Plant Physiology, August 1999, Vol. 120, pp. 951–959, www.plantphysiol.org © 1999 American Society of Plant Physiologists

Hydrophobic Protein Synthesized in the Pod Endocarp Adheres to the Seed Surface[1]

Mark Gijzen*, S. Shea Miller, Kuflom Kuflu, Richard I. Buzzell, and Brian L.A. Miki

Agriculture and Agri-Food Canada, Southern Crop Protection and Food Research Centre, 1391 Sandford Street, London, Ontario, Canada N5V 4T3 (M.G., K.K.); Eastern Cereals and Oilseeds Research Centre, Ottawa, Ontario, Canada K1A 0C6 (S.S.M., B.L.A.M.); and Greenhouse and Processing Crops Research Centre, Harrow, Ontario, Canada N0R 1G0 (R.I.B.)

Soybean (*Glycine max* [L.] Merr.) hydrophobic protein (HPS) is an abundant seed constituent and a potentially hazardous allergen that causes asthma in persons allergic to soybean dust. By analyzing surface extracts of soybean seeds with sodium dodecyl sulfate-polyacrylamide gel electrophoresis and amino-terminal microsequencing, we determined that large amounts of HPS are deposited on the seed surface. The quantity of HPS present varies among soybean cultivars and is more prevalent on dull-seeded phenotypes. We have also isolated cDNA clones encoding HPS and determined that the preprotein is translated with a membrane-spanning signal sequence and a short hydrophilic domain. Southern analysis indicated that multiple copies of the *HPS* gene are present in the soybean genome, and that the *HPS* gene structure is polymorphic among cultivars that differ in seed coat luster. The pattern of *HPS* gene expression, determined by in situ hybridization and RNA analysis, shows that HPS is synthesized in the endocarp of the inner ovary wall and is deposited on the seed surface during development. This study demonstrates that a seed dust allergen is associated with the seed luster phenotype in soybean and that compositional properties of the seed surface may be altered by manipulating gene expression in the ovary wall.

Angiosperm seeds are composite structures that develop from fertilized ovules. The essential components, the embryo, endosperm, and seed coat, each have separate developmental origins and fates in the reproductive cycle. Although these features are common to all angiosperms, seeds from different species follow distinct developmental patterns that produce a vast array of sizes, shapes, colors, textures, and compositions.

The development of complex, highly differentiated seed coats is a general feature of legumes and is a characteristic that is often used as an aid for their identification and classification (Corner, 1951; Esau, 1977). At maturity, the seed coat tissues of soybean (*Glycine max* [L.] Merr.) consist of several cell layers that together constitute 4% to 8% of the seed mass. The color, luster, and permeability of the seed and its resistance to seed-borne diseases are all properties that may be determined by the seed coat and associated tissues. The composition, texture, and nutritional value of food and feed products derived from the seed are also influenced by the presence of the seed coat. For these reasons, we are interested in identifying genes that control seed coat traits.

Seeds of *Glycine* spp. are highly variable in their surface texture and appearance. In many wild species the seed coat is completely obscured by the adherence of endocarp to the seed surface (Wolf and Baker, 1972; Newell and Hymowitz, 1978). Specifically, it is the membranous inner epidermis of the endocarp that detaches from the other tissues of the pericarp, or pod wall, to cover the seed. The presence of adhering endocarp on the seed also occurs in the domesticated soybean and influences the luster of the seed surface. As shown in Table I, many different seed luster phenotypes have been described, including shiny, intermediate, dull, light bloom, bloom, and dense bloom (Bernard and Weiss, 1973; Juvik et al., 1989). Three complementary genes, *B1*, *B2*, and *B3*, have been proposed to control the development of bloom (Woodworth, 1933; Goudong et al., 1987), but there is no genetic model to account for all of the different luster phenotypes observed. For example, most soybean cultivars are described as having either dull or shiny seed coats, yet genetic and biochemical control of this trait remains undetermined.

To begin to resolve these uncertainties, we have compared proteins occurring on the surface of seeds with different luster phenotypes. We show that a previously characterized allergen, HPS, is an abundant seed surface protein associated with dull-seeded phenotypes. We isolated clones encoding cDNA copies of HPS to study the expression and structure of the *HPS* gene in different seed luster phenotypes. We show that HPS is synthesized in the endocarp and deposited on the seed surface. Furthermore, there is variation in the amount of HPS present among different soybean lines that arises from polymorphic *HPS* gene structure. Overall, our results suggest a functional role for HPS in influencing the physical properties of the seed surface, and illustrate how seed phenotype and allergenicity are linked.

[1] This research was supported in part by a grant from the Ontario Soybean Grower's Marketing Board.

* Corresponding author; e-mail gijzenm@em.agr.ca; fax 519–457–3997.

Abbreviations: HPS, hydrophobic protein from soybean; SEM, scanning electron microscope(y).

Table I. *Seed coat luster phenotypes for soybean*[a]

Soybean seeds show variation in light-reflective and surface properties.

Abbreviation	Phenotype	Description
B	Bloom	Heavy coating of powdery substance adhering to seed coat
Db	Dense bloom	Heavy bloom
Lb	Light bloom	Slight bloom
D	Dull	Trace amounts of bloom
I	Intermediate	Between dull and shiny
S	Shiny	Absence of bloom

[a] Juvik et al. (1989).

MATERIALS AND METHODS

Plant Materials

Soybean (*Glycine max* [L.] Merr.) seed was from the collections at Agriculture and Agri-Food Canada in Harrow and Ottawa, Ontario. Plants were grown in field plots outdoors or in glass-enclosed greenhouses. The Clark isoline L69-4544 is a self-colored (*i/i*) bloom (*B1/B1*) genotype, hereafter referred to simply as "Clark *B1*". This isoline originated from the U.S. Department of Agriculture Soybean Germplasm Collection and was derived through backcrossing L67-3469 (6) × cv Sooty, where L67-3469 is a self-colored (*i/i*) Clark mutant.

Seed Surface Protein Analysis

Seed surface proteins of different soybean cultivars were compared by SDS-PAGE analysis. A single seed was placed in a 2-mL plastic-capped test tube, and surface proteins were extracted by adding 0.5 mL of a buffer-detergent solution containing 10 mM Tris-Cl, pH 7.5, 0.5% (v/v) SDS, and 20 mM DTT and placing the tube in a boiling water bath for 2 min. The contents of the tube were mixed and an aliquot was withdrawn and centrifuged for 5 min at 14,000*g*. Freshly prepared loading buffer containing 20 mM DTT was added to the sample and proteins were electrophoretically separated on 15% acrylamide gels in the presence of SDS using a modified Laemmli system, as described by Fling and Gregerson (1986). The DTT was omitted from the extraction solution but included in the loading buffer at a range of concentrations to determine its effect on protein migration. Fixation and visualization of the proteins by silver staining followed the method of Blum et al. (1987). Amino-terminal microsequencing of blotted proteins was according to the method of Moos et al. (1988).

Isolation of HPS cDNA Clones and DNA Sequencing

A seed coat cDNA library was constructed from poly(A^+) mRNA isolated from soybean cv Harosoy 63 seeds in the mid to late stage of development (Gijzen, 1997). A sample of the total amplified library was used to subclone inserts from the original vector (Lambda ZAP, Stratagene) into pBK-CMV (Stratagene). Random clones were chosen from this mass excision for plasmid purification and DNA sequencing to establish an expressed sequence tag database of seed coat genes. Automated sequencing of DNA was performed (model 377, Applied Biosystems) using dye-labeled terminators. These DNA sequences were searched for open reading frames encoding HPS by using the BLASTX program at the National Center for Biotechnology Information site (http://www.ncbi.nlm.nih.gov/).

DNA and RNA Hybridizations

Soybean genomic DNA was isolated from frozen, lyophilized tissue according to the method of Dellaporta et al. (1983). Restriction enzyme digestion of 30 μg of DNA, separation on 0.5% agarose gels, and blotting to nylon membranes followed standard protocols (Sambrook et al., 1989). Digoxigenin-labeled cDNA was prepared and used to probe DNA blots according to the instructions provided by the manufacturer (Boehringer Mannheim). Hybridization was carried out at 65°C for 16 h in 0.25 M Na_2HPO_4, pH 7.2, 20% (w/v) SDS, 1 mM EDTA, and 0.5% (w/v) blocking reagent (Boehringer Mannheim). Filters were then washed four times for 15 min each at 22°C in high-stringency wash solution (20 mM Na_2HPO_4, pH 7.2, 1% [w/v] SDS, and 1 mM EDTA), followed by three 15-min washes in the same solution at 68°C.

Total RNA was isolated from roots, stems, leaves, flowers, pods, seed coats, and embryos dissected from soybean plants at various stages of development according to published methods (Wang and Vodkin, 1994). Samples of total RNA (10 μg each) were electrophoretically separated in formaldehyde gels and briefly stained with ethidium bromide to ensure equal loading of samples prior to blotting to nylon membranes. Filters were preincubated at 65°C for 4 h in 0.25 M Na_2HPO_4, pH 7.2, 1% (w/v) BSA, 7% (w/v) SDS, and 1 mM EDTA. The hybridization solution was identical to that used for preincubation, except that 2.5 ng mL^{-1} [^{32}P]cDNA probe was added. Filters were hybridized for 16 h at 65°C, and then washed several times at 68°C and at 22°C.

In Situ Hybridization Analysis of *HPS* Gene Expression

Tissue samples were fixed in a solution of 50% ethanol, 5% acetic acid, and 3.7% formaldehyde (all solutions v/v) for 3 h at room temperature, dehydrated in an ethanol series (50%, 60%, 70%, 80%, 90%, 95%, and 100%), and infiltrated with *t*-butyl alcohol in a stepwise series. Samples were then embedded in paraffin embedding medium (Paraplast, Sigma), placed in blocks, and allowed to harden. Sections of 8 to 10 μm were cut on a rotary microtome and affixed to glass slides. Prior to hybridization, sections were dewaxed in xylene and rehydrated in an ethanol series (100%, 95%, 85%, 70%, 50%, 30%, 15%, and 0% ethanol in distilled, RNase-free water). Sections were then treated with proteinase K and acetylated with acetic anhydride in triethanolamine. Antisense [^{35}S]RNA probes were generated from the *HPS* cDNA clone. Hybridization methods followed published protocols (Cox and Goldberg, 1988). Sections were hybridized overnight at 42°C, then washed and dehydrated in an ethanol series before appli-

cation of track emulsion (NTB-2, Kodak). After 1 week at 4°C, slides were developed in (D-19 developer, Kodak), fixed (Kodak), and briefly stained in Toluidine Blue O. They were then dehydrated in an ethanol and xylene series and placed in synthetic mounting medium (Permount, Fisher Scientific). Slides were photographed on slide film (EPL 400, Kodak) using dark-field optics.

SEM and Droplet-Surface Analysis

Whole, fully mature seeds were mounted to stages with conductive adhesive, sputter coated with gold, and examined using an ISI-DS-130 (International Scientific Instruments, Tokyo, Japan) or a field emission SEM (model S-4500, Hitachi, Tokyo). For the contact angle analysis, seeds were analyzed using a contact angle goniometer (model 100, Ramè-Hart, Mountain Lake, NJ) equipped with a microsyringe attachment. A random sample of four or five individual seeds were measured for each cultivar using water as a probe liquid. To measure static angles, 4 μL of water was deposited on the seed surface. More water was added to the drop to measure the advancing angle.

RESULTS

HPS Occurs on the Seed Surface

To determine the composition of proteins deposited on the soybean seed surface, seeds were washed with a detergent-buffer solution and the extracted peptides were separated by SDS-PAGE. Protein extracts from the seed coat and embryo were also prepared for comparison. These results are shown in Figure 1A. The embryo and seed coat extracts contained many proteins covering a wide range of molecular masses. In contrast, extracts from the seed surface were dominated by a few low-molecular-mass proteins. Initial inconsistencies in the quantity and composition of the surface-extracted proteins was found to result from two main factors: First, oxidation of DTT in the gel loading buffer caused striking changes in the peptides detected by this analysis (Fig. 1B); fresh solutions containing high concentrations of DTT were required to obtain consistent patterns. Second, the amount of protein detected in these extracts varied greatly among different soybean cultivars. Figure 1C shows that the presence of surface protein is correlated with the luster, or light-reflective, properties of the seed surface. Surface extracts from shiny-seeded phenotypes usually contained far less protein than dull-seeded extracts. Moreover, there were large differences in the amount of protein present on the seed surfaces of the two bloom phenotypes examined. To determine the connection between surface protein and seed phenotype, seeds of 80 F_2 plants developed from a cross of dull (OX281) and shiny (cv Mukden) parents were scored for luster and the presence of surface protein. This analysis clearly indicated that the presence of surface protein either contributes to the development of dull phenotypes or that corresponding genes controlling seed luster and surface protein are tightly linked in this cross. The genetics are being studied further and will be reported elsewhere.

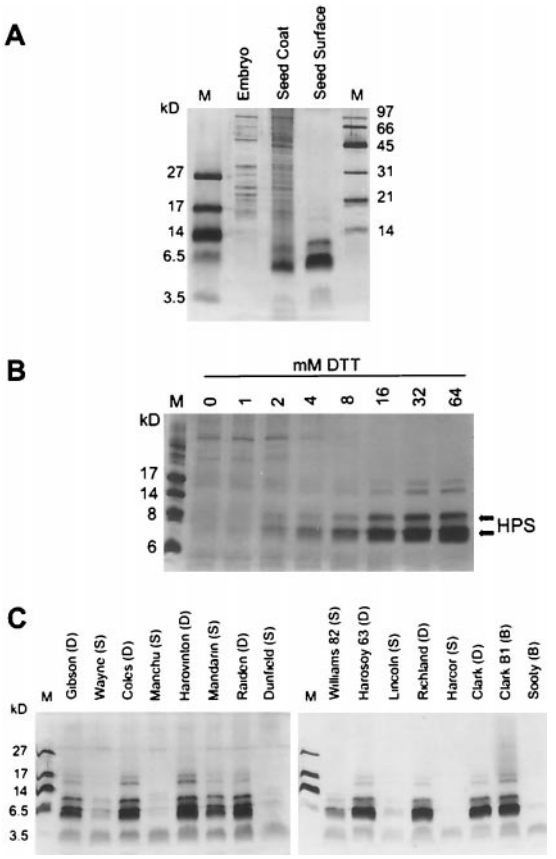

Figure 1. SDS-PAGE analysis of protein extracts from seed tissues and surface. Shown are silver-stained protein gels. Lanes marked "M" indicate protein standards, and their corresponding mass in kilodaltons is provided at left. A, Soluble protein extracts from the embryo, seed coat, and seed surface of a dull phenotype (cv Harosoy 63). Each sample was approximately 1 μg of total protein. B, Seed surface protein extracts of a dull phenotype (cv Harosoy 63) with different concentrations of DTT present in the sample loading buffer, as indicated at the top of each lane. C, Seed surface protein extracts of dull (D), shiny (S), and bloom (B) phenotypes, as indicated at the top of each lane.

Next we wanted to identify the most abundant of these seed surface proteins. Two peptides were purified and subjected to amino-terminal microsequencing, as indicated in Figure 1B. The resulting amino acid sequences were identical and matched existing sequences in the protein database for HPS (Odani et al., 1987; Baud et al., 1993) and soybean dust allergen (Gonzalez et al., 1995). Both peptides had alternative N-terminal residues of Ala or Ile, as has been previously noted for HPS. The different electrophoretic mobilities of the two peptides could not be accounted for from the microsequencing analysis, but may have been due to differences in glycosylation.

The HPS Preprotein Contains a Signal Sequence and a Short Hydrophilic Domain

To obtain the cDNA transcript of HPS, sequences in a seed coat EST database were searched for reading frames corresponding to the HPS amino acid sequence. Using this

strategy, several identical cDNA transcripts that included in their reading frames peptide sequences exactly matching HPS were isolated. A 700-bp transcript that was fully sequenced included 30 bp of 5′ untranslated region, an open reading frame of 119 amino acids, and 313 bp of 3′ untranslated region. The complete deduced amino acid sequence of HPS is shown in Figure 2A. The final 80 residues of this sequence correspond to the peptide sequence reported for the HPS (Odani et al., 1987). Thus, the cDNA transcript indicates that HPS is translated with a leader sequence of 39 amino acids that is cleaved during processing. Figure 2B shows that this long leader sequence consists of a hydrophobic membrane-spanning domain and a short hydrophilic domain. This is significant because similar structural features occur in a group of hybrid proteins identified from several plant species and in plant lipid transfer proteins.

The hybrid or bimodular proteins are so named because their deduced peptide sequences consist of two discrete domains, one hydrophobic and one hydrophilic. Examples

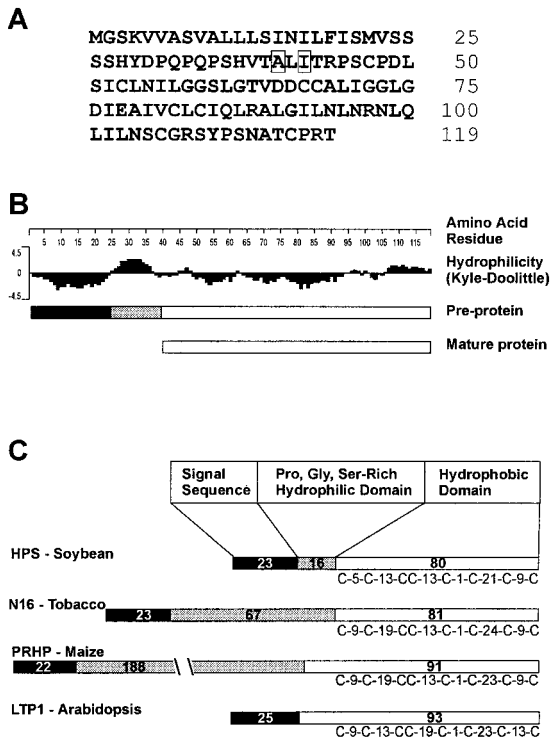

Figure 2. A, Deduced amino acid sequence of HPS preprotein. Alternate N-terminal residues, as determined by peptide microsequence analysis, are boxed. B, A Kyle-Doolittle hydrophilicity plot of HPS (LASERGENE software, DNASTAR, Madison, WI). In this plot, positive values indicate greater hydrophilic character. Also represented are the three domains of the HPS preprotein and the length of the mature peptide. C, A schematic comparison of HPS domain structure to three other plant proteins. Bold numbers indicate the length in amino acid residues for the domain segments. The pattern of spacing between the eight Cys residues within the hydrophobic domains is also shown below each protein. Sequences for the tobacco N16 polypeptide (accession no. D86629), the maize Pro-rich hydrophobic protein (PRHP) (accession no. X60432), and the Arabidopsis lipid transfer protein 1 (LTP1) (accession no. M80567) were retrieved from GenBank.

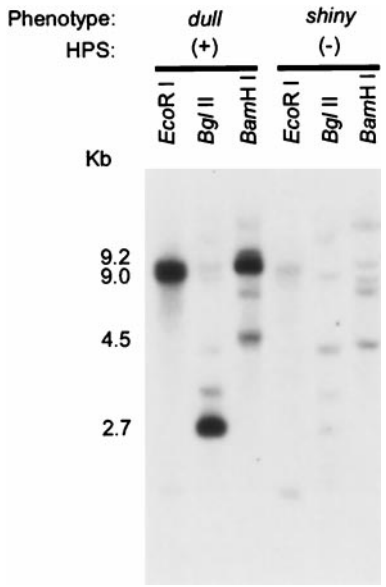

Figure 3. Restriction fragment length polymorphisms between dull and shiny phenotypes. Genomic DNA from dull (cv Harosoy 63) and shiny (cv Williams 82) soybeans with abundant (+) or trace (−) amounts of HPS on the seed surface was digested with restriction enzymes, electrophoretically separated, blotted, and hybridized to the HPS cDNA probe. The size of hybridizing fragments was estimated by comparison with standards and is shown on the left (in kb).

of two of these hybrid proteins and a lipid transfer protein are compared with HPS in Figure 2C. This comparison shows that all of these proteins possess an N-terminal membrane-spanning signal sequence and a 9- to 10-kD hydrophobic domain with eight regularly spaced Cys residues. However, in HPS and the hybrid proteins, the N-terminal signal sequence and the hydrophobic domain are interrupted by a hydrophilic domain. The hydrophilic domains of these proteins are highly variable in their length and in their amino acid sequence and compositions.

Different Seed Luster Phenotypes Show Polymorphic *HPS* Gene Structure

To compare *HPS* gene structure in two different seed luster phenotypes that were also different in the amount of HPS present on the seed surfaces, we hybridized genomic DNA blots with probes derived from the HPS cDNA sequence under high-stringency conditions. A typical result from such a Southern analysis is shown in Figure 3. Genomic DNA blots from cultivars that accumulated large amounts of HPS on the seed surface produced strong hybridization signals. These intensely hybridizing fragments were not present in genomic DNA from plants with only trace amounts of HPS on the seed surface. However, several fainter signals were also present in DNA blots from both types of plants. These results indicate that sequences related to the HPS cDNA are prevalent in the soybean genome, and that the *HPS* gene structure is polymorphic among soybean cultivars. Soybean types that accumulate large amounts of HPS on the seed surface possess additional copies of this gene.

High Expression of *HPS* Occurs in the Pod Endocarp

Developmental and tissue-specific expression patterns for *HPS* were determined by RNA analysis and in situ hybridization. Representative RNA blots probed with *HPS* cDNA are shown in Figure 4. These results show that *HPS* is highly expressed in the pod during the mid to late stages of seed development. Hybridization signals were also observed in seed coat RNA samples. No expression was evident in the flower, leaf, embryo, stem, or root. We also compared *HPS* transcript levels of two different seed luster phenotypes that differ in the amount of HPS present on their seed surfaces. Figure 4B shows that HPS mRNA levels are several times greater in dull-seeded plants that accumulate large amounts of HPS on the seed surface compared with shiny-seeded plants that have only trace amounts of HPS on the seed surface. Faint signals corresponding to low *HPS* transcript levels were detectable in shiny-seeded phenotypes after prolonged exposure times (not shown).

Localization of *HPS* gene expression by in situ hybridization is shown in Figure 5. At 6 DPA, the expression of *HPS* was limited to the membranous inner layer of the pericarp. By 12 DPA expression was very strong and the inner epidermis was showing signs of becoming detached from the rest of the pericarp (and in places was adhering to the seed surface). Tissue sections from this stage of development also showed strong hybridization signals in the sclerenchyma, indicating that *HPS* expression occurs throughout the endocarp.

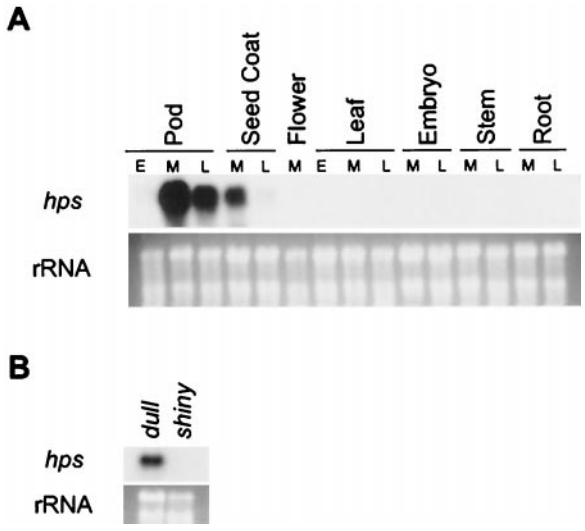

Figure 4. Analysis of *HPS* gene expression by RNA hybridization. Total RNA was isolated from leaf, flower, pod shells, seed coat, embryo, stem, or root tissue. Equal amounts of RNA (10 μg) were blotted to nylon and probed with *HPS* cDNA. rRNA, visualized by staining with ethidium bromide, is shown as control. A, RNA from tissues at early (E), mid (M), or late (L) stages of development were compared for *HPS* gene expression. All samples shown are from a dull-seeded phenotype (cv Harosoy 63). B, RNA from pod tissues of dull (cv Harosoy 63)- and shiny (cv Williams 82)-seeded soybeans were compared for *HPS* gene expression.

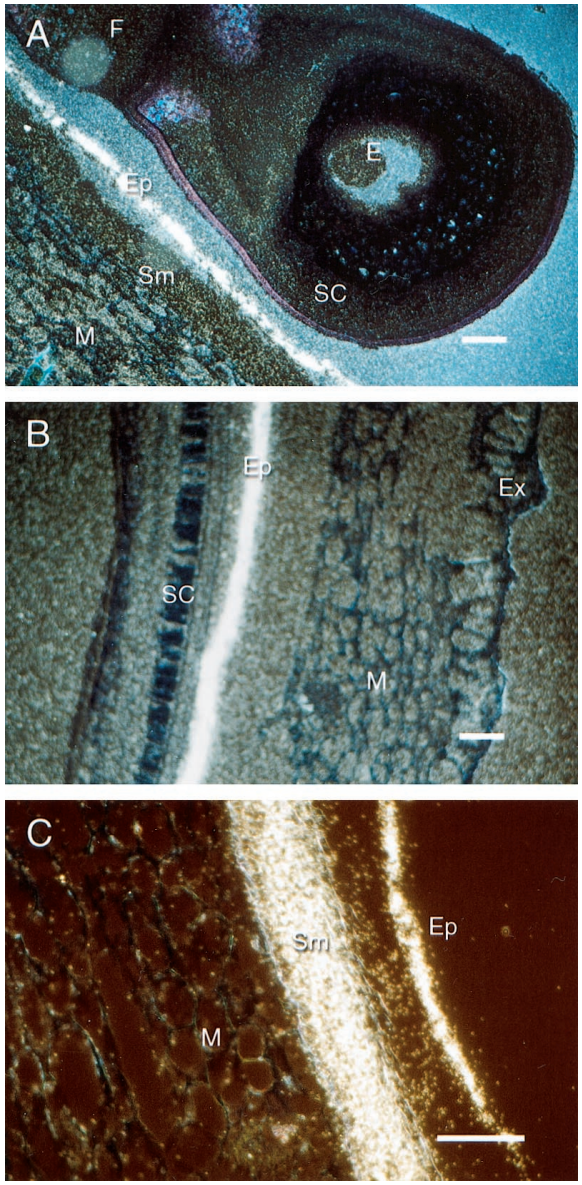

Figure 5. Localization of *HPS* mRNA transcript by in situ hybridization. Cross-sections of soybean pods containing immature seeds (dull phenotype, HPS [+], cv Maple Presto). Hybridization of ^{35}S-labeled *HPS* probe to complementary mRNA appears as a bright white signal in these dark-field microscopy images. E, Embryo; Ep, inner epidermal layer of endocarp; Ex, exocarp; F, funiculus; M, mesocarp; SC, seed coat; Sm, sclerenchyma layer of endocarp. Bar = 100 μm. A, Expression at 6 DPA. B and C, Expression at 12 DPA.

Physical Properties of the Seed Surface Are Affected by the Luster Phenotype

Figure 6 shows SEM images of the seed surfaces of four soybean cultivars. The four cultivars represent three distinct surface phenotypes: shiny, dull, and bloom. The dull-seeded cv Clark and its bloom isoline Clark *B1* accumulate large amounts of HPS on their surfaces, whereas bloom cv Sooty and shiny cv Williams 82 have only trace amounts of HPS. SEM analysis showed that the shiny seeded soybeans have a relatively smooth and undulating surface, whereas

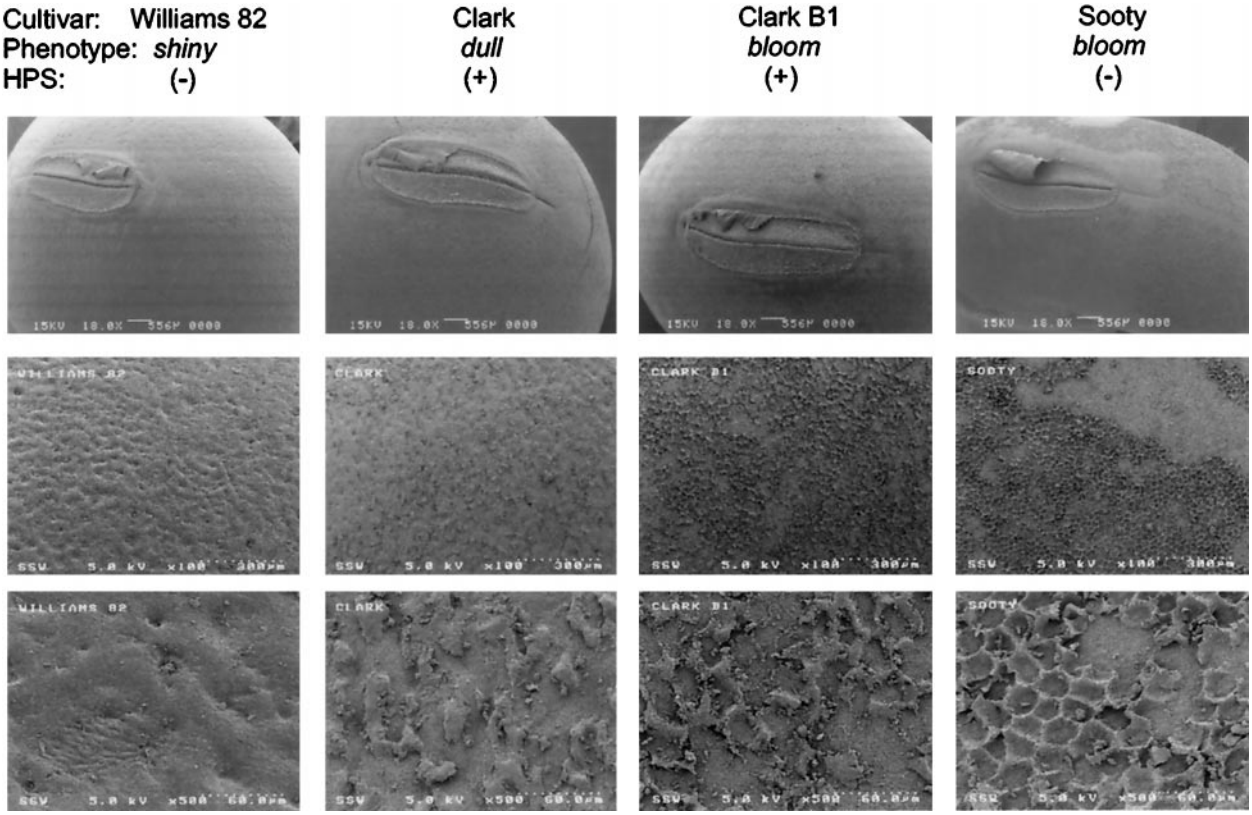

Figure 6. SEM micrographs of seed surfaces of shiny, dull, and bloom phenotypes. Four different combinations of phenotype and HPS content (−, trace; +, abundant) are shown at three magnifications. The lowest magnifications (top micrographs) show views of the whole seeds. The large, oval-shaped scar on the seed surface is the hilum, corresponding to the point of detachment of the mature seed from the funiculus. Higher magnifications are focused outside of hilum region. Lengths of scale bars or dashed lines are indicated in micrometers. Lengths across the horizontal field of view for each of the magnifications are: 7.1 mm (top); 1.1 mm (middle); and 0.2 mm (bottom).

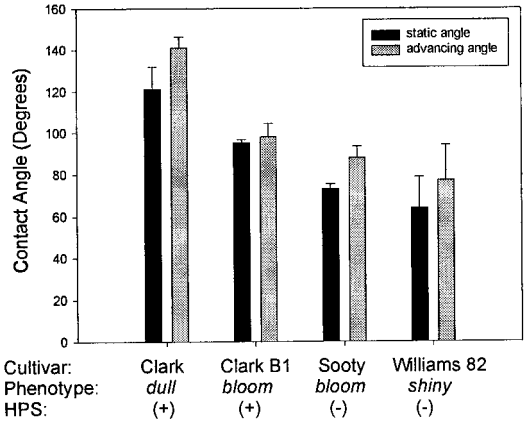

Figure 7. Surface droplet contact angles for seeds of shiny, dull, and bloom phenotypes. Four different combinations of phenotype and HPS content (−, trace; +, abundant), corresponding to the four cultivars shown in Figure 6, were compared for surface droplet contact angles. Values are means and SE values for four or five independent measurements.

dull types are uniformly covered with bits of adhering endocarp. Large patches of contiguous membranous endocarp produce a honeycomb-like pattern on the surface of bloom phenotypes, although this tissue appears more fragmented in Clark *B1* than in cv Sooty.

Static and advancing surface-droplet contact angles were also compared for the four soybean cultivars to determine how seed phenotype and HPS may affect surface hydrophobicity. In this analysis, high contact angles were characteristic of hydrophobic surfaces but may have also resulted from differences in surface topography. As shown in Figure 7, the highest contact angles were observed for seeds that accumulated large amounts of HPS on the surface. Dull-seeded phenotypes consistently displayed the greatest contact angles, higher than either bloom or shiny phenotypes.

DISCUSSION

Soybean seeds display a wide variety of phenotypes that differ in coloration, size, shape, luster, and permeability. For example, self-colored (black)-seeded phenotypes differ markedly from the commonly grown yellow seeded types.

This trait is in part determined by the *I* locus, a cluster of chalcone synthase genes that control anthocyanin biosynthesis in the seed coat (Todd and Vodkin, 1996). There is also variation in the composition of proteins from seed coats of different soybean varieties (Lindstrom and Vodkin, 1991; Gijzen et al., 1993), and corresponding genes encoding both structural and soluble seed coat proteins have been isolated (Schmidt et al., 1994; Gijzen, 1997). Despite these examples, there are no clear genetic or biochemical models to account for many of the observed phenotypes. Thus, we undertook a study comparing seed surface protein composition to the luster, or light-reflective, properties of the seed surface.

To differentiate proteins that are present in the tissues of the seed coat from those that are deposited on the surface of the seed, we prepared seed surface extracts without dissection or homogenization. This analysis resulted in the identification of HPS as an abundant seed surface protein and provided a link between HPS and dull phenotypes. Whereas HPS has been purified and characterized as a seed constituent and a potent allergen, there have been no studies on the expression, localization, or function of the protein or any description of the corresponding gene. Our initial results raised many questions that could only be addressed by a more extensive investigation of HPS.

The association of HPS and seed luster phenotypes was further tested by scoring different soybean cultivars and a segregating F_2 population for HPS and luster phenotype. This revealed a strong association between HPS and dull phenotypes in soybean cultivars and in the F_2 population. However, the quantity of HPS on the seed surface is not simply dependent upon the amount of adhering endocarp tissue, since the bloom phenotype cv Sooty possessed a heavy coating of endocarp tissue but only trace amounts of surface HPS. The integration of the *B1* gene from cv Sooty into Clark *B1* apparently occurred without loss of the abundant surface HPS present in the recurrent Clark parent. Although cv Sooty and Clark *B1* are both described as bloom phenotypes, SEM analysis showed that the endocarp is more fragmented in Clark *B1*. This fragmentation may result from higher levels of *HPS* gene expression in Clark *B1*.

The interrelationships among seed luster, adhering endocarp, and HPS are not entirely clear, but the present study did suggest the following. The amount of endocarp tissue adhering to the seed influences the luster of the surface in a quantitative manner. The progression from shiny to intermediate, dull, and bloom phenotypes seems to depend mostly on the amount of adhering endocarp tissue. However, the appearance of the underlying surface and the pattern of attachment of the endocarp may also be important contributing factors. Taken together, the evidence suggests that seed luster is a quantitative trait determined by several loci. Thus, the expression of HPS in the endocarp may be one factor of many that influence how this tissue clings to the seed surface and produces a spectrum of luster phenotypes. It is also possible that HPS does not have any role in the fragmentation or attachment of the endocarp to the seed, but that it is tightly linked to other genes that control this trait. Regardless, DNA and RNA analysis clearly shows that *HPS* gene structure and transcript levels are very different in plants that accumulate large amounts of HPS on the seed surface than in those that do not.

Isolation of cDNA clones encoding HPS has provided the complete sequence of the protein precursor to HPS and confirmed its relationship to a group of hybrid Pro-rich and extensin-like proteins from several other plant species. All of these proteins possess a distinct hydrophobic domain of 80 to 100 amino acids encompassing eight regularly spaced Cys residues. Transcripts encoding hybrid proteins have been isolated from many different plant species under conditions of cold (Castonguay et al., 1994), high salt (Deutch and Winicov, 1995), mechanical stress (Huang et al., 1998), or tissue-specific selection (Josè-Estanyol et al., 1992; Coupe et al., 1993; Yasuda et al., 1997). Ascribing functional roles to these proteins has been difficult and in no case has a protein of this type been associated with a phenotypic character.

Plant lipid transfer proteins also show similarity to HPS in size, hydrophobicity, and in the number and spacing of Cys residues in the peptide chain. These proteins are commonly found on leaf surfaces, where they are thought to participate in cuticle biosynthesis and possibly in defense and environmental adaptation (Kader, 1996). Another group of small, Cys-rich, hydrophobic proteins that occur on surfaces are the fungal hydrophobins, a group of secreted proteins that cover hyphae or reproductive structures and influence physical properties of the fungal surface (Wessels, 1997; Kershaw and Talbot, 1998). Thus, a common feature shared by HPS, many lipid transfer proteins, and fungal hydrophobins is surface localization. These compact, hydrophobic, and Cys-rich proteins offer properties that make them attractive for covering surfaces. For example, Sc3p is a hydrophobin from *Schizophyllum commune* that self assembles in vitro to form rodlet structures identical to those occurring on the surface of aerial hyphae (Wösten et al., 1994). The capacity of HPS to quickly crystallize out of solution (Odani et al., 1987) and the requirement for high concentrations of DTT to reduce soluble extracts of the protein to monomers demonstrates that HPS also has strong self-associative properties.

Results from RNA analysis suggest that *HPS* is highly expressed in both the pod and seed coat tissues during the mid to late stages of development. However, localization of *HPS* mRNA by in situ hybridization suggests that HPS expression is tightly restricted to the inner epidermis and sclerenchyma of the pod endocarp. Hybridization signals observed in seed coat RNA blots are likely due to contamination of the seed coat with the membranous inner epidermis of the pericarp, since this tissue sticks to surface of developing seeds and is difficult to completely remove. Thus, we conclude that *HPS* is specifically expressed in the endocarp. Proteins expressed in this tissue, or whole sections of the inner epidermis itself, adhere to the seed surface during development and become a component of the seed coat of mature, fully developed soybeans.

Odani et al. (1987) estimated the abundance of HPS to be in the range of 200 mg kg^{-1} whole seed. The presence of such large amounts of protein, restricted entirely to the seed surface, would alter the physical properties of the

surface and suggest a structural or defensive function for the protein. Results from contact angle analysis of surface droplets provide correlative evidence that HPS reduces the wettability of seed surfaces. The hydrophobicity and topography of the surface could affect pathogen attachment and penetration or influence the water-absorptive properties of the seed. It is also possible that HPS acts directly as a feeding deterrent or toxin against specific herbivores, pests, or pathogens. More experimentation is required to clarify the functional role of HPS.

The demonstration that large amounts of HPS are present on the seed surface is consistent with the localization of the soybean dust allergen to the seed hull fraction (Rodrigo et al., 1990; Swanson et al., 1991), since this allergen was subsequently identified as HPS (Gonzalez et al., 1995). Re-occurring, community-wide outbreaks of asthma in Barcelona and Cartagena (Spain) from 1981 to 1987 were caused by the release of soybean dust through the unloading of seed from container vessels (Antó et al., 1989). These epidemics affected hundreds of individuals and resulted in several deaths (Antó et al., 1993). Soybean dust is also the probable cause of earlier asthma outbreaks in other cities, including New Orleans (Weill et al., 1964), and is listed as a workplace hazard for food industry workers (Pepys, 1986).

Our work offers new opportunities for lessening the health hazard of seed dust exposure. For example, phenotypic or genetic screens may be devised to select plants with reduced amounts of HPS on the seed surface. More broadly, results presented here indicate that physical, textural, or compositional properties of the seed surface may be altered by manipulating gene expression in the ovary wall.

ACKNOWLEDGMENTS

We thank Ross Davidson, Mark Biesinger, and Mary Jane Walzak at Surface Science Western for electron microscopy and droplet-surface analysis; Pearl Campbell and Heather Schneider at the Robarts Research Institute for DNA sequencing; Aldona Gaidauskas-Scott and Lu-Ann Bowman for technical assistance; Dorothy Drew for library services; and the Biotechnology Service Centre at the University of Toronto for peptide microsequencing.

Received February 3, 1999; accepted May 12, 1999.
The accession number for the sequence reported in this article is AF100159.

LITERATURE CITED

Antó JM, Sunyer J, Reed CE, Sabrià J, Martínez F, Morrell F, Cordina R, Rodriguez-Roisin R, Rodrigo MJ, Roco J, et al. (1993) Preventing asthma epidemics due to soybeans by dust-control measures. N Engl J Med **329:** 1760–1763

Antó JM, Sunyer J, Rodriguez-Roisin R, Suarez-Cervera M, Vazquez L (1989) Community outbreaks of asthma associated with inhalation of soybean dust. N Engl J Med **320:** 1097–1102

Baud F, Pebay-Peyroula E, Cohen-Addad C, Odani S, Lehmann MS (1993) Crystal structure of hydrophobic protein from soybean: a member of a new cysteine-rich family. J Mol Biol **231:** 877–887

Bernard RL, Weiss MG (1973) Qualitative genetics. *In* BE Caldwell, ed, Soybeans: Improvement, Production, and Uses. American Society of Agronomy, Madison, WI, pp 117–154

Blum H, Beier H, Gross HJ (1987) Improved staining of plant proteins, RNA and DNA in polyacrylamide gels. Electrophoresis **8:** 93–99

Castonguay Y, Laberge S, Nadeau P, Vézina LP (1994) A cold-induced gene from *Medicago sativa* encodes a bimodular protein similar to developmentally regulated proteins. Plant Mol Biol **24:** 799–804

Corner EJH (1951) The leguminous seed. Phytomorphology **1:** 117–150

Coupe SA, Taylor JE, Isaac PG, Roberts JA (1993) Identification and characterization of a proline-rich mRNA that accumulates during pod development in oilseed rape (*Brassica napus* L.). Plant Mol Biol **23:** 1223–1232

Cox KH, Goldberg RB (1988) Analysis of plant gene expression. *In* CH Shaw, ed, Plant Molecular Biology: A Practical Approach. IRL Press, Oxford, pp 1–35

Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA minipreparation: version II. Plant Mol Biol Rep **1:** 19–21

Deutch CE, Winicov I (1995) Post-transcriptional regulation of a salt-inducible alfalfa gene encoding a putative chimeric proline-rich cell wall protein. Plant Mol Biol **27:** 411–418

Esau K (1977) Anatomy of Seed Plants. Wiley, New York

Fling SP, Gregerson DS (1986) Peptide and protein molecular weight determination by electrophoresis using a high-molarity Tris buffer system without urea. Anal Biochem **155:** 83–88

Gijzen M (1997) A deletion mutation at the *ep* locus causes low seed coat peroxidase activity in soybean. Plant J **12:** 991–998

Gijzen M, van Huystee R, Buzzell R (1993) Soybean seed coat peroxidase: a comparison of high-activity and low-activity genotypes. Plant Physiol **103:** 1061–1066

Gonzalez R, Varela J, Carreira J, Polo F (1995) Soybean hydrophobic protein and soybean hull allergy. Lancet **346:** 48–49

Guodong Z, Jinling W, Qingxi M (1987) Inheritance of bloom on seed coat in soybean. Soybean Genet Newsl **14:** 91–93

Huang YF, Jordan WR, Wing RA, Morgan PW (1998) Gene expression induced by physical impedance in maize roots. Plant Mol Biol **37:** 921–930

Josè-Estanyol M, Ruiz-Avila L, Puigdomènech P (1992) A maize embryo-specific gene encodes a proline-rich and hydrophobic protein. Plant Cell **4:** 413–423

Juvik GA, Bernard RL, Chang R, Cavins JF (1989) Evaluation of the USDA wild soybean germplasm collection: maturity groups 000 to IV (PI 65.549 to PI 483.464). USDA Technical Bulletin No. 1761

Kader JC (1996) Lipid transfer proteins in plants. Annu Rev Plant Physiol Plant Mol Biol **47:** 627–654

Kershaw MJ, Talbot NJ (1998) Hydrophobins and repellents: proteins with fundamental roles in fungal morphogenesis. Fungal Genet Biol **23:** 18–33

Lindstrom JT, Vodkin LO (1991) A soybean cell wall protein is affected by seed color genotype. Plant Cell **3:** 561–571

Moos M, Nguyen NY, Liu TY (1988) Reproducible high yield sequencing of proteins electrophoretically separated and transferred to an inert support. J Biol Chem **263:** 6005–6008

Newell CA, Hymowitz T (1978) Seed coat variation in *Glycine* Willd. subgenus *Glycine* (Leguminosae) by SEM. Brittonia **30:** 76–88

Odani S, Koide T, Ono T, Seto Y, Tanaka T (1987) Soybean hydrophobic protein: isolation, partial characterization and the complete primary structure. Eur J Biochem **162:** 485–491

Pepys J (1986) Occupational allergic lung disease caused by organic agents. J Allergy Clin Immunol **78:** 1058–1062

Rodrigo MJ, Morell F, Helm RM, Swanson M, Greife A, Anto JM, Sunyer J, Reed CE (1990) Identification and partial charac-

terization of the soybean-dust allergens in the Barcelona asthma epidemic. J Allergy Clin Immunol **85:** 778–784

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual, Ed 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Schmidt JS, Lindstrom JT, Vodkin LO (1994) Genetic length polymorphisms create size variation in proline-rich proteins of the cell wall. Plant J **6:** 177–186

Swanson MC, Li JTC, Wentz-Mutha PE, Trudeau WL, Fernadez-Caldas E, Greife A, Rodrigo MAJ, Morell F, Reed CE (1991) Source of the aeroallergen of soybean dust: a low molecularmass glycopeptide from the soybean tela. J Allergy Clin Immunol **87:** 783–788

Todd JJ, Vodkin LO (1996) Duplications that suppress and deletions that restore expression from a chalcone synthase multigene family. Plant Cell **8:** 687–699

Wang CS, Vodkin LO (1994) Extraction of RNA from tissues containing high levels of procyanidins that bind RNA. Plant Mol Biol Rep **12:** 132–145

Weill H, Ziskind MM, Dicherson RC, Derbes VJ (1964) Epidemic asthma in New Orleans. JAMA **190:** 11–14

Wessels JGH (1997) Hydrophobins: proteins that change the nature of the fungal surface. Adv Microbial Physiol **38:** 1–45

Wolf WJ, Baker FL (1972) Scanning electron microscopy of soybeans. Cereal Sci Today **17:** 124–130

Woodworth CM (1933) Genetics of the soybean. J Am Soc Agron **25:** 36–51

Wösten HAB, Ásgeirsdóttir SA, Krook JH, Drenth JHH, Wessels JGH (1994) The fungal hydrophobin Sc3p self-assembles at the surface of aerial hyphae as a protein membrane constituting the hydrophobic rodlet layer. Eur J Cell Biol **63:** 122–129

Yasuda E, Ebinuma H, Wabiko H (1997) A novel glycine-rich/hydrophobic 16 kDa polypeptide gene from tobacco: similarity to proline-rich protein genes and its wound-inducible and developmentally regulated expression. Plant Mol Biol **33:** 667–678