Cloning and Molecular Characterization of the Basic Peroxidase Isoenzyme from Zinnia elegans, an Enzyme Involved in Lignin Biosynthesis¹[w]

Carlos Gabaldón, Matías López-Serrano, María A. Pedreño, and A. Ros Barceló*

Department of Plant Biology, University of Murcia, E–30100 Murcia, Spain

The major basic peroxidase from Zinnia elegans (ZePrx) suspension cell cultures was purified and cloned, and its properties and sequence expression were characterized. The ZePrx was composed of two isoforms with an $M_r$ (determined by matrix-assisted laser-desorption ionization time of flight) of 34,700 (ZePrx34.70) and a $M_r$ of 33,440 (ZePrx33.44). Both isoforms showed absorption maxima at 403 (Soret band), 500, and 640 nm, suggesting that both are high-spin ferric secretory class III peroxidases. $M_r$ differences between them were due to the glycan moieties, and were confirmed from the total similarity of the N-terminal sequences (LSTTFYDTT) and by the 99.9% similarity of the tryptic fragment fingerprints obtained by reverse-phase nano-liquid chromatography. Four full-length cDNAs coding for these peroxidases were cloned. They only differ in the 5′-untranslated region. These differences probably indicate different ways in mRNA transport, stability, and regulation. According to the $k_{cat}$ and apparent $K_m$ values shown by both peroxidases for the three monolignols, sinapyl alcohol was the best substrate, the endwise polymerization of sinapyl alcohol by both ZePrxs yielding highly polymerized lignins with polymerization degrees ≥87. Western blots using anti-ZePrx34.70 IgGs showed that ZePrx33.44 was expressed in tracheary elements, roots, and hypocotyls, while ZePrx34.70 was only expressed in roots and young hypocotyls. None of the ZePrx isoforms was significantly expressed in either leaves or cotyledons. A neighbor-joining tree constructed for the four full-length cDNAs suggests that the organ expression were characterized. The ZePrx was composed of two isoforms with a $M_r$ (determined by matrix-assisted laser-desorption ionization time of flight) of 34,700 (ZePrx34.70) and a $M_r$ of 33,440 (ZePrx33.44). Both isoforms showed absorption maxima at 403 (Soret band), 500, and 640 nm, suggesting that both are high-spin ferric secretory class III peroxidases. $M_r$ differences between them were due to the glycan moieties, and were confirmed from the total similarity of the N-terminal sequences (LSTTFYDTT) and by the 99.9% similarity of the tryptic fragment fingerprints obtained by reverse-phase nano-liquid chromatography. Four full-length cDNAs coding for these peroxidases were cloned. They only differ in the 5′-untranslated region. These differences probably indicate different ways in mRNA transport, stability, and regulation. According to the $k_{cat}$ and apparent $K_m$ values shown by both peroxidases for the three monolignols, sinapyl alcohol was the best substrate, the endwise polymerization of sinapyl alcohol by both ZePrxs yielding highly polymerized lignins with polymerization degrees ≥87. Western blots using anti-ZePrx34.70 IgGs showed that ZePrx33.44 was expressed in tracheary elements, roots, and hypocotyls, while ZePrx34.70 was only expressed in roots and young hypocotyls. None of the ZePrx isoforms was significantly expressed in either leaves or cotyledons. A neighbor-joining tree constructed for the four full-length cDNAs suggests that the four putative paralogous genes encoding the four cDNAs result from duplication of a previously duplicated ancestral gene, as may be deduced from the conserved nature and conserved position of the introns.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: A. Ros Barceló (rosbarce@um.es).

¹ This work was supported by grants from the Fundación Séneca (project no. 00545/PI/04) and Ministerio de Ciencia y Tecnología (grant nos. BOS2002-03590 and BFU2005-06317). C.G. holds fellowships (Formación de Profesorado Universitario) from the Ministerio de Educación, Cultura y Deporte.

*w Corresponding author; e-mail rosbarce@um.es; fax 34–968–363–963.

The process of sealing plant cell walls through lignin deposition is known as lignification and provides mechanical strength to the stems, protecting cellulose fibers from chemical and biological degradation (Lewis and Yamamoto, 1990). In this context, plant cell wall lignification is one of the main restrictive factors in the use and recycling of plant biomass. Lignins are three-dimensional, amorphous heteropolymers that result from the oxidative coupling of three $p$-hydroxycinnamyl alcohols, $p$-coumaryl, coniferyl, and sinapyl alcohols, in a reaction mediated by both laccases and class III plant peroxidases (Ros Barceló, 1997). The cross-coupling reaction produces an optically inactive hydrophobic heteropolymer (Ralph et al., 2004) composed of $p$-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) units, respectively.

The spatial and temporal control of lignin biosynthesis is extremely important since lignification is a metabolically costly process that requires large quantities of carbon skeletons and reducing equivalents (Amthor, 2003). Plants do not possess a mechanism to degrade lignins (Lewis and Yamamoto, 1990), so any carbon invested in lignin biosynthesis is not recoverable. Consequently, lignified cells represent a significant carbon sink (Patzlaff et al., 2003) and, as such, plants must carefully balance the synthesis of lignin polymers against the availability of resources, and this means that the monolignol biosynthetic pathway is strongly regulated (Patzlaff et al., 2003).

The metabolic flux (carbon allocation) in the phenylpropanoid pathway is controlled at multiple enzymatic
levels (Boerjan et al., 2003). Studies using lignifying cell cultures of *Pinus taeda*, a gymnosperm, have established (Anterola et al., 1999, 2002) that both the carbon allocation to the pathway and its differential distribution into the two monolignols, 
*p*-coumaryl and coniferyl alcohol, are controlled by the rate of Phe supply and the differential modulation of cinnamate-4-hydroxylase (C4H) and 
*p*-coumarate-3-hydroxylase (C3H), respectively. In angiosperms, there is a novel branching point, the step catalyzed by coniferylaldehyde-5-hydroxylase (CAld5H), which diverts G backbones for the synthesis of S (i.e. sinapyl alcohol) moieties, the precursors of S lignins. CAld5H has not been studied in *P. taeda* cell cultures since they were derived from a gymnosperm, but one may extrapolate, in the absence of available data, that, like C4H and C3H in gymnosperms, CAld5H might also constitute a rate-limiting step in angiosperms. A similar role, as a rate-limiting step, has recently been proposed for the polymerization of monolignols catalyzed by basic pi class III plant peroxidases (Ipekci et al., 1999; Talas-Ogras et al., 2001; Blee et al., 2003). In this context, all the branching (also rate limiting) enzymes of the lignin biosynthetic pathway (C4H, C3H, and CAld5H) and the lignin-assembling enzyme (peroxidase) are hemoproteins. Any possible metabolic controls over these enzymes would lead to the regulation of not only the global monolignol pools in lignifying plant cells and the H/G/S ratio for carbon partitioning, but also of their rates of polymerization. Nevertheless, none of these enzymes has been cloned in *Zinnia elegans*, despite the frequent use of this plant as a model in lignification studies.

*Z. elegans* is a flowering plant belonging to the Asteraceae family. This species is commonly used as a model for studying the last step of lignin biosynthesis, i.e. the polymerization process, due to the simplicity and duality of the lignification pattern shown by stems and hypocotyls and also because of the nature of the peroxidase isoenzyme complement, which is almost completely restricted to the presence of a basic peroxidase isoenzyme (López-Serrano et al., 2004). Furthermore, *Z. elegans* (Fukuda, 1996) and, recently, Arabidopsis (*Arabidopsis thaliana*; Oda et al., 2005) offer the unique possibility of working with cell cultures that resemble differentiating xylem cells and constitute exceptionally useful models for monitoring the expression of enzymes from the lignin biosynthetic pathway, especially the segment that is concerned with the phenylpropanoid backbones (Demura et al., 2002; Milloni et al., 2002).

The lignification pattern of *Z. elegans* seedlings is unique in that, at a certain developmental stage, it offers simultaneously two models of lignification that closely resemble those occurring in gymnosperms and angiosperms. Thus, in 25- to 30-d-old plants, hypocotyl lignins are mainly composed of G/S units in a 42:58 ratio, while stem lignins contain significant amounts of H units in a H/G/S ratio of 22:56:22 (Ros Barceló et al., 2004). That is, S units predominate in the hypocotyl, while G units predominate in the stem. In this regard, *Z. elegans* hypocotyl lignins are typical of angiosperms, while the lignins of the young stem partially resemble that which occurs in gymnosperms, since (H + G) alone constitute 78% of the lignin building blocks.

Stems, hypocotyls, and transdifferentiating *Z. elegans* mesophyll cell cultures express the same basic peroxidase isoenzyme (López-Serrano et al., 2004). Molecular studies of this basic peroxidase isoenzyme are strongly hampered by the difficulty of obtaining the protein in large amounts (Sato et al., 1995). Transdifferentiating *Z. elegans* mesophyll cell cultures, leaves, stems, hypocotyls, and even roots are a limited source of the protein since it is partially covalently bound to cell walls (Masuda et al., 1983; Sato et al., 1995; Ros Barceló and Aznar-Asensio, 1999). A promising source of this isoenzyme may be *Z. elegans* plant cell cultures since the protein, being located in cell walls, should be rapidly secreted to the culture medium. Callus cultures derived from hypocotyls or stems grow from cambial cells, a quiescent region that also originates the vascular tissues in the shoot, including the xylem, and therefore constitutes a potential and valuable source of the protein. In fact, the Arabidopsis ATPA2 peroxidase, a plant peroxidase involved in lignification (Østergaard et al., 2000; Elhiting et al., 2005), constitutes the major peroxidase in the spent medium of Arabidopsis suspension cell cultures (SCCs) and, originally, was purified and cloned from this source (Østergaard et al., 1996).

In this article, we have purified, characterized, and cloned the major peroxidase from the spent medium of *Z. elegans* SCCs. The results showed that this peroxidase (ZePrx), which is expressed in *Z. elegans* SCCs under two distinctive forms differing in the glycosylation pattern, is coded by four full-length cDNAs differing in the 5'-untranslated regions (UTRs). These results suggest that this peroxidase isoenzyme is encoded by a complex multigene family in the genome of *Z. elegans*.

RESULTS

SCCs as a Source of the ZePrx Basic Peroxidase Isoenzymes

Isoelectric focusing (IEF) under nonequilibrium conditions was chosen to screen the presence of the ZePrx isoenzyme in the spent medium of SCCs. Results (Fig. 1) showed that transdifferentiating 3-d-old *Z. elegans* mesophyll cell cultures (tracheary elements [TE]), 26-d-old hypocotyls, 26-d-old stems, and SCCs expressed the same basic peroxidase isoenzyme. Quantitatively, although the level of peroxidase activity against coniferyl alcohol in SCCs was only 33% of that seen in TE when expressed on a cell basis (Table I), it was 113 times greater on a culture volume basis (Table I). This was undoubtedly due to the fact that
SCCs may grow to reach cell densities of $10^8$ cells mL$^{-1}$, while TE are generally grown at cell densities of $10^5$ cells mL$^{-1}$ (Fukuda and Komamine, 1982). Cell density in transdifferentiating *Z. elegans* mesophyll cell cultures cannot be increased since it is critical for the differentiation process (Motose et al., 2001). Therefore, both qualitatively and quantitatively, SCC from *Z. elegans* constitutes a valuable source of the enzyme.

### Purification of the ZePrx Isoenzymes

The basic ZePrx isoenzyme was purified to homogeneity from the spent medium of SCC in a four-step purification protocol, which includes adsorption chromatography on phenyl Sepharose (Fig. 2A), size-exclusion chromatography on Superdex 75 (Fig. 2B), cationic exchange chromatography on SP Sepharose (Fig. 2C), and, finally, affinity chromatography on concanavalin A (Fig. 2D). This last step resolves the ZePrx isoenzyme into two isoforms: one fully glycosylated and weakly retained by the column (Fig. 2D, peak a), and the other partially glycosylated and strongly retained by the column (Fig. 2D, peak b). The $M_r$ values estimated by SDS-PAGE for the two isoforms were 38,480 ± 1,490 and 36,470 ± 1,400, which contrast with that estimated by matrix-assisted laser-desorption ionization time-of-flight (MALDI-TOF) mass spectrometry (MS), which yielded values of 34,700 and 33,440, respectively. Table II and Figure 3 illustrate the progress of the purification of the two isoforms.

Starting from 18.5 L of spent medium, the partially glycosylated isoform (renamed ZePrx33.44 according to its $M_r$ value determined by MALDI-TOF MS) was obtained with a yield of 20% (5.2 mg), an $R_z$ (Reinheit Zahl value, ratio of absorbance at 403 and 280 nm) of 3.03, and a specific activity of 11.1 μkat mg$^{-1}$ protein when assayed against 4-methoxy-α-naphthol (Table II). The fully glycosylated isoform, ZePrx34.70, was obtained with a yield of 6% (2.5 mg), an $R_z$ value of 3.00, and a specific activity of 6.6 μkat mg$^{-1}$ protein (Table II). To judge from the spent medium protein fingerprint for the crude fraction (Fig. 3), both proteins together constitute about 40% of the total protein secreted to the medium by *Z. elegans* SCC. Both proteins were capable of oxidizing ascorbic acid, ferulic acid, and sinapyl alcohol in a reaction strictly dependent on H$_2$O$_2$ (Table III), but sinapyl alcohol was the best substrate, supporting the role of ZePrxs in lignin biosynthesis.

### Molecular Characterization

The two purified proteins showed absorption maxima in the visible (VIS) region at 403 (Soret band), 500, and 640 nm, which shifted to 435, 555, and 580 nm in the case of ZePrx33.44, and to 435, 560, and 580 in the case of ZePrx34.70, when the ferric enzyme was reduced by sodium dithionite. These spectral characteristics are typical and unequivocal for hemo-containing high-spin ferric secretory (class III) peroxidases (Yamazaki and Yokota, 1973).

The N-terminal amino sequences for both ZePrx33.44 and ZePrx34.70 were determined by subjecting the purified enzymes directly to Edman degradation. Microsequencing was only possible after deblocking the N terminus with pyrrolidone carboxyl peptidase, which suggests that the N terminus in both proteins is pyrrolidone carboxylic acid (Z, Pyr). After removing the N-terminal Z residue, the sequences obtained, LSTTFYDTTUPALTSTI for ZePrx33.44 (where U is undetermined) and LSTTFYDTTUTUTSTI for ZePrx34.70, suggest that both proteins share the same N terminus, being the LST motif conserved in most known class III plant peroxidase sequences (Tyson and Dhindsa, 1995).

Both ZePrx33.44 and ZePrx34.70 were subjected to deglycosylation with trifluoromethanesulfonic acid (TFMS) and further analysis of the deglycosylated proteins by SDS-PAGE. After deglycosylation, both proteins (d-ZePrx34.70 and d-ZePrx33.44) demonstrated the same mobility by SDS-PAGE (Fig. 4). The $M_r$ value estimated by SDS-PAGE for the two deglycosylated isoforms was 33,770, which contrasts with the value of 31,460 estimated by MALDI-TOF MS. That is, the VIS properties, the N-terminal amino acid sequence, and

---

**Table I. Coniferyl alcohol (CA) peroxidase activity in the spent medium protein fraction of *Z. elegans* suspension cell cultures and transdifferentiating 3-d-old *Z. elegans* mesophyll cell cultures**

<table>
<thead>
<tr>
<th></th>
<th><em>Z. elegans</em> SCCs</th>
<th>Transdifferentiating <em>Z. elegans</em> Mesophyll Cell Cultures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell density</td>
<td>(cells mL$^{-1}$)</td>
<td>1.4 (±0.4) $10^6$ (3) 4.0 (±0.5) $10^6$ (8)</td>
</tr>
<tr>
<td>CA peroxidase activity (nkat $10^6$ cells$^{-1}$)</td>
<td>0.27 ± 0.11 (3) 0.82 ± 0.12 (8)</td>
<td></td>
</tr>
<tr>
<td>CA peroxidase activity (nkat mL$^{-1}$)</td>
<td>37 ± 6 (3) 0.33 ± 0.08 (8)</td>
<td></td>
</tr>
</tbody>
</table>

Values are means ± se (n). n, Number of independent determinations.
Cloning of the Basic Peroxidase Isoenzyme from Z. elegans

To confirm this observation, both ZePrx33.61 and ZePrx34.70 were digested by trypsin, and tryptic fragments were analyzed by reverse-phase (RP) nano-LC and characterized by MALDI-TOF MS. Peaks from self-digested trypsin were subtracted from the spectra manually, and mean peptide masses were calculated for the remaining peaks. Both digestions provided similar RP nano-LC profiles with identical (within the experimental error) MALDI-TOF-estimated Ms values (Fig. 5), indicating amino acid sequence identity between the two. In fact, of 93 tryptic fragments obtained with mass-to-charge ratio (m/z) lower than 3,600, only fragments of m/z 689.4248, 734.3907, 756.5001, 975.3651, 1,014.2025, and 2,193.1426 (all of them with abundances below 5%) were specific for ZePrx33.44, and only fragments of m/z 859.5576 (9%), 1,423.7561 (<5%), and 2,197.1262 (<5%) were specific for ZePrx34.70.

All the peptides were entered in the PeptideSearch program (http://www.matrixscience.com), the tandem mass spectrometry (MS/MS) ion search being performed with the following restrictions: (1) Calculated masses were average masses; (2) Met residues were either unmodified (M) or oxidized (M + O); (3) Cys residues were carboxymethylated; (4) peptides were protonated; (5) mass values were monoisotopic; (6) protein mass was unrestricted; (7) peptide mass tolerance was ±0.33 Da; and (8) fragment mass tolerance was ±300 mmu. Searching the database with all the peptides gave only two matches. One of the matches was obtained for the peptide of m/z (M + H+) 1,401.7473 in ZePrx33.44 and 1,401.7467 in ZePrx34.70, whose MS2 fragment ion spectrum fits well with the MS3 fragment ion spectrum of the tryptic peptide of Mr (M + H+) (calc) 1,401.7211, DAS-VAVGPPSWTV (M) (calc) 1,556.7849, EMVALSGSHTIGQAR (M) (calc) 1,556.7992, in the tryptic digest of peroxidase 2 from Scutellaria baicalensis (accession no. AB024438; Morimoto et al., 1999). The other match was obtained for the peptide of m/z (M + H+) 1,556.8221 in ZePrx33.44 and 1,556.8193 in ZePrx34.70, whose Ms3 fragment ion spectrum fits well with the Ms3 fragment ion spectrum of the tryptic peptide of Mr (M + H+) (calc) 1,505.7992, in the tryptic digest of peroxidase 3 from S. baicalensis (accession no. AB024439; Morimoto et al., 1999).

Table II. Purification of ZePrxs

<table>
<thead>
<tr>
<th>Step</th>
<th>Protein Specific Activity</th>
<th>Yield (%)</th>
<th>Rz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ammonium sulfate precipitation</td>
<td>64</td>
<td>n.d.</td>
<td>100 n.d.</td>
</tr>
<tr>
<td>Phenyl Sepharose chromatography</td>
<td>28</td>
<td>5.0</td>
<td>78 0.65</td>
</tr>
<tr>
<td>Superdex 75 chromatography</td>
<td>19</td>
<td>n.d.</td>
<td>46 2.15</td>
</tr>
<tr>
<td>SP Sepharose chromatography</td>
<td>10</td>
<td>5.1</td>
<td>29 2.70</td>
</tr>
<tr>
<td>Concanavalin A Sepharose chromatography (ZePrx34.70)</td>
<td>2.5</td>
<td>6.5</td>
<td>6 3.00</td>
</tr>
<tr>
<td>Concanavalin A Sepharose chromatography (ZePrx33.44)</td>
<td>5.2</td>
<td>11.1</td>
<td>20 3.03</td>
</tr>
</tbody>
</table>

*a.n.d., Not determined.*
using total RNA isolated from 6-d-old Z. elegans hypocotyls. In the first step, two truncated 865-bp cDNAs were first synthesized by reverse transcription (RT)-PCR, using two (one degenerated and the other specific) primers, whose nucleotide sequences were designed to be complementary to the coding strand for the peptide sequences VRTLCGNP, present in the C-terminal sequence of both ZePrxs, and LSTTFYDT, present in the N-terminal sequence obtained for both ZePrxs.

The complete full-length sequence of the cDNAs, including the signal peptide and the 5′ and 3′ flanking regions, was obtained by 5′-RACE and 3′-RACE. 5′-RACE was performed using the RCPxZe-R primer while 3′-RACE PCR was performed using the RCPxZe-L primer. Both primers are shaded in green in Figure 6. Four full-length cDNAs of 1,331 bp (accession no. AJ880394) in the EMBL nucleotide sequence database; http://europe.cbl.leeds.ac.uk/l4, 1,329 bp (accession no. AJ880395), 1,304 bp (accession no. AJ880392), and 1,302 bp (accession no. AJ880393) were obtained (Fig. 6). Nucleotide sequences of the amplified full-length cDNAs were determined from both strands to ensure accuracy and were confirmed by PCR amplification of the DNA isolated from Z. elegans leaves. The four full-length cDNAs contained an identical 966-bp open reading frame (ORF). The deduced primary structure of ZePrxs is also shown in Figure 6, and contains the N-terminal amino acid sequence and the three tryptic fragments, determined experimentally. Furthermore, the uncertainty in the I/L residues of the peptide MSE(I/L)GVTVGTSG(I/L)VR, of m/z (M + H+), 1,505.8348 in ZePrx33.44 and 1,505.8392 in ZePrx34.70, was resolved in favor of I.

The ORFs of the four full-length cDNAs corresponded to a deduced polypeptide of 321 amino acids, including a signal peptide (N-terminal propeptide) of 30 amino acids, which directs the polypeptide chain to the ER membrane. The mature polypeptide therefore showed 291 amino acids. Like other class III plant peroxidases, the mature polypeptide (Fig. 6) started with a glutamyl (Q) residue, which probably generates the pyrroolidone carboxyl residue (Z) found in the purified enzyme. The predicted polypeptide also contained a peroxidase active site signature (33-AAL-VIRLLFHD C), a peroxidase proximal hemo-ligand signature (157-EMVALSGSHTL), eight conserved Cys (C11, C44, C49, C87, C93, C172, C198, C287), which probably yield the four disulfide bridges (C11-C87, C44-C49, C93-C287, C172-C198) common in most class III plant peroxidases (Welinder et al., 2002), two putative N-glycosylation sites (181-NSTL and 191-NRSL), and two Ca2+-binding sites, one distal (D43 and D50) and the other proximal (D211 and D219), characteristic for all the active class III peroxidases (Welinder et al., 2002). The position of these Ca2+-binding sites was deduced by alignment of the Z. elegans polypeptide with the polypeptide of horseradish peroxidase (HRP) C (Veitch, 2004).

**Biochemical Properties of ZePrxs**

Since glycosylation does not significantly affect the pH-dependence profile (Fig. 7A) nor the thermal stability of the ZePrxs (Fig. 7B), the oxidation of the three p-hydroxyccinnamyl (p-coumaryl, coniferyl, and sinapyl) alcohols by the two ZePrxs was studied to ascertain whether glycosylation modifies the catalytic properties of the enzymes. The kcat and apparent Km (KmRH) values for the three monolignols during this peroxidase-catalyzed reaction are reported in Table IV. It can be seen that sinapyl alcohol was the best substrate for both enzymes. Although no clear tendency was found for the effect of glycosylation on the kcat in the case of p-coumaryl and coniferyl alcohol, the full glycosylation that occurs in ZePrx34.70 reduces the kcat for sinapyl alcohol. Likewise, the full glycosylation reduces the affinity (increases the apparent KmRH value) of the enzyme for both p-coumaryl and coniferyl alcohol. In the case of sinapyl alcohol, the apparent KmRH value was not affected significantly by glycosylation (Table IV).

Since sinapyl alcohol appeared to be the best substrate for both ZePrxs, the oligomeric nature of the oxidation products of sinapyl alcohol oxidation by

**Table III. Specific activity of ZePrxs against different peroxidase substrates**

<table>
<thead>
<tr>
<th>Substrate</th>
<th>ZePrx34.70 μkat mg⁻¹ protein</th>
<th>ZePrx33.44 μkat mg⁻¹ protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-Methoxy-α-naphthol (100 μM)</td>
<td>6 ± 1</td>
<td>11 ± 1</td>
</tr>
<tr>
<td>Ascorbic acid (500 μM)</td>
<td>5 ± 1</td>
<td>6 ± 1</td>
</tr>
<tr>
<td>Ferulic acid (50 μM)</td>
<td>11 ± 1</td>
<td>14 ± 1</td>
</tr>
<tr>
<td>Sinapyl alcohol (50 μM)</td>
<td>30 ± 2</td>
<td>32 ± 1</td>
</tr>
</tbody>
</table>
both ZePrxs was studied. Endwise polymerization showed that both ZePrxs are able to oxidize sinapyl alcohol to yield highly polymerized lignins, which were resolved by gel permeation (GP)-HPLC (Fig. 8) in a polymer of $M_r \approx 18,226$ (polymerization degree, $n \approx 87$), and in oligomers with mean $M_r$ values of 4,052 ($n \approx 19$), 1,274 ($n \approx 6$), and 667 ($n \approx 3$). No differences were found in the relative abundance of the polymers and oligomers of sinapyl alcohol during its oxidation by both ZePrxs.

Organ Expression

To study the expression of both isoforms in different organs and in the in vitro culture systems of Z. elegans, polyclonal antibodies were prepared against the fully glycosylated form of ZePrx, ZePrx34.70. Anti-ZePrx34.70 IgGs not only recognize ZePrx34.70 but also ZePrx33.44 (Fig. 9A), confirming that both peroxidases share similar epitopes.

The anti-ZePrx34.70 IgGs recognized not only both the ZePrxs but also several other proteins in the spent medium fraction of Z. elegans SCC (Fig. 9B, lane a). This would be because antibodies against plant glycoproteins contain a population of antibodies that recognize antigenic groups in the polypeptide chain, and another population of antibodies, usually in greater abundance, that recognize the oligosaccharidic chain present in the glycoproteins. An example of this is the major fraction of anti-HRP IgGs, which specifically recognize $\alpha$-1,3-fucosylated N-linked glycans (Wilson, 2002). These glycan-derived antigens, such as the high-Man glycans, or the complex glycans rich in Fuc and Xyl allergens (Wilson, 2002) are unspecific since they are found in most plant secretory glycoproteins (Wilson, 2002). To circumvent this problem, blotted proteins on polyvinylidene difluoride (PVDF) membranes were treated with sodium periodate, which, under controlled conditions of time and temperature, only oxidizes vicinal hydroxyl groups of glycans without altering the structure of the polypeptide chains, and in this way is able to remove glycans from glycoproteins (Woodward et al., 1985). After this treatment, anti-ZePrx34.70 IgGs only recognized the two ZePrxs in protein fingerprints obtained from the spent medium fraction of Z. elegans SCCs (Fig. 9B, lane b).

Anti-ZePrx34.70 IgGs were also tested against periodate-treated protein fingerprints obtained from TE, cotyledons, leaves, roots, hypocotyls, and stems (Fig. 9, C and D). Western blots showed that only ZePrx33.44 was expressed in TE (Fig. 9C), both ZePrx33.44 and ZePrx34.70 were expressed in roots and young hypocotyls (Fig. 9D), and only ZePrx33.44 was expressed in old hypocotyls and stems (Fig. 9D). None of the ZePrxs was significantly expressed in either cotyledons or leaves (Fig. 9C).

DISCUSSION

Basic ZePrxs in SCC

Class III plant peroxidases involved in lignin biosynthesis are usually classified into acidic (pI below 7.0) and basic (pI above 7.0) peroxidases. Both types of peroxidases are capable of oxidizing $p$-coumaryl and coniferyl alcohol. However, this situation is not so clear as regards sinapyl alcohol, which possesses a $S$ moiety, and for which acidic peroxidases, with some exceptions (Christensen et al., 1998, 2001), are generally regarded as poor catalysts (Dean et al., 1994; Takahama, 1995; Bernards et al., 1999). This observation constitutes a central key for unraveling the specificity of lignin assembly since sinapyl alcohol is more prone to oxidation than either coniferyl alcohol or $p$-coumaryl alcohol (Russell et al., 1996) and suggests that, although peroxidase-catalyzed reactions are driven by redox thermodynamic forces (Folkes and Candeias, 1997), substrate accommodation in (exclusion from) the catalytic center of the enzyme determines the real role played by each peroxidase isoenzyme in lignin biosynthesis, as has been revealed from x-ray crystallographic studies (Östergaard et al., 2000).

In fact, oxidation of sinapyl alcohol by certain acidic and neutral peroxidases is sterically hindered due to unfavorable hydrophobic interactions between the sinapyl alcohol methoxy atoms and the conserved I-138 and P-139 residues at the substrate binding site of the enzyme (Östergaard et al., 2000). In the case of basic peroxidases, the capacity of these enzymes to oxidize $S$ moieties is universally accepted (Bernards et al., 1999; Quiroga et al., 2000; Ros Barceló and Pomar, 2001; Holm et al., 2003; Sasaki et al., 2004), and this observation would explain why antisense suppression of basic peroxidases in transgenic plants produces decreased levels of both $G$ and $S$ lignins (Blee et al., 2003), while antisense suppression of acidic peroxidases produces only decreased levels of $G$ lignins (Li et al., 2003).

In accordance with their key role in lignin biosynthesis, cationic (basic) peroxidases are differentially expressed during the transdifferentiation of Z. elegans mesophyll cell cultures, where they act as molecular markers of xylogenesis (Masuda et al., 1983; Sato et al., 1995; López-Serrano et al., 2004). These same peroxidase isoenzymes may also be recovered in the spent...
medium protein fraction from *Z. elegans* SSC (Fig. 1) and show a special ability for oxidizing sinapyl alcohol when compared with other putative peroxidase substrates (Table III).

**Heterogeneity of ZePrxs**

The basic peroxidase isoenzyme isolated from *Z. elegans* shows homogeneity as regards the pI under IEF (Fig. 1), but can be resolved by chromatography on concanavalin A (Fig. 2D) in two isoforms: one partially glycosylated (*M*, 33,440) and strongly retained by the column, and another fully glycosylated (*M*, 34,700) and weakly retained by the column. This heterogeneity has been described for some basic peroxidases (Rasmussen et al., 1991; Wan et al., 1994) and it is not surprising since full glycosylation, which leads to the formation of complex glycans in plant glycoproteins, initially involves the addition of Fuc and Xyl end moieties to the Man-rich core and, finally, substitution.

**Figure 5.** MALDI-TOF MS of tryptic (glyco-) peptides released from (A) the fully glycosylated isoform (ZePrx34.70) and (B) the partially glycosylated isoform (ZePrx33.44) of the basic *Z. elegans* peroxidase isoenzyme. Insets show in an amplified scale the tryptic fragments for both isoforms that were sequenced (arrowheads).
Figure 6. cDNAs encoding both ZePrx33.44 and ZePrx34.70 obtained from RNA isolated from 6-d-old *Z. elegans* hypocotyls and deduced amino acid sequence of ZePrxs. The complete full-length sequence of the cDNAs, including the 5’ and 3’ flanking regions, was obtained by 5’-RACE and 3’-RACE using the primers RCPxZe-R (5’-RACE) and RCPxZe-L (3’-RACE; shaded in green). These cDNAs isolated from *Z. elegans* may be categorized as full-length cDNAs from the unusual length of the 5’ tails (216–245 bp), and by the fact that DNA transcription usually starts at a purine base (i.e. A), coding T (shaded in red) in the corresponding cDNAs. They also contain an identical poly(A) tract added to the 3’ end, just downstream of the signal sequence.
of the Man end moieties by Gal units (Wilson, 2002). Undoubtedly, the coating of the Man-rich core with Fuc, Xyl, or Gal end units reduces the strength with which the glycoprotein will be retained by concanavalin A, and this allows both glycoproteins to be separated.

In any case, deglycosylation with TFMS of both ZePrxs33.44 and ZePrxs34.70 yielded the same deglycosylated protein (Fig. 4) with an estimated M_r (MALDI-TOF) of 31,460. That is, the heterogeneity of the glycan moieties, as has previously been reported for peroxidases from flax (Gaudreault and Tyson, 1988), barley (Rasmussen et al., 1991; Johansson et al., 1992), and peanut (Wan et al., 1994). This assumption was further strengthened by the total similarity of the N-terminal sequence determined for both peroxidases (LSTTFYDTT), and by the 99.9% similarity of the tryptic fragment fingerprints obtained by RP nano-LC and MALDI-TOF MS (Fig. 5).

Figure 6. (Continued.)

5'-AATAAAA-3' (shaded in gray), which is recognized by the poly(A) polymerase. The four full-length cDNAs contain an identical 966-bp ORF, whose nucleotide sequences translate into an identical amino acid sequence, despite the observed C/T, G/A, G/C, and A/C substitutions (Fig. 6); (3) a 120-bp 3'-UTR; and (4) an identical polyadenyllic acid [poly(A)] tract added to the 3' end, just downstream of the signal sequence 5'-AATAAAA-3' (Fig. 6; 5'-AAUAAA-3' in the mRNAs), which is recognized by the poly(A) polymerase. Although little is known about the precise function of the 5'-UTRs, it is accepted (Kozak, 1991) that these regions contain sequences that can form secondary structures and interact with proteins that regulate transport, translation, and stability of the mRNAs. Thus, the

Cloning and Primary Structure of ZePrxs

From the N-terminal amino acid sequence and a C-terminal tryptic fragment, four full-length cDNAs encoding the ZePrxs were isolated (Fig. 6). The four full-length cDNAs contain an identical 966-bp ORF encoding 321 amino acids, which coded for the primary structure of ZePrxs shown in Figure 6. These cDNAs isolated from Z. elegans may be categorized as full-length cDNAs may be deduced from the unusually great length of the 5' tails (216–245 bp), and by the fact that DNA transcription usually starts at a purine base (i.e. adenine [A]; Nishiyama et al., 2003), coding thymine (T) in the corresponding cDNAs, as was the case here (Fig. 6).

The four full-length cDNAs contain (1) a nonconserved 5'-UTR of variable length (245 bp in AJ880394, 243 bp in AJ880395, 218 bp in AJ880392, and 216 bp in AJ880393); (2) a 966-bp ORF, whose nucleotide sequences translate into an identical amino acid sequence, despite the observed C/T, G/A, G/C, and A/C substitutions (Fig. 6); (3) an identical 120-bp 3'-UTR; and (4) an identical polyadenyllic acid [poly(A)] tract added to the 3' end, just downstream of the signal sequence 5'-AATAAAA-3' (Fig. 6; 5'-AAUAAA-3' in the mRNAs), which is recognized by the poly(A) polymerase.

5'-binding sites (shaded in blue), one distal (D43) and one proximal (D211 and D219), characteristics for all the active class III peroxidases. The amino acid sequence common to the root RP5a peroxidase (Sato et al., 1995) and the nucleotide base sequence common to the expressed sequence tag (AJ504423) expressed by transdifferentiating Z. elegans mesophyll cell cultures, and described by Milloni et al. (2002), are underlined. Amino acid sequences matched by tryptic fragments and the experimentally determined N-terminal sequence are double underlined.
differences observed in the 5'-UTRs for the four full-length mRNAs probably indicate different ways in their transport, stability, and regulation.

Since translation of most eukaryotic mRNAs starts at the first (5' proximal) AUG codon, in accordance with the scanning behavior of the ribosomal 40S unit (Kozak, 1991), the start codon for the four ORFs was established at 246 bp in AJ880394, 244 bp in AJ880395, 219 bp in AJ880392, and 217 bp in AJ880393, downstream of the 5'-end. None of these cDNAs contains the start codon preference found in dicots (5'-AAAaugGC-3'), which would challenge the above rule. Besides, this putative position for the start codon is reinforced by the presence of an A at -3 (Fig. 6), which is known to lend translational weight to the first recognizable 5'-proximal AUG codon (Kozak, 1991), as has been shown in Arabidopsis (Østergaard et al., 1998).

The position of the start codon for the four ORFs described in Figure 6 suggests that the immature polypeptide contains a signal peptide (N-terminal propeptide) of 30 amino acids (MSYHKSSGTTLMVPLFMLISVNYFMSCNA), which directs the polypeptide chain to the endoplasmic reticulum membrane. In the absence of a hydropathy profile (Nielsen et al., 1997), the signal peptide cleavage site was predicted in dicots (5'-AAAuugGC-3'), which would challenge the above rule. Besides, this putative position for the start codon is reinforced by the presence of an A at -3 (Fig. 6), which is known to lend translational weight to the first recognizable 5'-proximal AUG codon (Kozak, 1991), as has been shown in Arabidopsis (Østergaard et al., 1998).

The position of the start codon for the four ORFs described in Figure 6 suggests that the immature polypeptide contains a signal peptide (N-terminal propeptide) of 30 amino acids (MSYHKSSGTTLMVPLFMLISVNYFMSCNA), which directs the polypeptide chain to the endoplasmic reticulum membrane. In the absence of a hydropathy profile (Nielsen et al., 1997), the signal peptide cleavage site was predicted in dicots (5'-AAAuugGC-3'), which would challenge the above rule. Besides, this putative position for the start codon is reinforced by the presence of an A at -3 (Fig. 6), which is known to lend translational weight to the first recognizable 5'-proximal AUG codon (Kozak, 1991), as has been shown in Arabidopsis (Østergaard et al., 1998).

Table IV. $k_{cat}$ and $K_m$ values for the oxidation of the three monolignols by ZePrxs

<table>
<thead>
<tr>
<th>Monolignol</th>
<th>$k_{cat}$</th>
<th>$K_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZePrx34.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZePrx33.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZePrx34.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZePrx33.44</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$K_m$ and $k_{cat}$ values were determined for a H$_2$O$_2$ concentration in the assay media of 10.2 µM. Values are means ± SE (n = 4).

### Figure 8

GP-HPLC profiles of the products of sinapyl alcohol polymerization by ZePrx34.70 (A) and ZePrx33.44 (B), using the endwise protocol, which resolved them into a polymer of $M_\text{r}$ of 30,863, which, with the additional contribution of hemo b (616), yields a putative mature protein of $M_\text{r}$, 31,479. The value fits well with the $M_\text{r}$ obtained for the deglycosylated forms of both ZePrx (31,460 – Z + Q = 31,477). The mature protein is of a basic nature and has a theoretical pI of 8.47. In the case of peroxidases, predicted pIs between 5 and 10 are generally two units lower than the experimental values because the two calcium ions and the hemo are not included in the calculation (Welinder et al., 2002). After this correction, the theoretical pI for both ZePrxs (8.47 + approximately 2.0) fits well
Figure 9. A, Molecular mass markers, whose values (×10^3) are drawn in the margin, and SDS-PAGE western blots using anti-ZePrx34.70 IgGs of ZePrx34.70 (1.0 μg) and ZePrx33.44 (1.0 μg). B, SDS-PAGE western blots using anti-ZePrx34.70 IgGs of total protein (10 μg) obtained from blotted membranes with sodium periodate. C, SDS-PAGE western blots using anti-ZePrx34.70 IgGs of total protein obtained from the spent medium fraction of 3-d-old Z. elegans SCC before (lane a) and after (lane b) treatment of blotted membranes with sodium periodate. D, SDS-PAGE western blots using anti-ZePrx34.70 IgGs of 3-d-old (110 μg protein) and 6-d-old (170 μg protein) roots; 3-d-old (80 μg protein) hypocotyls; and 6-d-old (200 μg protein) stems, after treatment of the blotted membranes with sodium periodate.

with the pI value of about 10.2 to 10.5 determined experimentally (López-Serrano et al., 2004).

The mature polypeptide also contains the following peptide motif around the eighth C (C287), TGTS-GIVRTLCGNPS, whose partial sequence, TGTS-GIVR, was also identified in the tryptic fragment of m/z 1,505.8348 in ZePrx 33.44 and 1,505.8392 in ZePrx34.70. The proximity (four amino acids) of the eighth C to the end codon suggests that this polypeptide has no C-terminal propeptide extension to target the proteins for vacuolar transport, which, when present in class III plant peroxidases (Welinder et al., 2002), usually adds a 20- to 30-amino acid tail downstream of the end codon. This cell wall localization of the protein is corroborated from the recovery of the enzyme in the spent medium fraction of Z. elegans SCC.

Possible Physiological Roles of ZePrx

The best way to ascertain for which particular metabolic reaction of the complex process of lignin assembly ZePrxs have been designed in the course of vascular plant evolution is to compare the apparent K_m values for the three monolignols. An example of the importance of these values is the observation that microsomal 5-hydroxylases, previously named ferulate-5-hydroxylases, hydroxylate coniferyl aldehyde (K_m = 1 to 2.7 μM), and coniferyl alcohol (K_m = 5 μM) with a greater affinity than ferulate (K_m = 286 to 1,000 μM; Humphreys et al., 1999; Osakabe et al., 1999), leading to the tendency to call them as coniferyl aldehyde-5-hydroxylases (CAld5H).

To cast light on this question, the apparent K_m values of the ZePrxs for the three monolignols (K_m RH) during the peroxidase reaction were examined in detail. From these data (Table IV), it can be concluded that ZePrxs show the lowest apparent K_m values (in the micromolar range) for sinapyl alcohol, illustrating the great affinity of ZePrxs for this monomeric lignin precursor. Besides, the full glycosylation occurring in ZePrx34.70 reduces the affinity (increases the apparent K_m RH value) of the enzyme for both p-coumaryl and coniferyl alcohol.

The apparent K_m RH values for sinapyl alcohol shown by the ZePrxs were in the same range as those shown by the peroxidase-preceding enzymes of the lignin biosynthetic pathway, cinnamyl alcohol dehydrogenase and microsomal 5-hydroxylases, which also use both cinnamyl alcohols and aldehydes as substrates (Humphreys et al., 1999; Osakabe et al., 1999). The micromolar nature of the K_m RH values for these enzymes implicitly indicates that the one-way high-

way of lignin macromolecule construction, which properly starts with the reduction of cinnamoyl-CoAs by cinnamoyl-CoA reductases, contains no metabolic potholes in which the lignin building blocks might be accumulated. This observation is in accordance with the predictions that may be made from the observed pools and metabolic fluxes of monolignols in lignifying cells (Anterola et al., 1999).

The overall kinetic properties during the peroxidase reaction of these hemo proteins suggested again that sinapyl alcohol is the better substrate for both ZePrxs, since it shows not only the lowest apparent K_m RH values but also the highest k_cat, which leads to the
highest catalytic efficacy \(k_{cat}/K_m^{RH}\); Table IV). This is in accordance with the oxido/reduction potentials for the three monolignols since sinapyl alcohol is more prone to oxidation than coniferyl alcohol and much more so than \(p\)-coumaryl alcohol, which is the least reactive substrate (Russell et al., 1996). Furthermore, the reactivity of ZePrxs for synapyl alcohol is such that sinapyl alcohol oxidation by both ZePrxs yields oxidation products of an oligomeric nature, with \(M_r \approx 18,226\) and polymerization degrees, \(n \approx 87\) (Fig. 8), which supports a role for these peroxidases in lignin biosynthesis. This observation led us to conclude that ZePrxs, unlike Arabidopsis ATP2 and HRP A2s (Østergaard et al., 2000; Nielsen et al., 2001), have no steric hindrance in the hemo crevice for actively discriminating among the three \(p\)-hydroxycinnamyl alcohols, and suggests a high degree of metabolic plasticity (versatility) for this basic peroxidase, making it one of the possible driving forces in the evolution of plant lignin heterogeneity.

We should mention, at this point, that the versatility of certain enzymes is one of the main driving forces in the evolution of land plants, and that there is a general consensus (Boerjan et al., 2003) that such enzymes confer high metabolic plasticity to the lignin biosynthetic pathway. The metabolic plasticity of this basic peroxidase, capable of catalyzing with high affinity and high catalytic activity the polymerization of the three \(p\)-hydroxycinnamyl alcohols, means that the heterogeneity of lignin precursors is not prohibitive, nor is its metabolic cost terribly expensive for the plant. In fact, the versatility of this basic peroxidase confers a certain sense to the heterogeneity (three \(p\)-hydroxycinnamyl alcohols) of the lignin biosynthetic pathway, as far as we know in angiosperms.

The cDNAs Coding for ZePrxs Are Both Expressed and Translated in Z. elegans TE

ZePrxs that differ in the glycosylation pattern not only show differential kinetic properties but also a differential organ expression when analyzed by western blot. Thus, when anti-ZePrx34.70 IgGs, which recognize both ZePrx33.44 and ZePrx34.70 (Fig. 9A), were tested against protein fingerprints obtained from SCC, TE, cotyledons, leaves, roots, hypocotyls, and stems, it was found that ZePrx33.44 was expressed in SCC, TE, roots, hypocotyls, and stems, while ZePrx34.70 was only expressed in SCC, roots, and young hypocotyls (Fig. 9, B–D).

The immunological reactivity of the root Z. elegans peroxidases against anti-ZePrx34.70 IgGs (Fig. 9D) is supported by the observation that a BrCN-derived peptide sequence, VALSGSHTLG (where U is undetermined), which is contained in the Z. elegans peroxidase, RP5a, purified from roots, and considered as a marker of xylogenesis in Z. elegans (Sato et al., 1995), is self-contained in the primary amino acid sequence of ZePrxs, concretely in the peptide 159-VALSGSHTLG (Fig. 6). These results suggest that both peroxidases are similar, if not the same.

Likewise, the immunological reactivity of the TE Z. elegans peroxidases against anti-ZePrx34.70 IgGs (Fig. 9C) is in accordance with the observation that an expressed sequence tag isolated from transdifferentiating Z. elegans mesophyll cells (AJ504423) by Milloni et al. (2002) is shared with the four full-length cDNA clones isolated from Z. elegans hypocotyls, the region comprised between \(-4\) to \(-221\) bp upstream the end codon (TAA; Fig. 6, underlined). These results suggest that the cDNAs coding for ZePrxs are both expressed and translated in Z. elegans TE.

ZePrxs Are Encoded by a Complex Multigene Family in the Genome of Z. elegans

A neighbor-joining tree (Kumar et al., 2001) constructed for the four full-length cDNAs isolated from hypocotyls suggests that the four putative paralogous genes encoding the four cDNAs could arise from the duplication of a previously duplicated ancestral gene, since AJ880394 and AJ880395, on the one hand, and AJ880392 and AJ880393, on the other hand, are clustered together. Preliminary results obtained in our laboratory suggest that the four genes codifying ZePrxs consisted of three exons and two introns, one of type 2 and the other of type 3, according to Tognolli et al.’s (2002) nomenclature. The nature and position (see Fig. 6) of these introns suggest that ZePrxs belong to the peroxidase gene subgroup of class a, already described in Oryza (Passardi et al., 2004a) and Arabidopsis (Tognolli et al., 2002). Interestingly, the position in which each exon is disrupted by the introns was the same in the four genes (Fig. 6, arrowheads, \(-389\) bp upstream the end codon for intron 3 and \(-734\) bp upstream the end codon for intron 2). Furthermore, the base pair sequence of intron 2 was totally conserved in the genes codifying AJ880393 and AJ880395 (accession no. AJ971430), and in the genes codifying AJ880392 and AJ880394 (accession no. AJ971431). Finally, search homologies suggest that both introns, AJ971430 and AJ971431, are paralogous, showing 94.7% similarity. These preliminary results suggest that duplications of the ZePrxs occurred by recombination within introns, as has been described for two HRP basic genes coded in tandem (Fujiyama et al., 1988).

In Arabidopsis, analysis of the genome sequence revealed the existence of pairs of chromosomal blocks that exhibit high-level synteny and a similar gene content (Arabidopsis Genome Initiative, 2000). These blocks span almost the whole of the genome without overlapping and are thought to be remnants from a recent genome-wide duplication. Statistical and phylogenetic studies suggest that these ancient duplication events occurred before the divergence of Arabidopsis from other dicots and may constitute an even earlier event predating the monocot-dicot divergence (Kim et al., 2005). Although it has been reported that, after a duplication event, degenerative mutations will probably eliminate many duplicated genes from the genome (Lynch and Conery, 2000), the remaining
duplicated genes are thought to be a source of biochemical diversity preserved by subfunctionalization (Lynch and Force, 2000).

Neofunctionalization (acquisition of a novel function in a protein owing to mutations or duplications at the DNA level during evolution) and subfunctionalization (acquisition of a new expression profile for a duplicated gene) are not new facts when we treat genes codifying class III plant peroxidases. In fact, class III plant peroxidases belong to a multigene family, whose evolution seems to be correlated with the increasing complexity of plant cell wall architecture and the diversification of their biotopes and pathogens (Duruoux and Welinder, 2003; Passardi et al., 2004b). To exemplify this, a high duplication rate in some plants has led to large multigene families, as suggested by the presence of 138 peroxidase-encoding genes in Oryza (Passardi et al., 2004a) and of 73 peroxidase-encoding genes in Arabidopsis (Tognolli et al., 2002; Welinder et al., 2002; Valério et al., 2004). Furthermore, in Arabidopsis, two identical peroxidase genes are represented by AtP11 (AtP11.1 and AtP11.2; Welinder et al., 2002). Either subfunctionalization or neofunctionalization could explain the conservation of these duplicate genes and the presence of such large multigene families in which each paralog (homologous genes produced by duplication within a genome) could become specialized for a determined task.

At this stage, it is not easy to trace a forward relationship (mRNAs versus proteins) between the four full-length cDNAs and the two ZePrxs, since the four full-length cDNAs code for an identical protein primary structure and the two ZePrxs only differ in their glycosylation pattern. However, since both ZePrxs differ in their organ expression (Fig. 9) and catalytic properties (Table IV), it is probable that we are dealing here with two new and very particular examples of subfunctionalization and neofunctionalization performed at the posttranslational level. Subfunctionalization and neofunctionalization are not new facts when talking of enzymes of the lignin biosynthesis pathway. In fact, retained copies of Phenol ammonia lyase and cytochrome P450 monooxygenases in plant genomes led to an increased gene dosage biosynthesis pathway. In fact, retained copies of Phenol ammonia lyase and cytochrome P450 monooxygenases in plant genomes led to an increased gene dosage.
loaded on 44 × 1.0-cm gel-bed column equilibrated with 50 mM CAPS, pH 9.5, at a flow rate of 1.0 mL min⁻¹. Fractions of 1.0 mL were recovered. The eluent chromatography program was as follows: from 0 to 35 min (100% A, 0% B), from 35 to 100 min (0%-100% B), and from 100 to 200 min (100% B), where buffer A was 50 mM CAPS, pH 9.5, and buffer B was 50 mM CAPS, pH 11.5.

The fourth step involved affinity chromatography on concanavalin A-Sepharose 4B (Sigma). For this, the peroxidase-rich fractions obtained from the ion-exchange chromatography were dialyzed against 50 mM Tris-HCl, pH 7.5, loaded on a concanavalin A-Sepharose 4B 26 × 1-cm gel-bed column, and chromatographed at a flow rate of 0.5 mL min⁻¹. Fractions of 1.0 mL were recovered. The eluent chromatography program was as follows: from 0 to 115 min (100% A), and from 115 to 240 min (100% B), where buffer A was 50 mM Tris-HCl, pH 7.5, containing 1 mM NaCl, 1 mM CaCl₂, 1 mM MgCl₂, 1 mM MnCl₂; and buffer B was 50 mM Tris-HCl, pH 7.5, containing 1 mM NaCl, 1 mM CaCl₂, 1 mM MgCl₂, 1 mM MnCl₂, and 0.5 M methyl-α-D-mannopyranoside. This chromatography yielded two peroxidase fractions (ZePrx33.44 and ZePrx34.70), according to their respective Mr values determined by MALDI-TOF (see below), which were again repurified by ion-exchange chromatography on SP Sepharose Fast Flow as described above.

Assay of Peroxidase Activity, Determination of Kinetic Properties, and Determination of Protein Content

The assay of peroxidase activity using 4-methoxy-α-naphthol, coniferyl alcohol, sinapyl alcohol, ascorbic acid, and ferulic acid was performed as described (Takahama, 1995; Ros Barceló and Pomar, 2001; Ros Barceló et al., 2005). The oxidation of p-hydroxycinnamyl (p-coumaryl, coniferyl, and sinapyl) alcohols by ZePrxs, in the presence of H₂O₂, was assayed spectrophotometrically at 25°C in a reaction medium containing 50 mM Tris-acetate buffer, pH 5.0, using a ΔA₂₅₀ = 14,756 M⁻¹ cm⁻¹ for p-coumaryl alcohol, a ΔA₁₂₀ = 9,750 M⁻¹ cm⁻¹ for coniferyl alcohol, and a ΔA₂₅₀ = 4,140 M⁻¹ cm⁻¹ for sinapyl alcohol. To obtain apparent Km and Vmax (Vmax/EC₅₀) values for the three monolignols, phenolic concentrations ranged from 6.25 to 100 μM, while H₂O₂ concentration was 10.2 μM. Reaction media contained 2.16 mM of the three in the case of coniferyl alcohol and sinapyl alcohol, and 19.25 mM (in the case of p-coumaryl alcohol). Protein was determined according to Bradford (1976).

Endwise Polymerization of Sinapyl Alcohol by ZePrxs and HPLC Identification of the Polymerization Products

The endwise polymerization of sinapyl alcohol by either ZePrx33.44 or ZePrx34.70 was performed in a 50 mM sodium phosphate buffer, pH 5.0, reaction medium (5 mL), which contained 5 mg sinapyl alcohol (dissolved in 50 μL acetone and added at a rate of 5 μL h⁻¹), and equivolumar amounts of H₂O₂ also added in 10 successive steps. Either ZePrx33.44 or ZePrx34.70 was added to a final concentration of 25 μM mL⁻¹. A brown-yellow precipitate appeared after incubation for 24 h at 30°C, which was separated from the solution by centrifugation at 3,400g for 15 min. This precipitate was washed twice with water, dried on silica gel at 4°C, and redissolved in N,N-dimethylformamide (DMF).

HPLC analyses were performed using a Waters system (Millipore, Waters Chromatography) comprising a model 600 controller, model 600 pump, Rheodyne 7725 manual injector, and 1900 photodiode array detector. The data was processed with the Waters Millennium 2010 LC version 2.10 software. GP-HPLC of the peroxidase-mediated sinapyl alcohol oxidation products was carried out at room temperature on a 30 cm × 7.8-mm i.d. TSK-Gel G2500HR column (TosoHaas; TOSOH) using DMF as eluent at a flow rate of 1 mL min⁻¹ (López-Serrano and Ros Barceló, 2001). Due to the different λmax of the different peroxidase-mediated sinapyl alcohol oxidation products, chromatographic profiles were recorded at the maximum wavelength (λmax chromatogram). Mr calibration in the GP column was performed using poly(stirene) standards (Aldrich) and hydrolytic lignin (Aldrich).

Absorption Spectrum of Peroxidase Species

The visible absorption spectrum of ZePrx33.44 and ZePrx34.70 was recorded in 200 mM potassium phosphate buffer, pH 7.0, using an enzyme concentration of 1.3 μM (εmax = 1.02 × 10³ M⁻¹ cm⁻¹). The visible spectrum of the ferrous forms was recorded in 50 mM Tris-HCl, pH 7.5, after the addition of sodium dithionite.

Deglycosylation with TFMS

Deglycosylation of ZePrx33.44 and ZePrx34.70 was performed as described by Tams and Welinder (1995), with minor modifications. Phenol (8.5 μL), heated at 50°C, was mixed with 1,000 μL TFMS and 40 μg dry powder protein and incubated at −20°C for 2 h in an O₂-free atmosphere. The reaction was stopped at room temperature for 40 min, and the reaction medium neutralized with 1,000 μL of 2 M Tris-HCl, pH 8.0. Samples were concentrated and dialyzed using the Ultrafree 4 sytem, equilibrated with 50 mM Tris-HCl, pH 7.5.

Electrophoretic Analysis

IEF under nonequilibrium conditions was performed as described by Pomar et al. (2002). Protein migration was followed at 4°C using cytochrome c as a migration marker. Peroxidase isoenzymes were stained with 4-methoxy-α-naphthol (López-Serrano et al., 2004). SDS-PAGE was performed on 10% (w/v) polyacrylamide gels using a MiniProtein 3 cell electrophoresis kit (Bio-Rad) and a pH 8.8 electrophoresis buffer composed of 192 mM Gly and 25 mM Tris containing 0.1% SDS. SDS-PAGE was performed at 200 V for 40 to 45 min at room temperature. Proteins were stained by either the ammoniacal silver method (Oakley et al., 1980) or the Coomassie Blue method (Hirano et al., 1993).

Preparation of Polyclonal Antibodies

New Zealand white rabbits were immunized three times at 2-week intervals by injection (200 μL) of the ZePrx34.70 nonglycosylated form (200 μg) in Freund’s complete adjuvant. Seventy days after the first inoculation, rabbits were immunized with 250 μL of Freund’s incomplete adjuvant containing the antigen. The rabbit serum titer was determined by ELISA after each injection. For assays, the final serum was directly diluted in phosphate-buffered saline (PBS) buffer.

Western Blots

Following SDS-PAGE, gels were placed in contact with a PVDF membrane (0.2 μm; Bio-Rad), and the proteins were electrobotted for 75 min at 100 V in a Bio-Rad mini trans-blot system. The blotting buffer was 192 mM Gly, 25 mM Tris-base, pH 8.3, containing 20% (v/v) methanol and 0.1% (w/v) SDS. Blotted PVDF membranes were blocked for 2 h in PBS buffer containing 0.2% (v/v) Tween 20 and 3% (w/v) decaeed milk. Blotted PVDF membranes were rinsed in 50 mM sodium acetate buffer, pH 4.5, three times. After this time, PVDF membranes were treated for 1 h in the dark at room temperature with 2 mM sodium periodate in the same buffer (Woodward et al., 1985). Blotted PVDF membranes were finally rinsed in 50 mM sodium acetate buffer and incubated for 35 min at room temperature in PBS buffer containing 50 mM sodium borohydride.

Periodate-treated PVDF membranes were washed with 0.2% (v/v) Tween 20 in PBS buffer and the blots incubated overnight at room temperature in the presence of anti-ZePrx34.70 rabbit IgGs, as primary antibodies, diluted 1:250 in PBS buffer containing 3% (w/v) decamed milk. Unbound primary antibodies were removed by washing in PBS-Tween 20 buffer and membranes were then incubated for 1 h at room temperature in PBS buffer containing 5% (w/v) skimmed milk and HRP-conjugated goat anti-rabbit IgG (diluted 1:12,000). Following the removal of unbound secondary antibody, peroxidase activity of HRP was revealed as described (Pomar et al., 2002).

Mr MALDI-TOF MS

Protein bands resolved by SDS-PAGE and stained with Coomassie Blue were excised, washed with 100 mM NH₄HCO₃, and treated with 3 mM DTT and 100 mM NH₄HCO₃, at 60°C for 30 min. Proteins were alkylated by the addition of 10 μL 100 mM iodoacetamide and further incubation for 30 min in the dark at room temperature (Kristensen et al., 1999). Gel pieces were washed in CH₃CN and 100 mM NH₄HCO₃, for 1 h, dehydrated in CH₃CN for 10 min, dried in vacuum, and reswelled with 10 μL of 25 mM NH₄HCO₃. Sinaipinic acid (Hewlett-Packard) was used as a matrix, and the mass spectra were recorded.
up to $10^8$ D in a linear positive mode, using an acceleration voltage of 28 kV in a MALDI-TOF MS system (G2025A; Hewlett-Packard) with a sampling rate of 200 MHz and a pressure of less than 30.39 x $10^{-3}$ Pa.

N-Terminal Microsequencing

The N-terminal amino acid sequences of both ZePrx33.44 and ZePrx34.70 were determined by transferring the SDS-PAGE purified bands onto a PVDF membrane, and deblocking the N terminus with pyrrolidine carboxyl peptide (pyrrolutamyl peptide, EC 3.4.19.3), which removes amino-terminal l-pyrrolidyl acid from peptides and proteins (Hirano et al., 1993; Tsuchasawa et al., 1998). For this, the PVDF membrane was excised and pretreated with 200 µL of 0.5% (w/v) polyanilypyrrolidone-40 in 100 mM acetic acid at 37°C for 30 min to block unbound protein-binding sites on the membrane. The membrane was washed 10 times with 1 mL of deionized water, and soaked of 0.1 M phosphate buffer, pH 8.0, containing 5 mM DTT and 10 mM EDTA. Pyrrolidine carboxyl peptide (5 µg) was added, and the reaction medium incubated at 30°C for 24 h (Hirano et al., 1993). The membrane was washed with deionized water and dried. ZePrxs were also deblocked at 30°C for 24 h in a homogenous aqueous phase containing 0.1 µM phosphate buffer, pH 8.0, 5 mM DTT, and 10 mM EDTA, by direct incubation of the proteins with pyrrolidine carboxyl peptide at a ratio 80:1 (w/w). After this incubation period, the proteins were subjected to N-terminal sequencing either directly or after purification by SDS-PAGE. In all cases, the N-terminal sequence was determined by automated Edman degradation using an Applied Biosystems Procise Sequencer.

Trypsin Digestion and MALDI-TOF MS

SDS-PAGE-resolved proteins were alkylated with iodoacetamide as described above. Gel pieces were dehydrated in CH3CN for 10 min, dried in vacuum, and rewelled with 10 µL of 25 mM NH4HCO3 containing 0.2 µg trypsin (EC 3.4.21.4). The digestion was carried out overnight at 37°C, after which the supernatant was saved and the peptides were extracted from the gel slices twice with CH3CN and 0.1% (v/v) trifluoroacetic acid. The supernatant and extracts were combined and dried in vacuum. The peptides released were reconstituted in 0.1% (v/v) trifluoroacetic acid, to which the matrix (α-cyano-4-hydroxy-trans-cinnamic acid [MALDI grade; Hewlett-Packard]) was added. The mass spectra of the peptides were recorded up to $10^8$ D in a linear positive mode, with an acceleration voltage of 28 kV using a MALDI-TOF MS system (G2025A; Hewlett-Packard), as described above.

DNA Isolation and Amplification

DNA was isolated from Z. elegans leaves using the DNasey plant mini system (Qiagen) according to the manufacturer. For PCR amplification of the ZePrx coding region of the intact and spliced oligonucleotides were designed (Izasa), whose nucleotide sequences were complementary to a 5’ region and to a 3’ region of the sequenced cDNAs. The sequence of PxZe-INI primer (5’-ATGATGATCATATAACGAGTG-GAA-3’) corresponded to the sense orientation of the first cDNA coding region, while the sequence of the PxZe-END primer (5’-CGCTACCACTCAGCTTAAACG-3’) corresponded to the antisense orientation of the last cDNA coding region. For PCR amplification of intron of type 2, the following primers were used: 11L (5’-TAATTTATTTTTCTTTTCAAT-3’) and 11R (5’-CTCACTACAGCGCCT-3’). For PCR amplification of intron of type 3, the following primers were used: 11L (5’-GAATTTATTTTTCTTTTCAAT-3’) and 11R (5’-CTCACTACAGCGCCT-3’). For PCR amplification of intron of type 4, the following primers were used: 11L (5’-GAATTTATTTTTCTTTTCAAT-3’) and 11R (5’-CTCACTACAGCGCCT-3’).

DNA was isolated from Z. elegans leaves using the DNasey plant mini system (Qiagen) according to the manufacturer. For PCR amplification of the ZePrx coding region of the intact and spliced oligonucleotides were designed (Izasa), whose nucleotide sequences were complementary to a 5’ region and to a 3’ region of the sequenced cDNAs. The sequence of PxZe-INI primer (5’-ATGATGATCATATAACGAGTG-GAA-3’) corresponded to the sense orientation of the first cDNA coding region, while the sequence of the PxZe-END primer (5’-CGCTACCACTCAGCTTAAACG-3’) corresponded to the antisense orientation of the last cDNA coding region. For PCR amplification of intron of type 2, the following primers were used: 11L (5’-TAATTTATTTTTCTTTTCAAT-3’) and 11R (5’-CTCACTACAGCGCCT-3’). For PCR amplification of intron of type 3, the following primers were used: 11L (5’-GAATTTATTTTTCTTTTCAAT-3’) and 11R (5’-CTCACTACAGCGCCT-3’). For PCR amplification of intron of type 4, the following primers were used: 11L (5’-GAATTTATTTTTCTTTTCAAT-3’) and 11R (5’-CTCACTACAGCGCCT-3’).

DNA Isolation and Amplification of ZePrx cDNAs

For PCR amplification of truncated ZePrx cDNAs (865 bp), two oligonucleotide (one degenerated and one specific, both shaded in yellow in Fig. 6) probes were designed and synthesized (Izasa). The nucleotide sequences were designed to be complementary to the coding strand for the peptide sequences VRLCNGPN, present in the sequence near the C terminus, and LSTFYD, present in the N-terminal sequence of ZePrx. The sequence of the N-Ze primer was (5’-TTTACCACTTTTACAG-3’) and 11R (5’-CTCACTACAGCGCCTA-3’), which is desoxyinosine, S is C/G, W is A/T, Y is C/T, and Y-Q is 3’ C/T), and corresponded to the sense orientation of LSTFYD. The sequence of Ze266r (5’-GATTACGGCAAA-GAGTCTTC-3’) corresponded to the antisense orientation of VRLCNGPN.

The mass spectra of the peptides were recorded up to $10^4$ D in a linear positive mode, using an acceleration voltage of 28 kV in a MALDI-TOF MS system (G2025A; Hewlett-Packard), as described above.

The proteolysis conditions involved an initial denaturation step at 94°C for 2 min, followed by 37 cycles with a 1-min denaturing step at 94°C and a 1-min elongation step at 72°C. The annealing temperatures were 51°C for two cycles, from 51°C to 48°C in six cycles, and 48°C for the remaining 29 cycles. A final extension step at 72°C for 5 min was included. Eight hundred- to 900-bp PCR products were cloned into pCRII-TOPO vector (Invitrogen), transformed into Escherichia coli competent cells (TOP10F®), and sequenced on an Applied Biosystems 373S DNA sequencer with the SP6 (5’-GATTITAGGGCTACTAGTG3’) and 17 (5’-TAAATACGCTCTATAGCG-3’) primers (Izasa).

The 5’ and 3’ flanking regions of the complete ZePrx cDNAs were obtained by 5’-RACE and 3’-RACE using the RACE kit from CLONTECH, according to the manufacturer. 5’-RACE PCR was performed using the RCPxZe-R primer (5’-AGGTGTCGCAAGGCTACTGTTAC-3’, shaded in green in Fig. 6), and the universal primer A mix (UPM, 5’-CTATACGACTC-CAATAGCGCAACGGCTTGATGGGGG-3’). The kit’s 3’-RACE PCR was performed using the RCPxZe-L primer (5’-GTTGACTCCGGTTCCGGTTGTC-3’, shaded in green in Fig. 6) and the UPM. The PCR conditions involved an initial denaturation step at 94°C for 2 min, followed by 38 cycles with a 15-s denaturing step at 94°C, and a 1.5-min elongation step at 72°C. The annealing temperatures were 66°C for two cycles and 64°C for the remaining 34 cycles. A final extension step at 72°C for 5 min was included. Select PCR products were cloned into pCR-II-TOPO vector (Invitrogen), transformed into E. coli competent cells (TOP10F®), and sequenced on an Applied Biosystems 373S DNA sequencer, with the SP6 and M13D3r (5’-CCGACTACGTTTACGATGACGAG-3’) primers (Izasa).

Sequence homology analysis was carried out using algorithms of ClustalW, ALING, and Transeq (http://www.ebi.ac.uk), the PROSITE (http://www.expasy.org/prosite), the software Molecular Evolutionary Genetics Analysis (version 2.1), Masatoshi Nei (http://www.megasoftware.net) (Kumar et al., 2001), and Chromas (version 1.45; http://www.technelysium.com.au).
Ze666 (5′-TCGACAAACAATCTACTAGG-3′), RCPxZe-L2 (5′-GCAAAGTACGATCACACGAGC-3′), RCPxZe-R3 (5′-CATCGGAGCCTGTGCTCCTGGA-3′), RCPxZe-R2 (5′-GCAAATGAGCATTACACGAGC-3′), SP6, and T7 primers.

Chemicals
Coniferyl alcohol, sinapyl alcohol, HRP-conjugated goat anti-rabbit IgGs (A9169), and 1-pyrrolinecarboxyl peptide (EC 3.4.19.3) were purchased from Sigma. The rest of the chemicals were obtained from various suppliers and were of the highest purity available. p-Coumaryl alcohol was generously provided by Hoon Kim and John Ralph (U.S. Dairy Forage Research Center).

Sequence data from this article can be found in the GenBank/EMBL data libraries under accession numbers P84332, P84333, AJ880392 to AJ880395, AJ971430, and AJ971431.

Received August 9, 2005; revised September 1, 2005; accepted September 12, 2005; published October 28, 2005.

LITERATURE CITED


Plant Physiol. Vol. 139, 2005

Downloaded from on September 15, 2017 - Published by www.plantphysiol.org
Copyright © 2005 American Society of Plant Biologists. All rights reserved.
Oda Y, Mimura T, Hasezawa S (2005) Regulation of secondary cell wall
Ros Barcelo´ A, Ostergaard L, Teilum K, Mirza O, Mattsson O, Petersen M, Welinder KG,
Osakabe K, Tsao CC, Li L, Popko JL, Umezawa T, Carraway DT, Smeltzer
Rasmussen SK, Welinder KG, Hejgaard J (1991) cDNA cloning, charac-
Roberts K, McCann MC, Ostergaard L, Abelskov AK, Mattsson O, Welinder KG
Quiroga M, Guerrero C, Botella MA, Barcelo´ A, Amaya I, Medina MI,
Gabaldón et al.

Osakabe K, Tsao CC, Li L, Popko JL, Umezawa T, Carraway DT, Smeltzer