Journal of Cereal Science 74 (2017) 155–164

Contents lists available at ScienceDirect

Journal of Cereal Science

journal homepage: www.elsevier.com/locate/jcs

The *Gsp-1* genes encode the wheat arabinogalactan peptide

Mark D. Wilkinson [a, *], Paola Tosi [b], Alison Lovegrove [a], Delia I. Corol [a], Jane L. Ward [a], Richard Palmer [a], Stephen Powers [a], Diana Passmore [a], Gina Webster [a], Susan E. Marcus [c], J. Paul Knox [c], Peter R. Shewry [a, b, **]

[a] Plant Biology and Crop Science Department, Rothamsted Research, Harpenden AL5 2JQ, UK
[b] School of Agriculture Policy and Development, University of Reading, Reading RG6 6AH, UK
[c] Centre for Plant Sciences, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK

ARTICLE INFO

Article history:
Received 25 January 2017
Accepted 14 February 2017
Available online 16 February 2017

Keywords:
Wheat
Arabinogalactan peptide
Grain softness protein
Immunocytochemistry

ABSTRACT

Western blotting, ELISA and ^{1}H-NMR spectroscopy showed that RNAi down-regulation of the wheat *Gsp-1* gene resulted in reduced contents of both arabinogalactan peptide (AGP) and grain softness protein (*GSP-1*) in mature wheat grains confirming that these components are encoded by the same gene. A small increase in grain hardness and decrease in the viscosity of aqueous extracts of the transgenic lines also indicated small effects on functional properties. Immunolocalisation using a novel wheat AGP monoclonal antibody in conjunction with confocal microscopy showed that the major form of AGP which was eliminated in knockout lines is located within the cell, probably in the vacuole, and not in the plasma membrane or cell wall. However, clear localisation of the AGP epitope to the plasma membrane was observed in both control and transgenic lines and probably resulted from the presence of one or more separate forms of arabinogalactan protein. The existence of such additional form(s) was also indicated by ^{1}H-NMR spectroscopy which showed that the ratio of arabinose to galactose differed between the control and transgenic lines.

1. Introduction

The *Gsp-1* genes of wheat are proposed to encode precursor proteins that are post-translationally processed to give two components: the arabinogalactan peptide (comprising 15 amino acids) and the grain softness protein (*GSP-1*) (the major form comprising 113 amino acids with 10 cysteine residues) (Elmorjani et al., 2013). However, it has not been directly demonstrated that the AGP is derived from *Gsp-1* genes. *Gsp-1* genes appear to be ubiquitous in grasses, being identified in 65 species from the five major grass subfamilies, with over 20 different *Gsp-1* alleles being reported (Wilkinson et al., 2013). The GSP protein belongs to a large family of seeds storage proteins called the "prolamin superfamily" (Shewry and Halford, 2002) and has been predicted to have a similar structure to other members of this family, including the 2S storage albumins which occur in dicotyledonous seeds and the puroindolines (Pins) of cereal seeds (Elmorjani et al., 2013).

The *Gsp-1* genes are present at the *Ha* (Hardness) locus on chromosome 5D (Chantret et al., 2005), which also comprises two genes encoding the puroindolines Pin a and Pin b. Allelic variation in the expression and/or sequences of the Pin proteins has been shown to account for 60–80% of the variation in endosperm texture (hardness) in bread wheat (*T. aestivum* L.) cultivars (Turner et al., 2004). Grain hardness has a major impact on the processing properties of bread wheat with hard texture being preferred for bread making and soft texture for biscuits, cakes and pastries (Morris and Rose, 1996). Although Pins have been studied in considerable detail, little is known of their structures and biological roles, nor about the mechanisms that determine grain texture. Although GSP was initially reported to affect grain texture (Jolly et al., 1996), this has since been suggested to result from the tight linkage of the *Gsp-1* and *Pin* genes with no direct effect of GSP itself (Tranquilli et al., 2002).

Fincher and Stone (1974) showed that the wheat arabinogalactan peptide (AGP) has a mass of about 22,000 and consists of 80% polysaccharide. More recent studies have shown that wheat AGP comprises about 0.4% of the dry weight of the wheat endosperm

* Corresponding author.
** Corresponding author. Plant Biology and Crop Science Department, Rothamsted Research, Harpenden AL5 2JQ, UK.
E-mail addresses: mark.wilkinson@rothamsted.ac.uk (M.D. Wilkinson), peter.shewry@rothamsted.ac.uk (P.R. Shewry).

http://dx.doi.org/10.1016/j.jcs.2017.02.006

(Loosveld et al., 1997) with similar contents of AGPs in related cereals, namely barley (0.28%), rye (0.21%), durum wheat (0.28%) and triticale (0.32%) (Van den Bulck et al., 2005). Van den Bulck et al. (2005) showed that the peptide moiety of AGP was identical in sequence to the first 15 amino acids predicted for the protein precursor of *GSP-1* (allowing for post-translational hydroxylation and glycosylation of the 3 proline residues) (Van den Bulck et al., 2005). The cleavage between the AGP and *GSP-1* sequences is at the C-terminal side of an aspartic acid residue, which is consistent with the action of a legumin-type cysteine endopeptidase similar to those responsible for the post-translational processing of storage globulins and 2S albumins in the vacuoles of dicotyledonous seeds. It can therefore be hypothesised that the AGP/GSP precursor protein is synthesised on the endoplasmic reticulum (ER), transferred into the lumen (a signal sequence of 19 amino acids also being predicted) and post-translationally processed (proline hydroxylation, O-glycosylation) in the ER and Golgi before being proteolytically cleaved in the vacuole (Wilkinson et al., 2013). In support of this hypothesis, Elmorjani et al. (2013) showed that the GSP precursor protein is further proteolytically cleaved at the *N*- and *C*-termini to give four forms comprising between 110 and 113 residues.

Although substantial amounts of both AGP and GSP are present in the starchy endosperm (white flour) of the mature grain nothing is known about their locations in the cells, or their biological functions, with no evidence for the widely held assumption that AGP is located in the cell wall. Furthermore, nothing is known about their effect on grain processing or human health, although AGP is present in similar amounts to water-soluble fibre components.

We have therefore generated and analysed transgenic wheat lines with RNAi-mediated suppression of *Gsp-1* transcripts. ^{1}H-NMR analysis, ELISA and western blotting showed decreases in the contents of both AGP and GSP, demonstrating for the first time that these two components are indeed encoded by the same genes. Immunolabelling shows that the major form of AGP encoded by *Gsp-1* has an intracellular location, probably within the vacuoles that give rise to protein bodies, while analysis of mature grain shows small but statistically significant effects on both grain texture and on the viscosity of aqueous extracts of endosperm flour.

2. Materials and methods

2.1. RNAi construct preparation, transformation and confirmation of transgenic wheat lines

An RNAi construct to down-regulate *Gsp-1* under the control of the starchy endosperm-specific *HMW1Dx5* promoter was created using a *Bgl*II/*BamH*1 cloning strategy as described by Nemeth et al. (2010), using 344-bp fragments (+1 to +344bp of the sequence) based on a cDNA sequence from *T. aestivum* cv. Cadenza AGP1; (Wilkinson et al., 2013). This fragment was chosen as it corresponds to the most highly expressed *Gsp-1* sequence. The fragment was generated by PCR (Phusion *Taq* polymerase from Thermo Scientific, Fisher Scientific - UK Ltd Bishop Meadow Road, Loughborough, Leicestershire, LE11 5RG United Kingdom), using the following PCR primers GSPRNAiF and GSPRNAiR (Supplementary Table S1). The full nucleotide and protein sequences of the *Gsp-1* sequence are shown in Supplementary Figs. S1a and S1b respectively.

Sequencing of the constructs was carried out using the BigDye Terminator Version 3.1 Cycle Sequencing Kit (Applied Biosystems, Life Technologies Ltd, 3 Fountain Drive, Inchinnan Business Park, Paisley, PA4 9RF, UK), with construct specific primers i.e. M13F, Rab1 and adh3R (Supplementary Table S1) used to confirm correct orientation of RNAi fragments. All reactions were analysed at Source Bioscience (Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK).

Wheat transformation was carried out by particle bombardment (PDS1000; Bio-Rad) of immature scutella (10–14 days after anthesis) of cv. Cadenza, co-bombarded with the pAHC20 plasmid, containing the selectable marker gene *bar* driven by the constitutive ubiquitin promoter from maize, as described by Sparks and Jones (2009).

Genomic DNA was extracted from leaf material using a Promega Wizard kit (Promega (UK) Ltd, Southampton, Hampshire, SO16 7NS, UK). PCR analysis to confirm the presence of the transgenes was carried out using the Rab1 and GSPRNAiR primers described above for the RNAi wheat lines (Supplementary Table S1). The reactions were performed in 25 μl using a 1.1× ReddyMix™PCR Master Mix (1.5 mM $MgCl_2$) from Thermo Scientific (ABgene House, Blenheim Road, Epsom, Surrey, KT19 9AP, UK), also containing ~200 ng of genomic DNA and 0.8 μM of each primer. The cycling conditions were 96 °C for 5 min followed by 32 cycles of 96 °C for 30 s; 56 °C for 30 s; 72 °C for 1 min 30 s and the extension of 72 °C for 10 min for PCR reactions. PCR products were analysed on 1.0% (w/v) agarose gels, stained with ethidium bromide and visualised by UV light.

Zygosity analysis of T1 plant material was carried out as described by Nemeth et al. (2010) to identify transgenic lines with segregation patterns consistent with a single insertion locus.

2.2. Plant growth

Homozygous and azygous (null) T2 segregants descended from the same original RNAi transformants were grown in four replicate pots, with four plants per pot, in a four block design (one replicate of each line i.e. transgenic and null per block) in temperature controlled GM glasshouse rooms with 18 °C–20 °C day and 14 °C–16 °C night temperatures and a 16-h photoperiod provided by natural light supplemented with banks of Son-T 400 W sodium lamps (Osram, Ltd.) giving 400 to 1000 μmol m^{-2} s^{-1} photosynthetically active radiation.

2.3. RNA extraction, quantitative reverse transcription-PCR analysis and semi-quantitative reverse transcription-PCR analyses of RNAi wheat lines

Tissue samples enriched in starchy endosperm cells were isolated from T3 developing grain of wheat (*T. aestivum* cv. Cadenza) RNAi lines at 14 d.p.a by gentle squeezing to remove the pericarp. RNA was extracted as reported by Wilkinson et al. (2013) and DNase (Promega (UK) Ltd, Southampton, Hampshire, SO16 7NS UK) used to remove contaminating DNA. First-strand cDNA synthesis was carried out with a Superscript III Reverse Transcriptase Kit (Invitrogen, Life Technologies Ltd, 3 Fountain Drive, Inchinnan Business Park, Paisley, PA4 9RF, UK). The cDNA synthesis step was carried out as follows in a final 20 μl volume: 2 μg equivalent of DNase treated RNA adjusted to a 12 μl volume with sterile distilled water, 1 μl of 100 μM Oligo-dT primer, 1 μl of 10 mm dNTPs; 65 °C for 5 min followed by 1 min incubation on ice. This was followed by the addition of 4 μl of a 5× first strand cDNA synthesis buffer, 1 μl of 0.1M DTT and 1 μl of Superscript III Reverse Transcriptase; 50 °C for 60 min and 70 °C for 15 min.

Down regulation of the transcript was measured using the SYBR Green Jumpstart™ Taq Readymix™ for quantitative reverse transcription-PCR (Sigma-Aldrich Company Ltd. Dorset, UK). The reaction was carried out in a 20 μl mixture containing 1× Green Jumpstart™ Taq Readymix™, 5 μl 0.4 μM of each primer with the ROX reference dye provided, 5 μl cDNA working volume (1:15 dilution of original cDNA reactions). The following temperature profile was used: an initial denaturation step of 95 °C for 10 min,