requiring an extremely high number of enzymes (Minic et al, 2004; 2006).
Fig. 2. Distribution of families of glycoside hydrolases (GH) and carbohydrate esterases (CE) in the A. thaliana stem sub-proteome. Proteins are listed in Table 2. Proteins have been classified according to the CAZy nomenclature (Henrissat et al., 1998; http://www.cazy.org/). Glycoside hydrolase (GH) families from 1 to 79 are on the left whereas carbohydrate esterase (CE) families 8 and 13 are on the right. White bars correspond to families that might participate in cell wall modification and reorganization.

Many GH families described here could have other functions than cell wall modifications. Predicted extracellular GHs such as β-D-glucuronidase (GH 79), α-D-mannosidase (GH 38) and acetyl-N-hexosaminase (chitinase-like enzymes, GH 19) could be involved in post-translational modifications of glycoproteins. Recently, an A. thaliana β-D-glucuronidase (AtGUS) was shown to hydrolyze glucuronic acids from carbohydrate chains of AGPs (Eudes, personal communication). Kinetic and structural analyses of Ginkgo α-D-mannosidase acting on a pyridylamino derivative of oligo mannosides strongly suggested its involvement in the catabolism and turnover of N-linked glycoproteins (Woo et al., 2004). The pumpkin endo-β-N-acetylglucosaminidase, partially purified from cotyledons, was highly active towards high-mannose type glycans (Kimura et al., 2002).

Among other GHs, one thioglucosyl hydrolase (GH 1), 3 β-1,3-D-glucanase (GH 17) and 3 chitinase-like enzymes (GH 19 and 20) were identified. Thiogluosidoses, also known as myrosinases, play diverse roles in cruciferous plant during growth, development and defence against microorganisms and insects (Rodman, 1991). Chitinase-like enzymes are able to degrade chitin in cell walls of fungal pathogens. However, the substrates and functions of most chitinase-like enzymes are not completely known. For example, a mutation in the chitinase-like gene classified in GH 19 family (AtCTL1/At1g05850) caused a cellulose deficiency as well as aberrant patterns of lignification with incomplete cell walls in the stem pith (Zhou et al., 2002; Rogers et al., 2005).

The second largest class of identified proteins comprises proteases. Seventeen putative proteinases including aspartyl and serine type proteases were found (Table 2). Fifteen of them were predicted to be secreted (Supplementary Table S1 at JXB online). The abundance of proteases in mature stem suggests that these enzymes may be actively involved in secondary
wall formation. Proteases may play various roles in plant development and during plant pathogen interactions through maturation of CWPs and generation of active peptides. It has been shown that the extracellular subtilisin-like serine protease SDD1 (STOMATAL DENSITY AND DISTRIBUTION 1) is involved in the regulation of stomatal density and distribution in *A. thaliana* (Berger and Altmann, 2000). ALE1 (ABNORMAL LEAF EPIDERMIS1) is also predicted to encode a subtilisin-like serine protease and is assumed to produce a peptide required for proper differentiation of epidermis (Tanaka et al., 2001). CDR1 (CONSTITUTIVE DISEASE RESISTANT 1) encodes a putative aspartic protease (Xia et al., 2004). Overexpression of CDR1 causes dwarfing and resistance to virulent *Pseudomonas syringae*. It was shown that CDR1 generates a small mobile signal (3-10 kDa) sensitive to heating and to protease. The substrates of these three proteases are yet unknown. On the contrary, it was shown that the CLE (CLV3/ESR-related) basic secreted proteins are processed at their C-terminus to generate 14 amino-acid peptides that carry a biological activity (Ito et al., 2006; Kondo et al. 2006). In those cases, the proteases have not yet been identified. Finally, several plant cell wall proteomic analyses show a large discrepancy between the observed and the expected molecular masses of proteins (Boudart et al., 2005; Kwon et al., 2005; Zhu et al., 2006). Proteases could be involved in processing and/or turnover of cell wall proteins.

Several oxido-reductases such as multicopper oxidase-like (6 proteins), peroxidases (4 proteins), germin-like protein (1 protein) and a homolog to berberine bridge enzyme were identified. In contrast to previously characterized cell wall proteomes, this study allowed the identification of numerous multicopper oxidase-like proteins. They catalyse full, four-electron reduction of dioxygen (O\(_2\)) to water (H\(_2\)O) using a variety of substrates (Solomon et al., 1997). They belong to a large gene family of 19 members in *A. thaliana* (Jacob and Roe, 2005). Only two members of the family have been previously studied, SKU5 (At4g12420) and SKS6 (At1g41830). It was shown that SKU5 is involved in the control of root growth (Sedbrook et al., 2002) and that SKS6 contributes to cotyledon vascular patterning during development (Jacob and Roe, 2005). Peroxidases are involved in many physiological and developmental processes that have been reviewed recently (Passardi et al., 2004). They can be involved in both cell elongation processes and in their arrest. In the latter case, they catalyze the formation of bridges across phenolic residues of lignins and between lignins and adjacent cell wall proteins or polysaccharides.

Three extracellular acid phosphatases were identified in this work. The presence of phosphorylated proteins and phosphatases in plant cell wall has been reported in several proteomes (Chivasa et al., 2002; Kwon et al., 2005; Jamet et al., 2006). However, no extracellular kinase has yet been found (Chivasa et al., 2005). Acid phosphatases may participate in extracellular signalling events or in regulation of cell wall proteins.

Some proteins contained interacting domains, such as LRRs. The polygalacturonase-inhibiting protein (PGIP2/At5g06870) plays a role in plant defence (Di Matteo et al., 2006). Two proteins identified as fasciclin-like AGPs (AtFLA8/At2g45470, AtFLA13/At5g45130) can participate in cell-to-cell adhesion in plant (Johnson et al., 2003; Groover and Robischon, 2006). Finally, 10 proteins of unknown function were found.

**Concluding remarks**

This study demonstrates the effectiveness of the purification procedure to isolate cell wall glycoproteins. This includes novel GHs, multicopper oxidases and proteases. In contrast to these analyses, we did not identify homologs to protease or pectin methylesterase inhibitors, non-specific lipid transfer proteins, blue copper binding proteins, expansins and structural
proteins. The abundance of GHs suggests a great plasticity of polysaccharides in cell walls, even in well-differentiated tissues such as mature stems. Finally, the presence of phosphatases, proteases and GHs suggests a complex regulation of cell wall proteins involving various types of post-translational modifications such as de-phosphorylation and hydrolytic processing by proteases or glycosidases.

**Supplementary data**

Supplementary data are available at *JEB* online. Table S1: Proteins identified in *Arabidopsis thaliana* mature stems. Table S2: LC-MS/MS data allowing identification of proteins in spots 1 to 15.

**Acknowledgements**

This work was partly funded by the Génoplante program Af2001-009. We thank Bruno Letarnec for growing *A. thaliana* plants, Dr Christain Malosse for mass spectrometry analyses, Drs Jorun Johansen and Herman Höfte for improving this manuscript. **References**

**References**


18


