PERGAMON

Phytochemistry 54 (2000) 99–106

PHYTOCHEMISTRY

www.elsevier.com/locate/phytochem

Gum arabic glycoprotein contains glycomodules of both extensin and arabinogalactan-glycoproteins

Leslie J. Goodrum[a], Amar Patel[b], Joseph F. Leykam[c], Marcia J. Kieliszewski[d,*]

[a]*Ohio State University School of Ophthalmology, Columbus, OH, USA*
[b]*Medical College of Georgia, Augusta, GA 30909, USA*
[c]*Biochemistry Department, Michigan State University, East Lansing, MI 48824, USA*
[d]*Department of Chemistry, Ohio University, Athens, OH 45701, USA*

Received 30 September 1999; received in revised form 6 December 1999

Abstract

Gum arabic glycoprotein (GAGP) is a large molecular weight, hydroxyproline-rich arabinogalactan-protein (AGP) component of gum arabic. GAGP has a simple, highly biased amino acid composition indicating a repetitive polypeptide backbone. Previous work (Qi, W., Fong, C., Lamport, D.T.A., 1991. Plant Physiology 96, 848), suggested small (~11 residue) repetitive peptide motifs each with three Hyp-arabinoside attachment sites and a single Hyp-arabinogalactan polysaccharide attachment site. We tested that hypothesis by sequence analysis of the GAGP polypeptide after HF-deglycosylation. A family of closely related peptides confirmed the presence of a repetitive 19-residue consensus motif. However, the motif: Ser-Hyp-Hyp-Hyp-Thr-Leu-Ser-Hyp-Ser-Hyp-Thr-Hyp-Thr-Hyp-Hyp-Leu-Gly-Pro-His, was about twice the size anticipated. Thus, judging by Hyp-glycoside profiles of GAGP, the consensus motif contained six Hyp-arabinosides rather than three and two Hyp-polysaccharides rather than one. We inferred the glycosylation sites based on the Hyp contiguity hypothesis which predicts arabinosides on contiguous Hyp residues and arabinogalactan polysaccharides on clustered non-contiguous Hyp residues, i.e. the GAGP motif would consist of arabinosylated contiguous Hyp blocks flanking two central Hyp-polysaccharides. We predict this rigidifies the glycoprotein, enhances the overall symmetry of the glycopeptide motif, and may explain some of the remarkable properties of gum arabic.

Keywords: *Acacia senegal*; Leguminosae; Gum arabic glycoprotein; Hydroxyproline-rich glycoprotein; Arabinogalactan-protein; Extensin; Exudate gum

1. Introduction

Gummosis is a common wound response that results in the exudation of a plastic gum sealant at the site of cracks in bark (Stephen et al., 1990). The exudate is a composite of polysaccharides and glycoproteins structurally related to cell wall components such as galactans (Aspinall, 1970) and hydroxyproline-rich glycoproteins (Anderson and McDougall, 1988; Anderson and Weiping, 1990; Anderson et al., 1990). Gum arabic is probably the best characterized of these exudates and appears to be a complex mixture (Osman et al., 1993, 1995; Islam et al., 1997). In essence, however, it consists of two major components, a microheterogeneous rhamnoglucuronoarabinogalactan polysaccharide (Osman et al., 1993, 1995) and a higher molecular weight hydroxyproline-rich glycoprotein (Qi et al., 1991; Akiyama et al., 1984; Connolly et al., 1987) of commercial interest because of its flavor emulsifying and stabilizing properties (Islam et al., 1997, Ray et al., 1995; Prakash and Mangino, 1990; Randall et al., 1988). This raises questions about the underlying molecular basis for such different phenomena as plastic

* Corresponding author. Tel.: +1-740-593-9466; fax: +1-740-593-0148.
E-mail address: kielisze@helios.phy.ohiou.edu (M.J. Kieliszewski).

0031-9422/00/$ - see front matter
PII: S0031-9422(00)00043-1

gum sealants and flavor emulsification, hence the need to elucidate the molecular structure of the gum arabic glycoprotein (GAGP), in order to understand its structural biology.

From the highly biased amino acid composition of GAGP, Qi et al. (1991) derived a simple empirical formula, Hyp_4 Ser_2 Thr Pro Gly Leu His, indicating a periodic structure for the polypeptide backbone. In the proposed model, GAGP consists of many small (~11 residues) repetitive peptide units, each containing several arabinosylation sites and a single polysaccharide attachment site. As an essential first step towards testing the model, we generated and sequenced peptides of GAGP, and determined the glycosyl and linkage analysis of an isolated Hyp-polysaccharide. From these data, we propose a revised model for the major consensus repeat of GAGP.

2. Results and discussion

Using the method of Qi et al. (1991) GAGP was isolated via preparative Superose-6 gel filtration and subsequently deglycosylated using HF for chymotryptic peptide mapping. This gave a major peak (designated dGAGP) when further fractionated by reversed phase HPLC (Fig. 1). Amino acid analysis showed dGAGP had a highly biased but constant amino acid composition in fractions sampled across the peak (Table 1), indicating that dGAGP was a single polypeptide component sufficiently pure for sequence analysis. This was confirmed by the isolation of peptides (Table 2) similar in composition to one other and to the parent GAGP (Table 1).

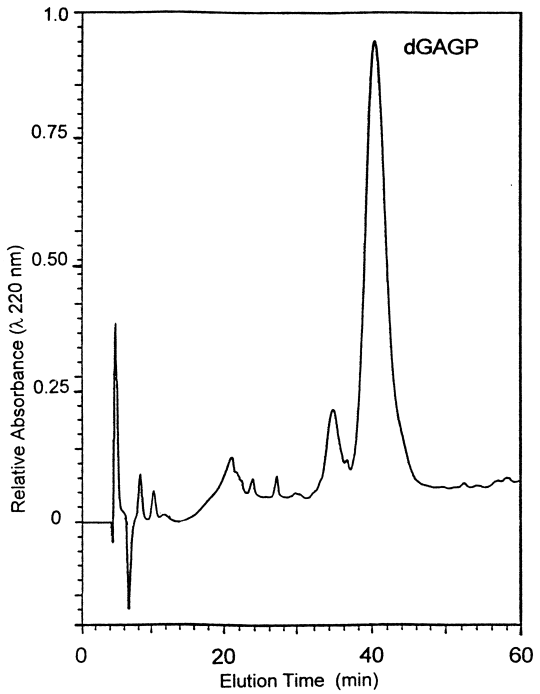

Fig. 1. Isolation of dGAGP by reverse phase of HPLC. Approximately 1–5 mg dGAGP was injected onto a Hamilton PRP-1 column and fractionated by gradient elution. The component at 35 min was a Hyp-poor contaminant.

Although native GAGP resists pronase digestion (Akiyama et al., 1984; Chikamai et al., 1996), yielding only large fragments of ~200 kDa (Connolly et al., 1988) preliminary work in Lamport's laboratory showed that exhaustive digestion with pronase effectively cleaved dGAGP to small peptides (Delonnay, 1993). However, the peptides lacked some of the amino acids present in the empirical formula: Hyp_4 Ser_2 Thr Pro Gly Leu His of the repeat motif suggested by Qi et al. (1991), most notably His (Table 2, peptide PH3G2.) Therefore we performed a partial pronase digestion of dGAGP. This gave two large major peptides P1 and P3 (Fig. 2) with partial sequences (Table 2) containing all of the amino acids in the empirical formula. We also digested dGAGP with chymotrypsin, which slowly cleaved leucyl and histidyl bonds, followed by a two-stage HPLC fractionation scheme. Initial separation of the chymotryptides on a PolySULFOETHYL A™ (PolyLC, Ellicott

Table 1
Amino acid compositions of glycosylated GAGP (GAGP) and deglycosylated GAGP (dGAGP) fractions obtained by reversed phase HPLC compared to dGAGP isolated by Qi et al. (1991)

		dGAGP Peak fractions[a]			
Amino acid[b]	GAGP	Ascending	Center	Descending	GAGP[c]
Hyp	40.0	38.4	36.7	36.3	36.9
Asx	0.0	0.0	0.0	0.0	1.6
Ser	22.2	21.6	21.6	22.5	19.4
Glx	0.0	0.0	0.0	0.0	1.9
Gly	4.5	4.8	4.4	4.3	6.4
His	6.6	8.7	8.2	8.4	7.1
Arg	0.0	0.0	0.0	0.0	0.0
Thr	10.2	10.6	12.2	11.4	8.8
Ala	1.2	0.7	0.8	1.0	1.3
Pro	8.0	7.6	8.3	8.1	6.8
Tyr	0.0	0.0	0.0	0.0	0.3
Val	0.0	0.0	0.0	0.0	0.8
Met	n.d.[d]	n.d.[d]	n.d.[d]	n.d.[d]	n.d.[d]
Lys	0.0	0.0	0.0	0.0	1.0
IIe	0.2	0.0	0.0	0.0	0.4
Leu	6.4	7.6	7.8	8.1	6.4
Phe	0.5	0.0	0.0	0.0	0.9
Trp	n.d.[d]	n.d.[d]	n.d.[d]	n.d.[d]	n.d.[d]
Cys	0.0	0.0	0.0	0.0	0.0

[a] To check peak homogeneity, we analyzed three consecutive fractions across the dGAGP peak, designated Ascending, Center, and Descending.
[b] Represented as mole percent.
[c] Data from Qi et al. (1991).
[d] Not determined.

Table 2
Pronase and chymotryptic peptide sequences from the dGAGP polypeptide backbone

Peptide	Sequence[a]
Pronase peptide	
P1	SOOOTLSOSOTOTOOOGPHSOOO(O)-
P3	SOOO(T/S)LSOSOTOTXOO-
PH3G2[b]	SOSOTOTOOOGP
Chymotryptic peptide	
S1P2	SOOOSLSOSOTOTOOTGPH
S1P3	SOOOOLSOSOTOTOOOGP-
S1P4	SOLPTLSOLP(A/T)OTOOOGPH
S1P5	SOOOOLSOSLTOTOOLGP-
S2P1	SOSOTOTOOOGPH
S2P2a	SOSOAOTOOLGPH
S2P2b	SOLPTOTOOLGPHS
S2P3	SOSOTOTOOLGPH
S2P4	SOOLTOTOOLLPH
Consensus[c]	SOOO(O/T/S)LSOSOTOTOO(O/L)GPH

[a] O denotes hydroxyproline in the peptide sequences; X denotes a blank cycle.
[b] From Delonnay, 1993.
[c] Derived from the major peptides P1, P3, S1P3, S1P5, S2P1 S2P3 and PH3G2.

City, MD) cation exchanger yielded two major fractions designated S1 and S2 (Fig. 3). Further chromatography on a Hamilton PRP-1 reversed phase column resolved fractions S1 and S2 into several peptides (Fig. 4(a) and (b)). Edman degradation showed that these chymotryptides were closely related to each other and to the pronase peptides (Table 2). These peptides reflect the overall amino acid composition of GAGP and can be related to a single 19-residue consensus sequence (Table 2). Thus, we conclude that GAGP possesses a highly repetitive polypeptide, albeit with minor variations in the sequence. Based on a linear GAGP molecule of 150 nm (Qi et al., 1991) and presuming the extended polyproline II helix present in both extensins and AGPs (Kieliszewski and Lamport, 1994; Nothnagel, 1997), we estimate that GAGP contains about 20 peptide repeats with occasional partial repeats. Partial repeats of the consensus sequence may account for the somewhat higher serine content in native GAGP compared to that in the consensus sequence. Addressing the possibility that polymorphic forms of GAGP occur with variations in the polypeptide backbone awaits the outcome of DNA cloning.

The 19-residue GAGP consensus motif contains ~nine Hyp residues and is roughly twice the size of that previously postulated (Qi et al., 1991) with only a single polysaccharide attachment site. Judging from the Hyp-glycoside profile of GAGP (Table 3) (Qi et al., 1991), about one in every five Hyp residues is polysaccharide-substituted. Thus, there are approximately two Hyp-polysaccharide sites in the consensus sequence, which raises the question as to the nature Hyp residues involved in polysaccharide attachment. The Hyp contiguity hypothesis (Kieliszewski and Lam-

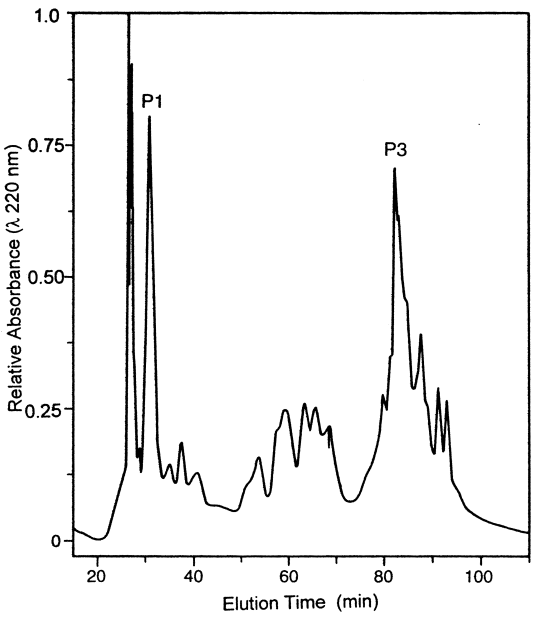

Fig. 2. Fractionation of the dGAGP incomplete pronase digest by reversed phase HPLC. An incomplete digest of dGAGP fractionated on the Hamilton PRP-1 reversed phase column yielded two major peptide fractions, designated P1 and P3. P3 was subjected to further Hamilton PRP-1 reversed phase column chromatography a second time in order to purify it for sequencing (not shown).

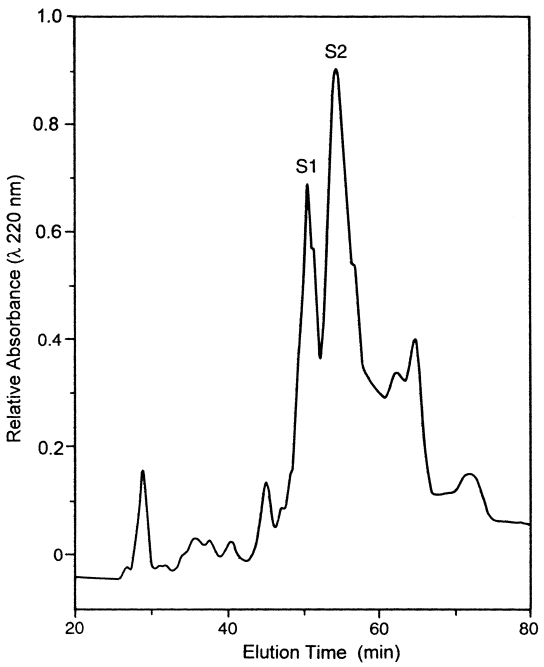

Fig. 3. Chymotryptic digest of dGAGP fractionated on a polysulfoethyl aspartamide cation exchange column. The major chymotryptic fractions, S1 and S2, were collected for further fractionation by reversed phase column chromatography.

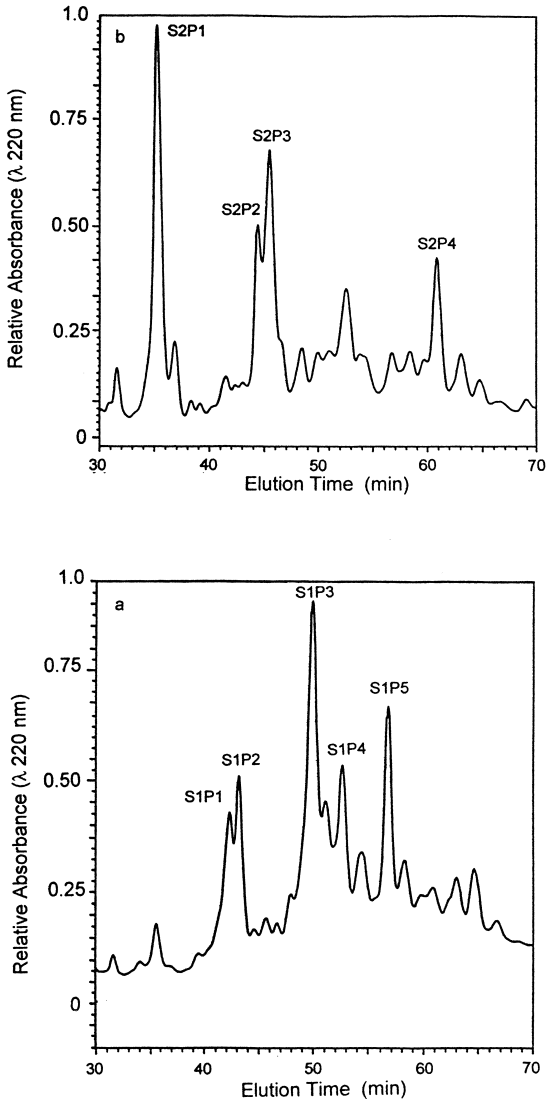

Fig. 4. Purification of dGAGP chymotryptic peptides by reversed phase column chromatography. Polysulfoethyl aspartamide fractions S1 and S2 were further fractionated by reversed phase chromatography on a Hamilton PRP-1 column. (a) Reversed phase fractionation of S1 resolved five major peptides labeled S1P1– S1P5. A sequencer failure precluded sequencing of S1P1. (b) Reversed phase fractionation of S2 resolved four major peptides, designated S2P1–S2P4, which were sequenced.

port, 1994; Kieliszewski, 1999; Kieliszewski et al., 1995) predicts arabinosylation of contiguous Hyp residues and arabinogalactan-polysaccharide addition to clustered non-contiguous Hyp residues, such as the X-Hyp-X-Hyp modules common in AGPs (Nothnagel, 1997; Kieliszewski, 1999). This hypothesis was corroborated by arabinosylation site mapping of a proline-rich protein, which indicated that extensive arabinosylation only occurs on contiguous Hyp residues while single non-contiguous Hyp residues were only rarely arabinosylated with a single residue (Kieliszewski et al., 1995). The hypothesis is also consistent with Lamport's early work (Lamport, 1977) which showed that each Hyp residue is extensively arabinosylated in the Ser-Hyp_4 motifs of glycopeptides isolated from tomato extensins. The Hyp contiguity hypothesis is also consistent with our recent demonstration that the sequence $(Ser\text{-}Hyp)_{32}$ is glycosylated exclusively with arabinogalactan-polysaccharide in tobacco (Shpak et al., 1999). In light of this, we predict that the GAGP consensus motif contains ~two polysaccharide attachment sites in the clustered non-contiguous Hyp motif: Ser-Hyp-Ser-Hyp-Thr-Hyp which is flanked by arabinosylated contiguous Hyp residues (Fig. 5).

Hydroxyproline-*O*-glycosidic linkages are stable in base (Lamport, 1967; Miller et al., 1972; Pope, 1977), in contrast to other *O*-glycosylated hydroxyamino acids such as serine and threonine, which undergo β-elimination (Lamport et al., 1973). We therefore used alkaline hydrolysis to isolate and characterize Hyp-arabinogalactan polysaccharides from GAGP, as demonstrated earlier (Qi et al., 1991).

Compositional analysis of the small Hyp-polysaccharides isolated from GAGP after fractionation of the alkaline hydrolysate on Sephadex G-50 (Fig. 6; Table 4) indicated a content of 5158 nM sugar. Corresponding quantitative Hyp assays showed a total of 220 nm Hyp in the peak isolated and analyzed (Fig. 6). The molar ratio of 220 nm Hyp : 5156 nm sugar indicated a ~23-residue rhamnoglucuronoarabinogalactan Hyp-polysaccharide substituent in this fraction. Methylation analysis of the polysaccharide (Table 5) showed linkages consistent with the model featured in Fig. 5, but containing 21–22 sugar residues rather than the 23

Table 3
GAGP hydroxyproline glycoside profile

Hydroxyproline glycoside	Percent of total hydroxyproline
Hyp-polysaccharide	20
Hyp-Ara_4	5
Hyp-Ara_3	27
Hyp-Ara_2	27
Hyp-Ara	10
Non-glycosylated Hyp	11

Table 4
Glycosyl compositions of intact GAGP and a GAGP Hyp-polysaccharide isolated from GAGP base hydrolysates

Glycosyl residue	GAGP[9] (Mol%)	GAGP Hyp-polysaccharide (Mol%)
Ara	36	38
Gal	36	34
Rha	10	13
GlcUA	9	15

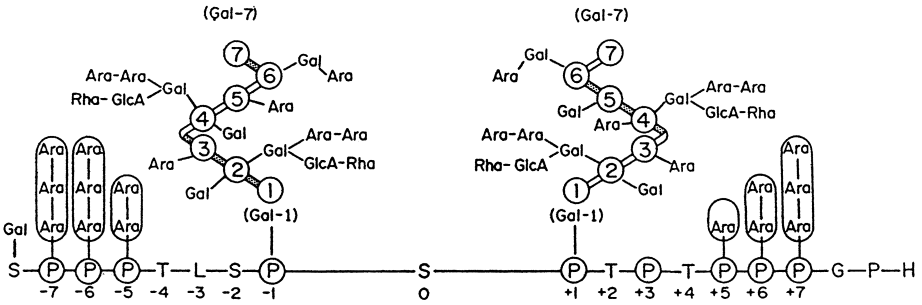

Fig. 5. Model of the GAGP consensus glycopeptide. The figure uses the standard single letter code for amino acids except for Hyp which is denoted by (Du et al., 1994), and the standard three letter code for sugars, except for glucuronic acid which is denoted as GlcA. This model depicts a symmetrical distribution of arabinosides and polysaccharide substituents which is directed by the palindrome-like arrangment of the Hyp residues in the peptide backbone; Ser-0 is the palindromic center. However, degenerate variations occur (Table 2). We based this structure on compositional and linkage analyses of the isolated Hyp-polysaccharide fraction (Tables 3 and 4) (Qi et al., 1991) and on the pentasaccharide side-chain structure elucidated for crude gum arabic by Defaye and Wong (1986) (corresponding to Rha_t, Ara_t, 3-Ara, 4-GlcA, and 2,3,6-Gal in Table 5). The three-dimensional arrangement of the saccharide side chains, the details of their microheterogeneity, and the precise placement along the galactan backbone remain to be elucidated.

featured in Fig. 5. We postulate that each small polysaccharide contains two pentasaccharide side chains (Gal, Ara_2, GlcA, Rha) arranged along a ~7-residue (1–3)-β-D-galactan backbone helix which also contains monosaccharide side chains of Ara and Gal.

Significantly, the linkage analyses of both Hyp-polysaccharide and GAGP (Table 5) are similar. This suggests that the larger Hyp-polysaccharides (Fig. 6) may be comprised of repeating units containing ~12 galactose residues/repeat. Polysaccharide and linkage analyses also relate GAGP to gum arabic polysaccharides, which are also built from regularly repeating blocks containing a total of ~13 galactose residues (Churms et al., 1983), and combine a β-1,3-galactan core with the rhamnoglucuronoarabinogalactose pentasaccharide side chains elucidated by Defaye and Wong (1986) (Fig. 5). Defaye and Wong (1986) found one rhamnose residue per gum arabic oligosaccharide side-chain and Qi et al. (1991) indicated there are five rhamnose residues in the larger polysaccharides of GAGP. Hence, we estimate as many as five side chains (~40 sugars) occur in the larger arabinogalactan moieties which eluted in fractions 18–26 from the G-50 Sephadex column (Fig. 6). However, the Hyp-polysaccharides of GAGP are very much smaller than the postulated 200 kDa carbohydrate blocks attached to a polypeptide backbone (Fig. 5 of Islam et al., 1997). Those workers explained the discrepancy between their

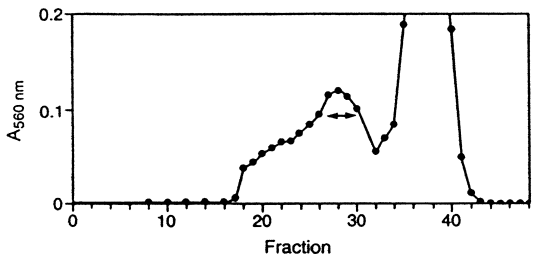

Fig. 6. Fractionation of the GAGP base hydrolysate by Sephadex G-50 gel permeation chromatography. Assay of Hyp across the recovered fractions indicated a broad size range for the Hyp-polysaccharide (fractions 17–32). We collected fractions 27–30 for linkage and composition analyses. Hyp-arabinosides and non-glycosylated Hyp eluted in fractions 33–42.

Table 5
Glycosyl linkages of intact GAGP and a GAGP Hyp-polysaccharide isolated from the GAGP base hydrolysate

Glycosyl linkage	GAGP (Mol%)	GAGP Hyp-polysaccharide (Mol%)
t-Rha	6.7	10.4 (2)[a]
2,3,4-Rha	3.3	0.0
t-Ara (f)	13.3	16.2 (4)
t-Ara (p)	1.7	2.3 (0–1)
2-Ara (f)	2.5	0.0
3-Ara (f)	8.3	11.0 (2–3)
4-Ara (p) or 5-Ara (f)	1.7	0.0
2,4-Ara or 2,5-Ara (f)	0.8	0.0
2,3,4-Ara or 2,3,5-Ara (f)	2.5	0.0
t-Gal	5.8	11.8 (3)
2-Gal	0.8	0.0
3-Gal	2.7	4.5 (1)
4-Gal	0.8	0.5
6-Gal	2.5	2.4 (0–1)
3,4-Gal	2.5	7.7 (2)
3,6-Gal	11.7	12.7 (3)
3,4,6-Gal	10.0	9.4 (2)
2,3,6-Gal	3.3	0.0
2,3,4,6-Gal	5.8	0.0
t-GlcUA	1.7	0.9
4-GlcUA	7.5	10.2 (2)
3,4-GlcUA	1.7	0.0
2,4-GlcUA	0.8	0.0
2,3,4-GlcUA	0.8	0.0
4-Glc	0.8	0.0

[a] Estimated number of residues/polysaccharide.

work (Osman et al., 1995; Islam et al., 1997) and the Hyp-polysaccharide described earlier (Qi et al., 1991) by suggesting that alkali had degraded the carbohydrate component of gum arabic. Our data corroborate those of Qi et al. (1991). The terminal rhamnose content (Table 4) suggests that the Hyp-polysaccharides have not undergone appreciable β-eliminative degradation. Thus, the simplest explanation for GAGP and other AGP sensitivity to alkaline degradation (Osman et al., 1995) involves peptide bonds rather than glycosidic linkages as described below.

Partial alkaline hydrolysis under very mild conditions preferentially cleaves GAGP into large glycopeptides containing short peptide sequences (L. Goodrum and M. Kieliszewski, unpublished data). Edman degradation of those glycopeptides indicated that the base-labile bonds occur in the Ser-Hyp-Ser-Hyp-Thr-Hyp regions of the GAGP polypeptide backbone. This indicates that peptide bonds involving alternate serine and hydroxyproline residues in the polysaccharide attachment region are exceptionally labile to alkali. Indeed, even a brief treatment of GAGP with 0.2 M NaOH at 50°C as described earlier (Akiyama et al., 1984) released Hyp from the polypeptide backbone as a significant percent of the total bound hydroxyproline, i.e. 1.5% after 1 h; 3.4% after 3 h; 6.4% after 6 h. Significantly, no evidence of *O*-glycosylated serine or threonine residues appeared during Edman degradation of the glycopeptides generated by base hydrolysis (L. Goodrum and M. Kieliszewski, unpublished data.)

Given these result, how can we therefore relate GAGP structure and function? Remarkably, fifteen residues of the GAGP peptide consensus sequence are ''quasi-palindromic''. That is, the side chain sequence is almost the same whether read from either the N-terminus or C-terminus. Such peptide symmetry, which occurs frequently in extensins and AGPs, may enhance molecular packing (Kieliszewski and Lamport, 1994), recognition (Ohno, 1994; Giralt et al., 1994), and self-assembly (Giralt et al., 1994). Indeed, palindromic symmetry rigidified by contiguous Hyp blocks in the motifs: Ser-Hyp-Hyp-Hyp-(Hyp) and Thr-Hyp-Hyp-(Hyp), may impart self-ordering properties in GAGP and other HRGPs. Thus, the repeating glycopeptide symmetry of two central polysaccharides flanked by Hyp arabinosides may be related to gum arabic's remarkable properties. These properties include: an anomalously low viscosity (Churms et al., 1983) the ability to act as a flavor emulsifier and stabilizer, and GAGP's biological role as a component of a plastic sealant. The ''high molecular mass protein component'' (Islam et al., 1997; Ray et al., 1995) of gum arabic (GAGP) also binds to hydrophobic surfaces. The suggested mechanism involves adsorption of a hydrophobic polypeptide chain (Islam et al., 1997). However, the GAGP polypeptide is extremely hydrophilic being comprised mainly of hydroxyamino acids. Instead, we suggest that GAGP properties depend considerably on the polysaccharide substituents. The β-1,3 galactan backbone (Aspinall, 1970) is probably an amphipathic helix (Raman et al., 1992; Rees, 1977) the inside of the helix comprising a hydrophobic surface with the bulk of the galactan hydroxyl groups oriented toward the outer surface of the helix. Thus β-1,3-linked galactan polysaccharides may be analogous to the cyclodextrins which form inclusion complexes with small hydrophobic molecules (Raman et al., 1992; Rees, 1977).

Finally, it is instructive to compare the structural features of other HRGPs with those of GAGP. The presence of the repetitive X-Hyp-X-Hyp-X-Hyp motif (Nothnagel, 1997) and arabinogalactan polysaccharide, and co-precipitation with the Yariv reagent (Qi et al., 1991) typify the classical AGPs, although the virtual absence of alamine does not. On the other hand rigid Ser-Hyp_3 and Ser-Hyp_4 blocks (Table 2) and their Hyp-arabinosides typify the extensins. Thus GAGP, like some other HRGPs (Kieliszewski et al., 1992; Lind et al., 1994; Bucher et al., 1997) combines features of AGPs and extensins. Both are characterized by repetitive glycopeptide modules (i.e., glycomodules) of functional significance. As such, gum arabic glycoprotein exemplifies just one of the many ways that plants generate supramolecular structure by exploiting the hydroxyproline-rich polypeptide as a template for the orderly arrangement of saccharide sidechains.

3. Experimental

Nodules of gum arabic (Kordofan Province, Sudan) were a gift from Gary Wine of AEP Colloids (Ballston Spa, NY). We ground the nodules to a fine flour (ca. 2 min) in a Tekmar A-10 mill.

3.1. Size fractionation of gum arabic

Samples of gum arabic (100 mg/ml) were dissolved in water then diluted to 50 mg/ml in 0.2 M sodium phosphate buffer (pH 7). Samples were spun to pellet insoluble material and 1 ml aliquots were injected onto a semi-preparative Superose-6 gel filtration column (1.6 cm i.d. × 50 cm, Pharmacia), eluted isocratically as described earlier (Qi et al., 1991). The protein peaks corresponding to GAGP were dialyzed against water to remove salt and then freeze-dried.

3.2. HF-deglycosylation

The Superose-6 fractionated GAGP was deglycosylated in anhydrous hydrogen fluoride (HF) (20 mg

powder/ml HF for 1 h at 4°C) as described earlier (Qi et al., 1991). The procedure was repeated twice to ensure complete deglycosylation, the resultant preparation was designated dGAGP.

3.3. Purification of size-fractionated GAGP and dGAGP by reversed phase HPLC

Superose-fractionated GAGP was purified either for glycoside analyses, or dGAGP samples for peptide mapping on a Hamilton PRP-1 semi-preparative column (10 μm, 250 × 4.1 mm) equilibrated with Buffer A (0.1% TFA, aqueous) and eluted with Buffer B (0.1% TFA, 80% acetonitrile, aqueous) by gradient elution (0–100% B/80 min: 0.5 ml/min flow rate). The eluate was monitored at 220 nm. The collected peaks were evaporated to dryness with N_2 (g), redissolved in ddH_2O, then freeze-dried.

3.4. Proteolysis of deglycosylated GAGP with chymotrypsin or pronase

The 2–9 mg samples of dGAGP were digested with pronase or chymotrypsin as detailed earlier (Kieliszewski et al., 1992). The digests were then freeze-dried.

3.5. Fractionation of dGAGP chymotryptic peptides by cation exchange HPLC

The dGAGP chymotryptic peptides (400 μg/injection) were fractionated on a PolySULFOETHYL A™ cation exchange column (9.4 mm i.d. × 200 mm; PolyLC, Ellicot City, MD) equilibrated with Buffer A (5 mM potassium phosphate/phosphoric acid buffer, pH 3, containing 25% v/v acetonitrile) and eluted with Buffer B (Buffer A containing 1 M KCl) using programmed gradient elution. The elution gradient was 0–4% Buffer B in 45 min, 4–8% Buffer B from 45 to 50 min, and 8–30% Buffer B from 50–65 min. The flow rate was 0.4 ml/min and we monitored the absorbance at 220 nm. The collected peaks were pooled, evaporated with N_2 (g), redissolved in ddH_2O, then freeze dried.

3.6. Peptide isolation via reversed phase HPLC

The partial pronase digest of dGAGP and major peaks S1 and S2 PolySULFOETHYL Aspartamide column were dissolved in Buffer A (0.1% TFA, aqueous) and injected onto a Hamilton PRP-1 analytical reverse phase column (4.1 mm i.d. × 150 mm) which was eluted at 0.5 ml/min with a Buffer B (0.1% TFA and 80% v/v acetonitrile) gradient of 0–50% in 100 min. The effluent was monitored at 220 nm and collected peaks were evaporated with N_2(g), re-dissolved in ddH_2O, and then freeze dried prior to sequencing. For increased resolution of pronase peptide P3 (Fig. 2), P3 was subjected to further chromatography using the PRP-1 column a second time, this being eluted with a 0–30% Buffer B gradient.

3.7. Automated Edman degradation of dGAGP chymotryptic peptides

The dGAGP peptides were sequenced at the Michigan State University Macromolecular Facility on a 477A Applied Biosystems (Foster City, CA) gas phase sequencer.

3.8. Amino acid analysis

Amino acid compositions were determined by pre-column derivatization of amino acids with 6-aminoquinolyl-*N*-hydroxysuccinimidyl carbamate followed by reverse phase HPLC (Nova-Pak™ C_{18} column) using the Waters AccQ-Tag Chemistry. Package and the gradient recommended by Waters for analyzing collagen hydrolysates (Crimmins and Cherian, 1997; van Wandelen and Cohen, 1997).

3.9. Hydroxyproline glycoside profile

The distribution of GAGP hydroxyproline glycosides was determined after alkaline hydrolysis (105°, 18 h, 0.22 N $Ba(OH)_2$) and neutralization followed by chromatography on a 75 × 0.6 cm Technicon Chromobeads C2 cation exchange resin as described earlier (Lamport and Miller, 1971).

3.10. Isolation of the Hyp-polysaccharide

Alkaline hydrolysates (see above) of Superose-6 and PRP-1 purified GAGP were applied to a G-50 Sephadex gel permeation column eluted isocratically with 100 mM ammonium acetate buffer, pH 6.8, at a flow rate of 0.3 ml/min. One ml fractions were collected and 40 μl aliquots of each fraction were assayed for Hyp as described earlier (Kivirikko and Liesmaa, 1959; Kieliszewski et al., 1990). The fractions were freeze-dried, then weighed, and the amounts of Hyp and sugar in the fractions were calculated from the recovered weights, Hyp assays, and monosaccharide composition analyses.

3.11. Partial alkaline hydrolysis of GAGP

Superose-fractionated GAGP (10 mg/ml) was dissolved in 0.2 N $NaOH/NaBH_4$ and resulting solution was at 50°C as described earlier by Akiyama et al. (1984). A 200 μl aliquot was removed immediately (time zero control) and hourly thereafter up until 6 h, and, cooled in ice, to which 20 μl glacial acetic acid

was added (final pH = 5.8). Each sample was assayed for Hyp as described earlier (Kivirikko and Liesmaa, 1959; Kieliszewski et al., 1990).

3.12. Saccharide composition and linkage analysis

Monosaccharide compositions and linkage analyses were determined at the Complex Carbohydrate Research Center, University of Georgia following the methods of York et al. (1985), and Merkle and Poppe (1994).

Acknowledgements

We thank AEP Colloids and Mr. Gary Wine (Ballston Spa, NY) for providing gum arabic nodules, Ms. Carole Hahn and Marti Dolata for the graphic art and D. Lamport for comments on the manuscript. This work was supported by the United States Department of Agriculture NRI Competitive Grants Program (Grant Nos. 97-01299 and 95-37304-2375) and in part by a grant from the National Institutes of Health (No. 2-P41-RR05351-06) to the Complex Carbohydrate Research Center, a NIH-funded Resource Center for Biomedical Complex Carbohydrates.

References

Akiyama, Y., Eda, S., Kato, K., 1984. Agricultural and Biological Chemistry 48, 235.
Anderson, D.M.W., McDougall, F.J., 1988. Food Hydrocolloids 2, 329.
Anderson, D.M.W., Weiping, W., 1990. Phytochemistry 29, 1193.
Anderson, D.M.W., Weiping, W., Lewis, G.P., 1990. Biochemical Systematics and Ecology 18, 39.
Aspinall, G.O., 1970. The Carbohydrates 2B, 522.
Bucher, M., Schroeer, B., Willmitzer, L., Riesmeier, J., 1997. Plant Molecular Biology 35, 497.
Chikamai, B.N., Banks, W.B., Anderson, D.M.W., Weiping, W., 1996. Food Hydrocolloids 10, 309.
Churms, S.C., Merrifield, E.H., Stephen, A.M., 1983. Carbohydrate Research 123, 267.
Connolly, S., Fenyo, J.-C., Vandevelde, M.-C., 1987. Food Hydrocolloids 1, 477.
Connolly, S., Fenyo, J.-C., Vandevelde, M.-C., 1988. Carbohydrate Polymers 8, 23.
Crimmins, D.L., Cherian, R., 1997. Analytical Biochemistry 244, 407.
Defaye, J., Wong, E., 1986. Carbohydrate Research 150, 221.
Delonnay, C.L., 1993. Masters Thesis, Michigan State University, MI.
Du, H., Simpson, R.J., Moritz, R.L., Clarke, A.E., Bacic, A., 1994. Plant Cell 6, 1643.
Giralt, E., Albericio, F., Garcia-Echeverria, C., Pons, M., Royo, M., Ruiz-Gayo, M., 1994. Biochemistry Society Transactions 22, 1045.
Islam, A.M., Phillips, G.O., Sljivo, A., Snowden, M.J., Williams, P.A., 1997. Food Hydrocolloids 11, 493.
Kieliszewski, M.J., 1999. Synthetic genes for plant gums containing repeat motifs. PCT Int. Appl. Ohio University. CA. JP. (WO 9903978). CODEN: PIXXD2, p. 72.
Kieliszewski, M.J., Kamyab, A., Leykam, J.F., Lamport, D.T.A., 1992. Plant Physiology 99, 538.
Kieliszewski, M.J., Lamport, D.T.A., 1994. Plant Journal 5, 157.
Kieliszewski, M.J., Leykam, J.F., Lamport, D.T.A., 1990. Plant Physiology 92, 316.
Kieliszewski, M.J., O'Neill, M., Leykam, J., Orlando, R., 1995. Journal of Biological Chemistry 270, 2541.
Kivirikko, K.I., Liesmaa, M., 1959. Scandinavian Journal of Clinical Laboratories 11, 128.
Lamport, D.T.A., 1967. Nature 216, 1322.
Lamport, D.T.A. 1977. In: Loewus, F.A., Runeckles, V.C. (Eds.), Recent Advances in Phytochemistry. Plenum Press, New York, p. 79.
Lamport, D.T.A., Katona, L., Roerig, S., 1973. Biochemical Journal 133, 125.
Lamport, D.T.A., Miller, D.H., 1971. Plant Physiology 48, 454.
Lind, J.L., Bacic, A., Clarke, A.E., Anderson, M.A., 1994. Plant Journal 6, 491.
Merkle, R.K., Poppe, I., 1994. Methods in Enzymology 230, 1.
Miller, D.H., Lamport, D.T.A., Miller, M., 1972. Science 176, 918.
Nothnagel, E.A., 1997. International Review of Cytology 174, 195.
Ohno, S., 1994. Animal Genetics 25, 5.
Osman, M.E., Menzies, A.R., Martin, B.A., Williams, P.A., Phillips, G.O., Baldwin, T.C., 1995. Phytochemistry 38, 409.
Osman, M.E., Menzies, A.R., Williams, P.A., Phillips, G.O., Baldwin, T.C., 1993. Carbohydrate Research 246, 303.
Pope, D.G., 1977. Plant Physiology 59, 894.
Prakash, A., Mangino, M.E., 1990. Food Hydrocolloids 4, 177.
Qi, W., Fong, C., Lamport, D.T.A., 1991. Plant Physiology 96, 848.
Raman, B., Sundari, C.S., Balasubramanian, D., 1992. Indian Journal of Biochemistry and Biophysics 29, 143.
Randall, R.C., Phillips, G.O., Williams, P.A., 1988. Food Hydrocolloids 2, 131.
Ray, A.K., Bird, P.B., Iacobucci, G.A., Clark Jr, B.C., 1995. Food Hydrocolloids 9, 123.
Rees, D.A. 1977. In: Ashworth, M. (Ed.), Polysaccharide Shapes. Chapman & Hall, London, p. 41.
Shpak, E., Leykam, J.F., Kieliszewski, M.J., 1999. Proceedings of the National Academy of Sciences, US (in press).
Stephen, A.M., Churms, S.C., Vogt, D.C., 1990. Methods Plant Biochemistry 2, 483.
van Wandelen, C., Cohen, S.A., 1997. Journal of Chromatograaphy A 763, 11.
York, W.S., Darvill, A.G., McNeil, M., Stevenson, T.T., Albersheim, P., 1985. Methods in Enzymology 118, 3.