

Amino Acid Sequence of Wheat Flour Arabinogalactan-Peptide, Identical to Part of Grain Softness Protein GSP-1, Leads to Improved Structural Model

Katie Van den Bulck,^{1,2} Anne-Marie A. Loosveld,¹ Christophe M. Courtin,¹ Paul Proost,³ Jozef Van Damme,³ Johan Robben,^{1,4} Andrew Mort,⁵ and Jan A. Delcour¹

Cereal Chem. 79(3):329–331

Arabinogalactan-proteins (AGPr) are widely distributed in the plant kingdom. They are minor constituents in almost all tissues (Fincher et al 1983) and are often secreted in copious quantities by suspension cultured cells (Clarke et al 1979). AGPr, with a relative molecular weight (M_r) of typically 60,000 to 300,000, consist of a polypeptide core *O*-glycosylated with large branched glycan chains through Hyp or Ser residues. The glycan chains make up at least 90% of the AGPr molecular mass, consist mainly of galactose and arabinose, and can contain more than 50 saccharide residues. The polypeptides have an N-terminal signal sequence for entry into the secretory pathway (Sommer-Knudsen et al 1997) and contain a domain rich in Pro/Hyp, Ala, Ser, and Thr (Showalter 1993). However, the other domains of the polypeptide diverge. Based on these domains, AGPr were separated into three classes (Du et al 1996; Sommer-Knudsen et al 1997). The first class consists of the classical AGPr, which contains a hydrophobic transmembrane domain at the C-terminus that, in the mature AGPr, is replaced by a glycosylphosphatidylinositol (GPI) lipid anchor. The second class contains the nonclassical AGPr, with either a Cys-rich C-terminal domain or one or two Asn-rich domains linked to the Pro/Hyp-, Ala-, Ser-, and Thr-rich domain. The mature AGPr have neither a hydrophobic domain nor a GPI lipid anchor. The third class consists of chimeras of AGPr and extensins, which are proteoglycans *O*-glycosylated through Hyp and Ser with short arabinosyl chains and single galactosyl residues, respectively (Sommer-Knudsen et al 1998). The chimeras contain the large arabinogalactan polysaccharides of AGPr as well as the short Hyp-oligoarabinosides of extensins.

In spite of the wide occurrence of AGPr in cells and tissues, no general function has been ascribed to AGPr. Wound healing, frost hardiness, and drought resistance may result from the AGPr water-holding capacity (Clarke et al 1979). AGPr may be involved in developmentally regulated functions such as cell expansion (Willats and Knox 1996), cell death (Gao and Showalter 1999), cell proliferation (Serpe and Nothnagel 1994), somatic embryogenesis (Kreuger and van Holst 1993), pollen tube growth (Cheung et al 1995; Wu et al 1995), pollen germination (Sommer-Knudsen et al 1998), cell communication (Fincher et al 1983), and cell adhesion (Schultz et al 2000).

In contrast to the previously mentioned AGPr, arabinogalactan-peptide (AGP) of wheat (*Triticum aestivum* L.) endosperm is a group of rather small molecules with $M_r \approx 25,000$ (Loosveld et al 1997) of which the carbohydrate part received most attention (Fincher et al 1974; Neukom and Markwalder 1975; Strahm et al 1981; Loos-

veld et al 1998). Research on the peptide backbone to date has been limited to the study of its amino acid composition. The peptide backbone was estimated to consist of 15–20 amino acids, mainly Hyp, Ala, Glu/Gln, Ser, Thr, and Val (Fincher et al 1974; Strahm et al 1981; Meuser and Suckow 1986). We here report, for the first time, the full amino acid sequence of wheat endosperm AGP. It is identical to the sequence of grain softness protein (GSP) and allows for a more accurate structural model for AGP, revealing three glycosylation sites.

AGP was purified from commercial wheat flour as described by Loosveld and Delcour (2000). The procedure involved extraction with water, removal of starch by amylase and amyloglucosidase, removal of protein by adsorption onto silica, dialysis, and precipitation of the target components by stepwise addition of ethanol. The yield of purified AGP (139 mg of AGP/100 g of flour) was about half of the level occurring in wheat flour (290 mg/100 g of flour) (Loosveld et al 1998). AGP consisted of 90.9% arabinogalactan and 9.1% protein moiety. It was deglycosylated with hydrogen fluoride containing 10% methanol to prevent alkylation of the peptides by sugar fluorides (Sanger and Lamport 1983). The mixture of protein and sugar components was separated on a gel permeation column (ToyoPearl HW-40S, 50.0 × 1.0 cm, Supelco, Bellefonte, PA). The fraction containing the target peptide was further subjected to reversed-phase chromatography (Microsorb-MV column, 25.0 × 0.46 cm, Rainin Instruments, Woburn, MA) with 0.1% trifluoroacetic acid in water using a linear gradient of acetonitrile (0–60% in 60 min). Three major peaks were found.

A first peak had M_r 1,534 in ion trap mass spectrometry (Esquire-LC, Bruker-Daltonics, Bremen, Germany). The M_r of the second and third peaks were 1,548 and 1,562, respectively, indicating methylated forms of the first peak peptide. The determined amino acid sequence (477A/120A and 491 cLC protein sequencers, Applied Biosystems, Foster City, CA) is given in Fig. 1A. The peptide backbone consists of 15 amino acids as expected from the M_r and protein content of the purified AGP fraction. The amino acid composition data are in agreement with earlier data (Fincher et al 1974; Strahm et al 1981; Meuser and Suckow 1986). Amino acid sequences for peptides in the two other peaks yielded identical results apart from the presence of methyl groups on the third (Glu) and on the third (Glu) and tenth (Gln) amino acids for the second and third peaks, respectively, as determined by capillary degradation (Edman 477A protein sequencer, Applied Biosystems) and MS/MS mass spectrometry (ion trap ESQUIRE-LC).

To determine whether these methyl groups are naturally present or artefacts, the amino acid sequence of the native glycosylated AGP peptide was analyzed as well, using an Applied Biosystems 477A protein sequencer. A nonmethylated sequence identical to the sequence of the first peak was found, except for three amino acid residues that could not be identified (Fig. 1B). Compared with the deglycosylated AGP peptide sequence, the nonidentified residues correspond to Hyp residues, indicating that these Hyp residues were *O*-glycosylation sites as stated by Fincher et al (1974), who believed that there were three or four such residues in the back-

¹ Laboratory of Food Chemistry, Katholieke Universiteit Leuven, Kasteelpark Arenberg 20, B-3001 Leuven, Belgium.

² Corresponding author. E-mail: katie.vandenbulck@agr.kuleuven.ac.be

³ Laboratory of Molecular Immunology, Katholieke Universiteit Leuven, Minderbroedersstraat 10, B-3000 Leuven, Belgium.

⁴ Biomedical Research Institute, Limburgs Universitair Centrum, Universitaire Campus, B-3590 Diepenbeek, Belgium.

⁵ Department of Biochemistry and Molecular Biology, Oklahoma State University, Noble Research Center 246, Stillwater, OK 74078-3035.

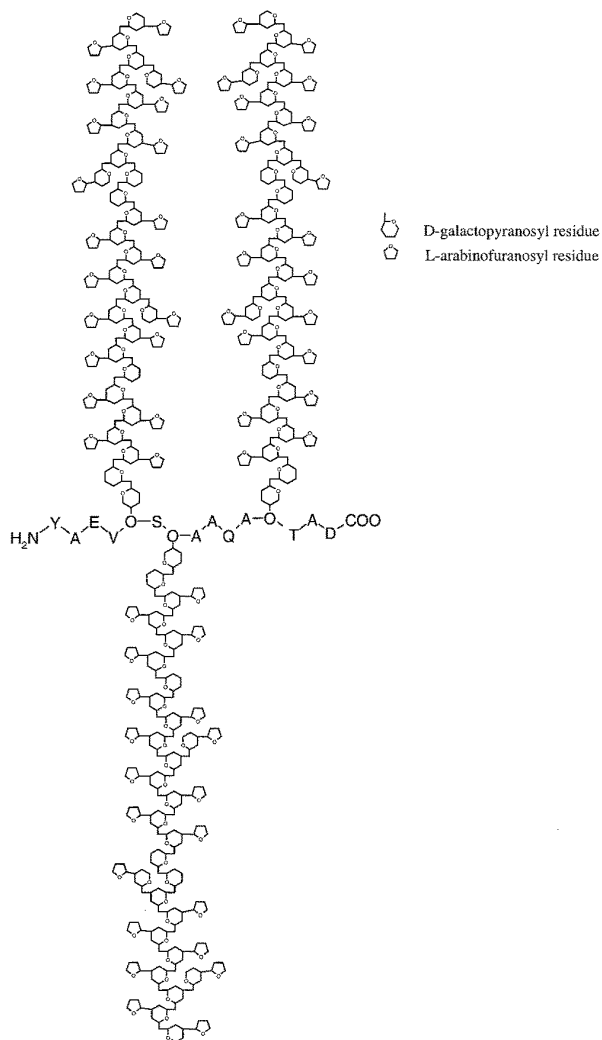


Fig. 2. Structural model for wheat flour arabinogalactan-peptide based on current data. Carbohydrate part consists of a β -1,6-D-linked galactopyranosyl backbone branched at the 3-position with either an L-arabinofuranosyl or a D-galactopyranosyl residue (Loosveld et al 1998). In line with the data by these authors, the total number of L-arabinose and D-galactose residues, and the relative proportion of the different linkages in the three chains correspond to the approximate total level of such units and linkages in an average AGP.

- Krishnamurthy, K., and Giroux, M. J. 2001. Expression of wheat puroindoline genes in transgenic rice enhances grain softness. *Nature Biotechnol.* 19:162-166.
- Lenard, J., and Hess, G. P. 1964. An approach to the specific cleavage of peptide bonds. II. Specific cleavage of peptide chains based on the hydrogen fluoride induced nitrogen to oxygen acyl shift. *J. Biol. Chem.* 239:3275-3291.
- Loosveld, A.-M. A., and Delcour, J. A. 2000. The significance of arabinogalactan-peptide for wheat flour bread-making. *J. Cereal Sci.* 32:147-157.
- Loosveld, A.-M. A., Grobet, P. J., and Delcour, J. A. 1997. Contents and structural features of water-extractable arabinogalactan in wheat flour fractions. *J. Agric. Food Chem.* 45:1998-2002.
- Loosveld, A.-M. A., Maes, C., van Casteren, W. H. M., Schols, H. A., Grobet, P. J., and Delcour, J. A. 1998. Structural variation levels of water-extractable arabinogalactan-peptide in European wheat flours. *Cereal Chem.* 75:815-819.
- Meuser, F., and Suckow, P. 1986. Non-starch polysaccharides. Pages 42-61 in: *Chemistry and Physics of Baking*. J. M. V. Blanshard, P. J. Frazier, T. Galliard, eds. R. Soc. Chem.: London.
- Neukom, H., and Markwalder, H. 1975. Isolation and characterization of an arabinogalactan from wheat flour. *Carbohydr. Res.* 39:387-389.
- Sanger, M. P., and Lamport, D. T. A. 1983. A microapparatus for liquid-hydrogen fluoride solvolysis—Sugar and amino sugar composition of *Erysiphe graminis* and *Triticum aestivum* cell-walls. *Anal. Biochem.* 128:66-70.
- Schultz, C. J., Johnson, K. L., Currie, G., and Bacic, A. 2000. The classical arabinogalactan protein gene family of arabidopsis. *Plant Cell* 12:1751-1767.
- Serpe, M. D., and Nothnagel, E. A. 1994. Effects of Yariv phenylglycosides on *Rosa* cell suspensions: Evidence for the involvement of arabinogalactan-proteins in cell proliferation. *Planta* 193:542-550.
- Showalter, A. M. 1993. Structure and function of plant cell wall proteins. *Plant Cell* 5:9-23.
- Sommer-Knudsen, J., Clarke, A. E., and Bacic, A. 1997. Proline- and hydroxyproline-rich gene products in the sexual tissues of flowers. *Sex. Plant Reprod.* 10:253-260.
- Sommer-Knudsen, J., Clarke, A. E., and Bacic, A. 1998. Hydroxyproline-rich plant glycoproteins. *Phytochemistry* 47:483-497.
- Strahm, A., Amado, R., and Neukom, H. 1981. Hydroxyproline-galactoside as a protein-polysaccharide linkage in a water-soluble arabinogalactan-peptide from wheat endosperm. *Phytochemistry* 47:483-497.
- Turner, M., Mukai, Y., Leroy, P., Charef, B., Appels, R., and Rahman, S. 1999. The Ha locus of wheat: Identification of a polymorphic region for tracing grain hardness in crosses. *Genome* 42:1242-1250.
- Willats, W. G. T., and Knox, J. P. 1996. A role for arabinogalactan-proteins in plant cell expansion: Evidence from studies on the interaction of β -glucosyl Yariv reagent with seedlings of *Arabidopsis thaliana*. *Plant J.* 9:919-925.
- Wu, H.-M., Wang, H., and Cheung, A. Y. 1995. A pollen tube growth stimulatory glycoprotein is deglycosylated by pollen tubes and displays a glycosylation gradient in the flower. *Cell* 82:395-403.

[Received December 18, 2001. Accepted February 13, 2002.]