



**HAL**  
open science

## **Proteomics of loosely bound cell wall proteins of *Arabidopsis thaliana* cell suspension cultures: a critical analysis.**

Gisèle Borderies, Elisabeth Jamet, Claude Lafitte, Michel Rossignol, Alain Jauneau, Georges Boudart, Bernard Monsarrat, Marie-Thérèse Esquerré-Tugayé, Alain Boudet, Rafael Pont-Lezica

### ► **To cite this version:**

Gisèle Borderies, Elisabeth Jamet, Claude Lafitte, Michel Rossignol, Alain Jauneau, et al.. Proteomics of loosely bound cell wall proteins of *Arabidopsis thaliana* cell suspension cultures: a critical analysis.. *Electrophoresis*, 2003, 24 (19-20), pp.3421-32. <10.1002/elps.200305608>. <hal-00119657>

**HAL Id: hal-00119657**

**<https://hal.science/hal-00119657v1>**

Submitted on 11 Dec 2006

**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.



HAL Authorization

## **Proteomics of loosely bound cell wall proteins of *Arabidopsis thaliana* cell suspension cultures: a critical analysis.**

Gisèle Borderies\*<sup>1</sup>, Elisabeth Jamet\*<sup>1</sup>, Claude Lafitte\*<sup>1</sup>, Michel Rossignol\*<sup>1</sup>, Alain Jauneau<sup>2</sup>, Georges Boudart<sup>1</sup>, Bernard Monsarrat<sup>3</sup>, Marie-Thérèse Esquerré-Tugayé<sup>1</sup>, Alain Boudet<sup>1</sup> and Rafael Pont-Lezica<sup>1</sup>

(\*) G. Borderies, E. Jamet, C. Lafitte and M. Rossignol contributed equally to this work.

<sup>1</sup>UMR 5546 CNRS-Université Paul Sabatier, Pôle de Biotechnologie Végétale, 24, Chemin de Borde Rouge, 31326 Castanet-Tolosan, France

<sup>2</sup>IFR 40, CNRS, Pôle de Biotechnologie Végétale, 24, Chemin de Borde Rouge, 31326 Castanet-Tolosan, France

<sup>3</sup>IPBS, CNRS, 250 route de Narbonne, 31400 Toulouse, France.

**Running title:** Arabidopsis cell wall proteins

Correspondence: Rafael Pont-Lezica  
POB 17, Auzeville  
31326 – Castanet-Tolosan, F  
Phone: 33+ 5 6219 3517  
FAX: 33+ 5 3219 3502  
E-mail: lezica@smcv.ups-tlse.fr

**Keywords:** *Arabidopsis thaliana*, cell suspension culture, cell wall, mass spectrometry, proteomics.

### **Summary**

The cell wall plays a major role in plant growth and development and in cell-to-cell interactions. However, our knowledge of cell wall components, mainly the proteins, is still fragmentary. The complete sequencing of the *Arabidopsis thaliana* genome allows the use of the recently developed mass spectrometry techniques to identify the cell wall proteins (CWPs). Most proteomic approaches depend on the quality of sample preparation. Extraction of CWPs is particularly complex since the proteins may be free in the apoplast or are embedded in a polysaccharide matrix where they are retained by Van der Waals interactions, hydrogen bonds, hydrophobic or ionic interactions, or cross-linked by covalent bonds. Specific and sequential extraction procedures thus need to be developed. We report on the sequential extraction of loosely bound CWPs from living *A. thaliana* cells in culture. Different salts and chelating agents were used for releasing the proteins from the wall. Their effects on the extraction of CWPs and on the integrity of the plasma membrane were evaluated. Bioinformatic software was used to identify proteins and to predict their sub-cellular localization. The obtained data show that the plasma membrane of cells in culture was easily damaged by some steps of the extraction procedure, leading to the release of increasing amounts of intracellular proteins. Nevertheless, we identified fifty CWPs among which thirteen were new proteins for the cell wall. In addition, 76 % of these CWPs were basic proteins not resolved in 2-D gel electrophoresis. The existence of two hypothetical proteins was confirmed. The structure of three proteins could be confirmed using mass spectrometry data.

## 1. Introduction

The plant cell wall is a dynamic structure playing essential roles all along the life of the plant, including growth, development, response to environmental factors, and interactions with pathogens or symbionts [1-6]. In addition, the cell wall is a source of signals molecules involved in self and non-self recognition [7, 8]. Two main types of cell walls are found in plants: the primary cell wall assembled during cell division and growth, which is able to elongate; and the secondary cell wall produced after elongation, conferring mechanical support to the whole plant.

Polysaccharides make up the greater part (90 %) of the primary cell wall, a framework of cellulose microfibrils embedded in a matrix of hemicelluloses and pectins. Polysaccharides represent a relatively small number of different molecules, some having a very ordered structure like cellulose, and others presenting a very complex composition like non-cellulosic polymers. Cell wall proteins (CWPs) correspond to about 10 % of the cell wall mass. However, they comprise several hundreds of different molecules with various functions [9]. According to the present knowledge, CWPs may be grouped in three main functional categories: cell wall modifying proteins [10], structural proteins [11], and defense proteins synthesized in response to biotic or abiotic stresses. The possibility might arise that other, as yet unidentified functional classes, do exist in the cell wall

The achievement of the sequencing of the *Arabidopsis thaliana* genome in the year 2000 opened new questions concerning the role of genes with unpredicted functions [12]. The assignment of a gene product (protein) to a specific cell compartment through proteomics may be considered as a step forward in the framework of functional genomics. Indeed, the technical advances in mass spectrometry now allow the direct identification of proteins providing tools for systematic analysis of the proteome of a given organ, tissue, cell compartment, or for a given physiological condition [13, 14]. Proteomic approaches have recently been used for the analysis of various subcellular organelles and compartments of *A. thaliana*, notably the mitochondria, the chloroplast envelope, the plasma membrane [15-18], and for profiling the proteins of wild type and gibberellin-deficient seeds during germination [19].

Proteomic analyses of yeast and of *A. thaliana* cell walls [20-21] have been recently published. However, the main challenges for the study of the cell wall proteome in higher plants remain the difficulty to solubilize many CWPs, the diversity of linkages by which proteins remain trapped or cross-linked in the matrix, and the post-translational modifications occurring in most of them. Extraction of CWPs is still a poorly solved problem for which it is necessary to use a set of different extractive procedures depending on their attachment to the wall.

Of the about 2000 genes estimated to participate in cell wall biosynthesis, assembly and modification [9], only forty-three were identified until now by proteomic analysis [21]. In this paper, we specifically analyzed the loosely bound CWPs extracted from 7-day old living cell suspension cultures of *A. thaliana*. Actively dividing cell suspension cultures were chosen for reasons of homogeneity since only the primary cell wall is present in these cells. We present a critical analysis of the effect of different salts and chelating agents on the extractability of CWPs and on membrane integrity by proteomic and microscopic approaches. The identification of CWPs and of their corresponding genes was achieved and the contribution of proteomics to the correct annotation of the genome is discussed. Finally, a tentative functional classification of the identified CWPs is presented.

## 2. Material and methods

### 2.1. Plant material

A cell suspension culture of *Arabidopsis thaliana* var. Columbia was grown on Gamborg liquid medium [22]. From this culture, 50 mL (25 g) were routinely transferred to 250 mL fresh medium in 1 L Erlenmeyer flasks every 2 weeks and shaken at 70 rpm in an orbital shaker, under continuous fluorescent light ( $30 \mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ ) at 22°C.

### 2.2. Photonic and electronic microscopy

Samples of cell suspensions were used for microscopy observations in an inverted microscope (DMIRBB, Leica, Germany). Images were acquired with a CDD camera (Colour Coolview, Photonic Sciences, UK) and treated by Image Pro-Plus (Media Cybernetics, USA).

For electronic microscopy, samples were prepared according to Quentin *et al.* [23]. Briefly, samples were fixed in a solution of 2.5 % glutaraldehyde, 50 mM cacodylate buffer pH 7.1 for 2 h at room temperature. They were rinsed in distilled water and then post-fixed in an aqueous solution of 1 % osmium tetroxide for 1 h at room temperature. After dehydration in a graded ethanol series, they were embedded in Spurr epoxy resin. Ultra-thin sections (90 nm in thickness) were prepared using an UltraCut E microtome (Reichert-Leica, Germany) and collected on gold grids. They were submitted to the periodic acid thiocarbohydrazide silver proteinate reaction (PATAg) [24] for polysaccharide visualization and observed using a transmission electronic microscope (EM 600, Hitachi, Japan).

### 2.3. Protein extraction and separation

Cells of 7-day old *A. thaliana* suspension culture were washed with water and pelleted by centrifugation at 200 x g. They were then plasmolyzed by successive immersion in 25 % glycerol, 50 % glycerol for 10 min each, and finally washed in 50 % cold glycerol. All subsequent extractions were performed at 0°C except otherwise stated.

**Protocol A** – Prior to protein extraction, the cells were washed with 50 mM sodium acetate buffer pH 6.5, 0.15 M NaCl, 10 mM DTT, 1 mM PMSF, 1 % ethanol, and 50 % glycerol. Proteins were then extracted by successive washing of the plasmolyzed cells under gentle stirring (30 min) in the proportion of 25 ml of pelleted cells per 50 ml of solution. The extraction solutions were (a) 1 M NaCl, (b) 0.2 M CaCl<sub>2</sub>, and (c) 50 mM EDTA in a solution containing 50 mM Na acetate pH 6.5, 10 mM DTT, 1 mM PMSF, 1 % ethanol, and 50 % glycerol; (d) 0.2 M Na<sub>2</sub>BO<sub>4</sub> pH 7.5, 1 mM PMSF, 1 % ethanol, and 50 % glycerol at 20°C. Between each step, the cells were washed with the same extraction solution, then with 50 % glycerol before centrifugation at 200 x g for 5 min. The extracts (NaCl, CaCl<sub>2</sub>, EDTA, borate) were dialyzed against 20 L H<sub>2</sub>O in Spectrapor 6 cellulose ester MWCO 2 kDa cut bags. The dialyzed extracts were adjusted to 10 mM Tris-HCl pH 7.5, and retained by a bimodal ion exchange column tandem Hi-Trap (Amersham Biosciences, Sweden). SP-Sepharose and Q-Sepharose were equilibrated with 10 mM Tris-HCl pH 7.5. Each column was eluted with 2 M NaCl, 10 mM Tris-HCl pH 7.5, and the fractions containing acidic or basic proteins were concentrated by overnight precipitation in cold 20 % TCA. The precipitate was recovered by centrifugation at 10 000 x g for 25 min, and washed three times with 80 % methanol at 4°C.

**Protocol B** – In this protocol, protein extracts were sequentially recovered by washing the cells with: (a) 0.15 M NaCl, (b) 1 M NaCl, (c) 50 mM EDTA, and (d) 2 M Li Cl, in a solution

containing 50 mM Na acetate pH 6.5, 10 mM DTT, 1 mM PMSF, 1 % ethanol, and 50 % glycerol at 0°C. The extracts (0.15 M NaCl, 1 M NaCl, EDTA, Li Cl) were processed into acidic and basic fractions, as described in protocol A.

The protein content of each extract was measured by the bicinchoninic acid (Interbiotech, France) method using BSA as standard [25]. Basic proteins were resuspended in sample buffer, 62 mM Tris-HCl pH 6.8, 2 % SDS, 10 % glycerol, and 5 %  $\beta$ -mercaptoethanol, and separated by SDS-PAGE in 12.5 % polyacrylamide [26]. Acidic proteins were resuspended in 2-D sample buffer composed of 8 M urea, 2 % CHAPS, 0.3 % DTT, 2 % IPG buffer pH 4-7, and loaded directly in pH 4-7 IEF 7 cm gel strips (Amersham Biosciences) [27]. Proteins were focused using a Multiphor II system (Amersham Biosciences) for 15-20 h at 40 kVh. After focusing, the proteins of the sample were reduced for 15 min in 0.1 M Tris-HCl pH 7.5, 2 % SDS, 6 M urea, 30 % glycerol, and 50 mM DTT, and then alkylated for another 15 min in the same buffer containing 250 mM iodoacetamide (Sigma Chemical, USA). The gel strip was loaded on top of a 12.5 % polyacrylamide gel 80 x 60 x 1.5 mm for SDS PAGE in a Bio-Rad system (France). After electrophoresis, 1-D and 2-D gels were fixed in ethanol/acetic acid/water (45/5/50), stained with 0.1 % CBB in ethanol/acetic acid/water (25/8/67), and numerized with an Image scanner (Amersham Biosciences).

#### **2.4. In gel digestion and protein identification by mass spectrometry**

The stained spots were excised with a 2 mm inner diameter Pasteur pipette. Each gel piece was washed twice with 75  $\mu$ L of 25 mM ammonium bicarbonate, 50 % ACN, for 15 min. After drying *in vacuo*, the gel pieces were rehydrated with 10  $\mu$ L of 10  $\mu$ g/mL modified trypsin (Promega, France sequencing grade) in 25 mM ammonium bicarbonate and digested overnight at 37°C. Digestion was stopped by acidification with 40  $\mu$ L of 10 % formic acid, 62.5 % ACN. Gel pieces were incubated at 37°C for 10 min, and then sonicated for 5 min. Additional 5 min sonication was made after addition of 25  $\mu$ L ACN. The liquid extract containing the peptide mixture was evaporated to dryness and resuspended in 4  $\mu$ L of 0.1 % TFA, 50 % ACN.

An aliquot of 0.75  $\mu$ L from this peptide mixture was spotted on the sample plate of the mass spectrometer with 1  $\mu$ L of the matrix solution (6 g L<sup>-1</sup> of  $\alpha$ -cyano-4-hydroxycinnamic acid in 50 % ACN / 0.1 % trifluoroacetic acid). Analysis was performed with a MALDI-TOF mass spectrometer (Voyager-DE<sup>TM</sup> STR, Perseptive Biosystems, USA) which was operated in positive reflector mode at the following parameters: accelerating voltage 20 kV, grid voltage 68 %, and extraction delay time 200 ns. Acquisition mass<sup>+</sup> was between 750 and 3000 Da. Internal mass calibration was performed using trypsin autolysis peaks (monoisotopic MH<sup>+</sup> 842.51 and 2211.10).

Whenever it was necessary to improve the ionization efficiency of MALDI-TOF MS and to perform MALDI-TOF post-source decay (PSD) analysis, the remaining peptide mixture was desalted using Zip Tips C<sub>18</sub> (Millipore, France). It was directly eluted on the sample plate in 70 % acetonitrile and 1  $\mu$ L of the above matrix solution was added. MALDI-TOF MS acquisition method was the same as above. To improve identification with MALDI-TOF PSD, fragmentation was induced using high laser energy; accelerating voltage 20 kV, grid voltage 75 % and extraction delay time 100 ns. The mass selection of the precursor ion was achieved with a timed ion selector of +/- 10 Da. Fragment mass spectra were acquired in 12 segments with a decrement ratio of 0.75. Typically, 500 shoots were averaged for each production segment. The Voyager software assembled the final PSD spectra.

The peptide mass fingerprinting data was analyzed by MS-FIT and the PSD results by MS-TAG (Protein Prospector, <http://prospector.ucsf.edu>). The NCBI nr database was used to identify the proteins. The retained parameters were mass tolerance 20 ppm and one missed cleavage. Identification was considered positive when the difference in MOWSE score between first and second ranked proteins was more than two orders of magnitude.

## **2.5. Bioinformatic analysis of protein sequences**

DNA sequences related to each protein were collected using data available in the Unigene database (<http://www3.ncbi.nlm.nih.gov/Entrez/>). All predicted proteins were compared using CLUSTALW (<http://npsa-pbil.ibcp.fr>) [28]. The most probable protein was chosen according to two criteria: (i) the frequency of amino acid sequence, (ii) the relevance to experimental data. Eventually, the location of the translation initiation codon was checked using ATGpr (<http://www.hri.co.jp/atgpr/>) and NetStart (<http://www.cbs.dtu.dk/services/NetStart/>) [29]. Sub-cellular localization as well as length of signal peptides were determined using PSORT (<http://psort.nibb.ac.jp/>) and TargetP (<http://www.cbs.dtu.dk/services/TargetP/>) [30]. Prediction of transmembrane domains was done with TMHMM (<http://www.cbs.dtu.dk/services/TMHMM/>) and TopPred (<http://bioweb.pasteur.fr/seqanal/interfaces/toppred.html>). Molecular mass and pI values were calculated using the aBi program ([http://www.up.univ-mrs.fr/~wabim/d\\_abim/compo-p.html](http://www.up.univ-mrs.fr/~wabim/d_abim/compo-p.html)). Homologies to other proteins were searched using BLAST programs (<http://www.ncbi.nlm.nih.gov/BLAST/>) [31]. Identification of protein families and domains was performed using PROSITE (<http://www.expasy.org/prosite/>) [32], and PFSCAN (<http://hits.isb-sib.ch/cgi-bin/PFSCAN>). The Arabidopsis Genome Initiative (AGI) nomenclature was used (<http://mips.gsf.de/proj/thal/>; <http://www.tigr.org/tdb/tgi/agi/>).

## **3. Results and discussion**

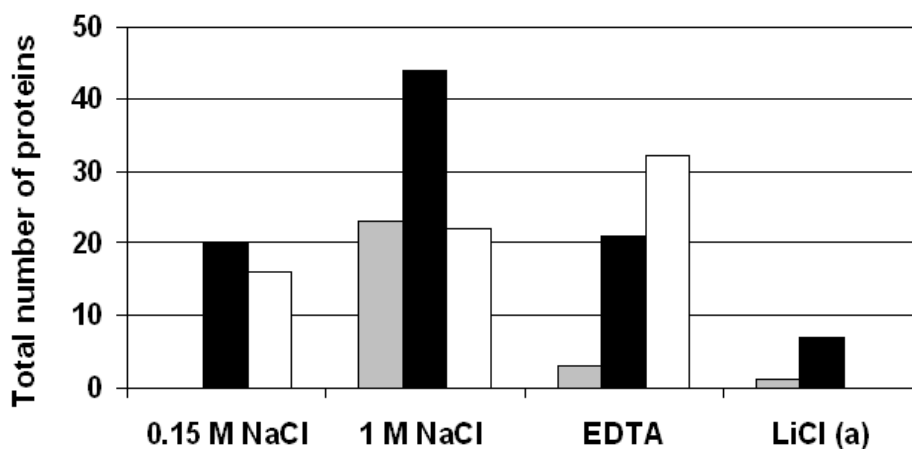
A protocol to extract loosely bound CWPs from cell suspension cultures was initially proposed by Robertson et al. [33] in 1997. At that time, the complete sequence of the *A. thaliana* genome was not available and the authors used N-terminal sequencing for identification of CWPs from various plants. Recently, Chivasa et al. [21] took advantage of the new advances in mass spectrometry and the availability of the complete *A. thaliana* genome to analyze proteins sequentially extracted from cell wall preparations. The method they used included extensive washes of the cell wall residue to eliminate contamination by intracellular proteins. The main drawback of this strategy is that many CWPs may be lost during washings. In the present work, we focused on loosely bound CWPs extracted from the walls of living cells. To limit potential contamination by intracellular proteins and to recover a maximum number of CWPs, we adapted to living cells the sequential extraction techniques described previously [21, 33]. To achieve this goal, all extraction steps were done on cells plasmolyzed in 50 % glycerol. The challenge was that the various salts or chaotropic agents mentioned thereafter could act freely on the walls while maintaining the cellular content confined within the protoplast. NaCl is usually used for extraction of proteins retained by ionic interactions in the cell wall. CaCl<sub>2</sub> has been reported as an efficient salt for the extraction of CWPs from purified cell walls [34]. EDTA, a well-known calcium chelant, might solubilize proteins associated with pectins [35]. Borate was used to disrupt interactions between the side chains of glycoproteins and wall polysaccharides [33]. LiCl was also retained according to its potential to extract hydroxyproline-rich glycoproteins from intact cells in *Chlamydomonas reinhardtii* [36].

The extracts were concentrated, separated, and stained as described in Material and Methods. The peptide mass fingerprints and PSD obtained by MALDI-TOF allowed the identification of most of the proteins stained with CBB. A few criteria were retained to discriminate between intracellular and extracellular proteins through bioinformatic analyses, namely: i) occurrence of a cleavable N-terminal signal sequence but no more than one transmembrane domain, ii) absence of the KDEL or HDEL endoplasmic reticulum retention motifs.

### 3.1. Evaluation of the protocols retained for CWP extraction

The first extractions were performed according to protocol A. Analysis of the NaCl extract showed 60 % of CWPs. However, the CaCl<sub>2</sub> one contained 86 % of intracellular proteins (data not shown). Membrane permeability seemed to be strongly altered by CaCl<sub>2</sub>, a salt also used to introduce DNA in bacteria [37]. This first set of results and particularly the high number of intracellular contaminants in the CaCl<sub>2</sub> extract clearly showed that the plasma membrane might have lost its integrity and was leaking.

To overcome these drawbacks, a new set of extractions was performed according to protocol B (Fig. 1). The first two extracts 0.15 M and 1 M NaCl released 107 proteins obtained from the acidic and basic fractions. Among these proteins, 65 % were CWPs. The EDTA extract was not very satisfactory since intracellular contaminants increased by 50 % and only three new CWPs were obtained. This result was surprising, since EDTA is used to extract pectins, and it was expected to release proteins associated with this polysaccharide. Finally, the LiCl extract contained just one new CWP.



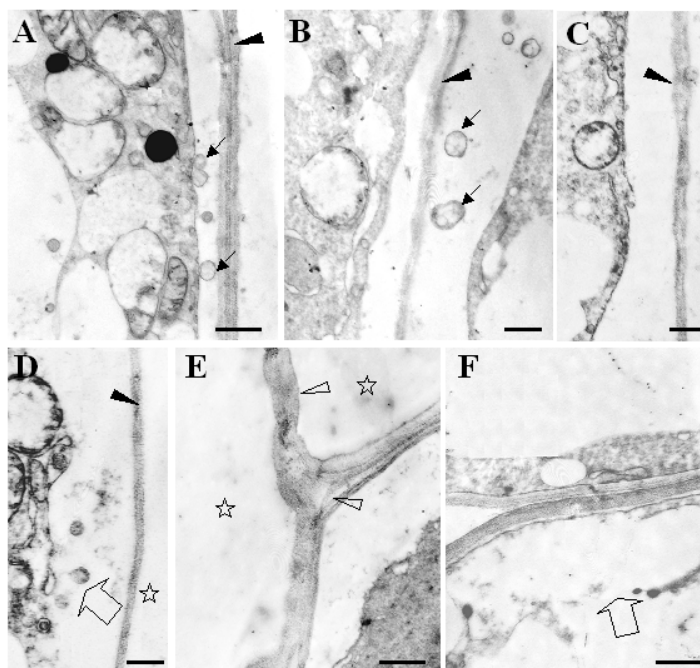
**Figure 1.** Comparison of the total number of extracellular (black bars) and intracellular (white bars) proteins obtained by sequential extraction of living cells with protocol B. The gray bars represent new CWPs that were not extracted during the previous step. Bars represent the total number of different proteins obtained from the acidic and basic fractions, with the exception of extracts marked (a), where only the basic fractions were analyzed. Serial extractions (from left to right) of 7-day old *A. thaliana* cell suspension cultures were done on plasmolyzed cells, as described in Materials and Methods.

It was noticed that the number of released intracellular proteins increased along extractions (Fig. 1). Their molecular mass also increased, ranging from 11 to 23 kDa in the 0.15 M NaCl extract, from 13 to 32 kDa in the 1M NaCl extract, and from 11 to 62 kDa in the EDTA extract. Moreover, intracellular proteins were cytosolic in the two first extracts, whereas several peroxysomal proteins were identified in the third one. All these data pointed to deleterious effects of the sequential extractions and suggested that the plasma membrane was

first damaged, followed by other organelles. Living cells do not tolerate more than two extractions without severe alterations of their structures. Proteins released into the culture medium were analyzed in the same way, and thirteen CWPs were identified in that fraction, without contamination by intracellular proteins. Ten of these proteins were also found in the NaCl extracts (Table 1).

### 3.2. Microscopic survey

To visualize the integrity of cellular compartments, cell suspensions were treated as in protocol B before being prepared for electron microscopy analysis. The control consisted of unextracted, glycerol-plasmolyzed cells. As expected, the plasma membrane was detached from the cell wall (Fig. 2A), and the various organelles were well preserved. Some small vesicles were observed in the space between the plasma membrane and the wall, in agreement with a current artifact due to aldehyde fixation [38]. The PATAg reagent allowed to clearly visualize the middle lamella in controls as well as in cells treated with 0.15 M NaCl (Fig. 2B). After additional treatments, the results were quite different. Samples treated with 1 M NaCl showed cells with a conserved structure (Fig. 2C), while others exhibited dramatic damages. Integrity of their plasma membrane was altered, and strong vesicularization was observed (Fig. 2D). After treatment with EDTA, an important number of cells had completely lost their cellular content and appeared empty (Fig. 2E). The cell walls were swelled and parts of them were lightly cleared out because of solubilization of some pectic components. After treatment with CaCl<sub>2</sub> as in protocol A, the electron microscopy micrograph illustrates the rupture of the membrane in some samples, as well as the presence of empty cells in others (Fig. 2F). As seen by electron microscopy, integrity of the cell structure was preserved only after mild treatment of the cells, i.e. 50 % glycerol alone or 0.15 M NaCl. Other treatments led to dramatic cell structure damages, which explained the extent of contaminant proteins.

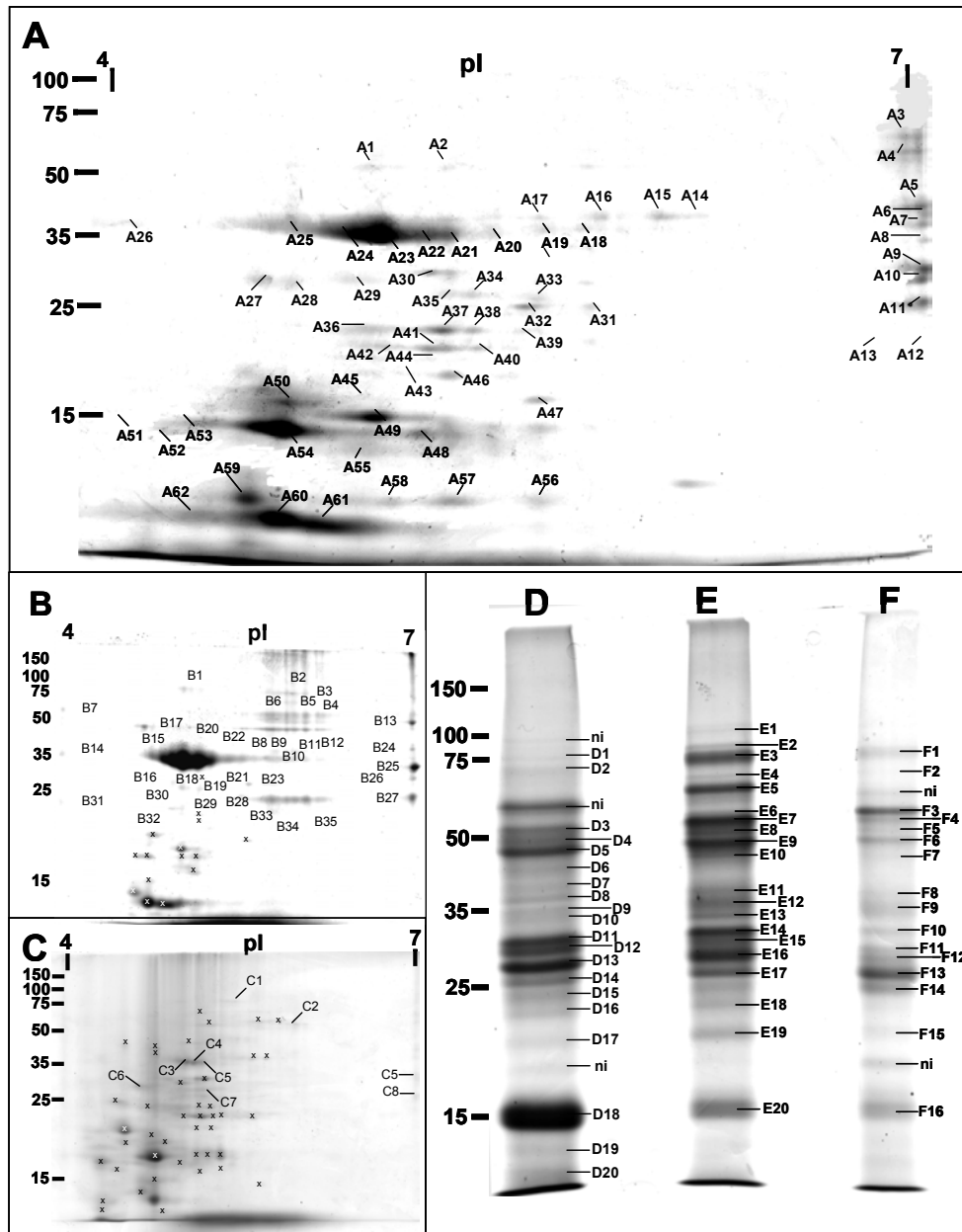


**Figure 2.** Transmission electron micrographs of control plasmolyzed cells (A), and of cells treated with 0.15 M NaCl (B), 1 M NaCl (C and D), EDTA (E) as in protocol B or CaCl<sub>2</sub> (F) as in protocol A. Arrows indicate vesicles between the cell wall (arrowheads) and the plasma membrane. Empty cells are indicated by stars. Large open arrows indicate cell membrane damages (D and F). Open arrowheads indicate the cleared out area within a tri-cellular junction and the swelling of the wall in EDTA treated samples (E).

### 3.3. Identification of proteins

Figure 3 shows the patterns of proteins obtained after electrophoresis and CBB staining of 0.15 M NaCl, 1 M NaCl and EDTA respectively. The acidic fractions were separated by 2-D

electrophoresis (Fig. 3A, B, C) whereas the basic fractions were best resolved by 1-D SDS PAGE (Fig. 3D, E, F). Proteins released into the culture medium were separated in the same way (data not shown). All detectable spots were analyzed by MALDI-TOF. We identified ninety-five proteins from all the fractions from cell suspensions. Fifty proteins were CWPs and among these, 76.5 % were basic proteins. Most contaminants were acidic intracellular proteins (95 %). Only two basic cyclophilins, coded by At4g34870 and At2g16600 were found in the basic fractions (Table 2).



**Figure 3.** Analysis of acidic and basic proteins extracted according to protocol B. 2-D electrophoresis of acidic proteins from 1 M NaCl (A), 0.15 M NaCl (B), and EDTA (C) extracts. The numbers indicate proteins spots identified by MALDI-TOF mass spectrometry. In A, all proteins (CWPs and intracellular proteins) have been given a number. In gels B and C, only CWPs have a number, (x) indicate identified intracellular proteins. 1-D SDS PAGE analysis of basic proteins from 0.15 M NaCl (D), 1 M NaCl (E), and EDTA (F) extracts. Numbers indicate the bands that have been analyzed by MALDI-TOF mass spectrometry. Numbers on the top of 2-D gels (4, 7) indicate range of pH for IEF. Numbers on the left of 2-D and 1-D gels indicate sizes of molecular markers in kDa. n.i. stands for not identified.

The identified proteins were classified as CWPs (Table 1) and intracellular proteins (Table 2). CWPs were tentatively assigned to functional classes as discussed below. In Fig 3A all detectable spots were numbered, but only those corresponding to CWPs were numbered in Fig. 3B and C. Spots indicated by crosses correspond to intracellular proteins.

The identification process deserves a few additional comments:

- i) In spite of the low resolution of 1-D gels, it was possible to identify several proteins within the same band. For example, in band E6 (Fig. 3E, and Table 1), a pectin methylesterase (At1g41830), a homologue to L-ascorbate oxidase (At5g21105), and a homologue to beta D-exoglucanase (At5g20950) were identified. Similarly, band E13 was found to contain the products of genes At1g78830, At5g06870, and At1g74000 (Fig. 3E, and Table 1). Interestingly, members of the same gene family were identified in band E15, i.e. the three peroxidases AtP3 (At64100), AtP15 (At64120) and AtP31 (At36430) (Fig 3E, and Table 1).
- ii) In many cases, several isoforms of the same protein were found, such as for peroxidase AtP13 (At3g32980) identified in spots A14-A17, resulting in the ranges of molecular masses and pI indicated in Tables 1 and 2. Basic proteins could be found both on the right side of a 2-D gel and on a 1-D gel. As an example, spots A12 and A13 (Fig. 3A) contain the same protein as band E19 (Fig.3E).
- iii) It was also noticed that proteins not completely extracted by one solvent were present in several fractions, as it is the case for the protein showing homology to lectins (At1g78830). This protein was also found in the culture medium (Table 1).

### **3.4. Bioinformatic analysis of proteomic data**

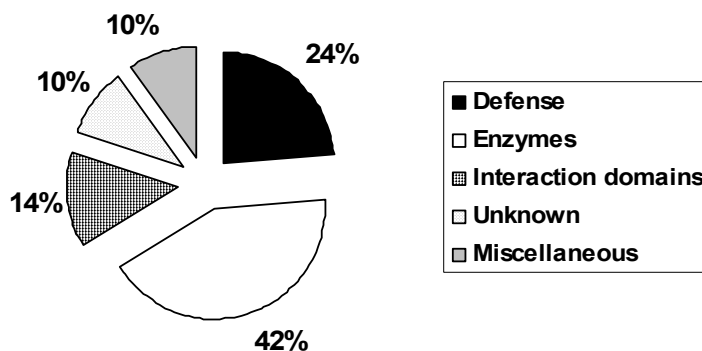
We introduced a careful bioinformatic analysis of the protein sequences. This proved to be important, first to identify different members of large multigene families; second, to predict subcellular localization according to the criteria already stated; and third, to improve the annotation of the *A. thaliana* genome, as illustrated hereafter.

- i) Four pectin methylesterases and ten peroxidases were clearly identified (Table 1). In the case of peroxidases, it is generally difficult to link the results of transcription with those of the protein content for most of the members of the peroxidase class III family. Thus, protein extracts from *A. thaliana* plants or cell suspension cultures showed few peroxidases after IEF [39, 40], possibly due to the fact that most of them are basic proteins, and IEF does not resolve those proteins satisfactorily. Only the acidic peroxidases AtP36, AtP16, and AtP49 were well resolved after 2-D electrophoresis in our work.
- ii) Bioinformatic-based proteomics can also contribute to determine the final destination of a protein. This is the case for some peroxidases predicted to be targeted to the vacuole via a C-terminal extension, the so-called CX propeptide [39]. In barley grains, peroxidase BP1 cDNA contains an additional C-terminal peptide of 22 residues that is absent from the mature protein [41, 42]. A close homologue, BP 2, was immunolocalized to the vacuole [43]. We have identified a peroxidase with such a predicted C-terminus, AtP16. After verification of the peptide fingerprint, we found that the complete C-terminal peptide was present in the secreted protein which means that no cleavage was observed, contrary to the seed-specific barley BP 1.
- iii) The contributions to the annotation of the *A. thaliana* genome are multiple since around 10 % of the protein sequences we analyzed had errors at different levels such as wrong intron prediction, or errors in predicted function. We have confirmed the existence of two “hypothetical proteins”. At present, neither mRNAs, nor ESTs were found for the corresponding genes in databases. These proteins were encoded by At2g18140 and At4g36350. Both were CWPs (see Table 1), one was a peroxidase and the other had no clearly defined function. The structure of three genes encoding intracellular proteins could be confirmed. In two cases (At5g02240 and At5g24400), there was a discrepancy between the

predictions of protein sequences from BAC sequences (<http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi>) and those found in the RefSeq database (<http://www.ncbi.nlm.nih.gov/RefSeq/>). Our results were in agreement with the RefSeq data, which also take into account biological data such as mRNA and EST sequences when they are available [44]. Finally, proteins identified by MALDI-TOF and encoded by At5g02240 and At5g24400 fitted with proteins described under accession numbers NP\_568098 and NP\_568445 respectively. In the case of At3g17240, two different proteins were predicted and could be confirmed by translation of distinct mRNAs probably resulting from alternative splicing. The protein identified in this study was NP\_683572.

### 3.5. Functional classification of CWPs

CWPs identified in this work were classified in functional categories. This classification is only tentative, since the biological role of many of the proteins identified has not been established experimentally. Proteins are ordered according to this classification in Table 1 and Fig. 4 presents the proportion of each functional class. The most represented class is that of proteins acting on other cell wall components, which are all enzymes with the exception of one expansin. Well-known cell wall enzymes like pectin methylesterases, peroxidases, and glycosyl transferases are represented by several members of the family.



**Figure 4** Functional classification of CWPs. The extracellular proteins were classified in the following functional groups: Defense: defense proteins; Enzymes: cell wall modifying proteins; Interaction domains: proteins with domains interacting with polysaccharides or proteins; Unknown: unknown function; Miscellaneous: proteins with homologies to known proteins.

The second more abundant group contains defense proteins as expected from the fact that cell cultures are a non-physiological situation. In this group, we ranged classical defense proteins like chitinases and proteases as well as three proteins with homology to the S-reticulins oxidoreductases whose homologous enzymes from *Papaver* and *Berberis* are implicated in the synthesis of alkaloids, and are localized in the vacuole [45, 46]. Through bioinformatic analyses, the predictions were consistent with a vacuolar localization for the *Papaver* and *Berberis* proteins whereas the *Arabidopsis* proteins were predicted to be extracellular. In addition, we identified two out of the three members of the *A. thaliana* S-reticulins oxidoreductase family in the culture medium, which was not contaminated with intracellular proteins.

The group of proteins containing a domain for interaction with polysaccharides or proteins collects proteins with unknown biochemical or biological function, but with a prediction derived from their sequence. Two lectins belonging to different families were found: At3g15356 encodes a member of the legume type, while At1g78830 belongs to the *Hevea* family of lectins. Lectins are extremely abundant in the storage vacuoles from seeds, but they are extracellular in most vegetative tissues [47]. Four proteins (At1g33590, At5g23400, At5g06860, and At5g06870) present several LRR motifs that are usually involved in protein-

protein interactions. However, the functionality of these domains has not been established experimentally and it remains a theoretical class.

The miscellaneous group presents some very interesting proteins like EXORDIUM or SKU5, which are involved in development. EXO is targeted to the secretory pathway, but it has not yet been shown whether it is localized in the cell wall. EXO functions as a negative regulatory system of cell division control for meristem maintenance [48]. This seems to be the first evidence of EXO as a CWP. On the other hand, SKU5 is an extracellular protein that may be free in the cell wall or lipid-anchored by a GPI motif. This protein is a member of a multigene family and it seems to be involved in directional growth and participate in cell wall expansion [49]. These two examples show the importance of CWPs in plant development. Finally, we ranged in a group corresponding to unknown functions all the proteins with no evident homology to known proteins from other organisms.

#### **4. Concluding remarks**

In order to obtain loosely bound CWPs from cell suspension cultures, several procedures may be used. Culture medium may be analyzed to get the secreted proteins. Alternatively, a mild extraction of CWPs with salts or chelators may be performed on living cells. However, this study shows that this procedure cannot be repeated more than twice because the plasma membrane is easily damaged.  $\text{CaCl}_2$  as well as EDTA appear to be particularly deleterious to the plasma membrane. A first wash with 0.15 M NaCl, followed by a second one with 1M NaCl, seems to be the suitable method. In this study, all but three proteins from the culture medium were found in the NaCl extracts. Since the more drastic extractions required to release proteins more strongly embedded in the extracellular matrix cannot be done with living cells, the logical procedure is to prepare purified cell walls as proposed by Chivasa et al. [21]. The proteins extracted by their procedure essentially using  $\text{CaCl}_2$  and urea were in their great majority different from those identified in this work. Indeed, the total number of CWPs identified by both approaches is ninety-six, whereas only eleven proteins were found to be common. Thus, the two approaches proved complementary. Finally, 26 % of the CWPs found in this study were novel proteins for this compartment. This includes all the proteins with unknown functions and the miscellaneous class. However, among these loosely bound proteins, certain families are under-represented or absent. For example, only one expansin was found and arabinogalactan proteins [11] are missing since the protocol used for protein preparation is not adapted to such heavily glycosylated proteins. At present, attempts to extract such proteins as well as strongly bound structural proteins are in progress.

This study improved our knowledge of gene expression in cell cultures. From information available in databases, only eleven out of the fifty CWP genes and eight out of the forty-five intracellular protein genes identified were known to be transcribed in actively-dividing cell suspension cultures.

The precise identification of particular members of large multigene families, like peroxidases, showed the importance of proteomic approaches as well as the contribution of bioinformatic analysis to the correct annotation of the *A. thaliana* genome.

#### **Acknowledgements**

We are indebted to Génoplante (Contract N° NO 2001027), the Conseil National de la Recherche Scientifique and the Université Paul Sabatier, France, for supporting this work.

The authors are grateful to Prof. Pierre Rougé (UMR 5546 CNRS-UPS, Toulouse) for the characterization of the cell-wall lectins identified in this study.

## References

- [1] Roberts, K., *Curr. Opin. Cell Biol.*, 1989, 1:1020-1027.
- [2] Roberts, K., *Curr. Opin. Cell Biol.*, 1994, 6:688-694.
- [3] Carpita, N., Gibeaut, D., *Plant J.*, 1993, 3: 1-30.
- [4] Bolwell, G., *Int. Rev. Cytol.*, 1993, 146:261-324.
- [5] Bolwell, G., in: Greppin, H. (Ed.), *Travelling shot on plant development*, University of Geneva Press: Geneva.1997, pp. 255-273.
- [6] Sakurai, N., *J. Plant Res.*, 1998, 111:133-148.
- [7] Brownlee, C., *Curr. Opin. Cell Biol.*, 2002, 5:396-401.
- [8] Esquerré-Tugayé, M. T., Boudart, G., Dumas, B., *Plant Physiol. Biochem.*, 2000, 38: 157-163.
- [9] Carpita, N., Tierney, M., Campbell, M., *Plant Mol. Biol.*, 2001, 47:1-5.
- [10] Darley, C., Forrester, A., McQueen-Mason, S., *Plant Mol. Biol.*, 2001, 47:179-195.
- [11] Cassab, G. I., *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 1998, 49: 281-309.
- [12] Arabidopsis Genome Initiative, *Nature*, 2000, 408: 796-815.
- [13] Yates, J. r., *J. Mass Spectrom.*, 1998, 33: 1-9.
- [14] Kuster, B., Mortensen, P., Andersen, J., Mann, M., *Proteomics*, 2001, 1: 641-650.
- [15] Millar, A., Sweetlove, L., Giege, P., Leaver, C., *Plant Physiol.*, 2001, 127: 1711-1727.
- [16] Prime, T., Sherrier, D., Mahon, P., LC, P., Dupree, P., *Electrophoresis*, 2000, 21: 3488-3499.
- [17] Ferro, M., Salvi, D., Rivière-Rolland, H., Vermat, T., Seigneurin-Berny, D., Grunwald, D., Garin, J., Joyard, J., Rolland, N., *Proc. Natl. Acad. Sci. USA*, 2002, 99: 11487-11492.
- [18] Santoni, V., Rouquié, D., Doumas, P., Mansion, M., Boutry, M., Degand, H., Dupree, P., Packman, L., Sherrier, J., Prime, T., Bauw, G., Posada, E., Rouzé, P., Dehais, P., Sahnoun, I., Barlier, I., Rossignol, M., *Plant J.*, 1998, 16: 633-641.
- [19] Gallardo, K., Job, C., Groot, S., Puype, M., Demol, H., Vandekerckhove, J., Job, D., *Plant Physiol.*, 2002, 129: 823-837.
- [20] Pardo, M., Ward, M., Bains, S., Molina, M., Blackstock, W., Gil, C., Nombella, C., *Electrophoresis*, 2000, 21: 3396-3410.
- [21] Chivasa, S., Ndimba, B., Simon, W., Robertson, D., Yu, X.-L., Knox, J., Bolwell, P., Slabas, A., *Electrophoresis*, 2002, 23: 1754-1765.
- [22] Axelos, M., Curie, C., Mazzolini, L., Bardet, C., Lescure, B., *Plant Physiol. Biochem.*, 1992, 30: 123-128.
- [23] Quentin, M., Jauneau, A., Morvan, O., Mareck, A., Gaffé, J., Morvan, C., *Plant Physiol. Biochem.*, 1997, 35: 475-482.
- [24] Thiéry, J., *J. Microsc.*, 1967, 6: 987-1018.
- [25] Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., Klenk, D. C., *Anal. Biochem.*, 1985, 150: 76-85.
- [26] Laemmli, U. K., *Nature*, 1970, 227: 680-685.
- [27] Rabilloud, T., *Proteome Research: Two dimensional electroporesis and identification methods*. Springer Berlin, Heidelberg, New York, 2000, pp. 1-242.
- [28] Thompson, J., Higgins, D., Gibson, T., *Nucleic Acids Res.*, 1994, 22: 4673-4680.
- [29] Pedersen, A., Nielsen, H., *Proc. Int. Conf. Intell. Sys. Mol. Biol.*, 1997, 5: 226-233.

- [30] Emanuelsson, O., Nielsen, H., Brunak, S., von Heijne, G., *J. Mol. Biol.*, 2000, 300: 1005-1016.
- [31] Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D., *J. Mol. Biol.*, 1990, 215: 403-410.
- [32] Falquet, L., Pagni, M., Bucher, P., Hulo, N., Sigrist, C., Hofmann, K., Bairoch, A., *Nucleic Acids Res.*, 2002, 30: 235-238.
- [33] Robertson, D., Mitchell, G., Gilroy, J., Gerrish, C., Bolwell, G., Slabas, A., *J. Biol. Chem.*, 1997, 272: 15841-15848.
- [34] Smith, J., Muldoon, E., Lamport, D., *Phytochemistry*, 1984, 23: 1233-1239.
- [35] Selvendran, R. R., O'Neil, M. A., *Methods Biochem. Anal.*, 1987, 32: 25-153.
- [36] Voigt, J., *Planta*, 1985, 164: 379-389.
- [37] Oishi, M., Cosloy, S., *Nature*, 1974, 248: 112-116
- [38] McGee-Russell, S. M., De Bruijn, W. C., in: S.M. McGee-Russell, S. M., K.F.A. Ross, K. F. A. (Eds.), *Cell structure and its interpretation*, Edward Arnold Ltd., London. 1968, pp. 115-134.
- [39] Welinder, K. G., Justesen, A. F., Kjaersgard, I. V. H., Jensen, R. B., Rasmussen, S. K., Jespersen, H. M., Duroux, L., *Eur. J. Biochem.*, 2002, 269: 6063-6081.
- [40] Østergaard, L., Pedersen, A., Jespersen, H. M., Brunak, S., Welinder, K. G., *FEBS Lett.*, 1998, 433: 98-102.
- [41] Welinder, K. G., *Eur. J. Biochem.*, 1979, 96: 483-502.
- [42] Johansson, A., Rasmussen, S. K., Harthill, J. E., Welinder, K. G., *Plant Mol. Biol.*, 1992, 18: 1151-1161.
- [43] Theilade, B., Rasmussen, S. K., Rosenkrands, I., Frokjaer, H., Hejgaard, J., Theilade, J., Pihakaski-Maunsbach, K., Maunsbach, A. B., in: Welinder K. G., Rasmussen, S., Penel, C., Greppin, H.(Eds.), *Plant Peroxidases, Biochemistry and Physiology*, University of Geneva: Geneva, 1993, pp. 321-324.
- [44] Pruitt, K., Maglott, D. R., *Nucleic Acids Res.*, 2001, 29: 137-140.
- [45] Bird, D., A., Facchini, P. J., *Planta*, 2001, 213: 888-897.
- [46] Bock, A., Wanner, G., Zenk, M. H., *Planta*, 2002, 216: 57-63.
- [47] Vasta, G., Pont-Lezica, R., in Mecham, R. P., Adair, W. S. (Eds), *Organization and Assembly of Plant and Animal Extracellular Matrix*, Academic Press Inc: Orlando 1990, pp. 173-245.
- [48] Farrar, K., Evans, M., Topping, J. F., Souter, M. A., Nielsen, J. E., Lindsey, K., *Plant J.*, 2003, 33: 61-73.
- [49] Sendbrook, J. C., Carroll, K. L., Hung, K. F., Masson, P. H., Somerville, C. R., *Plant Cell*, 2002, 14: 1635-1648.

**Table 1.** Cell wall proteins identified in the sequential extracts of *A. thaliana* cell cultures.

Protein identification	Spot <sup>(a)</sup>	Gene Acc. N <sup>o</sup> <sup>(b)</sup>	Signal peptide	Experimental MM (kDa)	pI	Theoretical MM (kDa)	pI
<b>Defense proteins</b>							
Homology to berberine-bridge enzyme (S)-reticulín:oxygén oxidoreductase <sup>(c)</sup>	B5-7, D1, D4-5, E7-8, F5	At4g20830	1-30	83.8-49.6	4.0-n.d.	56.8	9.9
Homology to berberine-bridge enzyme (S)-reticulín:oxygén oxidoreductase	E8	At5g44380	1-32	50.7	n.d.	57.2	10
Homology to berberine-bridge enzyme (S)-reticulín:oxygén oxidoreductase <sup>(c)</sup>	c.m.	At2g34790	1-27	62.0-55.0	5.2-5.4	56.8	5.7
Homology to aspartyl protease	A4, C2, E9	At3g52500	1-18	56.0	n.d.	49.0	9.5
Homology to aspartyl protease <sup>(c)</sup>	D7, E10	At3g54400	1-19	44.1	n.d.	43.3	10.3
Cysteine proteinase RD21A <sup>(c)</sup>	A27-28, C6	At1g47128	1-21	28.3	4.6	48.8	5.0.
Subtilisine-like serine-protease	E4, F2	At2g05920	1-26	69.9	n.d.	77.3	9.4
Homology to L-ascorbate oxidase	E3, E6, F1	At5g21105	1-19	80.1-56.5	n.d.	61.9	9.3
Homology to germin subfamily 2 member 1	A12-13, E19, F15	At1g09560	1-23	23.4	n.d.	20.5	9.7
Chitinase class I	C9, E16-17	At4g01700	1-27	29.1	n.d.	28.4	9.9
Chitinase class I <sup>(c)</sup>	c.m.	At3g54420	1-28	35.0-32.0	4.2	26.5	4.6
Ribonuclease RNS1 <sup>(c)</sup>	B32	At2g02990	1-22	22.4	4.9	23.0	4.7
<b>Cell wall modifying proteins</b>							
Homology to pectinacetylésterase	A5, B24	At4g19410	1-23	36.5	n.d.	39.7	10.2
Pectin methylesterase	D3	At1g76160	1-23	58.7	n.d.	57.8	9.1
Pectin methylesterase	E6	At1g41830	1-24	56.5	n.d.	58.1	9.7
Pectin methylesterase	F3-4, F6, F14	At5g66920	1-30	55.2-23.6	n.d.	57.4	9.7
Pectin methylesterase	F11	At1g11580	1-38	27.9	n.d.	57.6	9.4
Peroxidase AtP7	A5-6, E12	At3g21770	1-27	36.5	n.d.	32.9	10.5
Peroxidase AtP13	A9-10, C9	At5g17820	1-22	30.0	n.d.	31.6	11.4
Peroxidase AtP36 <sup>(c)</sup>	A26, B20-21, C3-5	At2g18150	1-22	34.9-32.5	5.4-4.1	34.6	5.0
Peroxidase AtP3 <sup>(c)</sup>	A8, B25, D11-12, D16, E14-15, E18, F10	At5g64100	1-23	30.8-22.7	n.d.	33.2	10.7
Peroxidase AtP15	D11-12, D15, D17, E15, E19	At5g64120	1-23	31.1-20.2	n.d.	32.4	10.1

Table 1 (continued)

Peroxidase AtP31	D12, E15	At4g36430	1-22	29.5	n.d.	33.7	9.7
Peroxidase AtP16	A14-17	At3g32980	1-29	35	5.6-6.1	35.7	6
Peroxidase <sup>(c)</sup>	B26, D13-14, E16, F12	At5g05340	1-20	27.0	n.d.	32.1	9.7
Peroxidase AtP49 <sup>(c)</sup>	c.m.	At2g18140	1-16	35.0	5.2	35.4	5.8
Peroxidase AtCb	D8-9, E11	At3g49120	1-30	40.0-37.7	n.d.	35.7	9.4
Homology to $\beta$ -glucosidase	D4, E7, F6	At2g44450	1-22	54.4-46.8	n.d.	54.5	8.7
$\beta$ -galactosidase	E2	At5g63810	1-30	88.9	n.d.	79.8	9.2
Homology to $\beta$ -D-glucan exohydrolase	E5-6	At5g20950	1-19	65.7-56.5	n.d.	66.0	9.8
Homology to endo-1,4- $\beta$ -glucanase	E10	At1g71380	1-21	44.7	n.d.	51.0	9.4
Expansin-like 1 precursor At(EXPL1)	E17, F13-14	At3g45970	1-20	26.1-23.6	n.d.	26.2	9.5
Xyloglucan endo-1,4- $\beta$ -D-glucanase 6	B30	At4g25810	1-24	25.6	5.0	29.6	4.8
<b>Proteins containing domains for interaction with other proteins or polysaccharides</b>							
Homology to lectin <sup>(c)</sup>	A11, A18-26, B1-4, B14-23, B27, B31, D9-10, E13, F7-9, F14	At1g78830	1-22	77.2-23.5	4.0-n.d.	48.1	9.4
Homology to lectin	E16	At3g15356	1-19	28.1	n.d.	27.6	9.5
Homology to carrot EDGP <sup>(c)</sup>	B8-9, B10-12, D6-7, E9, F56	At1g03220	1-22	49.6-43.6	5.8-6.1	43.4	10.0
Disease resistance protein <sup>(c)</sup>	B12-13	At1g33590	1-24	48.5-45.3	6.2-n.d.	48.1	9.9
Unknown with LRR motifs	B12-13	At1g33590	1-24	48.5-45.3	6.2-n.d.	48.1	9.9
Polygalacturonase inhibiting protein 1	E12	At5g06860	1-21	35.4	n.d.	34.3	9.6
Polygalacturonase inhibiting protein 2	A7, E13	At5g06870	1-21	34.0	n.d.	34.8	9.8
<b>Miscellaneous</b>							
EXORDIUM	A9	At4g08950	1-21	29.4	n.d.	31.2	9.9
Homology to purple acid phosphatase	A1, A2	At4g36350	1-21	50.0	5.0-5.3	50.6	6.3
Homology to strictosidine synthase	E13	At1g74000	1-23	33.3	n.d.	32.1	10.1
Apospory-associated like protein	B25, D10-12, E14-15, F10-11	At4g25900	1-20	35.3-27.9	n.d.	33.9	9.9
SKU5	D2	At4g12420	1-20	75.0	n.d.	63.3	9.6
<b>Unknown function</b>							
Unknown	A5	At2g41800	1-18	37.9	n.d.	38.3	9.9

Unknown	B28	At4g34180	1-24	26.4	5.6	25.8	5.6
Unknown	E1	At5g18860	1-21	100.0	n.d.	96.7	7.9
Unknown	C1	At4g24890	1-20	77.9	5.3	59.3	5.5
Unknown	A33-35, B29, C7	At5g24460	1-46	27.1-25.5	5.2-4.6	33.5	9.73

---

CWPs were ordered according to putative functions. (a) Spot number in gels from Fig. 3. (b) Gene accession number according to the MIPS database. (c) Proteins found also in the culture medium. n.d.: non determined. c.m.: culture medium

**Table 2.** Intracellular proteins identified in the sequential extracts of *A. thaliana* cell cultures.

Protein identification	Spot <sup>(a)</sup>	Gene Acc. N <sup>o</sup> <sup>(b)</sup>	Experimental		Theoretical	
			MM (kDa)	pI	MM (kDa)	pI
Homology to aldose-1-epimerase	A17	At3g17940	4.8	5.6	37.2	5.8
Malate dehydrogenase	A18	At5g43330	33.7	5.7	35.7	6.3
Homology to jacalin	A29	At3g16450	28.5	4.9	32.0	4.9
Homology to jacalin	A30	At3g16420	29.1	5.2	32.2	5.6
Unknown function	A31-32	At5g02240	26.1	5.5-5.8	27.1	6.0
Triose phosphate isomerase	A36-39	At3g55440	24.2	5.0-5.5	27.2	5.1
Homology to glutathione S-transferase (GST)	A40-42	At1g19570	22.8	5.0-5.4	23.6	5.4
Gly-rich protein 2	A43, A45-46	At4g38680	21	5.0-5.5	19.1	5.4
Gly-rich protein 8	A55	At4g39260	16.4	5.0	16.6	5.1
Unknown function	A44	At2g40600	22.0	5.2	20.7	5.4
Glutathione peroxidase	A 47	At4g11600	19.3	5.6	18.6	7.6
Peroxiredoxin type 2	A48	At1g65970	17.3	5.2	17.4	5.1
Peroxiredoxin type 2	A49, A51	At1g65980	18.1	4.2-5.0	17.4	4.9
Homology to avirulence induced gene2 (AIG2)	A50	At5g39730	19.2	4.7	20.0	4.8
Homology to AIG2	A53-54	At3g28940	17.7	4.4-4.7	19.5	4.7
Unknown function	A52	At1g63220	17.3	4.3	16.3	4.1
Homology to Zn binding protein	A5-58	At3g56490	13.9	5.1-5.6	16.0	6.7
Profilin 1	A59	At2g19760	14.1	4.6	14.3	4.4
Profilin 2	A60-62	At4g29350	13.2	4.4-4.7	14.0	4.7
Thioredoxin H-type 5	A61	At1g45145	12.9	4.8	13.1	4.9
Thioredoxin H-type 3	A62	At5g42980	13.2	4.4	13.1	4.8
Cyclophilin	D16-17, D20, E20, F16	At4g34870	20.2-11.8	n.d.	18.4	9.9
Cyclophilin (ROC3)	D18-20, F16	At2g16600	14.8-11.8	n.d.	18.5	9.7
Unknown function		At5g11680	22.1	5.8	23.1	6.3
Copper chaperone (CCH)		At3g56240	19.2	4.8	13.0	4.6
Homology to sterol methyltransferase 3 (SMT) 3		At5g55160	14.6	4.7	11.6	5.0
Homology to 2,3-bisphosphoglycerate independent phosphoglycerate mutase		At1g09780	65.2	5.1	60.6	5.1
ATP synthase ( $\beta$ subunit)		At5g08670/80/90 <sup>(c)</sup>	55.0	5.2	59.7	5.9-6.1
Dihydrolipoamide dehydrogenase precursor		At3g17240	57.1	5.7	54.0	6.5
S-adenosyl methionine synthetase		At2g36880	46.3	5.5	42.5	5.6
Homology to putidaredoxin reductase		At5g03630	43.5	5.1	47.5	5.0
Homology to translationally controlled tumor protein-like protein		At3g16640	42.4	4.5	18.9	4.3
Annexin (AnnAt1)		At1g35720	34.9	5.0	36.2	5.0
Unknown function		At3g60450	30.0	4.9	30.1	4.9
Homology to phosphogluconolactonase		At5g24400	25.7	5.2	28.0	5.2
Unknown function		At2g32520	23.7	5.0	25.9	5.1
Homology to thiol methyl transferase		At2g43910	23.9	4.8	25.3	4.5
Homology to glutathione S-transferase 6		At2g47730	22.2	5.6	24.1	6.0

Cyclic phosphodiesterase	At4g18930	19.9	4.8	20.5	4.7
Homology to initiation factor 5A	At1g26630	17.5	5.2	17.1	5.3
Homology to actin depolymerizing factor 3	At5g59880	16.8	5.0	15.9	5.8
Nucleoside diphosphate kinase 1	At4g09320	14.6	5.7	16.5	6.2
Homology to calmodulin	At1g18210	13.0	4.3	18.3	4.0
Homology to cysteine protease inhibitor	At5g12140	12.2	4.8	11.3	5.1
Calmodulin camX <sup>(d)</sup>	<sup>(d)</sup>	12.4	4.3	16.8	3.9

(a) Spot number in gels from Fig. 3A, B and C were not labeled. (b) Gene accession number according to the MIPS database. (c) The peptides from MALDI-TOF matched all the proteins of the same family and do not allowed to discriminate between the three genes. (d) There are six very conserved genes for the classical calmodulins, the peptides obtained matched all the six deduced proteins and it was not possible to discriminate between the different proteins and genes (At2g27030, At2g41110, At3g56800, At5g37780, At1g66410, At3g43810). n.d.: non determined.