Plant Physiol. Biochem., 2000, **38** (1/2), 97−108 / © 2000 Éditions scientifiques et médicales Elsevier SAS. All rights reserved
S0981942800001650/REV

Plant cell wall glycoproteins and their genes

Matilde Josè-Estanyol, Pere Puigdomènech*

Departament de Genètica Molecular, Institut de Biologia Molecular de Barcelona, Centre d'Investigació i Desenvolupament, CSIC, Jordi Girona Salgado 18-26, 08034 Barcelona, Spain

* Author to whom correspondence should be addressed (fax +34 93 204 59 04; e-mail pprgmp@cid.csic.es)

(Received July 29, 1999; accepted October 12, 1999)

Abstract — At least two main groups of glycoproteins can be distinguished in plant cell walls: extensins which are insoluble cell wall proteins; and soluble arabinogalactan proteins (AGPs) which have a high carbohydrate content such that protein content constitutes in some cases only 5 % of the glycoprotein weight. These two groups of proteins together with other cell wall proteins more or less glycosylated, such as proline-rich proteins (PRPs), hybrid PRP (HyPRPs) and expansins, are reviewed and compared with similar proteins present in other cell compartments. Different patterns of N- or O-glycosylation are analysed. In some cases, these cell wall proteins or proteins related to them present patterns of glycosylation that act as epitopes recognizable by IgE in allergic responses. © 2000 Éditions scientifiques et médicales Elsevier SAS

Allergens / AGPs (arabinogalactan proteins) / cell wall / expansins / extensins / glycoproteins / PRPs (proline-rich proteins)

AGP, arabinogalactan proteins / CELP, Cys-rich extensin-like protein / HRGP, hydroxyproline-rich glycoproteins / HSP, hydrophobic soybean protein / HyPRP, hybrid proline-rich proteins / PRP, proline-rich protein / TTS, transmitting tissue specific

1. INTRODUCTION

When cell wall proteins are analysed, many examples of complex glycoproteins are found. In fact, proteins form a low proportion of cell walls that are mostly formed by carbohydrates. How these different components interact giving rise to the dynamic structure of the plant cell wall is an important question in plant biology. To define the specific modifications present in glycoproteins is a complex and often tedious work as it is necessary to purify the protein, to obtain different peptides and then characterize the carbohydrates attached to them. By using recombinant DNA techniques, it has been possible to identify cDNA or genomic clones coding for the polypeptides of plant cell wall glycoproteins although this is only a part of the information since it gives no data on the sugar moiety of the protein. However, molecular data are very valuable not only in terms of sequence information but also in terms of the expression of the corresponding genes. Indirectly, they allow the location of the protein by immunological means.

Plant cell wall proteins are mainly constituted by structural and enzymatic proteins. Although both groups of proteins can be glycosylated, we will focus mainly on structural proteins. Two main groups of glycosylated structural proteins have been described in plants, arabinogalactan proteins (AGPs) and extensins. Arabinogalactan proteins (AGPs) may be involved in the control of cell differentiation, development and cell-cell interactions. Carbohydrate epitopes in these proteoglycans correlate with the formation of specific anatomical patterns [65]. Extensins or hydroxyproline-rich glycoproteins (HRGPs) may be detected in most tissues of the plants analysed so far; the accumulation of their mRNA is developmentally regulated. They are supposed to be structural proteins but they also have a role in defence response and oxidative cross-linking mediated by extensin peroxidases [42, 50, 65].

Cell wall glycoproteins can be N- or O-glycosylated. N-glycosylation requires the presence of -N-X-S(T)- as the consensus sequence where X is any amino acid. However, there are examples of these sequences in proteins where no carbohydrates are found. This may be due to the lack of elements of the

enzymatic machinery for carbohydrate assembly in a given cell, or as a consequence of the tertiary folding of the protein that could protect the target sequence for modification [35]. In N-glycosylation, the linkage to the amide group of Asn is a N-glycosylamine bond and it always involves N-acetylglucosamine. In general, the carbohydrate component is assembled as lipid linked intermediates and transferred as a whole to the nascent polypeptide chain [35]. Cell wall glycoproteins such as AGPs and extensins are mainly O-glycosylated, although some cases of N-glycosylation have also been described. O-glycosylation of serine with galactose and hydroxyproline with arabinose have been described for extensins. Serine O-glycosylation needs a Ser-Hyp_4 environment [46]. Previous hydroxylation of proline to hydroxyproline is necessary for its glycosylation. Hydroxylation has been shown to be sequence dependent, recognized by prolyl hydroxylases specific for AGPs and extensins. Sequences such as Lys-Pro, Tyr-Pro and Phe-Pro are not hydroxylated in any known HRGP, while Pro-Val is always hydroxylated. Once Pro is hydroxylated, Hyp contiguity enhances the arabinosylation [46]. In O-glycosylation, all the sugars are added successively by the transfer from their nucleotide derivates and by the activity of specific glycosyltransferases in the Golgi apparatus [35].

Other classes of cell wall proteins, such as PRPs, HyPRPs and expansins that are assumed to be located in the cell wall as deduced from their protein sequence and from the few protein studies available, are characterized by a low level of glycosylation. In this case, proteins with related sequence, sometimes with different functions or with different cellular localization, and locally glycosylated may become important allergens. Examples are the soybean hydrophobic protein [53] and an insect amylase inhibitor [30] which have protein sequences related to those of HyPRPs, and Group I allergen which are related to expansins [18]. Recent advances in the study of these cell wall proteins and related ones will also be summarized.

2. EXTENSINS

Extensins, or HRGPs, are cell wall glycoproteins, whose sequence contains a signal peptide allowing them to be exported to the wall, and a repetitive region rich in proline where the main repeated motif is $SerPro_4$. Genomic or cDNA HRGP sequences have been cloned from both monocotyledonous and dicotyledonous plants, conifers and algae, and they have been the object of a number of reviews [42, 65]. Expression of the genes coding for most of them is developmentally regulated, and induced by wounding, elicitors or fungal attack [42]. They have been related to the cell cycle in tobacco [71]. Induction by nematode infection [70] and mycorrhizae fungi (*Glomus versiforme*) [3] has also recently been reported.

In a monocotyledonous species such as maize, two extensins have been described. Pex 1, containing the $SerPro_4$ element as the main repetitive motif [59], is specifically expressed in maize pollen and it has not been shown to be induced by wounding in any tissue. Maize HRGP [67] is encoded by a gene expressed in meristematic and vascular tissues and is induced by wounding. One of the tissues of maize plants where HRGPs are more abundant is the seed pericarp (*figure 1*). In fact, this was one of the tissues where the protein was first isolated from maize [39]. The protein sequence of the maize HRGP has only one $SerPro_4$ motif and the protein has been shown to be 27 % carbohydrates, occurring exclusively as arabinosylated Hyp, predominantly as the monosaccharide (15 %) and trisacccharide (25 %) forms with 48 % Hyp and all Ser non-glycosylated [46]. Sequences related to maize HRGP have also been described in rice, sorghum and teosinte. In the case of rice, the unique $SerPro_4$ motif is not present [42].

Although many extensin genes have been cloned, purification of their respective protein has not always been performed in tissues where molecular information is available. For this reason, information on post-transcriptional modifications such as hydroxylation, glycosylation or intra- and inter-molecular isodityrosine linkages, as well as interactions with other cell wall components such as pectins, is not always available. Only by purification of extensin, the Pro/Hyp and Ser content of glycosylated amino acids as well as the sugars involved can be determined.

Extensin proline hydroxylation is performed by sequence specific prolyl hydroxylases and their glycosylation occurs as a function of hydroxyproline continuity (Hyp_4). Extensin proteins with $Hyp_{3/2/1}$ have in every case, a lower degree of glycosylation [46]. Glycosylation is also known to stabilize the polyproline II extensin structure [42]. Once in the cell wall, extensin monomers are insolubilized, probably by the ability of tyrosines to form isodityrosine intermolecular bridges [28] with other extensin molecules or with other cell wall components. Extensin-specific peroxidases that could mediate this reaction have been isolated from tomato cell suspension cultures [62], and ascorbate and tyrosine have been shown to inhibit cell

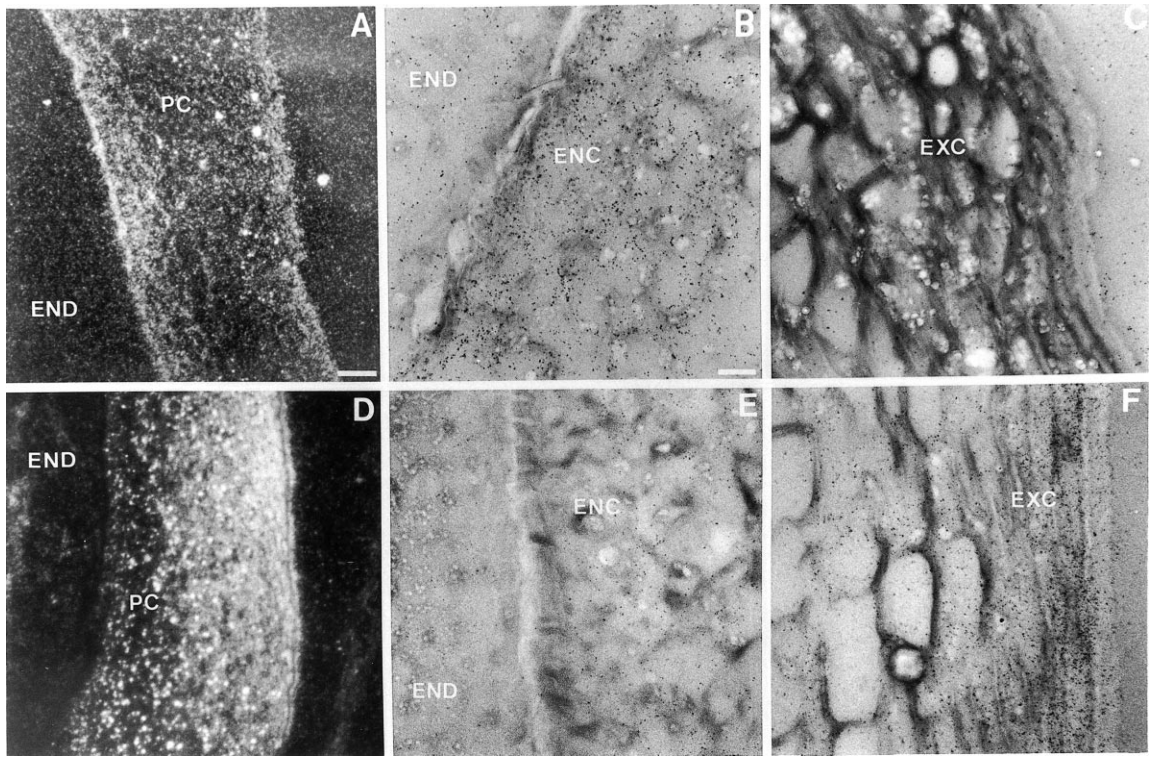

Figure 1. Cellular localization of HyPRP and mRNA in longitudinal sections from 8 DAP (days after anthesis) maize kernel pericarp. **A-C**, In situ hybridization with HyPRP antisense probe; **D-F**, in situ hybridization with HRGP antisense probe. EXC, Exocarp; ENC, endocarp; PC, pericarp; END, endosperm. Sections were examined under light field microscopy. Bars = 250 μM (**A**, **D**); 25 μM (**B**, **C**, **E**, **F**).

wall insolubilization in *Chlamydomonas reinhardtii* [73]. Unfortunately, as yet, only intramolecular isodityrosine cross-links have been isolated from partly hydrolysed cell wall tryptic peptides from tomato cell suspension cultures [46]. Evidence for extensin interactions with other cell wall components such as pectins have been described by Qi et al. [56].

The sequences controlling the expression of extensin genes have been studied in different species such as maize (HRGP), carrot (pdc5A1), tobacco (HRGPnt3) and oilseed rape (extA) [42]. Experiments to determine the regulatory elements responsible for extensin tissue specific expression and wound response have been carried out for the oilseed rape extA promoter [25]. Successive deletions of the extA promoter linked to GUS in tobacco transformed plants shows that there is an activator region between nucleotides –159 and –433, a repressor region between –664 to –789 and a negative regulator region between –433 to –664 that represses expression in non-phloem cells. This repression may be overcome under tensile stress which allows expression in all cell types. Sequences responsible for the wounding response were localized between nucleotides –940 and –3500. Using 0.9 kb of the SbHRGP3 promoter [1], a gene expressed mainly in soybean mature root and the hypocotyl of germinated seedlings induced expression in wounded leaves and stems of tobacco transformed plants in the presence of sucrose was observed. Expression was localized in phloem tissues and cambium cells. In germinated transformed seedlings, expression was observed in the epidermis of mature root and hypocotyls of 3-d-old seedlings and then inactivated. At 8 d, re-activation was observed only in a layer of the epidermal cells in the zone from which a lateral root was to be initiated, and not in the root tip. Wycoff et al. [76] studied the promoter activity of 0.964 kb at the 5'-end of the bean HRGP 4.1 gene fused to GUS in tobacco. Results show that HRGP 4.1 tissue specificity and wound inducibility lies in a region between nucleotides –94 and –251, while the sequence responsible for activation by infection lies outside this region.

3. AGPs

AGPs can be detected in a number of different plant organs by staining with the synthetic phenylglycoside known as the Yariv reagent which binds to and precipitates AGPs [26]. Most of the AGP molecule are carbohydrates (90–95 %), mainly galactose and arabinose residues, although rhamnose, glucuronic acid and other monosaccharide residues are present as minor components in some cases. The sugar groups are O-linked to hydroxyproline, serine and threonine amino acids in the core protein [26, 50, 65]. Two different structural models 'wattle-blossom' [26] and 'rodlike molecule' [55] have been proposed for AGPs. The 'wattle-blossom' model has been proposed for acidic and alanine-rich AGPs and predicts a few large polysaccharide substituents as ovoid arabinogalactans anchored to a polypeptide core. In solution, it is consistent with a spheroidal shape for the molecule [26]. The 'rodlike molecule' model [55] has been proposed for AGPs not especially rich in Ala, with a polypeptide chain more similar to extensins but with polysaccharide sidechains similar to AGPs where numerous small polysaccharide substituents will maximize intramolecular hydrogen bonding forming a twisted hairy rope. More information about the precise location of carbohydrates is necessary to define consensus sequences for glycosylation in these proteins. AGPs have been shown to be soluble proteins. They are highly heterogeneous as deduced from studies with anti-AGP monoclonal antibodies raised against AGP carbohydrate epitopes present in each organ and cell type. Some of them are glycoproteins and appear to be associated with the plasma membrane and intravesicular bodies, whereas others, proteoglycans, are secreted in large amounts both in the whole plant and by cultured cells, and are present in cell walls of vascular bundles, sclerenchyma cells, pollen and pollen tubes [65].

Many different functions have been attributed to AGPs. Their properties of aggregation and stickiness could allow them to contribute to the adhesion between pollen and stigma, or have a nutritive role in the style, providing carbohydrate precursors for the growing pollen tube wall. They may also take part in the control of water balance in the cell. A possible role in the processes of cell differentiation has also been proposed, as the addition of AGPs secreted by an embryogenic carrot cell line to a non-embryogenic line can induce embryogenic potential [65].

Cloned AGPs (*table I*) appear to be multidomain proteins with a signal peptide, a central proline-rich domain with generally Pro-Ala repeats, and a C-terminal domain that may be: (a) hydrophobic and probably absent in the mature protein for classical AGPs, as in the case of *Nicotiana alata* AGPNa1 [24], *Pyrus communis* AGPPc1 [10], and *Lycopersicon esculentum* LeAGP1g [47]; (b) non-hydrophobic as in the case of *N. alata* AGPNa2, which has a hydrophobic proline-rich central domain [49], and *P. communis* AGPPc2 [49]; (c) formed by interspersed hydrophobic and hydrophilic regions containing six Cys in a conserved pattern, as for *N. alata* TTS-1,2 [12] – (NaPRP4) [9]. In the same group, we include NaPRP5, a 120-kD protein with two Pro-rich domains: the first one extensin-like and the second one AGP-like, with a C-terminal domain having the same six Cys conserved pattern and 57 % identity to the NaPRP4 C-terminal domain [48, 63]; (d) a non-hydrophobic domain with eight Cys in a different distribution pattern and without sequence similarity to NaPRP4, as in the case of *N. alata* AGPNa3 [23].

Post-transcriptional processing has been proposed for AGPNa1 and AGPPc1 as the purified native proteins are shorter than the deduced sequence. In these cases, the C-terminal amino acid of the hydrophobic domain has been substituted by ethanolamine and glucosamine. Inositol is also present in the native AGPs [79]. These results have been taken as evidence for C-terminal processing involving glycosylphosphatidyl-inositol (GPI) membrane anchors. The same mechanism has been proposed for other classical AGPs such as LeAGP1g [29, 79].

Transmitting tissue-specific proteins (TTS) [12] were first classified as PRPs (HyPRPs domain A) [43] as their Pro-rich domain contains Lys-Pro-Pro interspersed repeats. They are not particularly rich in Ala, and they have a six Cys conserved pattern in the C-terminal domain with 50 % identity to another HyPRP domain A, bean PvPRP [64]. Since their purification, TTS proteins have been considered AGPs because they strongly react with β-glucosyl Yariv dye even though sugar residues constitute only 35 % of the total TTS protein molecular mass [13]. Protein purification also allowed corroboration of the predicted O-glycosylation of Hyp as well as N-glycosylation of one or both of the sites predicted by sequence data [13, 66].

4. PRPs

PRPs are proline-rich proteins with a signal peptide for processing to the cell wall followed by PPVYK as

Table I. cDNA and genomic clones encoding for arabinogalactan proteins (AGPs).

Name	Nature	Plant	aa	Reference
Classic AGPs				
(SP + proline-rich + C-terminal hydrophobic region)				
AGP*Na*1	complete cDNA	*N. alata* styles	132	[24]
AGP*Pc*1	complete cDNA	*P. communis* suspension cells	145	[10]
LeAGP1g	complete cDNA/genomic	*L. esculentum*	215	[47]
Non-classic AGPs				
SP + proline-rich + C-terminal non-hydrophobic				
AGP*Na*2	complete cDNA	*N. alata* suspension cells	461	[49]
AGP*Pc*2	complete cDNA	*P. communis* suspension cells	294	[49]
SP + proline-rich + C-terminal with six Cys (HyPRP Cys-rich domain A-like)				
TTS-1 (NaPRP4)	complete cDNA	Tobacco floral buds	257	[12]
TTS-2	complete cDNA	*N. alata* style	261	[12]
NaPRP5 (120 kD)	complete cDNA	*N. alata* style	461	[63]
SP + proline-rich + C-terminal with eight Cys				
AGP*Na*3	complete cDNA	*N. alata* pistils	169	[23]

the most frequent repetitive element. These proteins lack the $SerPro_4$ repetitive element defined for extensins [42]. For this reason, they are probably less glycosylated than extensins. One example is 1A10-2, a protein with less than 3 % glycosylation, mainly arabinose [42]. The Tyr content of PRPs make them good candidates for the oxidative stress response to wounding, elicitor treatment and the elicitor hypersensitive response [5, 6]. Rapid insolubilization may strengthen the wall and hinder pathogen invasion. PRP expression is mainly associated with protoxylem and xylem structures in contrast to extensins which are related to the primary cell wall. They can also, in some cases, act as nucleation sites for lignin deposition. PRPs have been described in dicots such as soybean [38], carrot [11] and alfalfa [34], and in monocots such as maize [72] and wheat [57], and related to nodulation in the dicot soybean [27] and *Medicago truncatula* [74]. They have also been described in a fungus (glycoprotein C1H1 of *Colletotrichum*) specific to the biotrophic interface formed in the fungus-bean interaction. It is proposed that C1H1 could prevent the plant from recognizing the fungus and act as a barrier to host defence molecules [54].

Maize PRP is mainly expressed in the mature hypocotyl and root of germinated seeds and in the pericarp of the kernel [72]. Soybean PRP1, one of the best studied PRPs, is expressed in germinated seeds and also in the seed coat but its expression is decreased in soybean lines with anthocyanin mutations [42, 52]. In addition, the purified PRP1 protein from the soybean variety Blackhawk is smaller than in wild type varieties. Cloning of PRP1 from this variety showed that the size reduction was the result of in-frame omissions of 30 bp that correspond to an integral number of repeat units in the protein structure [61]. The SbPRP1 protein tolerates loss of integral repeat units, but evolutionary pressure has avoided deletions within a repeat. Similar results were previously described for HRGP genes cloned from different maize varieties [42]. The regulatory elements and factors responsible for SbPRP1 expression have been studied by stable transformation in tobacco and cowpea. For maximal expression, sequences between nucleotides –1080 and –623 have been shown to be necessary, while the first –262 bp are sufficient for root expression [69].

5. HYBRID PRPs

A subset of PRPs are multidomain proteins and have been named HyPRPs. A signal peptide is followed by a central proline-rich domain and a C-terminal non-repetitive domain, not specifically rich in proline or glycine but with cysteines. The number and specific distribution of cysteines in this C-terminal domain allows classification of these proteins in two different groups [43].

The first of those groups is HyPRP domain A. In this group, the four-six Cys residues are present in a specific pattern (...C..C...C...C...C...C...). Since protein purification, one of its original member TTS protein has been considered as an AGP, as stated previously [13, 43, 66]. Another member of this group of proteins is PvPRP1 [64] whose expression is reduced after fungal infection as result of mRNA destabilization by binding of PvPRP1 mRNA binding protein

```
zmHyPRP   AVRT C PIDTLKLNA   C VDVLSGLIHLV IGQEARS  K CC PLVQGVADLDAAL CLC TTIRARLLNINIYLPIALNLLI T  C GKHAPS GFQ  C PPLYD        301
tPRP-F1   .QP. . ...A...G.   . ....G....IG ..GS.KQ  T .. ..LG.LV.....I ... ....LK......I.....QV..DD  . ..YP.K D.K  . .ST         346
CrHyPRP   KQP. . ...A.....   . ..L.G....IG ..RS.KD  T .. .VLG.L.G...GI ... ...K.K......I.....QV..DD  . .MIP.A ...  . .VDG        329
Msprp2    .QQ. . S..A...G.   . ....G....IG ..GS.KQ  T .. ..L..LV.....I ... ....LK.....LVI.L..QV.. D  . ..TP.E ..K  . .A .        381
Bnprp     KPE. . .......G.   . ....G....IGLG.SS.KK  E .. .VLG.LV.....V ... ...K.K..IVDLII....E... D  . ..TP.P ..K  . .S         376
MsACIC    TSQK . .T.....GV   . A...GLVNVI. GSPA S.  . .. T.I..L......V ... .A.K.NI.G..LNV..T.S..LSA  . E.SI.N ...  . S          166
PVR5      SGSG . .EN...FGV   . A... ...GVE L.KPPKT  P .. N.I..L...E..V ... .AL..NV.G..LNV..K....LNY  . ..KT.K D.V  . Y          127
ADR11-2   .QA. . .......G.   . ..L.G..V.IG L.DPVAN  Q .. .VL..LVEVE..V ... ..LKLK...L...V.L..Q..V .  . .. S.PP.YT  . SL         151
Rcc3      .FGR . .R.A...GV   . AN.. ...KAKVGVPP.E   P .. ..LE.LV..E..V ... .A..GNI.G..LN...D.S.ILNY  . ..TV.T ..K  .            133
Cr14kda   PKA. . .R.A...GV   . A.L.G...SA. ..APPKT  P .. SLIE.L...E..V ... .A.K.NV.G..LNV.VS.T..LNV  . S.KV.E ..I  . A          138
ZRP3      SHGR . ...A...KV   . AK..GLVKVGL PQY.Q      .. ..LE.LV..... ... .A.K.NV.G.HLNV.LS..FILNN  . .RIC.E D.T  . .N         129
SAC51     VTAK . .R.A...GV   . AN.....NIT L.KPPVK  P .. T.IK.L...E..A ... .ALK.NI.G..LNI..S.S..LNV  . S.KV.P ...  .            147
Rcc2      GHGR . ...A...RV   . AN..N.ALGVNVGHGPYD     .. ..LA.L..A...V ... .AVK.NV.GV.LNV.VE.K.ILNK  . ..TC.. D.T  .            146
DC 2.15    SAGK . .R.A...GV   . A...N .V.N.V..  SPPTLP .. S.LE.LVN.E..V ... .A.K.NI.GK.LN.....S.VLNN  . ..QV.N ..E  . T          137
NT16      GQGR . .R.A...GV   . ANL.G..VGVIVG   SPPTLP .. S.IA.L...E..V ... .A...NV.G..LNV.LS.S.VLNN  . .RNP.T ..T  .            170
HSP       TRPS . . .  .SI   . LNI.G.SLGT. DD          .. A.IG.LG.IE.IV ... IQL..  .G. LN.NRN.Q.ILNS  . .RSY.. NAT  . .RT         80
CELP5     PSSS . SASDEAKIYK  . MFNETK .    DP          .. . .FKSI .GTS  .P. YKY .EN.GNQE . ...QAY     . DID  . P    . KG.QVIKLSREEE160
```

Figure 2. Comparison of amino acid sequence of zmHyPRP C-terminal hydrophobic domain with thoses of related polypeptides. Only different amino acids are shown. Gaps have been introduced to maximize identity. Possible glycosylation consensus sites are in italics and underlined. The polypeptides are reported in *table II*. Proteins are ordered from more to less % identity.

(PRP-BP) activated protein to a U-rich sequence present in the 3' non-coding region of PvPRP1 mRNA [80].

The second group corresponds to HyPRP domain B which has eight cysteines in a specific pattern in the C-terminal domain (...C...C...CC...CXC...C...C...) associated with different proline-rich repeats and to a signal peptide [43]. They have been cloned from different mono- and dicotyledonous plants (*table II*). Homology in the C-terminal domain is found mainly around the six first cysteines (*figure 2*). In addition to the eight Cys in the C-terminal domain, all HyPRPs have two more Cys in the N-terminal region following the signal peptide processing site. The Pro-rich domain has different lengths, from 292 amino acids for MsPRP2 [22] to nineteen residues with only four Pro for RCc3 [77]. The most similar protein to the maize embryo specific zmHyPRP [45] is the TPRP-F1 protein, mainly expressed in young tomato fruits (67.85 % identity to the maize sequence) [60]; other examples are CrHyPRP expressed in the subapical region of *Cuscuta reflexa* vine (66.66 %) [68], MsPRP2, induced in alfalfa salt tolerant plants (64.7 %) [22], BNPRP, oilseed rape cold induced (63.9 %) [32] and alfalfa MsACIC, induced in cold tolerant plants (60 %) [8]. The most distantly related sequences are the soybean hydrophobic protein (HSP) (40 %) [45, 53] and tobacco CELPs (20 % or less) [75] (*table II*). The same C-terminal domain has also been reported to be fused to a central glycine-rich domain in tobacco NT16 [78].

The expression of HyPRP cloned genes is different for each gene but is always very specific to a defined plant development stage or is the result of hormonal or stress induction or repression (*table II*). Two HyPRPs, MsACIC [8] and NT16 [78] have been shown to be induced by wounding of alfalfa crowns and tobacco leaves respectively. Expression at the cellular level has been characterized by in-situ hybridization for some of them. For example, zmHyPRP is expressed in the scutellum and in the axis cortical cells of developing maize embryos and in the inner pericarp of maize developing kernels (*figure 1*) [44, 45], and PVR5 [16] and ZRP3 [41] in root cortex cells, 2 mm behind the root tip. CELPs are expressed in tobacco vascular bundles and epidermis of petals and sepals, in vascular bundles of pistils, in a narrow region between the transmitting tissue and the cortex of the style, and in a narrow row of cells lining the placenta in the ovary [75]. The regulatory elements responsible for HyPRPs expression have been studied by stable expression of GUS fused to DC 2.15 [37] and RCc2 [77] promoters in transformed carrot and rice plants respectively. DC 2.15 gene expression is specific to carrot somatic embryogenesis [2] and RCc2 is specifically expressed in maize roots [77]. Results indicate that DC 2.15 promoter, containing a GATA element, was able to direct the expression of GUS mainly at the torpedo stage of carrot somatic embryogenesis, while 1 656 bp of the RCc2 promoter were not enough to allow root specific expression.

None of these hybrid PRP protein domain B has been purified for the moment with the exception of HSP [53] and their glycosylation and hydroxylation patterns can only be theoretically predicted. One N-glycosylation consensus sequence is present in the C-terminal domain of the soybean hydrophobic protein (HSP) [53] and *Brassica napus* SAC51 [19] (*figure 2*). The general absence of $SerPro_4$ repeats make them poor candidates for O-glycosylation. The formation of isodityrosine linkages in response to oxidative stress could be predicted for zmHyPRP, as Tyr is uniformly expressed in its repeat (PPYV) [45]. This suggests an inherent protective or defence function as

Table II. Hybrid PRPs with Cys-rich C-terminal domain. A, Cys-rich domain A (...C..C...C...C...C...C...); % identity C-terminal domain A is in relation to PvPRP1; B, Cys-rich domain B (...C...C...CC...CXC...C...C...); % identity C-terminal domain B is in relation to zmHyPRP. nd, Not determined.

Name	Cys	Nature	Plant	aa	Expression	% Identity C-terminal	Repeats	Reference
Hybrid PRPs Cys-rich domain A								
PvPRP1	4	complete cDNA	elicitor-treated bean cells	297	bean cells and tissues	100	PVHPPVKPP & related	[64]
NaClassIII	6	complete cDNA	tobacco stigma, style	426	style transmitting tissue	42.8	SPPPP	[21]
Hybrid PRPs Cys-rich domain B								
MSACIC	8	complete cDNA	alfalfa	166	cold induced	59.75	SP_x;TP_x	[8]
MsPRP2	8	genomic	alfalfa	381	salt inducible	64.70	PPXVPZ; PPVVPVT; PPVH; PPXVZ	[22]
DC 2.15	8	complete cDNA	carrot cell culture growth in 2,4-D free medium	137	initial somatic embryogenesis	47.94	PX	[2]
CR14KDA	8	complete cDNA	*Catharanthus roseus*	138	cell culture induced by sucrose in auxin-free medium	51.80	PKP	[40]
CrHyPRP	8	complete cDNA	*Cuscuta reflexa* vine with haustorial monds	329	plant subapical region	66.66	PPXYP; PPHYP; PPXYK	[68]
SAC51	8	complete cDNA	oilseed rape dehiscence zone	147	dehiscence zone pod	51.22	$(PX)_n$	[19]
BNPRP	8	genomic & cDNA	oilseed rape	376	cold induced in leave	63.09	KPPK; KPPTVKPPP; STPKPPT; PPVVTPPT	[32]
ADR11-2	8	uncomplete cDNA	soybean mature hypocotyl	151	elongating hypocotyl	57.31	PXX; PPXXX	[20]
PvR5	8	complete cDNA	bean roots	127	roots	54.93	P_nX; P_nK	[16]
CELP1	8	complete cDNA	tobacco flower	209	tobacco flower	–	$XP_{(3-7)}$; SPPPP	[75]
CELP2	8	complete cDNA	tobacco flower	196	tobacco flower	21.53	$XP_{(3-7)}$; SPPPP	[75]
CELP3	8	complete cDNA	tobacco flower	166	tobacco flower	14.063	$XP_{(3-7)}$; SPPPP	[75]
CELP4	8	uncomplete cDNA	tobacco flower	159	tobacco flower	22.38	$XP_{(3-7)}$; SPPPP	[75]
CELP5	8	complete cDNA	tobacco flower	160	tobacco flower	18.90	$XP_{(3-7)}$; SPPPP	[75]
NT16	8	complete cDNA	gall tumour tobacco	170	gall tumour, root, wounded leave	45.88	GGGSGN	[78]
TPRP-F1	8	genomic	tomato	346	young tomato fruit	67.85	PPIV; PPST; PTPP	[60]
HSP	8	protein	soybean	80	soybean seeds	40.0	nd	[53]
zmHyPRP	8	genomic	maize	301	immature embryo and ovary	100.00	PPTPRPS; PPYV	[45]
ZRP3	9	complete cDNA	maize roots	129	maize roots	51.25	PVVPTP	[41]
RCc2	9	genomic & cDNA	rice roots	146	rice roots	48.14	PXVV; PX	[77]
RCc3	8	complete cDNA	rice roots	133	rice roots	53.08	PTPSTPTP; PX	[77]

an inhibitor of external attacks for this particular protein. In this sense, zmHyPRP is present in the first cell layer of the endocarp during kernel development (*figure 1*). Internal Cys bridges are also proposed for the C-terminal hydrophobic region as confirmed for the soybean hydrophobic protein crystal structure [43].

The function of HyPRPs is not clear at the moment, although different theories have been put forward. The C-terminal domain may be processed as reported for classical AGPs [79] and thionins [4]. Another possibility is that the hydrophobic C-terminal domain allows the protein to anchor in the cell membrane, orientated towards the cell wall. Hydrophobic-cluster analysis [36] has established similarities between the C-terminal domain of HSP, LTPs and a group of proteins with sequence similarity, including prolamin

storage proteins such as β- and γ-zeins of maize and 2S storage proteins, and soluble α-amylase/trypsin inhibitor proteins of cereal endosperm. This homology together with the similarity of HSP with HyPRP allows to connect all these proteins. Altogether, they have the same domain with eight Cys in the same specific pattern, with variable and conserved subdomains located along the sequence at precise positions. The conserved subdomains probably reflect a common ancestor where limited evolutionary differentiation has allowed functional divergence for each group of proteins maintaining the basic structure of the protein [36].

6. EXPANSINS, ALLERGENS AND GLYCOSYLATION

Expansins are cell wall proteins able to induce acid growth. Acid-extension appears to depend only on expansin at acidic pH. Some evidence indicates that expansin causes wall creep by loosening non-covalent association between wall polysaccharides. This loosening effect has also been found in pure cellulose paper in the absence of hydrolysis of cellulose. Expansins act in a different way to the degradative process carried out by cellulases. Cellulose-hemicellulose complexes appear to be the expansin substrate, where it can disrupt polysaccharide associations within the wall by disrupting hydrogen bonding between cellulose fibres. An alternative or accompanying mechanism could be a glycosyl hydrolase activity. This activity has not been detected in expansins but sequence similarities with the catalytic centre of fungal glycosyl hydrolases have been observed [17].

Expansins have been cloned and purified from cucumber [17], rice [14] *Arabidopsis thaliana* seedlings [17], pea petals [17], ripening tomato fruit [17], tomato apex meristems [58] and pine hypocotyls [17].

The expansin coding region has a highly conserved sequence. Expansins have a hydrophobic signal peptide, a central region containing eight conserved cysteines that could form intra-molecular bridges and where similarity to the glycosyl hydrolase catalytic centre is observed (HFD domain); a basic region and a C-terminus containing four conserved tryptophans with spacing similar to that found in some cellulose binding domains. Expansin sequence divergence appears at the 3' and 5' untranslated regions [17]. Expansin sequence divergence at the promoter level may allow amplification and development of a more complex control of expansin gene expression, but evolution has worked against changes in the amino acid sequence of the coding region as this sequence may be important to conserve expansin function. The purified protein does not appear to be very abundant even in rapidly growing tissues, is not soluble in solutions of low ionic strength, is not highly glycosylated and is tightly bound to cell walls [17].

In tomato, expansin genes are expressed in expanding and meristematic tissues [58], in green expanding and maturing fruit pericarp [7], and in the pericarp of ripening fruit under ethylene control [17] where it is proposed to contribute to the fruit softening process by disrupting polysaccharide associations. Expression at the cellular level has been defined by in situ hybridization in rice and tomato. Rice Os-EXP1, Os-EXP2, Os-EXP3 and Os-EXP4 have been shown to be expressed in the epidermis and around vascular cells in the internode and the root tip. They are also expressed in adventitious and lateral root [15]. Expression in leaf primordia has been defined for tomato LeExp18. Its expression predicts the site of primordium formation before histological changes can be detected [58].

Expansins share 20–25 % similarity with Group I allergens, the major allergenic component of grass pollen that produce hay fever, seasonal asthma, and related immune reactions in humans [17, 18]. To differentiate between them, expansins expressed in vegetative tissues have been named α-expansins and the pollen allergens, β-expansins. Homology is observed in the cysteine-rich region, the HFD domain (Hys-Phe-Asp) and the tryptophan C-terminal domain, probably responsible for functional activity of expansins. Group I allergens can be distinguished from expansins by their solubility, glycosylation and are highly abundant pollen proteins that do not bind tightly to the pollen cell wall [18]. The allergen has also expansin activity in cell wall loosening of stigma and style. The functions of expansins and Group I allergens are similar but the cell wall substrates may be different. As the allergic response is the result of protein carbohydrate containing IgE binding epitopes, they provide an example of how glycosylation can differentiate two groups of similar proteins making the glycosylated form sensitive to the immunological system.

Another protein which is able to provoke an allergic response when glycosylated and is reported here is the soybean hydrophobic seed protein responsible for soybean hull allergy that can produce death by asphyxia [31, 53]. Other examples include the barley glycosylated monomeric inhibitor of insect-amylase responsible for baker's asthma disease [51], such as

HyPRP domain B [43], and ParhI, an extensin-like soluble protein purified from *Partherium hysterophorus* pollen responsible for rhinitis allergy in India, the USA and Australia [33]. In each case, specific patterns of glycosylation make them different from their respective counterparts, becoming strong allergens. It is interesting to note that most of the glycoproteins described are abundant in two typical allergenic tissues, pollen and pericarp. Their study is therefore interesting not only for the function of the proteins themselves in plants but also for the possibility they open to reducing allergy problems produced in increasing numbers by cultivated crops.

7. CONCLUSION

Cell wall proteins have different levels of glycosylation. Cell wall glycoproteins, such as extensins and arabinogalactan proteins, are the main glycoprotein components of the cell wall while PRPs with different degrees of glycosylation have also been described. Other components such as expansins are not glycosylated. Glycosylation may be important to allow interaction of proteins with other cell wall components, to define cellular specific epitopes and to develop the allergen response.

Some of these proteins, such as AGPs or HyPRPs, are multidomain proteins resulting from the association of modular functional cassettes. Usually they are composed of a signal hydrophobic peptide to allow cell wall processing; a proline-rich, usually hydrophilic, domain involved in the structural role of the protein or in cell wall deposition; and a C-terminal domain more or less hydrophobic, with or without cysteines that could be involved in the specific function of the protein or in its membrane anchorage and that can be processed. Particularly interesting is the C-terminal domain present in the HyPRP domain B. This domain is also present in proteins including a central proline-rich domain, but localized in different cellular compartments such as vacuoles or protein bodies. Evolution of this large family of proteins has conserved the eight Cys pattern but a high degree of flexibility in the evolution of other amino acid residues has allowed their different functions and cellular compartment signalling. Another example is offered by expansins and Group I allergen proteins. In this case, the cysteine pattern and also the elements needed to allow cell wall expansion have been conserved but the final protein sequences differ in their glycosylation level and degree of solubility. Future studies will help to better understand the relationship between the different structure of the proteins described here and their specific function in the cell walls where they are inserted.

REFERENCES

[1] Ahn J.H., Choi Y., Kwon Y.M., Kim S.G., Choi Y.D., Lee J.S., A novel extensin gene encoding a hydroxyproline-rich glycoprotein requires sucrose for its wound-inducible expression in transgenic plants, Plant Cell 8 (1996) 1477–1490.

[2] Aleith F., Richter G., Gene expression during induction of somatic embryogenesis in carrot cell suspensions, Planta 183 (1990) 17–24.

[3] Balestrini R., Josè-Estanyol M., Puigdomènech P., Bonfante P., Hydroxyproline-rich glycoprotein mRNA accumulation in maize root cells colonized by an arbuscular mycorrhizal fungus as revealed by in situ hybridization, Protoplasma 198 (1997) 36–42.

[4] Bohlmann H., Apel K., Thionins, Annu. Rev. Plant Physiol. Plant Mol. Biol. 42 (1991) 227–240.

[5] Bradley D.J., Kjellbom P., Lamb C.J., Elicitor- and wound- induced oxidative cross-linking of a proline-rich plant cell wall protein: a novel, rapid defense response, Cell 70 (1992) 21–30.

[6] Brisson L.F., Tenhaken R., Lamb C., Function of oxidative cross-linking of cell wall structural proteins in plant disease resistance, Plant Cell 6 (1994) 1703–1712.

[7] Brummell D.A., Harpster M.H., Dunsmuir P., Differential expression of expansin gene family members during growth and ripening of tomato fruit, Plant Mol. Biol. 39 (1999) 161–169.

[8] Castonguay Y., Laberge S., Nadeau P., Vézina L.P., A cold-induced gene from *Medicago sativa* encodes a bimodular protein similar to developmentally regulated proteins, Plant Mol. Biol. 24 (1994) 799–804.

[9] Chen C.G., Mau S.L., Clarke A.E., Nucleotide sequence and style-specific expression of a novel proline-rich protein gene from *Nicotiana alata*, Plant Mol. Biol. 21 (1993) 391–395.

[10] Chen C.G., Pu Z.Y., Moritz R.L., Simpson R.J., Bacic A., Clarke A.E., Mau S.L., Molecular cloning of a gene encoding an arabinogalactan-protein from pear (*Pyrus communis*) cell suspension culture, Proc. Natl. Acad. Sci. USA 91 (1994) 10305–10309.

[11] Chen J., Varner J.E., Isolation and characterization of cDNA clones for carrot extensin and a proline-rich 33-kDa protein, Proc. Natl. Acad. Sci. USA 82 (1985) 4399–4403.

[12] Cheung A.Y., May B., Kawata E.E., Gu Q., Wu H.M., Characterization of cDNAs for stylar transmitting tissue-specific proline-rich proteins in tobacco, Plant J. 3 (1993) 151–160.

[13] Cheung A.Y., Wang H., Wu H.M., A floral transmitting tissue-specific glycoprotein attracts pollen tubes and stimulates their growth, Cell 82 (1995) 383–393.

[14] Cho H.T., Kende H., Expression of expansin genes is correlated with growth in deepwater rice, Plant Cell 9 (1997) 1661–1671.

[15] Cho H.T., Kende H., Tissue localization of expansins in deepwater rice, Plant J. 15 (1998) 805–812.

[16] Choi D.W., Song J.Y., Kwon Y.M., Kim S.G., Characterization of a cDNA encoding a proline-rich 14 kDa protein in developing cortical cells of the roots of bean (*Phaseolus vulgaris*) seedlings, Plant Mol. Biol. 30 (1996) 973–982.

[17] Cosgrove D.J., Relaxation in a high-stress environment: The molecular bases of extensible cell walls and cell enlargement, Plant Cell 9 (1997) 1031–1041.

[18] Cosgrove D.J., Bedinger P., Durachko D.M., Group I allergens of grass pollen as cell wall-loosening agents, Proc. Natl. Acad. Sci. USA 94 (1997) 6559–6564.

[19] Coupe S.A., Taylor J.E., Isaac P.G., Roberts J.A., Identification and characterization of a proline-rich mRNA that accumulates during pod development in oilseed rape (*Brassica napus* L.), Plant Mol. Biol. 23 (1993) 1223–1232.

[20] Datta N., LaFayette P.R., Kroner P.A., Nagao R.T., Key J.L., Isolation and characterization of three families of auxin down-regulated cDNA clones, Plant Mol. Biol. 21 (1993) 859–869.

[21] de S. Goldman M.H.S., Pezzotti M., Seurinck J., Mariani C., Developmental expression of tobacco pistil-specific genes encoding novel extensin-like proteins, Plant Cell 4 (1992) 1041–1082.

[22] Deutch C.E., Winicov I., Post-transcriptional regulation of a salt-inducible alfalfa gene encoding a putative chimeric proline-rich cell wall protein, Plant Mol. Biol. 27 (1995) 411–418.

[23] Du H., Simpson R.J., Clarke A.E., Bacic A., Molecular characterization of a stigma-specific gene encoding an arabinogalactan-protein (AGP) from *Nicotiana alata*, Plant J. 9 (1996) 313–323.

[24] Du H., Simpson R.J., Moritz R.L., Clarke A.E., Bacic A., Isolation of the protein backbone of an arabinogalactan-protein from the styles of *Nicotiana alata* and characterization of a corresponding cDNA, Plant Cell 6 (1994) 1643–1653.

[25] Elliott K.A., Shirsat A.H., Promoter regions of the extA extensin gene from *Brassica napus* control activation in response to wounding and tensile stress, Plant Mol. Biol. 37 (1998) 675–687.

[26] Fincher G.B., Stone B.A., Clarke A.E., Arabinogalactan-proteins: Structure, biosynthesis, and function, Annu. Rev. Plant Physiol. 34 (1983) 47–70.

[27] Franssen H.J., Nap J.P., Gloudemans T., Stiekema W., van Dam H., Govers F., Louwerse J., van Kammen A., Bisseling T., Characterization of cDNA for nodulin-75 of soybean: A gene product involved in early stages of root nodule development, Proc. Natl. Acad. Sci. USA 84 (1987) 4495–4499.

[28] Fry S.C., Cross-linking of matrix polymers in the growing cell walls of angiosperms, Annu. Rev. Plant Physiol. 37 (1986) 165–186.

[29] Gao M., Kieliszewski M.J., Lamport D.T.A., Showalter A.M., Isolation, characterization and immunolocalization of a novel, modular tomato arabinogalactan-protein corresponding to the LeAGP-1 gene, Plant J. 18 (1999) 43–55.

[30] Garcia-Olmedo F., Salcedo G., Sanchez-Monge R., Hernandez C.L., Carmona M.J., Lopez-Fando J.J., Fernandez J.A., Gómez L., Royo J., Garcia Maroto F., Castagnaro A., Carbonero P., Alphaamylase/trypsin inhibitors and thionins. Possible defence proteins from barley, in: Shewry P.R. (Ed.), Barley: Genetics, Molecular Biology and Biotechnology, CAB International, Wallingford, 1992, pp. 335–350.

[31] Gonzalez R., Varela J., Carreira J., Polo F., Soybean hydrophobic protein and soybean hull allergy, Lancet 346 (1995) 48–49.

[32] Goodwin W., Pallas J.A., Jenkins G.I., Transcripts of a gene encoding a putative cell wall-plasma membrane linker protein are specifically cold-induced in *Brassica napus*, Plant Mol. Biol. 31 (1996) 771–781.

[33] Gupta N., Martin B.M., Metcalfe D.D., Rao P.V., Identification of a novel hydroxyproline-rich glycoprotein as the major allergen in Parthenium pollen, J. Allergy Clin. Immunol. 98 (1996) 903–912.

[34] Györgyey J., Németh K., Magyar Z., Kelemen Z., Alliotte T., Inzé D., Dudits D., Expression of a novel-type small proline-rich protein gene of alfalfa is induced by 2,4-dichlorophenoxiacetic acid in dedifferentiated callus cells, Plant Mol. Biol. 34 (1997) 593–602.

[35] Han K.K., Martinage A., Possible relationship between coding recognition amino acid sequence motif or residue(s) and post-translational chemical modification of proteins, Int. J. Biochem. 24 (1992) 1349–1363.

[36] Henrissat B., Popineau Y., Kader J.C., Hydrophobic-cluster analysis of plant protein sequences. A domain homology between storage and lipid-transfer proteins, Biochem. J. 255 (1988) 901–905.

[37] Holk A., Kaldenhoff R., Richter G., Regulation of an embryogenic carrot gene (DC 2.15) and identification of its active promoter sites, Plant Mol. Biol. 31 (1996) 1153–1161.

[38] Hong J.C., Nagao R.T., Key J.L., Characterization of a proline-rich cell wall protein gene family of soybean. A comparative analysis, J. Biol. Chem. 265 (1990) 2470–2475.

[39] Hood E.E., Hood K.R., Fritz S.E., Hydroxyproline-rich glycoproteins in cell walls of pericarp from maize, Plant Sci. 79 (1991) 13–22.

[40] Hotze M., Waitz A., Schröder J., cDNA for a 14-kilodalton polypeptide from Madagascar Periwinkle (*Catharanthus roseus*). Eight conserved cysteines in six related proteins from different plants suggest common functional elements, Plant Physiol. 104 (1994) 1097–1098.

[41] John I., Wang H., Held B.M., Wurtele E.S., Colbert J.T., An mRNA that specifically accumulates in maize roots delineates a novel subset of developing cortical cells, Plant Mol. Biol. 20 (1992) 821–831.
[42] Josè M., Puigdomènech P., Structure and expression of genes coding for structural proteins of the plant cell wall, New Phytol. 125 (1993) 259–282.
[43] Josè M., Puigdomènech P., Hybrid-proline rich and related proteins, Curr. Top. Mol. Genet. (Life Sci. Adv.) 2 (1994) 109–126.
[44] Josè-Estanyol M., Puigdomènech P., Developmental and hormonal regulation of genes coding for proline-rich proteins in female inflorescences and kernels of maize, Plant Physiol. 116 (1998) 485–494.
[45] Josè-Estanyol M., Ruiz-Avila L., Puigdomènech P., A maize embryo-specific gene encodes a proline-rich and hydrophobic protein, Plant Cell 4 (1992) 413–423.
[46] Kieliszewski M.J., Lamport D.T.A., Extensin: repetitive motifs, functional sites, post-translational codes, and phylogeny, Plant J. 5 (1994) 157–172.
[47] Li S.X., Showalter A.M., Cloning and developmental/stress-regulated expression of a gene encoding a tomato arabinogalactan protein, Plant Mol. Biol. 32 (1996) 641–652.
[48] Lind J.L., Bacic A., Clarke A.E., Anderson M.A., A style-specific hydroxyproline-rich glycoprotein with properties of both extensins and arabinogalactan, Plant J. 6 (1994) 491–502.
[49] Mau S.L., Chen C.G., Pu Z.Y., Moritz R.L., Simpson R.J., Bacic A., Clarke A.E., Molecular cloning of cDNAs encoding the protein backbones of arabinogalactan-proteins from the filtrate of suspension-cultures cells of *Pyrus communis* and *Nicotiana alata*, Plant J. 8 (1995) 269–281.
[50] McNeil M., Darvill A.G., Fry S.C., Albersheim P., Structure and function of the primary cell walls of plants, Annu. Rev. Biochem. 53 (1984) 625–663.
[51] Mena M., Sanchez-Monge R., Gomez L., Salcedo G., Carbonero P., A major barley allergen associated with baker's asthma disease is a glycosylated monomeric inhibitor of insect α-amylase: cDNA cloning and chromosomal location of the gene, Plant Mol. Biol. 20 (1992) 451–458.
[52] Nicholas C.D., Lindstrom J.T., Vodkin L.O., Variation of proline rich cell wall proteins in soybean lines with anthocyanin mutations, Plant Mol. Biol. 21 (1993) 145–156.
[53] Odani S., Koide T., Ono T., Seto Y., Tanaka T., Soybean hydrophobic protein. Isolation, partial characterization and the complete primary structure, Eur. J. Biochem. 162 (1987) 485–491.
[54] Perfect S.E., O'Connell R.J., Green E.F., Doering-Saad C., Green J.R., Expression cloning of a fungal proline-rich glycoprotein specific to the biotrophic interface formed in the *Colletotrichum*-bean interaction, Plant J. 15 (1998) 273–279.
[55] Qi W., Fong C., Lamport D.T.A., Gum arabic glycoprotein is a twisted hairy rope. A new model based on O-galactosylhydroxyproline as the polysaccharide attachment site, Plant Physiol. 96 (1991) 848–855.
[56] Qi X., Behrens B.X., West P.R., Mort A.J., Solubilization and partial characterization of extensin fragments from cell walls of cotton suspension cultures. Evidence for a covalent cross-link between extensin and pectin, Plant Physiol. 108 (1995) 1691–1701.
[57] Raines C.A., Lloyd J.C., Chao S., John U.P., Murphy G.J.P., A novel proline-rich protein from wheat, Plant Mol. Biol. 16 (1991) 663–670.
[58] Reinhardt D., Wittwer F., Mandel T., Kuhlemeier C., Localized upregulation of a new expansin gene predicts the site of leaf formation in the tomato meristem, Plant Cell 10 (1998) 1427–1437.
[59] Rubinstein A.L., Broadwater A.H., Lowrey K.B., Bedinger P.A., Pex1, a pollen-specific gene with an extensin-like domain, Proc. Natl. Acad. Sci. USA 92 (1995) 3086–3090.
[60] Salts Y., Kenigsbuch D., Wachs R., Gruissem W., Barg R., DNA sequence of the tomato fruit expressed proline-rich protein gene TPRP-F1 reveals an intron within the 3' untranslated transcript, Plant Mol. Biol. 18 (1992) 407–409.
[61] Schmidt J.S., Lindstrom J.T., Vodkin L.O., Genetic length polymorphisms create size variation in proline-rich proteins of the cell wall, Plant J. 6 (1994) 177–186.
[62] Schnabelrauch L.S., Kieliszewski M., Upham B.L., Alizedeh H., Lamport D.T.A., Isolation of pI 4.6 extensin peroxidase from tomato cell suspension cultures and identification of Val-Tyr-Lys as putative intermolecular cross-link site, Plant J. 9 (1996) 477–489.
[63] Schultz C.J., Hauser K., Lind J.L., Atkinson A.H., Pu Z.Y., Anderson M.A., Clarke A.E., Molecular characterisation of a cDNA sequence encoding the backbone of a style-specific 120 kDa glycoprotein which has features of both extensins and arabinogalactan proteins, Plant Mol. Biol. 35 (1997) 833–845.
[64] Sheng J., D'Ovidio R., Mehdy M.C., Negative and positive regulation of a novel proline-rich protein mRNA by fungal elicitor and wounding, Plant J. 1 (1991) 345–354.
[65] Sommer-Knudsen J., Bacic A., Clarke A.E., Hydroxyproline-rich plant glycoproteins, Phytochemistry 47 (1998) 483–497.
[66] Sommer-Knudsen J., Clarke A.E., Bacic A., A galactose-rich, cell wall glycoprotein from styles of *Nicotiana alata*, Plant J. 9 (1996) 71–83.
[67] Stiefel V., Ruiz-Avila L., Raz R., Vallés M.P., Gómez J., Pagés M., Martínez-Izquierdo J.A., Ludevid M.D., Langdale J.A., Nelson T., Puigdomènech P., Expression of a maize cell wall hydroxyproline-rich glycoprotein gene in early leaf and root vascular differentiation, Plant Cell 2 (1990) 785–793.
[68] Subramaniam K., Ranie J., Srinivasa B.R., Sinha A.M., Mahadevan S., Cloning and sequence of a cDNA encoding a novel hybrid proline-rich protein associated with cytokinin-induced haustoria formation in *Cuscuta reflexa*, Gene 141 (1994) 207–210.

[69] Suzuki H., Fowler T.J., Tierney M.L., Deletion analysis and localization of SbPRP1, a soybean cell wall protein gene, in roots of transgenic tobacco and cowpea, Plant Mol. Biol. 21 (1993) 109–119.

[70] van der Eycken W., de Almeida Engler J., Inzé D., van Montagu M., Gheysen G., A molecular study of root-knot nematode-induced feeding sites, Plant J. 9 (1996) 45–54.

[71] Vera P., Lamb C., Doerner P.W., Cell-cycle regulation of hydroxyproline-rich glycoprotein *HRGPnt*3 gene expression during the initiation of lateral root meristems, Plant J. 6 (1994) 717–727.

[72] Vignols F., Josè-Estanyol M., Caparrós-Ruiz D., Rigau J., Puigdomènech P., Involvement of a maize proline-rich protein in secondary cell wall formation as deduced from its specific mRNA localization, Plant Mol. Biol. 39 (1999) 945–952.

[73] Waffenschmidt S., Woessner J.P., Beer K., Goodenough U.W., Isodityrose cross-linking mediates insolubilization of cell walls in *Chlamydomonas*, Plant Cell 5 (1993) 809–820.

[74] Wilson R.C., Long F., Maruoka E.M., Cooper J.B., A new proline-rich early nodulin from *Medicago truncatula* is highly expressed in nodule meristematic cells, Plant Cell 6 (1994) 1265–1275.

[75] Wu H.M., Zou J., May B., Gu Q., Cheung A.Y., A tobacco gene family for flower cell wall proteins with a proline-rich domain and a cysteine-rich domain, Proc. Natl. Acad. Sci. USA 90 (1993) 6829–6833.

[76] Wycoff K.L., Powell P.A., Gonzales R.A., Corbin D.R., Lamb C., Dixon R.A., Stress activation of a bean hydroxyproline-rich glycoprotein promoter is superimposed on a pattern of tissue-specific developmental expression, Plant Physiol. 109 (1995) 41–52.

[77] Xu Y., Buchholz W.G., DeRose R.T., Hall T.C., Characterization of a rice gene family encoding root-specific proteins, Plant Mol. Biol. 27 (1995) 237–248.

[78] Yasuda E., Ebinuma H., Wabiko H., A novel glycine-rich/hydrophobic 16 kDa polypeptide gene from tobacco: similarity to proline-rich protein genes and its wound-inducible and developmentally regulated expression, Plant Mol. Biol. 33 (1997) 667–678.

[79] Youl J.J., Bacic A., Oxley D., Arabinogalactan-proteins from *Nicotiana alata* and *Pyrus communis* contain glycosylphosphatidylinositol membrane anchors, Proc. Natl. Acad. Sci. USA 95 (1998) 7921–7926.

[80] Zhang S., Mehdy M.C., Binding of a 50 kD protein to a U-rich sequence in an mRNA encoding a proline-rich protein that is destabilized by fungal elicitor, Plant Cell 6 (1994) 135–145.