Hydroxyproline (Hyp) O-glycosylation characterizes the hydroxyproline-rich glycoprotein (HRGP) superfamily of the plant extracellular matrix. Hyp glycosylation occurs in two modes: Arabinosylation adds short oligoarabinosides (Hyp-arabinosides) while galactosylation leads to the addition of larger arabinogalactan polysaccharides (Hyp-polysaccharides). We hypothesize that sequence-dependent glycosylation of small peptide motifs results in glycomodules. These small functional units in combination with other repetitive peptide modules define the properties of HRGPs. The Hyp contiguity hypothesis predicts arabinosylation of contiguous Hyp residues and galactosylation of clustered noncontiguous Hyp residues. To determine the minimum level of Hyp contiguity that directs arabinosylation, we designed a series of synthetic genes encoding repetitive (Ser-Pro)n, (Ser-Pro)n, and (Ser-Pro)n. A signal sequence targeted these endogenous substrates to the endoplasmic reticulum/Golgi for post-translational proline hydroxylation and glycosylation in transformed Nicotiana tabacum cells. The fusion glycoproteins also contained green fluorescence protein, facilitating their detection and isolation. The (Ser-Pro)n and (Ser-Hyp)n fusion glycoproteins yielded Hyp-arabinosides but no Hyp-polysaccharide. The motif (Ser-Pro)n was incompletely hydroxylated, yielding mixed contiguous/noncontiguous Hyp and a corresponding mixture of Hyp-arabinosides and Hyp-polysaccharides. These results plus circular dichroic spectra of the glycosylated and deglycosylated (Ser-Pro)n, (Ser-Pro)n, and (Ser-Pro)n modules corroborate the Hyp contiguity hypothesis and indicate that Hyp O-glycosylation is indeed sequence-driven.

This paper is available online at http://www.jbc.org
Hydroxyproline O-Arabinosylation Codes

Fig. 1. Oligonucleotide sets used to construct the synthetic genes encoding (A) SP₂-EGFP and (B) SP₃-EGFP, and SP₄-EGFP. Internal repeat oligonucleotide sets were polymerized head-to-tail in the presence of the 5'-end linker sets. After ligation, the 3'-end linker sets were added. The genes were first subcloned into pUC18 as BamHI-SacI fragments (the restriction sites are boldfaced and italicized) and subsequently subcloned as XbaI-NcoI fragments into a pUC18-derived plasmid between a tobacco extensin signal sequence and the enhanced green fluorescence protein gene (EGFP, CLONTECH) as described earlier (12).

Hyp between the glycomodules of clustered noncontiguous Hyp also introduced arabinosides (12). To further test the predictive value of the Hyp contingency hypothesis and the likelihood that other small, conserved repeats direct Hyp glycosylation, such as the commonly occurring Xaa-Hyp motif of many AGPs, we designed another set of synthetic genes encoding the putative glycomodules. Here we describe the construction and expression of synthetic genes encoding the repetitive series: Ser-Pro-Pro (SP₂), Ser-Pro-Pro-Pro (SP₃), and Ser-Pro-Pro-Pro-Pro (SP₄), assuming that, targeted for secretion, they would be post-translationally hydroxylated in tobacco cells. Arabinosylation of about half of the Hyp residues in the dipetidyl Hyp blocks and almost 100% of the Hyp of the tetrapeptide Hyp blocks confirmed the predictive value of this simple contiguity code. However, the repetitive SP₃ motif of the tetrapeptide Hyp blocks confirmed the predictive value of this simple contiguity code. However, the repetitive SP₃ motif gave an expression product, nominally Ser-Hyp-Hyp-Hyp, but with incompletely hydroxylated Pro residues, which resulted in a mixture of contiguous and noncontiguous Hyp residues. Consistent with the Hyp contingency hypothesis, the corresponding Hyp-glycoside profile contained both Hyp-arabinosides and Hyp-polyarabinosides. Furthermore, circular dichroic spectra of the glycosylated and deglycosylated modules suggested that Hyp arabinosides facilitate the polyproline II conformation of HRGPs, whereas Hyp polyarabinosides favor a less ordered conformation.

Experimental Procedures

Synthetic Gene and Plasmid Construction—Construction of a given synthetic gene involved three sets of partially overlapping, complementary oligonucleotide pairs (Fig. 1) polymerized as described earlier (12, 13). The entire signal sequence-synthetic gene-enhanced green fluorescence protein (EGFP) constructs were then subcloned into the plant vector pBI121 (CLONTECH), as BamHI-SacI fragments in place of the tobacco extensin signal sequence and the enhanced green fluorescence protein gene (EGFP, CLONTECH) as described earlier (12).

Agrobacterium as described earlier (15). Transformed cells were grown on solid and liquid Schenk-Hildebrandt medium and selected for kanamycin resistance (12). EGFP fluorescence was visualized using a Molecular Dynamics Sarastro 2000 confocal laser-scanning fluorescence microscope equipped with an fluorescein isothiocyanate filter set comprising a 488-nm laser wavelength filter, a 510-nm primary beam splitter, and a 510-nm barrier filter.

Isolation of the Fusion Glycoproteins—Culture medium from transformed cells was harvested 16–21 days after subculture, concentrated by rotary evaporation, then dialyzed against water and concentrated again by rotary evaporation. Sodium chloride was added to a 2 M final concentration and 40–50 ml of the medium was injected onto a hydrophobic-interaction chromatography column (Phenyl-Sepharose 6 Fast Flow, 16 × 700 mm, Amersham Pharmacia Biotech) equilibrated in 2 M sodium chloride. We used a decreasing sodium chloride step gradient (2, 1, and 0 M, 150 ml of each) to elute the column at a flow rate of 1.3 ml/min. Collected fractions (4 ml) were monitored for fluorescence by a Hewlett-Packard 1100 Series flow-through fluorometer (488-nm excitation; 520–530 nm emission) or by a hand-held UV lamp (365 nm). The fluorescent fractions were pooled, concentrated by freeze-drying, redisolved in 1–2 ml of water, and injected onto a Hamilton semipreparative polymeric reverse phase column (Phenyl-Sepharose 6 Fast Flow, 305 mm) equilibrated with start buffer (0.1% aqueous trifluoroacetic acid). Proteins were gradient-eluted in 0.1% trifluoroacetic acid/80% (w/v) aqueous acetonitrile (0–70/120 min) at a flow rate of 0.75 ml/min. The fusion glycoprotein (Ser-Hyp)₄-EGFP and endogenous tobacco AGPs were isolated as described earlier (12). Protein sequence analysis was performed at the Michigan State University Macromolecular Facility on a 477-A Applied Biosystems Inc. gas phase sequencer.

Coprecipitation with β-Glucosyl Yariv Reagent—We assayed the ability of the fusion glycoproteins, as well as earlier reported (Ser-Hyp)₃-EGFP and tobacco AGPs (12), to coprecipitate with the β-glucosyl Yariv reagent (16). Absorbency was read at 420 nm.

Carbohydrate Analyses—Hyp-glycoside profiles were determined on 2–4 mg of isolated fusion glycoprotein as described earlier by Lamport et al. (17) and Shpak et al. (12). We monitored the automated postcolumn hydroxyproline assay at 560 nm. Neutral sugars were analyzed as alditol acetates (18) by gas chromatography using a 6-foot × 2-mm polyethylene glycol succinate 224 column programmed from 130 to 180° at 4 °C/min for neutral sugars. Data capture was achieved by Hewlett-Packard Chem station software. One hundred micrograms of glycoprotein was used for each analysis. We assayed the uronic acid content of 70% of each fusion glycoprotein via the specific colorimetric assay based on reaction with m-hydroxydiphenyl (19). Galacturonic acid was the standard.

Pronase Digestion of the Fusion Glycoproteins—Each fusion glycoprotein (10–20 mg, 10 mg/ml aqueous) was heat-denatured in boiling water for 2 min, cooled, and then incubated in an equal volume of...
freshly prepared 2% (w/v) ammonium bicarbonate containing 2.5 mM calcium chloride and Pronase (substrate:enzyme ratio 100:1, w/w). The digestion proceeded at room temperature overnight, then the peptides were freeze-dried, dissolved in 0.5 ml of Superoxide buffer, and separated via semipreparative Superoxide-12 gel filtration (see below).

**Superose-12 Gel Filtration Chromatography of Tobacco Cell Medium and Protolyzed Fusion Glycoproteins**—We injected 0.5 ml of tobacco medium from transformed cells (three lines of each construct) onto a Superose-12 analytical gel filtration column (10 × 300 mm) eluted at a flow rate of 0.2 ml/min and monitored via flow-through detection by a Hewlett-Packard 1100 series fluorometer (excitation, 488 nm; emission, 520 nm). The column was calibrated with molecular weight standards (bovine serum albumin, insulin, catalase, and sodium azide) and absorbency was read at 220 nm. Pronased fusion glycoproteins were injected onto a semipreparative Superose-12 column (16 × 500 mm) and eluted at a flow rate of 1 ml/min. For both the analytical and semipreparative Superose-12 columns, the elution buffer was 0.2 M sodium phosphate (pH 7.0) containing 0.01% sodium azide.

**Anhydrous Hydrogen Fluoride (HF) Deglycosylation**—We deglycosylated 4–5 mg of each fusion glycoprotein for 1 h in anhydrous HF as described earlier (20). The proteins were dialyzed against deionized, distilled water for 2 days at 4 °C and then freeze-dried. The glycomodules were isolated from Pronase digestion of the fusion glycoproteins as described above, then HF-deglycosylated 1 h at 4 °C. The HF was blown off under nitrogen gas, and the deglycosylated modules (designated dSP2, dSP3, dSP4) were then rerun on the reverse-phase column as described above.

**Circular Dichroism (CD)**—We recorded CD spectra of poly-L-hydroxyproline (5–20 kDa, Sigma Chemical Co.) and the isolated glycoprotein modules before and after deglycosylation on a Jasco-715 spectropolarimeter (Jasco Inc., Easton, MD). Spectra were averaged over two scans with a bandwidth of 1 nm, and step resolution was 0.1 nm. All spectra are reported in terms of mean residue ellipticity with the 180°-250-nm region using a 1-mm pathlength. The modules were dissolved in water at the following concentrations: poly-L-hydroxyproline, 18.4 mM; SP2, 9.2 mM; SP3, 19.6 mM; and SP4, 30.4 mM. Spectra were obtained from 18.4 μM of each deglycosylated module.

**RESULTS**

**Synthetic Gene and Plasmid Construction**—We built three plasmids, each encoding a tobacco signal sequence, a synthetic gene, and EGF placed in one transcriptional frame. The synthetic genes encoded an SP2-EGFP fusion protein containing 24 Ser-Pro repeats, an SP3-EGFP glycoprotein containing 15 Ser-Pro repeats, and an SP4-EGFP glycoprotein containing 18 Ser-Pro repeats. For brevity, from here forward the fusion glycoproteins encoded by the synthetic genes will be referred to as SP2-EGFP, SP3-EGFP, and SP4-EGFP, although most Pro residues were hydroxylated in the glycoproteins. The earlier reported fusion glycoprotein (Ser-Hyp)12-EGFP (12) will be referred to as SP6-EGFP.

**Tobacco Cell Transformation—Agrobacterium-mediated transformation of tobacco cell cultures gave stably transformed lines, judging by the fact that the cells continue to produce the gene products more than 1 year after transformation. We isolated three lines of each SP2-EGFP, SP3-EGFP, and SP4-EGFP by observing the characteristic fluorescence of EGF in both the growth medium and the cytoplasm of the transformed cells (not shown), which was similar to that published earlier (12). The three lines of each construct were assayed for the fusion glycoproteins by Superose-12 gel permeation chromatography. All lines produced fluorescent products that coeluted with the SP6-EGFP standard (Fig. 2A–D) as they eluted much earlier than predicted for the nonglycosylated products, which would have eluted near the EGF standards (E and F) (i.e. EGF is 27 kDa and the synthetic genes sans carbohydrate were estimated to be 34–37 kDa). Depending on the transgene expressed, the glycoproteins produced by separate culture lines yielded products that co-chromatographed. Controls included 10 μg of EGF standard from CLONTECH (F) and EGF targeted to the tobacco extracellular space by the tobacco extensin signal sequence (E).

**Carbohydrate Analyses**—Hyp-glycoside profiles of SP2-EGFP, SP3-EGFP, and SP4-EGFP showed that, although all the glycoproteins contained Hyp-arabinosides, SP2-EGFP glycoprotein also contained Hyp-arabinogalactan polysaccharides (Table II). Arabinose was the only saccharide component of SP2-EGFP and the major saccharide component of SP3-EGFP, although to a much lesser extent than endogenous tobacco AGPs or SP6-EGFP (Table I).
which contained a small amount of galactose as well (Table III). In contrast, the SP$_2$-EGFP contained mainly galactose and arabinose, with lesser amounts of rhamnose, glucose, and uronic acid (Table III). Saccharide accounted for 17% of SP$_2$-EGFP, 40% of SP$_3$-EGFP, and 41% of SP$_4$-EGFP on a dry weight basis.

**Anhydrous HF Deglycosylation—**Weight loss after HF deglycosylation agreed with the sugar analyses. Thus, from 3.9 mg of deglycosylated SP$_3$-EGFP (46% weight loss); from 4.8 mg of SP$_3$-EGFP we recovered 2.6 mg of deglycosylated SP$_3$-EGFP (46% weight loss); and 3.9 mg of SP$_4$-EGFP yielded 2.2 mg of deglycosylated SP$_4$-EGFP (44% weight loss).

**Amino Acid Analysis and Sequence Analysis—**Amino acid analyses (Table IV) showed that almost all the proline of the SP$_2$ and SP$_4$ glycomodules had been hydroxylated to form Hyp. However, 27% of the proline residues (20 mol % of the amino acids) in the SP$_3$ glycomodule remained nonhydroxylated. Edman degradation (Fig. 3) confirmed the identity of the gene products.

**Calculated Molecular Masses of the Fusion Glycoproteins—**Both SDS-polyacrylamide gel electrophoresis and gel permeation chromatography tend to underestimate HRGP molecular mass due to the biased amino acid compositions, extended structures, and extensive glycosylation (21). Therefore, we calculated the molecular masses of our fusion glycoproteins based on the gene sequences (Table II), amino acid compositions (Table IV), neutral sugar analyses (Table III), and protein sequences (Fig. 3). We calculated the following masses for the glycosylated and deglycosylated proteins: glycosylated SP$_2$-EGFP, 43.9 kDa; dSP$_2$-EGFP, 34.8 kDa; glycosylated SP$_3$-EGFP, 63–77 kDa; dSP$_3$-EGFP, 33.9 kDa; glycosylated SP$_4$-EGFP, 66.4 kDa; and dSP$_4$-EGFP, 37.3 kDa.

**Circular Dichroism of the Glycosylated and Deglycosylated Modules—**CD spectra of the deglycosylated modules (Fig. 4A) show that both negative (−206 nm) and positive (−226 nm) ellipticities increased in the order dSP$_2$ (light blue) < dSP$_3$ (lavender) < dSP$_4$ (dark green). The glycosylated SP$_2$ and SP$_3$ modules (orange and dark blue, respectively) had virtually identical spectra (Fig. 4B). Glycosylated SP$_3$ had a “random coil” conformation (Fig. 4, B and C, red), which adopted more structure after deglycosylation (Fig. 4C, light blue). The negative ellipticities of SP$_2$ and SP$_3$ increased markedly and shifted to a lower wavelength (−201 nm) when arabinose (Fig. 4, D and F, orange and light green, respectively), whereas SP$_3$ showed virtually no change in conformation after deglycosylation (Fig. 4E, dark blue). Samples containing Hyp-polsaccharide substituents had a second minimum at 180 nm (Fig. 4, A, B, C, and E, red and dark blue).

**DISCUSSION**

HRGP glycosylation involves proline-rich sequences targeted to the endoplasmic reticulum/Golgi for initial cotranslational hydroxylation of the proline residues followed by O-Hyp glycosylation. Generally, the multiplicity and similar properties of HRGPs make them difficult to purify. However, addition of the hydrophobic fluorescent EGFP reporter protein facilitated the chromatographic separation of HRGP-EGFP fusion proteins from the other endogenous, hydrophilic HRGPs. Thus, the design of new HRGPs composed of single glycomodule repeats and expressed as HRGP-EGFP fusion proteins provided bulk quantities of the glycomodules for structural analysis. This approach has allowed us to elucidate the major determinants of Hyp glycosylation and currently is enabling us to determine the detailed structures of the glycosyl substituents. Moreover, as endogenous substrates, these new HRGPs containing only simple repeats can define the enzyme specificity of prolyl hydroxylase(s) and glycosyl transferases under in vivo conditions and prior to isolation of the enzymes.

Because proline/hydroxyproline-rich polypeptides often adopt characteristic conformations (Fig. 4) one must ask whether it is primarily peptide conformation or peptide sequence that directs hydroxylation and glycosylation.

Prolyl hydroxylase does not hydroxylate all HRGP proline residues, the most notable example being the repetitive “insertion sequence” Val-Lys-Pro-Tyr-His-Pro of tomato P1 extensin. Indeed, in all known HRGP sequences, Lys-Pro always occurs nonhydroxylated, whereas His-Pro is hydroxylated in some HRGPs (22) but not in others (23), and Pro-Val is invariably hydroxylylated (1). The Pro residues in the SP$_2$ and SP$_4$ modules (Fig. 3 and Table IV) show almost complete hydroxylation of SP$_3$ proline residues indicates that the prolyl hydroxylase of *Nicotiana* has a low affinity for a tripeptidyl proline substrate. Not surprisingly, such sequences are rare in *Nicotiana* (24), although common in other species; for example, repetitive Ser-Hyp$_3$ occurs in the gum arabic glycoprotein of *Acacia senegal* (25) and in maize extensins (26). Likewise, Ser-Pro$_3$ motifs that may be completely hydroxylated occur frequently in potato cDNAs (27) and in an *Arabidopsis* extensin gene (28).

The above observations are consistent with sequence-specific hydroxylation rather than the earlier view of plant prolyl hydroxylase as specific for the polyproline II conformation (29). This included protocollagen (30), which has a polyproline II conformation, although collagen expressed in trans-
genic tobacco was not hydroxylated (31). Plant prolyl hydroxylase is therefore similar to that of animals in being sequence-specific, although it hydroxylates distinctly different sequences. It is also possible that plants, like some animals, have multiple prolyl hydroxylases (1) possessing a catalytic domain separate from a sequence-specific peptide substrate binding domain (32).

Other factors may also influence proline hydroxylation of a given substrate. These may include differences in the enzyme specificity of different plant species or the targeting of potential substrates to specific endoplasmic reticulum subdomains, as reported for rice storage proteins (33). This may explain why the repetitive Pro-Pro-Pro-Val-His-Leu motif of zein, the maize endosperm seed storage protein that shares sequence identity with the PRP HRGPs, is not hydroxylated (34).

Although the data show that Hyp is the major glycosylation site of HRGPs, it does not rule out the glycosylation of other hydroxyamino acids. Certainly single residues of galactose occur as O-galactosylserine in the Ser-Hyp ό glycomodules of extensin HRGPs (10, 35) and may account for the small amount of galactose in the SP4 glycoprotein (Table III), which contains Hyp-arabinosides but no Hyp-polysaccharide. Speculatively, one can suggest that, in concert with the Hyp-arabinosides, galactosylserine stabilizes the Ser-Hyp glycomodule.

Serine is often considered as a polysaccharide attachment site in the AGPs (36). Although some of the evidence is strongly suggestive (22, 37), it is not definitive (38) and is sometimes contradictory (38–41). On the other hand, addition of arabinogalactan polysaccharide to the clustered noncontiguous Hyp residues of (Ser-Hyp)$_n$, as previously demonstrated (12), is a natural corollary of the Hyp contingency hypothesis. Other clusters can also be defined by the design of synthetic gene constructs that test the efficacy of AGP motifs, such as (Ala-Pro)$_n$, (Thr-Pro)$_n$, and (Ala-Thr-Pro)$_n$, to direct polysaccharide addition. Identification of a sequence motif that directs polysaccharide addition to serine or threonine residues might also be a possible outcome of this work in progress.

Because hydroxyproline residues can exist in any one of three states, nonglycosylated, arabinosylated, and galactosylated, a sequence code seems more likely than a purely conformational control of glycosylation. The Hyp contingency hypo-

### Table III

<table>
<thead>
<tr>
<th>Glycosyl residue</th>
<th>SP-EGFP* (mol %)</th>
<th>SP$_2$-EGFP (mol %)</th>
<th>SP$_3$-EGFP (mol %)</th>
<th>SP$_4$-EGFP (mol %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ara</td>
<td>28</td>
<td>50</td>
<td>42</td>
<td>95</td>
</tr>
<tr>
<td>Gal</td>
<td>45</td>
<td>0</td>
<td>37</td>
<td>7</td>
</tr>
<tr>
<td>Rha</td>
<td>6</td>
<td>0</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Uronic acid$^a$</td>
<td>19</td>
<td>0</td>
<td>10</td>
<td>Trace</td>
</tr>
<tr>
<td>Glc</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

$^a$ From Ref. 12.

### Table IV

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>SP$_2$ line A (mol %)</th>
<th>SP$_3$ line A (mol %)</th>
<th>SP$_4$ line A (mol %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ser</td>
<td>33</td>
<td>25</td>
<td>20</td>
</tr>
<tr>
<td>Hyp</td>
<td>60</td>
<td>55</td>
<td>78</td>
</tr>
<tr>
<td>Pro</td>
<td>7</td>
<td>20</td>
<td>2</td>
</tr>
</tbody>
</table>

**Fig. 3.** N-terminal partial amino acid sequences of SP$_2$-EGFP, SP$_3$-EGFP, and SP$_4$-EGFP. A, SP$_2$-EGFP contained 47 Hyp residues and 3 Pro residues, one of which occurred at the N terminus judging by this sequence as well as the gene sequences (not shown) and the amino acid composition of the module (Table IV). B, Edman degradation of SP$_2$-EGFP indicated incomplete hydroxylation at each Pro residue except at the extreme N terminus. Thus the SP$_2$-EGFP was a population of molecules containing a mixture of contiguous and noncontiguous Hyp residues. X denotes a blank cycle that yielded no signal during Edman degradation. The SP$_3$ glycomodule contained about 28 Hyp residues and 22 Pro residues overall. C, Edman degradation of SP$_2$-EGFP indicated the N terminus was completely hydroxylated and the amino acid composition of the isolated SP$_3$ module indicated it contained 74 Hyp residues and 2 Pro residues.
thesis predicts the arabinosylation of contiguous Hyp residues, where contiguity begins with dipeptidyl Hyp (8), confirmed here by the arabinosylation mainly of a single Hyp residue of each SP2 module (Table II). Interestingly, not only the number of arabinosylated Hyp residues increase with the size of the contiguous Hyp block (about 50% in Ser-Hyp and nearly 100% in Ser-Hyp3), but also the size of the attached arabinooligosaccharide. Thus, the small amounts of penta-arabinoside occasionally reported (42) might be attributable to the uncommon Ser-Hyp6-6 motif.

Although both the SP3 and SP4 modules were arabinosylated, there was anomalous addition of polysaccharide to SP2, apparently due to incomplete hydroxylation of the proline residues resulting in some clustered noncontiguous Hyp. However, the polyproline II content of deglycosylated SP2 is higher than that of the deglycosylated SP4 and much higher than that of the deglycosylated SP (Fig. 4A). But SP2 contains both Hyp-Ara and Hyp-Gal, whereas SP4 contains Hyp-Ara exclusively, and SP (with the lowest polyproline II content) contains Hyp-Gal exclusively. Thus the CD data support sequence-directed Hyp-glycosylation rather than a simple conformational control based on polyproline II content.

The putative Hyp glycosyltransferases clearly distinguish between arabinosylation of contiguous Hyp and galactosylation of clustered noncontiguous Hyp residues. Thus it is now possible to predict the approximate glycosylation profile of an HRGP based on its primary structure, a crucial step toward predicting three-dimensional native structure. However, two exceptions suggest that certain flanking sequences may modify the simple code by suppressing the glycosylation of typical AGP motifs that occur in extensins. For example, the THRGP motif did not affect the polyproline II content, presumably because Hyp-arabinosylation favors polyproline II conformation whereas the Hyp-poly saccharide opposes it. Polysaccharide addition also resulted in spectra having a new minimum at 180 nm (Fig. 4, B, C, and E) contributed by the polysaccharide substituent itself (Fig. 4).

Overall, these CD data confirm that Hyp-arabinosides enhance the polyproline II helix in the Ser-Hyp glycomodules of extensin HRGPs (5, 46, 47) as they do in the Ser-Hyp2 and Ser-Hyp4 modules here. Presumably, this underlies the contribution of extensin to the tensional integrity of the cell wall itself. On the other hand, judging from the CD spectra, Hyp-polysaccharide does not contribute appreciably to the polypeptide conformation, although the 1–6-linked side chains should greatly enhance the water-holding capacity of AGPs (48). Hyp-polysaccharide is also a prominent constituent of glypiated AGPs anchored to the plasma membrane where they make a major contribution (49) as a periplasmic hydrophilic buffer between plasma membrane and cell wall.

Acknowledgments—We thank Li Tan for help with protein purification and Dr. Derek Lamport for comments on the manuscript.

REFERENCES

Hydroxyproline O-Arabinosylation Codes