A Gymnosperm Extensin Contains the Serine-Tetrahydroxyproline Motif

Cynthia Fong, Marcia J. Kieliszewski*, Renate de Zacks, Joseph F. Leykam, and Derek T. A. Lamport

Halfhollow Hills East High School, Dix Hills, New York 11746 (C.F.); Complex Carbohydrate Research Center, The University of Georgia, Athens, Georgia 30602 (M.J.K.); Michigan State University-Department of Energy Plant Research Laboratory (R.Z., D.T.A.L.) and Biochemistry Department, (J.F.L., D.T.A.L.), Michigan State University, East Lansing, Michigan 48823

ABSTRACT

The extensin family is a diverse group of hydroxyproline-rich glycoproteins located in the cell wall and characterized by repetitive peptide motifs glycosylated to various degrees. The origin of this diversity and its relationship to function led us earlier to compare extensins of the two major groups of angiosperms from which we concluded that the highly glycosylated Ser-Hyp motif was characteristic of advanced herbaceous dicots, occurring rarely or not at all in a representative gymnosperm monocot (Zea mays) and a chenopod (Beta vulgaris) representative of primitive dicots. Because these results could arise either from loss or acquisition of a characteristic feature, we chose a typical gymnosperm representing seed-bearing plants more primitive than the angiosperms. Thus, salt eluates of Douglas fir (Pseudotsuga menziesii) cell suspension cultures yielded two monomeric extensins differing in size and composition. The larger extensin reported earlier lacked the Ser-Hyp motif, was rich in proline and hydroxyproline, and contained peptide motifs similar to the dicot repetitive proline-rich proteins. The smaller extensin monomer reported here (Superose-6 peak 2 [SP2]) was compositionally similar to typical dicot extensins such as tomato P1, mainly consisting of Hyp, Thr, Ser, Pro, Val, Tyr, Lys, His, abundant arabinosine, and a small but significant galactose content. A chromatographic peptide map (on Hamilton PRP-1) of anhydrous hydrogen fluoride-deglycosylated SP2 yielded eight peptides sequenced after further purification on a high-resolution fast-size-exclusion column (polyhydroxyethyl aspartamide; Poly LC). Significantly, two of the eight peptides contained the Ser-Hyp motif, consistent with both the SP2 amino acid composition as well as the presence of hydroxyproline tetraarabinoside as a small (4% of total Hyp) component of the hydroxyproline arabinoside profile; thus, hydroxyproline tetraarabinoside corroborates the presence of Ser-Hyp, in agreement with our earlier observation that Hyp contiguity and Hyp glycosylation are positively correlated. Interestingly, other peptide sequences indicate that SP2 contains motifs such as Ser-Hyp-Thr-Hyp-Tyr, Ser-Hyp-Lys, and (Ala-Hyp), repeats that are related to and typify dicot extensins P1, P3, and arabinogalactan proteins, respectively. Overall, these peptide sequences confirm our previous prediction that Ser-Hyp is indeed an ancient motif and also strongly support our suggestion that the extensins comprise an extraordinarily diverse, but nevertheless phylogenetically related, family of cell wall hydroxyproline-rich glycoproteins.

Recently, we isolated a chenopod “split-block” extensin in which the Ser-Hyp motif characteristic of known dicot extensins was split by a short insertion sequence (18). Thus, we proposed that Ser-Hyp was not an immutable motif.

Subsequent work showed that THRGP2 and His-Hyp-rich glycoprotein extensins of the graminaceous monocot Zea mays (9–11) also lack the Ser-Hyp motif, with a single exception at the THRGP C-terminus (11, 27), but contain related motifs, such as Ser-Hyp-Lys-Pro-Hyp (11), Ser-Hyp-Hyp-His, and Ala-Hyp-Hyp-Hyp-Hyp (9). Hence, we mooted the possibility that extensins containing Ser-Hyp blocks might be more representative of advanced herbaceous dicots rather than primitive dicots and advanced graminaceous monocots. On the other hand, if we consider possible very early angiosperm origins and monocot-dicot divergence (20, 28), then the single C-terminal Ser-Hyp, of maize THRGP may mean that Ser-Hyp blocks are a conserved primitive feature; if that were so, we predicted their presence in gymnosperm extensins (11).

A cell suspension culture isolated from Douglas fir, Pseudotsuga menziesii (Mirbel), yielded two salt-elutable cell surface THRGP identified as monomeric extensins differing in size (8). The larger monomer was a PHRGp that lacked Ser-Hyp, but it did contain the general repetitive motif Pro-Hyp-X-Y-Lys similar to that of the repetitive Pro-rich proteins of dicots (2, 5, 7, 8, 19). Significantly, PHRGp differed by being lightly glycosylated (8). Because PHRGp lacked Ser-Hyp, we investigated the smaller extensin monomer, which we designate SP2 because of its elution position on a Superose-6 gel filtration column. Here, we report that a peptide map of HF-deglycosylated SP2 yielded peptides that do indeed contain the Ser-Hyp motif, which is therefore not exclusive to advanced dicots, but is a rather more general feature of seed-bearing plants.

MATERIALS AND METHODS

Preparation of Douglas Fir Crude HRGP by Intact Cell Elution

We prepared batches of crude Douglas fir (Pseudotsuga menziesii) HRGP from 6- or 7-d cultures involving salt elution of cell surface proteins as described earlier (25).

1 Supported by the U.S. Department of Agriculture grant No. 91–37304–6510 and the U.S. Department of Energy grant No. DE-FG02–90ER2–0021.

2 Abbreviations: THRGP, Thr-Hyp-rich glycoprotein; AGP, arabinogalactan protein; HF, anhydrous hydrogen fluoride; HRGP, Hyp-rich glycoprotein; IDT, isodiotyrosine; MeCN, acetonitrile; PHRGp, Pro-Hyp-rich glycoprotein; SP2, Superose peak 2.
Purification of SP2

We size fractionated the crude HRGP on a preparative Superose-6 (Pharmacia) gel filtration column as described earlier (8), collecting peak 2 (SP2, Fig. 1) for further fractionation on a semipreparative polysulfoethyl aspartamide column (200 × 9.4 mm, 5 μm, 200 Å pore; Poly LC). We applied 2 to 10 mg of Superose-fractionated SP2 dissolved in start buffer (10 mM NaH2PO4, brought to pH 3 with concentrated phosphoric acid, 10% MeCN), to a polysulfoethyl aspartamide column equilibrated in start buffer and eluted at 1.5 mL/min with a buffered (start buffer) 0 to 1 M NaCl gradient (Fig. 2). We further purified polysulfoethyl aspartamide SP2 on a Hamilton PRP-1 reverse phase column equilibrated in 0.1% TFA (buffer A) and eluted with a 0 to 50% gradient of buffer B (0.1% TFA/80% MeCN) at a flow rate of 0.5 mL/min, monitoring the absorbance at 220 nm (Fig. 3).

Hyp Assay and Hyp Arabinoside Profile

We determined SP2 Hyp content after hydrolysis (6 N HCl, 110°C, 18 h) of 100 μg SP2 preparations as described earlier (8), involving alkaline hypobromite oxidation and subsequent coupling with acidic Ehrlich’s reagent and monitoring at A560 nm (12). We determined Hyp arabinosides after alkaline hydrolysis (0.44 N Ba(OH)2, 18 h, 105°C) of Douglas fir SP2 and careful neutralization with concentrated H2SO4, followed by centrifugation and freeze drying of the supernatant fraction. We redissolved the lyophilate in distilled water and applied 200 μL, containing 100 to 200 μg Hyp, to a 75-× 0.6-cm Chromobeads C2 column equilibrated with water and eluted with a 0 to 0.5 N HCl gradient (17), and we monitored the reaction at A560 nm.

Amino Acid and Sugar Analysis

We determined amino acid compositions using a Pickering cation exchange column (3 mm i.d. × 150 mm) eluted by Pickering buffers A, B, and C. Postcolumn fluorometric detection involved NaOCl oxidation and O-phthalaldehyde coupling, which allow Hyp and Pro detection (29). Data capture was by a Compaq 386 with Nelson Turbochrom II software. We analyzed sugars as their alditol acetates by GLC (1) using a 6-foot × 2-mm i.d. PEG-succinate 224 column programmed from 130 to 180° at 4°/min for neutral sugars, with data capture via Turbochrom II.

HF Deglycosylation

We deglycosylated 19 mg SP2 in a microapparatus (22) containing 1 to 2 mL HF and 200 μL anhydrous methanol (21) for 1 h at 0°C, and quenched the reaction by slowly...
pouring into 29 mL of stirred ice-cold water, followed by dialysis and freeze drying. We recovered 46.2% of the initial HRGP as the deglycosylated product.

**Digestion of Deglycosylated SP2 with Chymotrypsin**

We digested 8.5 mg of deglycosylated SP2 with chymotrypsin ( Worthington) overnight at room temperature (2% ammonium bicarbonate [pH 8], 10 mM CaCl2; substrate to enzyme ratio was 100:1).

**Peptide Mapping and Sequence Analysis**

We obtained chymotryptic peptide maps of HF-deglycosylated SP2 via reverse phase HPLC on a Hamilton PRP-1 (4.1 mm i.d. × 150 mm) column using programmed gradient elution (0.5 mL/min) with the following mobile phase solvents: A = 0.1% TFA and B = 0.1% TFA in 80% MeCN. The gradient began at 100% A and increased (0.5%/min) from 0 to 50% B in 100 min, as described previously (8). Absorbance was monitored at 220 and 280 nm on a Hewlett-Packard photodiode array spectrophotometer. After initial peptide fractionation on PRP-1, peptides were size fractionated on a polyhydroxyethyl aspartamide column in gel filtration mode: 200 mM Na2SO4, 5 mM KH2PO4, 12% (v/v) MeCN, pH 3. The size-fractionated peptides were desalted on PRP-1 before sequencing by Edman degradation (3).

**RESULTS AND DISCUSSION**

Purification of SP2 extensin monomers from the crude intact cell eluate of Douglas fir cell suspension cultures involved gel filtration on preparative Superose-6 followed by polysulfonetyl aspartamide and HPLC on Hamilton PRP-1 (Figs. 1-3). The cells yielded 430 to 588 μg TCA-soluble crude HRGP/g cells fresh weight, of which approximately 30% was SP2. Glycosylated SP2 gave a single fuzzy band (molecular mass approximately 43 kD) on SDS-PAGE (Fig. 4). Amino acid and sugar analyses of SP2 (Tables I-III) were, with the exception of the elevated Pro content, quite similar to those of P1-type dicot extensins. In view of the approximate equimolar Pro to Lys content and the frequent occurrence of Lys-Pro in some extensins, we selected chymotrypsin rather than trypsin for peptide mapping (Fig. 5) of the HF-deglycosylated SP2. Sequence analysis of eight major peptides (Table IV) shows that SP2 contains several motifs present in different members of the extensin family, supporting our suggestion (9) of a phylogenetic series ranging from basic relatively lightly glycosylated extensins to the acidic highly glycosylated AGPs (6), as follows.

**Table I. Amino Acid Compositions of Douglas Fir PHRGp, SP2, and Tomato P1**

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Douglas Fir PHRGp</th>
<th>Douglas Fir SP2</th>
<th>Tomato P1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mol %</td>
<td>mol %</td>
<td>mol %</td>
</tr>
<tr>
<td>Hyp</td>
<td>28.1</td>
<td>29.2</td>
<td>32.7</td>
</tr>
<tr>
<td>Asx</td>
<td>0.0</td>
<td>0.0</td>
<td>1.4</td>
</tr>
<tr>
<td>Thr</td>
<td>1.5</td>
<td>4.5</td>
<td>6.2</td>
</tr>
<tr>
<td>Ser</td>
<td>2.5</td>
<td>6.8</td>
<td>9.8</td>
</tr>
<tr>
<td>Glx</td>
<td>0.0</td>
<td>2.9</td>
<td>1.5</td>
</tr>
<tr>
<td>Pro</td>
<td>21.3</td>
<td>14.2</td>
<td>9.6</td>
</tr>
<tr>
<td>Gly</td>
<td>0.6</td>
<td>2.0</td>
<td>1.7</td>
</tr>
<tr>
<td>Ala</td>
<td>0.0</td>
<td>4.1</td>
<td>2.9</td>
</tr>
<tr>
<td>Val</td>
<td>17.3</td>
<td>6.5</td>
<td>8.3</td>
</tr>
<tr>
<td>Ile</td>
<td>8.9</td>
<td>0.9</td>
<td>1.0</td>
</tr>
<tr>
<td>Leu</td>
<td>0.0</td>
<td>1.8</td>
<td>1.0</td>
</tr>
<tr>
<td>Tyr</td>
<td>4.9</td>
<td>4.6</td>
<td>7.7</td>
</tr>
<tr>
<td>Phe</td>
<td>0.0</td>
<td>0.7</td>
<td>0.0</td>
</tr>
<tr>
<td>Lys</td>
<td>11.0</td>
<td>14.9</td>
<td>9.5</td>
</tr>
<tr>
<td>His</td>
<td>1.2</td>
<td>6.0</td>
<td>6.1</td>
</tr>
<tr>
<td>Arg</td>
<td>1.9</td>
<td>0.6</td>
<td>0.7</td>
</tr>
</tbody>
</table>

* From ref. 8.  † From ref. 26.

<table>
<thead>
<tr>
<th>Sugar</th>
<th>SP2</th>
<th>PHRGp</th>
<th>P1†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mol %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Galactose</td>
<td>12</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>Arabinose</td>
<td>88</td>
<td>88</td>
<td>90</td>
</tr>
</tbody>
</table>

* Derived from ref. 8.  † From ref. 25.
First, the presence of the Ser-Hyp₄ motif in at least two peptides (peptides 5 and 7, Table IV) confirms our prediction that the Ser-Hyp₄ block is indeed an ancient motif, assuming an angiosperm gymnosperm divergence approximately 360 million years ago (28). However, Ser-Hyp₄ is not the dominant motif in SP2, which in that respect differs markedly from some dicot extensins and raises the difficult problem of accounting for the evolutionary ascent of the Ser-Hyp₄ motif, whose precise function is still open to speculation. Control of molecular flexibility is the prime candidate because increased Hyp content tends to rigidify the molecule (due to restricted rotation around phi), whereas arabinosylated Hyp residues may further stabilize the extended polypeptide backbone (15). Furthermore, Hyp contiguity and Hyp arabinosylation are highly correlated (8): single (i.e., noncontiguous) Hyp residues of soybean repetitive Pro-rich protein 1 are not glycosylated (2, 8) and extensins with di- and tripeptidyl Hyp tend to contain the mono-, di-, and triarabinosides (10, 11), whereas we tentatively associate the tetraarabinoside with the presence of tetrapeptidyl Hyp (8). That is, a small (4%) but significant amount of Hyp-tetraarabinoside (Table III) is consistent with the presence of the Ser-Hyp₄ as a minor rather than major component of SP2, in contrast to the presence of the tetraarabinoside as a major component in those dicot extensins in which Ser-Hyp₄ dominates (8, 13). If increased Hyp contiguity and glycosylation increase molecular rigidity then, as previously suggested, glycosylation conceivably acts as the thread of a molecular screw assisting reption of extensin into the wall matrix (16).

Second, various SP2 peptide sequences show intriguing identity or near identity to motifs that typify the dicot extensins P1 (26), P3 (14, 23, 24, 26), and AGPs (6). Thus, peptide 7 (Table IV) contains the Ser-Hyp-Hyp-Lys sequence previously observed only in P3-type extensins (14, 23, 24, 26), whereas peptide 13a (Table IV) resembles the canonical sequence Ser-Hyp-Hyp-Val-Tyr-Lys that characterizes dicot P1 extensin of tomato (26). In other words, SP2 shares characteristics of both dicot P1 and P3 extensins, and this indicates one way in which two radically different extensins such as P1 and P3 might diverge from a common ancestor. SP2 may also possibly possess the P3 IDT (4) motif Tyr-Tyr-Tyr-Lys; certainly the C-terminal Tyr-Tyr of peptide 7 (Table IV) hints at the possible Tyr-Tyr-Lys dicot IDT motif, whereas peptide 13b (Table IV) may contain a new IDT motif: Tyr-Pro-Tyr-Lys. Finally, the Ala-Hyp sequences of peptides 7 and 12 are also characteristic features of an AGP (6) and the graminaeous monocot His-Hyp-rich glycoprotein, which also has AGP characteristics (9). Again, this strongly supports our suggestion that the extensins comprise an extraordinarily diverse, but nevertheless phylogenetically related family of cell wall HRGs (9).

ACKNOWLEDGMENT

We thank Ms. Melanie Corlew of the Michigan State University Macromolecular Facility for assistance with peptide sequencing.

LITERATURE CITED

Protein Sequence Determination. Springer, New York, pp 211–275


