Open access • Journal Article • DOI:10.1002/PMIC.200400882

Cell wall proteins in apoplastic fluids of Arabidopsis thaliana rosettes: identification by mass spectrometry and bioinformatics. — Source link

Georges Boudart, Elisabeth Jamet, Michel Rossignol, Claude Lafitte ...+4 more authors

Institutions: Paul Sabatier University, Centre national de la recherche scientifique

Published on: 01 Jan 2005 - Proteomics (John Wiley & Sons, Ltd)

Topics: Proteome, Signal peptide, Cytoplasm, Proteomics and Cell wall

Cell wall proteins in apoplastic fluids of Arabidopsis thaliana rosettes: identification by mass spectrometry and bioinformatics.

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

Published in: Proteomics 5, 212-221 (2005)

Cell wall proteins in apoplastic fluids of *Arabidopsis thaliana* rosettes: identification by mass spectrometry and bioinformatics

Georges BOUDART[1, 2], Elisabeth JAMET[1, 2], Michel ROSSIGNOL[1, 2], Claude LAFITTE[2], Gisèle BORDERIES[2], Alain JAUNEAU[3], Marie-Thérèse ESQUERRÉ-TUGAYÉ[2], Rafael PONT-LEZICA[2]

([1]) G. Boudart, E. Jamet and M. Rossignol contributed equally to this work.

([2]) UMR 5546 CNRS-Université Paul Sabatier, Pôle de Biotechnologie Végétale, 24, Chemin de Borde Rouge, 31326 Castanet-Tolosan, France
([3]) IFR 40, CNRS, Pôle de Biotechnologie Végétale, 24, Chemin de Borde Rouge, 31326 Castanet-Tolosan, France

Running title: Arabidopsis cell wall proteins of rosettes
Corresponding author: 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*, bioinformatics, cell wall, mass spectrometry, proteomics

Abbreviations: **CWP**, cell wall protein; **CWMP**, cell wall modifying protein

Abstract

Weakly bound cell wall proteins of *Arabidopsis thaliana* were identified using a proteomic and bioinformatic approach. An efficient protocol of extraction based on vacuum-infiltration of the tissues was developed. Several salts and a chelating agent were compared for their ability to extract cell wall proteins without releasing cytoplasmic contaminants. Of the 93 proteins that were identified, a large proportion (60 %) was released by calcium chloride. From bioinformatics analysis, it may be predicted that most of them (87 out of 93) had a signal peptide, whereas only six originated from the cytoplasm. Among the putative apoplastic proteins, a high proportion (67 out of 87) had a basic pI. Numerous glycoside hydrolases, and proteins with interacting domains were identified, in agreement with the expected role of the extracellular matrix in polysaccharide metabolism and recognition phenomena. Ten proteinases were also found as well as six proteins with unknown functions. Comparison of the cell wall proteome of rosettes with the previously published cell wall proteome of cell suspension cultures showed a high level of cell specificity, especially for the different members of several large multigenic families.

1 Introduction

Cell walls of higher plants are dynamic structures essential for development and morphogenesis [1, 2]. They are involved in cell expansion and division, as well as in cell-cell interactions and cell separation phenomena [3-6]. Cell walls contribute to the general morphology of the plant and are at the forefront of plant-microorganism interactions [7, 8].

Plant primary cell wall is essentially composed of carbohydrates, consisting in a cellulose and hemicellulose network embedded in a pectin matrix. The different polysaccharides account for 86% of the cell wall mass of leaves in *Arabidopsis thaliana* [9]. With the discovery of proteins within the cell wall by Lamport and Northcote [10], the concept of a dynamic structure, containing many more proteins than expected, emerged progressively. The refinement of biochemical fractionation techniques and the completion of the *Arabidopsis* genome sequence [11] indicate that cell walls contain a few hundreds of different proteins, which account approximately for 10 % of the wall dry weight. According to the present knowledge, cell wall proteins (CWPs) may be grouped into four main functional categories related to cell wall structure, remodeling, signaling, and defense [12-17]. However, most cell wall models lack CWPs because their precise partition in different tissues and their localization within the matrix are not known yet [1]. Through their likely interaction with polysaccharides and other proteins, they might be involved in macromolecular complexes whose structure, function, and physico-chemical properties remain to be discovered.
Analysis of expressed sequence tags (ESTs) databases and systematic transcriptome studies have been applied to gene families encoding proteins known to be involved in wall biosynthesis and modeling. This includes notably cellulose synthases, xyloglucan endo-transglycosilases, glycosyl transferases, expansins, and peroxidases [18-22]. It also concerns proteins related to recognition and signalling phenomena, such as GPI-anchored proteins among which are the arabinogalactan proteins (AGPs), and the LRR-extensin (LRX) hybrid proteins [23-25]. These studies and the data gained from genomic analyses of a few crop plants, confirmed and extended the notion that CWPs are encoded by multigenic families, whose expression appear to be tissue-specific and developmentally regulated. Knowledge of these gene families may be used to establish the phylogeny and chromosomal location of their various members, as well as for functional studies.

Proteomic approaches have the great advantage to give direct access to the expression pattern of individual members of CWP families. With the recent progresses in mass spectrometry technology and the availability of complete genome sequences, it has become feasible to directly identify the proteins of a given cell compartment, tissue or organ in a given physiological situation. Proteomics of *A. thaliana* cell wall compartment is currently under investigation and presents several difficulties. The challenge of the project lies, in part, in the solubilization of cell wall proteins, due to the numerous interactions between proteins and the polysaccharide matrix. Translational modifications, notably glycosylation, commonly found in CWPs represent an additional problem. The third obstacle is the possibility of contamination with proteins of different subcellular compartments. Recent studies on the cell wall proteome of fungi and higher plants have shown the presence of classical well known CWPs, of proteins not known to be addressed to the cell wall, and of known cytosolic proteins like glycolytic enzymes, transcription factors, etc. [26-29]. The presence of these non-canonical CWPs at the surface of the cell has two explanations: either they are simply contaminants, or there is an unknown secretion system for proteins devoid of a canonical signal peptide. The latter hypothesis has not yet been experimentally demonstrated.

The present work is the first one to present the cell wall proteome of a living organ, the rosette. A careful extraction procedure permits the extraction of loosely bound CWPs with minimal contamination by cytoplasmic proteins, and supports the discovery of novel CWPs. The effect of different salts and chelating compounds on the solubilization of CWPs is analyzed. Finally, a comparison of the proteomes from rosettes and from cell suspension culture is discussed.

2 Materials and methods

2.1 Plant material

Plants of *Arabidopsis thaliana* ecotype Col-0 were grown on compost in a growth chamber at 70% hygrometry, with a photoperiod of 9 h light at 110 µE $m^{-2}s^{-1}$ at 22°C, and 15 h dark at 20°C.

2.2 Rosettes infiltration and extraction of apoplastic proteins

Four week-old plants, at the rosette stage, were carefully removed from the pots and compost was washed off with deionized water. Plants were then treated for infiltration as follows. A small noose was made with a piece of string and the root was passed through the noose. The noose was tightened around the collar. The root was then twisted around the string, behind the noose and wrapped in parafilm. Rosettes were completely immersed in a solution of 0.3 M mannitol in a dessicator connected to a vacuum pump and rosettes were vacuum infiltrated with 0.3 M mannitol for 2 minutes at room temperature. The infiltrated plants were transferred to a centrifuge tube, with the collar at about 1 cm from the edge of the tube. The lower part of the root was pasted outside of the tube with adhesive tape. Three hundred µL of a solution containing 66 mM DTT, 0.33 M thiourea and 3.3% (v/v) protease inhibitor cocktail (Sigma, St. Louis) in 0.3 M mannitol was introduced at the bottom of the tube prior to centrifugation. Plants were centrifuged in swinging buckets at 200g for 17 minutes at 20°C. The apoplastic washing fluids were collected and the volumes were estimated with a micropipette. Vacuum infiltration and centrifugation were repeated once. In that case, rosettes were directly infiltrated in the centrifuge tubes and the infiltration solution was discarded before centrifugation.

Afterward, rosettes were vacuum infiltrated with 1 M NaCl, 0.2 M $CaCl_2$, 0.05 M CDTA or 2 M LiCl in 0.3 M mannitol. All saline infiltration solutions were adjusted to pH 6.9 with 0.1 M HCl or 5 N NaOH, depending on the chemical. Rosettes were centrifuged as previously described. Vacuum infiltration and centrifugation were repeated once. The protein content of each extract was measured by the bicinchoninic acid (Interbiotech, France) method using BSA as standard.

Apoplastic fluids were assayed for malate dehydrogenase activity (MDH, EC 1.1.1.37) in order to detect cytoplasmic contaminations. MDH was assayed at room temperature in 3 mL of 107.5 µmol Tris-HCl pH 7.8, 15 µmol $MgCl_2$, 15 µmol DTT, 1.5 µmol NADP, 9 µmol malic acid, and one twentieth of the volume of the recovered apoplastic fluid. The reduction of NADP was followed at 340 nm. Apoplastic washing fluids with no detectable malate dehydrogenase activity were pooled and processed for protein separation.

2.3 Microscopy

Samples were fixed with 2.5% (v/v) glutaraldehyde in 0.05 M cacodylate buffer, pH 7.0, in the presence or the absence of mannitol in the mixture overnight (1 h at room temperature and

then at 4°C). They were then dehydrated in a graded ethanol series (20, 40, 60, 80% ethanol 2x 15 min, and absolute ethanol 2x 30 min). They were infiltrated with Spur's epoxy resin and polymerized overnight at 70°C. Transverse sections (1µm in thickness) were stained with 0.5% toluidine blue (m/v) in 2.5% sodium carbonate, pH 11. Sections were observed using an inverted microscope (DMIRBE, Leica) and images were acquired with a CCD camera (Color CooledView, Photonic Science, UK).

2.4 Fractionation of 0.3 M mannitol-eluted apoplastic fluid

Apoplastic washing fluids from rosettes infiltrated with 0.3 M mannitol were dialyzed at 4°C against 20 L of deionized water in dialysis bags, low binding 2 kDa cut-off (Spectra Por CE, Merck eurolab Poly Labo, France), and freeze-dried. The dry residue was solubilized in 3 mL of a mixture containing 1 M thiourea, 10 mM DTT, 1% (v/v) protease inhibitor cocktail and desalted on a desalting column (Econo-Pac 10DG, Amersham Biosciences, Sweden), equilibrated with 0.2 M ammonium formate, for the complete removal of mannitol. The recovered material eluted at the void volume was immediately frozen in liquid nitrogen and freeze-dried. The dry residue was solubilized in the 2-DE sample buffer composed of 8 M urea, 2% CHAPS, 1% (v/v) protease inhibitor cocktail mixture, 2% IPG buffer, pH 4-7 , 0.3% DTT, and loaded directly in a pH 4-7 IEF 7 cm gel strip. Proteins were focused using a Multiphor II apparatus (Amersham Biosciences, Sweden) at 45,000 Vh. After focusing, the proteins were reduced and alkylated, and the gel strip was loaded on top of a 12.5% acrylamide gel for SDS-PAGE, as previously described [29]. After electrophoresis, the 2-D gels were fixed and stained with silver nitrate according to Schevchenko et al. [30] and numerized with an Image scanner (Amersham Biosciences, Sweden).

2.5 Fractionation of the salt- or chelator-eluted apoplastic fluids

The apoplastic washing fluids from leaves infiltrated with 1 M NaCl, 0.2 M $CaCl_2$, 2 M LiCl, or 50 mM CDTA in 0.3 M mannitol were exhaustively dialyzed against cold deionized water as described above and separated on a SP-Sepharose column (Hi-Trap, Amersham Biosciences, Sweden) equilibrated with 10 mM MES buffer pH 5.2. The retained basic proteins were eluted with 2 M NaCl in the same buffer, desalted on a 5 ml BioGel P6DG column (Amersham Biosciences, Sweden) equilibrated with 0.2 M ammonium formate and freeze-dried. Proteins were resuspended in 40 µL of the sample buffer, 62 mM Tris-HCl, pH 6.8, 2% SDS, 10% glycerol, and 5% mercaptoethanol, and separated according to size by a 10%-17% gradient SDS-PAGE (16.5 x 13.5 x 0.15 cm). After electrophoresis, the 1-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. The 1-D gels could be subsequently stained with silver nitrate [30].

The acidic and neutral proteins in the Hi-Trap SP effluents were freeze-dried. The dry residues were solubilized with a minimal volume of a solution containing 1 M thiourea, 10 mM DTT, 1% (v/v) protease inhibitor cocktail and desalted as previously described. The desalted proteins were then freeze-dried and prepared for 2-DE as described above, except for 50 mM CDTA acidic proteins that were focused in a 3-10 NL IEF gel strip.

2.6 Protein identification by mass spectrometry and bioinformatics

Protein spots of interest were excised from CBB-stained gels. Proteins were characterized after trypsin in-gel digestion, by peptide mass fingerprinting using a MALDI-TOF mass spectrometer (Voyager-DE^{tm} STR, Perseptive Biosystems USA), as previously described **[29].** Peptide mass fingerprints were compared to the database of *Arabidopsis* from NCBI (National Center for Biotechnology) non-redundant database using MS-FIT (Protein

Prospector, http://prospector.ucsf.edu). The searches were done with a tolerance of 20 ppm, and considered oxidation of methionine, modification of cysteine by carbamidomethylation as well as partial cleavage leaving a maximum of one internal site uncleaved by trypsin. PSD spectra were generated from selected peptides each time it was necessary to improve results of peptide mass fingerprinting. Results were analyzed with MS-TAG (Protein Prospector, http://prospector.ucsf.edu).

For unidentified spots by MALDI techniques, trypsin digests were separated and analyzed using a Q-TRAP™ (AppliedBiosystems/MDS Sciex, USA) LC/MS/MS system. Briefly, peptides from digest were injected (LC Packings Famos autosampler) and concentrated on a µ-Precolumn™ Cartdrige (PepMap™, LC Packings) C18, 5µm, 100A° in buffer A, where buffer A was 0.1% formic acid in water. Peptides were then separated on a 75µm ID X150 mm C18, 3 µm column (PepMap™, LC Packings) by a 30 min linear gradient of 0-60% buffer B, where buffer B was 90% acetonitrile and 0.1% formic acid in water. Flow rate was 200 nl/min using LC Packings Ultimate™ LC pump. Data were acquired on mass spectrometer by an enhanced MS survey scan followed by a dependent enhanced resolution scan and two dependent enhanced product ion scans giving a total cycle time of about 4 seconds.

Mass data collected during LC-MS/MS analysis were processed by the Analyst software (AppliedBiosystems/MDS Sciex, USA) and submitted to the search software MASCOT (Matrix Science, London, UK). Searches were done with a tolerance on mass measurement of 0.5 Da in MS mode and 0.3 Da in MS/MS mode. Protein identification was accomplished using the NCBI non-redundant protein database.
All these results are given in Tables 1, 2 and 3 (supplementary material).

3 Results

3.1 Extraction of rosette CWPs from apoplastic fluids

To limit potential contamination by cytoplasmic proteins and to recover a maximum number of CWPs, we adapted to the vacuum infiltration technique the sequential extraction protocol described previously [31, 32]. The challenge was that the various salts or chaotropic agents, used for extraction, could act on the walls while maintaining the cellular content confined within the protoplast. Previous studies on *A. thaliana* cell cultures indicated that membranes do not stand repeated extractions and become leaky [29]. To avoid this problem, a new protocol was developed with all extraction steps done on four week-old rosettes plasmolyzed in 0.3 M mannitol. The sequential elution of the apoplastic fluid was performed first with 0.3 M mannitol alone, and then with each salt or chelating agent diluted in 0.3 M mannitol in the presence of a protease inhibitor mixture.

3.1.1 Proteins solubilized by mannitol

About 500 µL of apoplastic fluid were collected per plant infiltrated with 0.3 M mannitol. Only proteins with acidic pI and MM ranging from 4 to 7, and 10 to 150 kDa respectively, could be visualized on 2-DE. Fig. 1 shows a characteristic pattern of the gels obtained, at least two gels were run for each sample. Twenty-two proteins were identified out of 46 silver nitrate-stained spots subjected to mass spectrometry analysis. Several isoforms of the same protein were found thus reducing the total number of proteins to eight: two β-1,3-D-glucosidases (At3g57260 and At3g57240), one α-mannosidase (At5g13980), one thaumatin (At1g75040), four proteins with homology to a β-xylanase (At4g33810), a subtilisin-like serine protease (At5g67360), a β-xylosidase (At5g64570) and a berberine-bridge S-

reticulin:oxygen oxidoreductase (At2g34790), respectively. The remainder corresponded to proteolytic digests of some of the previously cited proteins - with the exception of the only cytoplasmic protein found, i.e. the precursor of the RUBISCO 3b small subunit (Fig. 1, spot 22).

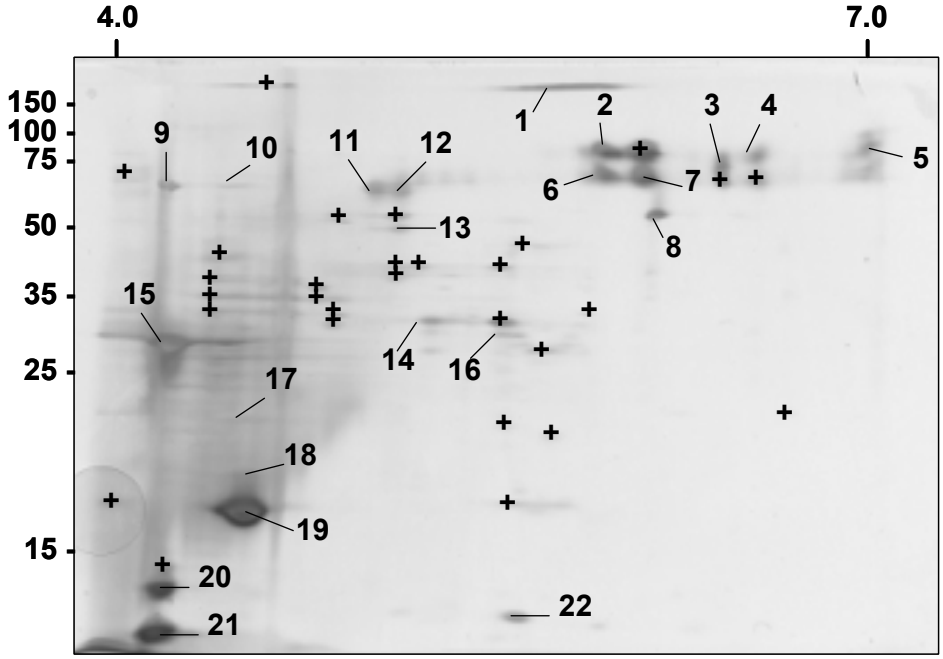

Figure 1. 2-D gel electrophoresis of rosette apoplastic proteins solubilized by 0.3 M mannitol. The numbers correspond to spots identified by MALDI-TOF-MS. + indicate non-identified proteins. The pH range used for IEF is indicated by numbers on top of the gel (4, 7). Numbers on the left indicate the sizes of molecular mass markers in kDa.

3.1.2 Proteins solubilized by salts or chelator

Basic proteins were separated from acidic ones by cation-exchange chromatography and submitted to gradient 1-D SDS PAGE, followed by CBB staining (Fig. 2). 1-D gels were highly reproducible and most of the proteins were identified on at least two gels. CBB staining was sufficient to reveal all protein bands, as checked by further staining of the 1-D gels with silver nitrate. The CBB staining patterns of the four extracts - 0.2 M $CaCl_2$, 1 M NaCl, 2 M LiCl and 50 mM CDTA - appeared to be quite different. Polypeptides of MM higher than 20 kDa were mainly found in $CaCl_2$, NaCl and CDTA extracts, whereas low MM polypeptides were more particularly detectable in the LiCl sample. Identification of CWPs and bioinformatics allowed assigning a basic pI to 80% of the mature proteins. In comparison, the amount of the non-retained acidic proteins was very low, whatever the salt extract. Very few spots were detected after 2-DE and silver nitrate staining (data not shown). The identification of only a few new proteins with pIs ranging from 5.0 to 8.9 and only little MM variations was possible in 2-D gels, namely one homolog to cysteine proteinase RD21A (At4g11310), one homolog to fimbrin protein (At2g04750), one homolog to glycoside hydrolase family 27 (At5g08380) and 7 isoforms of a germin-like protein (At5g20630).

Fig. 3 displays the respective efficiency of 0.2 M $CaCl_2$, 1 M NaCl, 2 M LiCl, or 50 mM CDTA in 0.3 M mannitol to remove weakly bound-cell wall proteins from the apoplast, as compared to that of 0.3 M mannitol alone. Only those proteins characterized as CWPs using bioinformatics analyses or data available in literature were taken into account. It should be noted that all these proteins were predicted to have putative N-glycosylation sites. Altogether, 87 different proteins were identified by mass spectrometry from the various salt and chelator extracts. The results clearly indicated that CWPs were differentially solubilized, depending on the chemical used for elution. Sixty per cent of all CWPs were released by $CaCl_2$, of which a large proportion (60 %) was solubilized in a $CaCl_2$-specific manner. Fewer CWPs were eluted

by NaCl, LiCl and CDTA than by $CaCl_2$, and even fewer in a specific manner, i.e. 17 %, 19 %, and 28.5 % respectively.

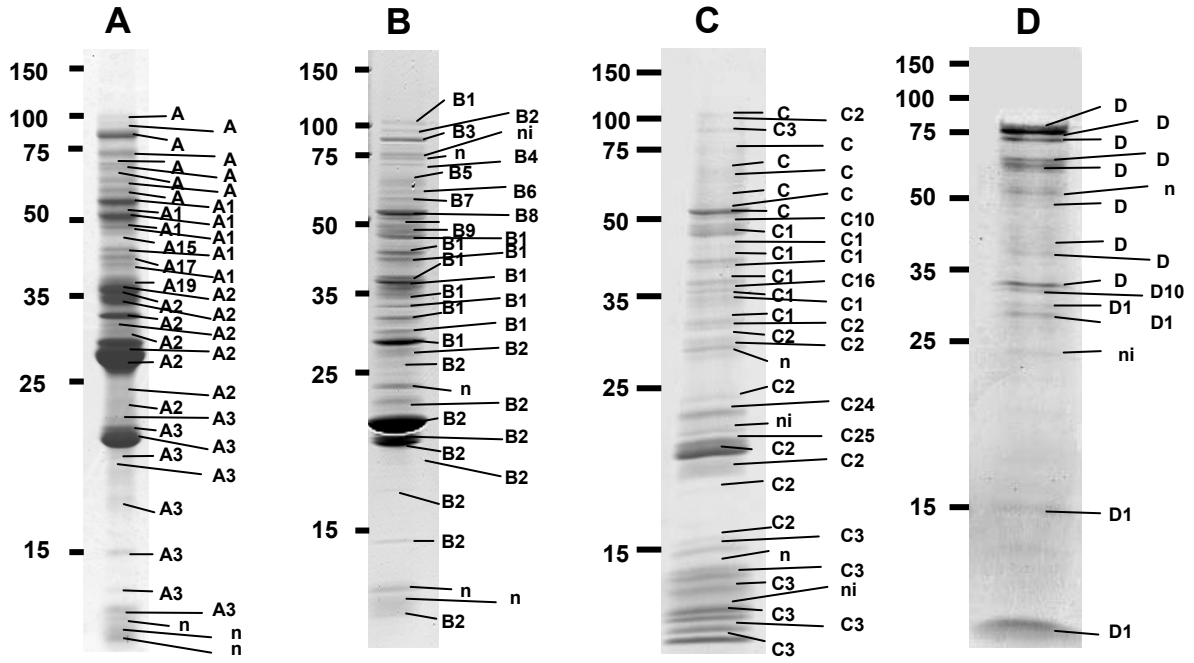

Figure 2. 1-D SDS-PAGE analysis of basic CWPs from 0.2 M $CaCl_2$ (A), 1 M NaCl (B), 2 M LiCl (C) or 50 mM CDTA (D) extracts. Numbers indicate the protein bands that were identified by MALDI-TOF-MS, as listed in Table 1 (supplementary material) n.i. corresponds to non-identified proteins. The size of molecular mass markers in kDa is indicated to the left of each lane.

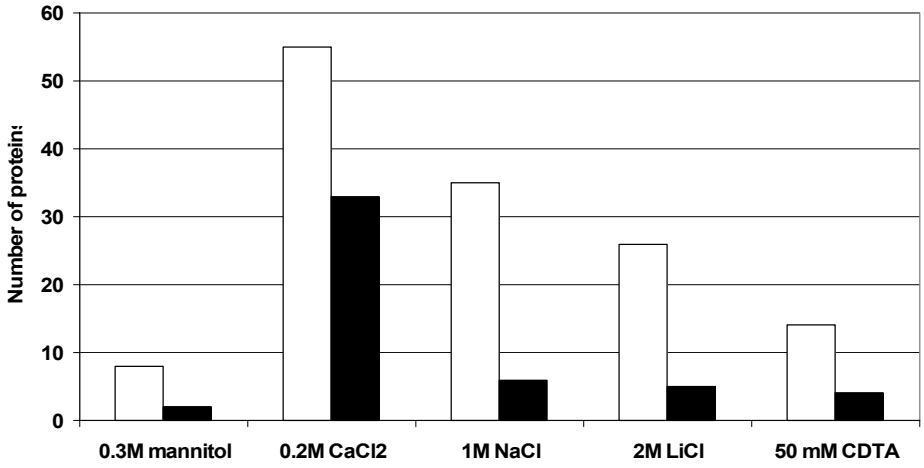

Figure 3. Comparison of the number of rosette apoplastic proteins solubilized by 0.3 M mannitol, 0.2 M $CaCl_2$, 1M NaCl, 2 M LiCl or 50 mM CDTA. Proteins were extracted, separated and identified as described in Material and Methods. The total number of proteins (empty boxes) and the number of proteins extracted in a salt-specific manner (black boxes) are indicated for each chemical. The average amount of proteins solubilized from 30 rosettes in each fraction was: mannitol, 70 µg; $CaCl_2$, 350 µg; NaCl, 130 µg; LiCl, 90 µg; and CDTA, 55 µg.

3.1.3 Microscopic survey

To control the integrity of the tissues submitted to plasmolysis and centrifugation, a microscopic analysis was done (Fig. 4). All the samples were fixed in solutions containing or not mannitol. In the absence of mannitol, the possibility for a plasmolyzed cell to recover its turgescence indicates that most of the cell compartments are functional. As expected, in the presence of mannitol in the fixative solution, the plasmolysis was clearly visible within the different tissues of the leaf (Fig. 4 A and B). Protoplasts were separated from the cell wall in the parenchyma cells of the upper (A) and of the middle (B) part of the leaf. Integrity of vacuoles and protoplasts appeared unaffected by the plasmolysis. When mannitol was omitted in the fixative solution, cells recovered their turgescence and thus exhibited a large central vacuole and chloroplasts were distributed all around cells (C and D). Figs E and F show $CaCl_2$ extracted samples fixed in the absence of mannitol. The overall tissue structure within the leaf was well conserved (E) and the main cell compartments were clearly visible including the nucleus (dotted arrow in F). The results obtained from CDTA- treated samples were similar, the only difference concerned the swelling of the epidermal and sub-epidermal cell walls (arrowheads in G) resulting from the well-known effect of this calcium chelator on solubilization of acidic pectin.

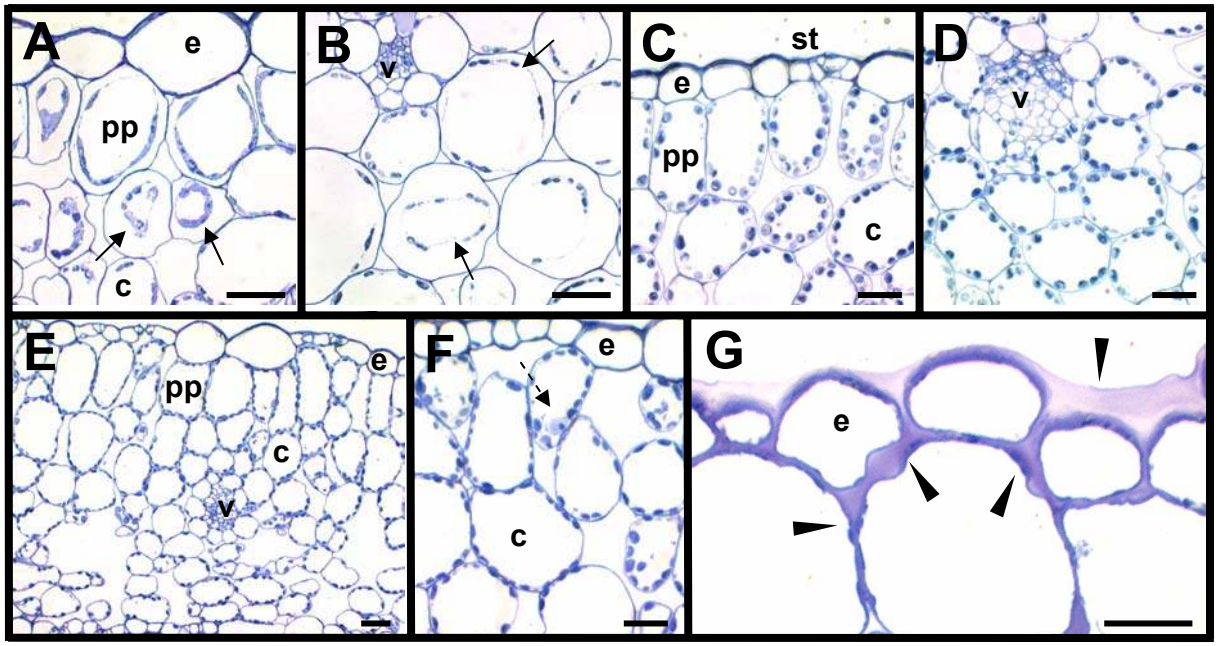

Figure 4. Optical microscopy on transverse sections of unextracted and extracted leaves. Unextracted plasmolyzed samples were fixed in the presence (A-B) or absence (C-D) of mannitol in the aldehyde solution and stained with toluidine blue. The upper and the middle part of the leaf are shown in A-C and B-D, respectively. E-F: General view and detail of the upper part of a leaf of $CaCl_2$-extracted samples fixed in the absence of mannitol. G: CDTA-treated sample, note the swelling of the epidermal and sub-epidermal cell walls (arrowheads). Arrows indicate plasmolyzed cells (A and B), dotted arrow, a nucleus and its nucleolus (F). e, epidermis; c, cortical parenchyma cell; pp, palisade parenchyma; v, vein. Bars: 20 µm.

3.2 Protein functional classification

The ascertained or putative functions of the CWPs identified in this study are listed in Tables 1 (supplementary material) and the number of proteins in each functional class is represented on Fig. 5. Proteins were classified into six classes: cell wall modifying proteins (CWMPs), defense-related proteins, proteins containing a domain for interaction with polysaccharides or proteins, proteinases, miscellaneous, and proteins with unknown function. Assignment of

several proteins was based on their supposed predominant function even if they may play different roles during development or in response to biotic or abiotic stresses.

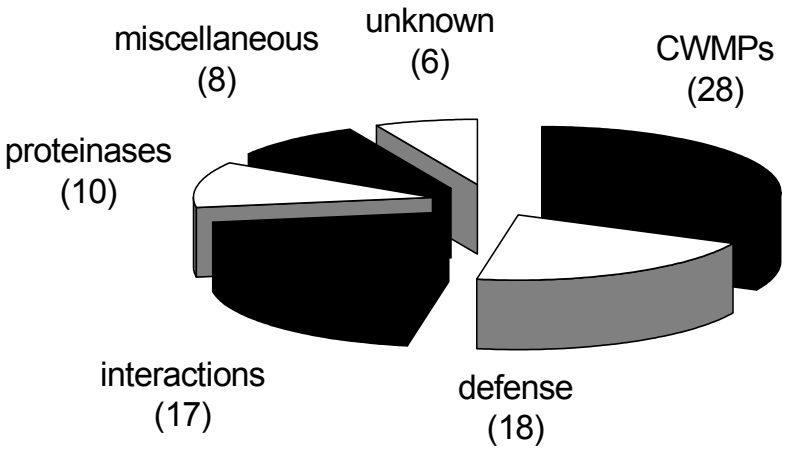

Figure 5. Functional classification of rosette apoplastic proteins. Proteins were assigned to six classes depending on their function. They were classified as: CWMPs: cell wall modifying proteins; defense: defense-related proteins; interaction: proteins with domain of interaction with polysaccharides and/or proteins; proteinases; miscellaneous; unknown: proteins with unknown function. The number of proteins present in each class is in brackets.

The three most represented classes are those of CWMPs, defense-related proteins and proteins with interacting domains, i.e. 28, 18 and 17 proteins, respectively. CWMPs are essentially composed of glycoside hydrolases - or proteins with homology to glycoside hydrolases - with the exception of two proteins showing homology to α-expansins (At1g20190 and At2g18660). Several families of glycoside hydrolases are represented, mostly including α- and β-glucosidases as well as α- and β- galactosidases. Other hydrolytic enzymes are related to pectin-modifying enzymes (At1g76160, At1g41830, At4g14310 and At5g45280) and xyloglucan endotransferase (At2g06850). The second class comprises three proteinase inhibitors (At1g17860, At4g16500, At5g47550) and one inhibitor of pectin methylesterase (At1g47960). Other well-known defense-related proteins are included in this class, notably the antifungal PR1 and PR5 proteins (At2g14610, At1g75040), several members of the germin family, one thaumatin-like (At1g75040) and one chitinase (At4g19810). The third class includes 17 CWPs containing a domain of interaction with other proteins and/or polysaccharides. Most of them have homology to lectins or disease resistance proteins having LRR domains. The biochemical or biological functions of these proteins are still unknown, except for the PGIP2 polygalacturonase inhibitor (At5g06870) [33].

The fourth class corresponds to ten proteolytic enzymes, including one subtilisin-like serine protease (At5g67360) as well as several members of the protease family with homology to serine- (At2g05920, At1g01900, At4g00230, At5g42240), aspartyl- (At1g09750, At5g07030) and cysteine- (At1g20850, At4g11310) proteinases, and peptidases (At5g27430, At3g05230). Miscellaneous proteins (8 proteins) most notably include a protein with homology to a lipid-transfer protein (At3g51600) and a homolog to MD-2-related lipid recognition domain protein (At3g44100). Four proteins with a predicted transmembrane domain were also deliberately included in this class, namely one homolog to CLAVATA1 (At5g05160), one homolog to fimbrin protein (At2g04750), IRX1 (At4g18780) and one homolog to FAD-linked oxidoreductase family (At2g46740). All these proteins were identified as proteolytic fragments belonging to their extracellular domains. Finally, six proteins with unknown function and without homology to other proteins were found and listed in the sixth class. They might correspond to plant-specific genes.

It should be noted that, presently, the function of only 16 proteins in total, has been demonstrated experimentally. The function of all the others (75%) was deduced from sequence or structural similarity with proteins from other organisms.

3.3 Distribution of classified proteins

CWPs were unevenly distributed within each of the functional classes defined above, depending on the chemical used for the extraction. For each class, the number of mannitol- or salt-solubilized proteins is shown on Fig. 6A. As already mentioned, $CaCl_2$ solubilized ionically bound-cell wall proteins more efficiently than the other salts whatever the protein class considered. Defense-related proteins and proteins having interaction domains were found to be preferentially eluted by NaCl and LiCl. CDTA extracts did not contain any miscellaneous protein or protein of unknown function. Only a few CWMPs and some defense-related proteins were solubilized by mannitol.

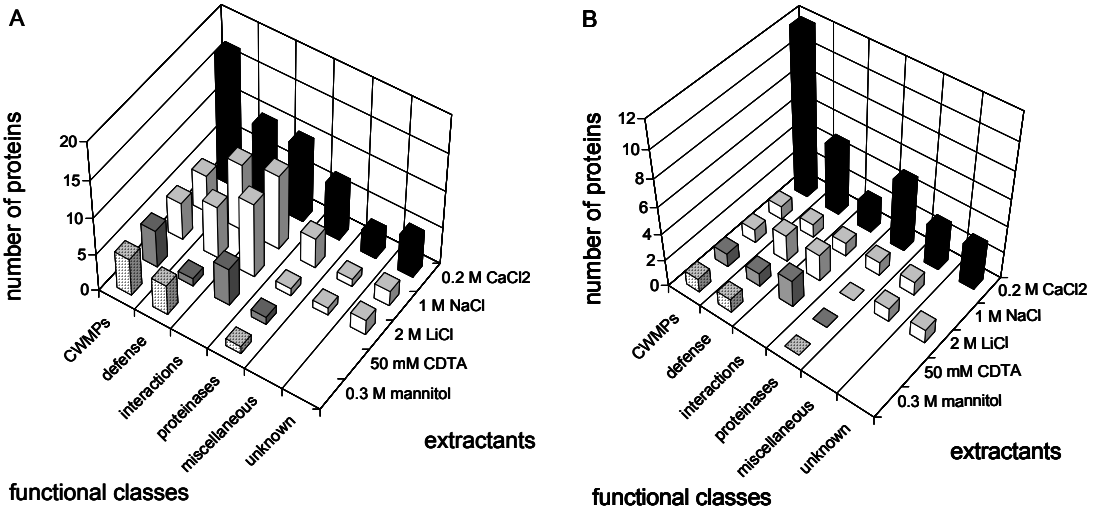

Figure 6. Distribution within six functional classes of the rosette apoplastic proteins solubilized by each extractant. (A) Total number of proteins extracted by each chemical. (B) Number of proteins extracted in a salt-specific manner. Apoplastic proteins extracted by each chemical were identified by MALDI-TOF-MS and classified according to their putative function: CWMPs: cell wall modifying proteins; defense: defense-related proteins; interaction: proteins with domain of interaction with polysaccharides and/or proteins; proteinases; miscellaneous; unknown: proteins with unknown function. The number of proteins in each class is indicated.

The number of CWPs extracted in a salt-specific manner is reported on Fig. 6B. Only $CaCl_2$ was able to substantially solubilize proteins from each functional class. Thus, 12 out of 17 CWMPs, 5 out of 7 proteinases, 3 miscellaneous proteins and 3 proteins of unknown function were solubilized by $CaCl_2$. By comparison, the number of defense-related proteins that were specifically extracted by NaCl was low, whereas NaCl solubilized only one CWMP and no protein having interaction domains.

4 Discussion

This work contributes to three different areas: (i) it describes a mild, reliable protocol to release loosely bound CWPs from *A. thaliana* rosettes; (ii) it evaluates the efficiency of several salts and one chelating agent to elute these CWPs; (iii) it results in the identification of 93 different proteins. In addition, it allows the comparison of CWPs from rosettes to the previously reported cell wall proteome of cell suspension cultures.

First, a mild extraction protocol was developed based on infiltration of rosette leaves followed by gentle centrifugation of the infiltrated material. In order to preserve cell integrity from damages provoked by the combination of vacuum and centrifugation, all infiltration steps were performed in the presence of an osmoticum, 0.3 M mannitol. Microscopic observations of the leaves confirmed that good shape and functionality of the cells were preserved. Finally, mass spectrometry combined to bioinformatics analysis allowed the identification of 93 proteins among which 87 were CWPs based on the presence of a signal peptide predicted by either PSORT or SignalP. It has been postulated that proteins without signal peptide may be secreted [26-28]. The evidence came from previous proteomic studies on cell walls prepared from fungi or *A. thaliana* where well-known cytosolic proteins such as glycolytic enzymes were reported as CWPs. The absence of such proteins and the low level of contamination of our CWP preparations do not lend support to this assumption.

Following mannitol washing, several salts and one chelating agent were assayed for their ability to release CWPs from apoplastic fluids. A solution of 0.3 M mannitol solubilized a few CWPs expected to be located only in intercellular spaces. Most identified proteins were acidic, suggesting no ionic interactions with negatively charged cell wall components. CDTA, a chelating agent solubilized Ca-pectin complexes. It released a small number of proteins having domains of interaction with polysaccharides, notably proteins showing homology to the curculin-like or legume lectin families. This suggests an interaction of these proteins with polysaccharides associated to pectins. By comparison, many more proteins were released from rosette cell walls by the salt solutions, particularly by $CaCl_2$. It has already been used for the elution of cell wall proteins, such as peroxidases and extensins, from intact cells in suspension cultures [34]. The effect of $CaCl_2$ appeared to rely not only on ion exchange since no relationship could be observed between ionic strength and ability to extract CWPs. The ability of calcium to strongly chelate acidic and neutral carbohydrates [35, 36] might explain, through a competition mechanism, that proteins or glycoproteins weakly bound to cell wall polysaccharides could be selectively solubilized by $CaCl_2$. Calcium chloride was very efficient since it allowed to recover 60 % of the total number of CWPs identified in this study, most of them being specifically eluted by this salt. This is particularly well exemplified by the large number of glycoside hydrolases (12) and of proteases (7) identified. It is interesting to note that no cytoplasmic proteins were recovered in $CaCl_2$-released apoplastic fluids. This situation is quite opposite to the severe damages caused by this salt to the plasma membrane of cells in suspension cultures which led to the release of a large number of cytoplasmic proteins [29]. It appears that cell surfaces may highly vary depending on the surrounding medium.

The fact that cell type and environment highly influenced the nature and number of released proteins is further illustrated in Table 4, which compares CWPs identified from rosettes (this study) to those from cell suspension cultures ([29], Borderies et al., unpublished results). Rosettes are composed of several types of differentiated cells. Cell suspension cultures are mainly composed of microcalli with undifferentiated cells dividing and enlarging. For each

material, proteins were assigned to the functional classes described above, each class being sub-divided into specific biological activities. Overall, it appears that only 17 out of 132 proteins identified from rosettes and cells were common to both materials. One could expect to find, *a priori*, identical proteins both in the 0.3 M mannitol apoplastic fluids and in the culture medium of cell suspension cultures. In fact, only one extracellular protein with homology to a berberine-bridge enzyme (S)-reticulin:oxygen oxidoreductase (At2g34790) was found to be common.

Through bioinformatic analysis of identified proteins, functions or putative functions could be deduced for most of them. Apart from a few common proteins, striking differences were found in the pattern of released proteins. A high proportion of hydrolytic enzymes and only a few oxidases characterized the CWPs from rosettes. In contrast, the reverse was observed with CWPs from cells, i.e. a high proportion of oxidases and a few hydrolases. Miscellaneous proteins and proteins of unknown functions appeared to be very specific to each material. Within this context, several protein families deserve particular comments.

Numerous glycoside hydrolases could be found in rosette extracts. The role of such proteins in cell walls might be to participate in rearrangements of polysaccharides during development [37-40]. It points out the importance of polysaccharide metabolism within the extracellular matrix allowing flexibility in cell wall structure even in mature cell walls like in rosette leaves.

Another surprising finding was the identification of numerous proteins showing homology to proteinases or to proteinase inhibitors in the rosette cell wall proteome. Despite the presence of cell wall proteinase inhibitors and the use of commercial proteinase inhibitors for protein extraction, it cannot be excluded that proteolysis occurs during the isolation procedure, especially in the $CaCl_2$ extract. Table 1 shows that many proteins were represented by their intact polypeptide plus a variable number of proteolytic products, suggesting an important protease activity. In particular, several proteins having a transmembrane domain could be identified through the partial proteolysis of their extracellular domain. Besides proteins of unknown functions, we found a protein homolog to the CLAVATA1 (At5g05160). CLAVATA1 is a putative receptor kinase presumably involved in a cell-signaling pathway that controls the size of shoot and floral meristems in *A. thaliana* [41]. The IRX1 cellulose synthase catalytic subunit (At4g18780) involved in secondary cell wall synthesis [42] was also identified in this study. However, it is likely that protein degradation by proteinases occurs in the extracellular matrix. The biological significance of protein cleavage by proteinases in the cell wall and its putative control by proteinase inhibitors is a field of investigation that remains largely unexplored [43]. However, the recent finding that an apoplastic subtilisin-like serine-protease regulates stomatal density has given a new insight into the significance of proteolysis in plant development [44].

All proteins except one having interaction domains with polysaccharides *via* lectin domains or with other proteins *via* LRR domains were either common to rosettes and cell suspension cultures or specific to rosettes. They represent 20 and 14 % of the total number of identified proteins in rosettes and in cell suspension cultures, respectively. The three lectins of the curculin-type presented homologies with the carrot-secreted EP1, a glycoprotein showing homology with the *Brassica* S-locus glycoproteins involved in self-incompatibility [45]. Four lectins were of the legume type. Seven proteins had LRR domains supposed to be involved in protein-protein interactions. Among these are PGIP2 and a protein homolog to PGIP1 that are inhibitors of polygalacturonase activity presently known as defense proteins [46]. Two

proteins were homolog to the carrot extracellular dermal glycoprotein (EDGP) [47]. Such a protein has recently been described in tomato as a xyloglucan-specific endoglucanase inhibitor [48]. Altogether, the high number of proteins having interaction domains highlights the multiplicity and likely complexity of perception and/or recognition phenomena taking place at the cell wall level. However, it is expected that many other interacting proteins require to be more strongly associated to the cell wall matrix to be functional [17, 49].

Defense-related proteins are more abundant and more diverse in the cell wall proteome of rosettes than in that of cell suspension cultures. Among the proteins exclusively characterized in the rosette cell wall are the antifungal PR1, a chitinase, a PR5 thaumatin-like, some proteinase inhibitors, and several members of the germin family. Germins and germin-like proteins constitute a large and diverse family of ubiquitous plant proteins that are likely to be associated with plant development and defense [43, 50]. The precise activities and functions of germins remain to be determined. Another defense-related protein exclusively identified in rosettes is a pectin methylesterase inhibitor (PMEI). PMEIs share homology with inhibitors of apoplastic invertases and are known to inactivate pectin methyl esterases as well as apoplastic invertase [51].

A striking difference between the rosette and the cell suspension proteomes was the failure to identify peroxidases in the rosette proteome. This is in apparent contradiction with previous data [52, 53]. Peroxidases have been successfully solubilized from segments of hypocotyls or leaf disks by the infiltration-centrifugation technique in conditions similar to those used in this study. However, these works were not performed on whole plants. Different explanations can be proposed. Rosette peroxidases may be tightly bound to the cell wall. It has been shown that an apoplastic peroxidase has a Ca^{2+}-pectate binding site [54]. Alternatively, peroxidases may be localized in a subset of rosette cells and thus not present in sufficient amount to be identified in our study. Such a tissue- or a cell-specific pattern of expression has been shown for several peroxidases [53, 55].

Most of the known cell wall proteins are encoded by large multigenic families with 20 members as average, and 70 for peroxidases. This redundancy is surprising, and this work is an example of the importance of proteomic studies of different organs, different developmental stages, or different stresses to indicate that specific genes are expressed in each situation. The same protein families may be found in both proteomes, but they can be represented by different members, which would indicate a high level of specificity in gene expression depending on cell type. This is the case for expansin-like proteins: At1g20190 and At2g18660 were found in rosettes whereas At3g45970 and At2g39700 were identified in cell suspension cultures.

This study has led to the identification for the first time of a large set of CWPs from rosette leaves. Among the CWPs identified, at least one class of proteins was unexpectedly absent, namely peroxidases. It suggests either that the vacuum-infiltration protocol with salts has a limited efficiency or that these proteins are more tightly bound to the cell wall than expected. To solubilize these proteins, the resort to harsher extraction procedures is now necessary.

We are grateful to Dr Zoran Minic (INRA, Versailles, France) for communication of unpublished results and identification of A. thaliana xylosidases. This work was supported by Génoplante (Contract N° NO 2001027), the Centre National de la Recherche Scientifique, and the Université Paul Sabatier, France.

5 References

[1] Carpita, N., Gibeaut, D., *Plant J.* 1993, *3,* 1-30.
[2] Roberts, K., *Curr. Opin. Cell Biol.* 1994, *6,* 688-694.
[3] Willats, W. G., Knox, J. P., *Plant J.* 1996, *9,* 919-925.
[4] Rose, J. K., Hadfield, K. A., Labavitch, J. M., Bennett, A. B., *Plant Physiol.* 1998, *117,* 345-361.
[5] Atkinson, R. G., Schroder, R., Hallett, I. C., Cohen, D., MacRae, E. A., *Plant Physiol.* 2002, *129,* 122-133.
[6] Hall, Q., Cannon, M. C., *Plant Cell* 2002, *14,* 1161-1172.
[7] Cote, F., Hahn, M., *Plant Mol. Biol.* 1994, *26,* 1379-1411.
[8] Esquerré-Tugayé, M. T., Boudart, G., Dumas, B., *Plant Physiol. Biochem.* 2000, *38,* 157-163.
[9] Zablackis, E., Huang, J., Muller, B., Darvill, A., Albersheim, P., *Plant Physiol.* 1995, *107,* 1129-1138.
[10] Lamport, D., Northcote, D., *Nature* 1960, *188,* 665-666.
[11] Arabidopsis Genome Initiative, *Nature* 2000, *408,* 796-815.
[12] Fry, S., Smith, R., Renwick, K., Martin, D., Hodge, S., Matthews, K., *Biochem. J.* 1992, *282,* 821-828.
[13] Cassab, G. I., *Annu. Rev. Plant Physiol. Plant Mo.l Biol.* 1998, *49,* 281-309.
[14] Cho, H., Cosgrove, D., *Proc. Nat. Acad. Sci. USA* 2000, *97,* 9783-9788.
[15] Majewska-Sawka, A., Nothnagel, E., *Plant Physiol.* 2000, *122,* 3-10.
[16] Powell, A., van Kan, J., ten Have, A., Visser, J., Greve, L., Bennett, A., Labavitch, J., *Mol. Plant Microbe Interact.* 2000, *13,* 942-950.
[17] Wagner, T., Kohorn, B., *Plant Cell* 2001, *13,* 303-318.
[18] Richmond, T., Somerville, C., *Plant Mol. Biol.* 2001, *47,* 131-143.
[19] Sarria, R., Wagner, T., O'Neill, M., Faik, A., Wilkerson, C., Keegstra, K., Raikhel, N., *Plant Physiol.* 2001, *127,* 1595-1606.
[20] Yokoyama, R., Nishitani, K., *Plant Cell Physio.l* 2001, *42,* 1025-1033.
[21] Cosgrove, D., Li, L., Cho, H., Hoffmann-Benning, S., Moore, R., Blecker, D., *Plant Cell Physiol.* 2002, *43,* 1436-1444.
[22] 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.
[23] Schultz, C., Johnson, K., Currie, G.,A, B., *Plant Cell* 2000, *12,* 1751-1767.
[24] Borner, G., Sherrier, D., Stevens, T., Arkin, I., Dupree, P., *Plant Physiol.* 2002, *129,* 486-499.
[25] Baumberger, N., Doesseger, B., Guyot, R., Diet, A., Parsons, R. L., Clark, M. A., Simmons, M. P., Bedinger, P., Goff, S. A., Ringli, C., Keller, B., *Plant Physiol.* 2003, *131,* 1313-1326.
[26] Pardo, M., Ward, M., Bains, S., Molina, M., Blackstock, W., Gil, C., Nombela, C., *Electrophoresis* 2000, *21,* 3396-3410.
[27] Pitarch, A., Sanchez, M., Nombela, C., Gil, C., *Mol. Cell Proteomics.* 2002, *1,* 967-982.
[28] Chivasa, S., Ndimba, B., Simon, W., Robertson, D., Yu, X.-L., Knox, J., Bolwell, P., Slabas, A., *Electrophoresis* 2002, *23,* 1754-1765.
[29] Borderies, G., Jamet, E., Lafitte, C., Rossignol, M., Jauneau, A., Boudart, G., Monsarrat, B., Esquerré-Tugayé, M., Boudet, A., Pont-Lezica, R., *Electrophoresis* 2003, *24,* 3421-3432.
[30] Shevchenko, A., Wilm, M., Vorm, O., Mann, M., *Anal. Chem.* 1996, *68,* 850-858.
[31] Pinedo, M., Segarra, C., Conde, R., *Physiologia Plantarum* 1993, *88,* 287-293.

[32] Robertson, D., Mitchell, G., Gilroy, J., Gerrish, C., Bolwell, G., Slabas, A., *J. Biol. Chem.* 1997, *272,* 15841-15848.
[33] Di Matteo, A., Federici, L., Mattei, B., Salvi, G., Johnson, K. A., Savino, C., De Lorenzo, G., Tsernoglou, D., Cervone, F., *Proc. Natl. Acad. Sci. U S A* 2003, *100,* 10124-10128.
[34] 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.
[35] Angyal, S., *Adv. Carbohyd.r Chem. Biochem.* 1989, *47,* 1-43.
[36] van Buren, J., in: Walter, R. H. (Ed.), *The Chemistry and Technology of Pectin*, Academic Press, New York 1991, pp. 1-22.
[37] Cassab, G., Varner, J. E., *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 1988, *39,* 321-353.
[38] Freshour, G., Clay, R. P., Fuller, M. S., Albersheim, P., Darvill, A. G., Hahn, M. G., *Plant Physiol.* 1996, *110,* 1413-1429.
[39] Nicol, F., Hofte, H., *Curr. Opin. Plant Biol.* 1998, *1,* 12-17.
[40] Rose, J. K., Bennett, A. B., *Trends Plant Sci.* 1999, *4,* 176-183.
[41] Clark, S. E., Williams, R. W., Meyerowitz, E. M., *Cell* 1997, *89,* 575-585.
[42] Scheible, W. R., Eshed, R., Richmond, T., Delmer, D., Somerville, C., *Proc. Natl. Acad. Sci. U S A* 2001, *98,* 10079-10084.
[43] Segarra, C. I., Casalongue, C. A., Pinedo, M. L., Ronchi, V. P., Conde, R. D., *J. Exp. Bot.* 2003, *54,* 1335-1341.
[44] Von Groll, U., Berger, D., Altmann, T., *Plant Cell* 2002, *14,* 1527-1539.
[45] van Engelen, F. A., Hartog, M. V., Thomas, T. L., Taylor, B., Sturm, A., van Kammen, A., de Vries, S. C., *Plant J.* 1993, *4,* 855-862.
[46] De Lorenzo, G., Ferrari, S., *Curr. Opin. Plant Biol.* 2002, *5,* 295-299.
[47] Satoh, S., Sturm, A., Fujii, T., Chrispeels, M. J., *Planta* 1992, *188,* 432-438.
[48] Qin, Q., Bergmann, C. W., Rose, J. K., Saladie, M., Kolli, V. S., Albersheim, P., Darvill, A. G., York, W. S., *Plant J.* 2003, *34,* 327-338.
[49] He, Z. H., Fujiki, M., Kohorn, B. D., *J. Biol. Chem.* 1996, *271,* 19789-19793.
[50] Bernier, F., Berna, A., *Plant Physiol. Biochem.* 2001, *39,* 545-554.
[51] Giovane, A., Balestrieri, C., Quagliuolo, L., Castaldo, D., Servillo, L., *Eur. J. Biochem.* 1995, *233,* 926-929.
[52] Bernal, M., Pedreno, M., Calderon, A., Munoz, R., Ros Barcelo, A., Merino De Caceres, F., *Ann. Bot.* 1993, *72,* 415-421.
[53] Carpin, S., Crevecoeur, M., Greppin, H., Penel, C., *Plant Physiol.* 1999, *120,* 799-810.
[54] Carpin, S., Crevecoeur, M., de Meyer, M., Simon, P., Greppin, H., Penel, C., *Plant Cell* 2001, *13,* 511-520.
[55] de Marco, A., Guzzardi, P., Jamet, E., *Plant Physiol.* 1999, *120,* 371-382.