

Protein sorting and expression of a unique soybean cotyledon protein, GmSBP, destined for the protein storage vacuole

Aaron Elmer¹, Wun Chao² and Howard Grimes^{3,*}

¹Program in Plant Physiology, School of Molecular Sciences, Washington State University, Pullman, WA 99164-4234, USA (*author for correspondence; e-mail grimes@wsu.edu); ²Biosciences Research Laboratory USDA-ARS, Fargo, ND 58105-5674, USA; ³School of Molecular Biosciences, Washington State University, Pullman, WA 99164-4234, USA

Received 18 November 2002; accepted in revised form 25 March 2003

Key words: immunolocalization, protein storage vacuole, seed storage proteins, soybean, sucrose-binding protein

Abstract

The initial biochemical characterization of the soybean sucrose-binding protein, GmSBP, within our lab and others produced several incongruous characteristics that required a re-characterization of GmSBP via sequence homology, cell biology, immunolocalization, and semi-quantitative analysis. The GmSBP proteins share amino acid sequence homology as well as putative structural homology with globulin-like seed storage proteins. A comparison to the major soybean seed storage proteins, glycinin and β -conglycinin established several storage protein-like characteristics for GmSBP. All three proteins were present in a prevacuolar compartment and protein storage vacuole. All three proteins increased in expression during seed development and are remobilized during germination. Quantitatively, the relative concentrations of GmSBP, β -conglycinin (α/α' subunits), and glycinin (acidic subunits) indicated that GmSBP contributes 19-fold less to the stored nitrogen. The quantitative differences between GmSBP and glycinin may be attributed to the unconserved order and spacing of cis-acting regulatory elements present within the promoter regions. Ultimately, GmSBP is transported to the mature protein storage vacuole. The biological function of GmSBP within the protein storage vacuole remains uncertain, but its localization is a remnant of its evolutionary link to a globulin-like or vicilin-like ancestor that gave rise to the 7S family of storage proteins.

Abbreviations: GmSBP, *Glycine max* sucrose-binding protein; GmSBP1, *Glycine max* sucrose-binding protein 1; GmSBP2, *Glycine max* sucrose-binding protein 2; GmSUT1, *Glycine max* H⁺-sucrose transporter; 6'-HABS, 6-deoxy-6'-(4-azido-2-hydroxy)-benzamidosucrose; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; VfSBPL, *Vicia faba* sucrose-binding protein-like protein

Introduction

The original soybean sucrose binding protein (GmSBP1) was identified on the basis of its ability to bind 6'-HABS, a competitive inhibitor of sucrose transport in soybean protoplasts. Partial purification of GmSBP1 allowed the production of polyclonal antibodies, which labeled epitopes present in the plasma membrane (Warmbrodt *et al.*, 1989, 1991), as well as unique structures such as multivesicular plasma membrane elaborations (Ripp *et al.*, 1988), regions of invaginated plasma membrane (Grimes *et al.*, 1992),

and cell wall ingrowths (Tegeder *et al.*, 1999). Furthermore, independent expression of GmSBP1 in yeast suggested that this protein is capable of inducing low rates of proton-independent sucrose transport across the yeast plasma membrane (Grimes and Overvoorde, 1996; Overvoorde *et al.*, 1996). Support for GmSBP-mediated sucrose transport was provided by Pedra *et al.* (1999), who used antisense repression and sense over-expression of a GmSBP1 homologue (GmSBP64) to demonstrate altered carbohydrate composition in tobacco leaf.

Despite the apparent role of GmSBP1 in proton independent sucrose transport, it is now well established that this protein is a member of a cupin superfamily (Dunwell *et al.*, 2000). This superfamily is composed of proteins with a modular Swiss roll β -barrel structure and includes the vicilin-like seed storage proteins in plants. Both GmSBP1 and the vicilin-like proteins contain N-terminal signal sequences and N-terminal CXXXC motifs (Kriz, 1989; McHenry and Fritz, 1992; Heck *et al.*, 1993; Braun *et al.*, 1996). Both have a pseudo-dyad primary structure in which the N-terminal and C-terminal domains arose from gene duplication (Argos *et al.*, 1985; Lawrence *et al.*, 1994; Shutov *et al.*, 1995).

Crystal structures have been elucidated for three vicilin-like proteins: phaseolin (Lawrence *et al.*, 1994), canavalin (Ko *et al.*, 1993), and β -conglycinin β -subunit (Maruyama *et al.*, 2001). Each of these crystallographic analyses produced homotrimeric structures with monomers that displayed N-terminal and C-terminal modules consisting of a core Swiss-roll β -barrel and a loop domain centered on a pseudo-dyad axis. Although GmSBP1 and the vicilins are only 20%–37% identical (Overvoorde *et al.*, 1997), models based on the phaseolin and canavalin support the formation of trimeric GmSBP with a canavalin-like structure (Ko *et al.*, 1993; Lawrence *et al.*, 1994).

The sequence and putative structural homology shared by GmSBP and the vicilin-like seed storage proteins proved incompatible with the known topology of integral membrane proteins. Vicilin-like proteins are exclusively localized to the protein storage vacuole, and are not present on the plasma membrane. A recently cloned GmSBP homologue from *Vicia faba* (VfSBPL) shares 58% homology with GmSBP1 at the amino acid level (Heim *et al.*, 2001). Heterologous expression of VfSBPL demonstrated that VfSBPL did not mediate sucrose uptake in yeast. *In situ* hybridization demonstrated that VfSBPL mRNA was localized to the epidermal transfer cell layer and around vascular strands during the prestorage/storage transition phase. Immunocytochemistry demonstrated specific localization of VfSBPL to the protein storage vacuoles (Heim *et al.*, 2001).

In our ongoing studies of the functional biology of GmSBP1 and sucrose transport in soybean cotyledons, we created a large number of site-directed mutants of GmSBP1 in an attempt to elucidate the mechanism of proton independent sucrose transport. Unfortunately, every mutant resulted in a loss of sucrose transporter activity (data not shown) and it was impossible to dis-

cern the kinetic and mechanistic outcomes of these mutations, especially in the context of the initial low rates of sucrose transport mediated by GmSBP1 (Chao and Grimes, unpublished).

Our lab has also cloned two additional proteins: GmSBP2 (91% identical to GmSBP1 within the amino acid sequence) and GmSUT1 (Chao and Grimes, unpublished). GmSUT1 is a member of the MFS superfamily of membrane transport proteins and mediates H^+ -dependent sucrose symport in soybean cotyledons. Kinetic comparisons of the rate of sucrose transport between GmSBP1, GmSBP2, and GmSUT1 strongly suggests that GmSUT1, and possibly other GmSUT1 related proteins, is primarily responsible for sucrose uptake into the soy cotyledon.

Our inability to formulate a mechanism for GmSBP1-mediated sucrose transport as well as the sequence homology shared by GmSBP and the globulin-like seed storage proteins, guided our interest in characterizing the seed storage protein-like characteristics of GmSBP1. To approach this, we generated additional antibodies to GmSBP1 using bacterial expression systems to avoid potential problems with highly antigenic co-purifying plasma membrane proteins. Our characterization of GmSBP extends its homology with seed storage proteins beyond sequence analysis to include seed-specific expression, developmental expression, localization within the endomembrane system, and localization within the protein storage vacuole. However, GmSBP differs quantitatively from the major soybean seed storage proteins, contributing only 2.4% to total cellular protein. These data indicate that while GmSBP1 and GmSBP2 are destined for the protein storage vacuole, their function in soybean seed may not involve nitrogen storage.

Materials and methods

Plant material

Soybean (*Glycine max* (L.) Merr.) seed were planted in 7.56-liter pots and grown as described in Overvoorde and Grimes (1994) with a 16/8 h photoperiod. Temperature was maintained at 24 °C/18 °C over the day/night cycle. Plants were fertilized once a week with 500 mg/l of Peters 20-20-20 fertilizer. Germination experiments utilized seed imbibed in soil, and harvested every 24 h for the duration of the experiment.

Cloning of GmSBP1 and GmSBP2 5'-flanking regions

Genomic DNA was extracted from 1-month-old soybean leaves (Sambrook et al., 1989). Separate aliquots of genomic DNA were digested with *Dra*I, *Eco*RV, *Pvu*II, *Sca*I, or *Stu*I. Each aliquot of genomic DNA was then ligated to the GenomeWalker adapter and used to generate GenomeWalker 'libraries' according to manufacture's instructions (Clontech).

The initial 5'-flanking regions of *GmSBP1* and *GmSBP2* were obtained using a *GmSBP1* gene-specific reverse primer, 5'-AAGAAAAAGATAGCT-AAAGAAAGTTTGG-3', and the AP2 adapter primer (Clontech). The resulting 0.8 kb (pSBP1-0.8P) and 0.9 kb (pSBP2-0.9P) fragments were ligated into the *Eco*RV site of pBluescript-SK(-). Sequencing revealed that pSBP1-0.8P contained a portion of the *GmSBP1* ORF, and pSBP2-0.9P fragment contained a related sequence that would later be used to clone *GmSBP2*.

Cloning of GmSBP2

The second soybean sucrose-binding protein, *GmSBP2*, was cloned from a cDNA library generated from developing cotyledons according to Grimes et al. (1992). The full-length *GmSBP2* cDNA was PCR amplified using the gene-specific forward primer, 5'-TTGTTCTCCTCACTGACCTCACCATGGCGACC-3', and the AP1 adapter primer from CLONTECH. The 1.7 kb PCR fragment was cloned into the *Eco*RV site of pBluescript-SK(-), and sequenced with an ABI373 DNA sequencer (Stretch).

Sequence analysis

Nucleotide and amino acids sequences were imported into Biology Workbench 3.2 (<http://workbench.sdsc.edu>), which is maintained at the University of California-San Diego Supercomputer Center. The non-redundant database was searched using BLASTP (Altschul et al., 1997). Multiple sequence alignments were obtained using CLUSTAL W (Thompson et al., 1994). A phylogenetic dendrogram was produced using PHYLIP 3.2 (Felsenstein, 1989). The following proteins were included in the phylogenetic analysis: *Glycine max* β -conglycinin α -chain (GmCG α), accession number P13916; *Glycine max* β -conglycinin α -prime (GmCG α'), P11827; *Glycine max* β -conglycinin β -chain (GmCG β), P25974; *Phaseolus vulgaris* phaseolin (PvPhas), 1PHS; *Canavalia ensiformis* canavalin (CeCan), JQ2264; *Canavalia gladiata*

canavalin 2 (CgCan2), S00281; *Canavalia gladiata* canavalin 1 (CgCan1), CAA33172; *Vicia narbonensis* convicilin (VnCon), Z71986; *Lens culinaris* convicilin (LcCon), AJ276683; *Pisum sativum* convicilin 1 (PsCon1), CAB82855; *Pisum sativum* convicilin 2 (PsCon2), S00566; *Pisum sativum* vicilin 1 (PsVic1), P13918; *Vicia faba* vicilin (VfVic1), S06309; *Vicia narbonensis* vicilin (VnVic), Z71987; *Pisum sativum* vicilin 2 (PsVic2), S35757; *Artocarpus heterophyllus* jacalin α -chain (AhJac), A46630; *Arachis hypogaea* major allergen (AraH1), AF432231; *Macadamia integrifolia* vicilin 1 (MiVic1), AAD54244; *Macadamia integrifolia* vicilin 2 (MiVic2), AAD54246; *Macadamia integrifolia* vicilin 3 (MiVic3), AAD54245; *Juglans regia* vicilin-like protein (JrVic), AF066055; *Gossypium hirsutum* α -globulin type B 1 (GoGlbB1), S08059; *Gossypium hirsutum* α -globulin type B 2 (GoGlbB1), P09801; *Gossypium hirsutum* α -globulin type A (GoGlbA), P09799; *Theobroma cacao* vicilin 1 (TcVic1), AAF13479; *Theobroma cacao* vicilin 2 (TcVic2), S22477; *Picea glauca* vicilin-like storage protein 2 (PgVic2), S24756; *Picea glauca* vicilin-like storage protein 1 (PgVic1), AAB01554; *Zamia furfuracea* vicilin (ZfVic), CAA90652; *Elaeis guineensis* 7S globulin (EgGlb), AF250228; *Zea mays* globulin-2 (ZmGlb2), S15675; *Hordeum vulgare* globulin (HvBeg1), S35221; *Zea mays* vicilin-like storage protein (ZmGlb1), B53234; *Zea mays* globulin S (ZmGlbS), P15590; *Sesamum indicum* globulin (SiGlb), AF240006; *Vicia faba* SBP-like protein (VfS-BPL), CAC27161; *Pisum sativum* SBP-like protein (PsSBP54), T06459; *Glycine max* sucrose-binding protein S-64 (GmSBP64), AAF05723; *Glycine max* binding protein 2 (GmSBP2), AY234869; *Glycine max* binding protein 1 (GmSBP1), Q04672.

RNAse protection assay

RNAse protection assays were performed using the RPA II kit (Ambion). Antisense *GmSBP1* and *GmSBP2* RNA probes were synthesized *in vitro* by T3 polymerase using linearized pSBP1-0.8P and pSBP2-0.9P vectors. The ³²P-labeled antisense RNAs were hybridized with 2 μ g of poly (A⁺) mRNA extracted from developing seeds. Developing seeds were categorized according to Chao et al. (1995): stage AB <4.0 mm, stage C = 5.0 mm, stage D = 7.0 mm, stage E = 8.0–9.0 mm, stage GI = 1.1–1.2 mm, stage S = initiation of desiccation. The tips from 3 mm, 5 mm, and 10 mm developing leaves were also used for mRNA extraction. Samples were run on a 6% se-

quencing gel for 1.5 h, dried, and exposed to Kodak X-ray film overnight.

Antibodies

The GmSBP C-terminal domain antibody (anti-SBP-CDA) was developed to the C-terminal 165 amino acids (T360 to A525) of GmSBP1 expressed as an MBP fusion protein in *Escherichia coli* BL21DE3 and purified using a dextrin column (New England Biolabs). The GmSBP C-terminal peptide antibody (anti-SBP-CTP) was a synthetic peptide designed from the C-terminal 13 amino acids of GmSBP1 (N'-CELPREERGRADA-'C). Columns used for affinity purification were made by coupling the SBP-CDA antigen or SBP-CTP antigen to SulfoLink resin (Pierce Chemical) according to manufacture's instructions. The protein concentration of the antibodies was determined using the BCA protein assay (Pierce Chemical). When necessary, the antibodies were concentrated by precipitation with 50% ammonia sulfate. Monoclonal actin antibodies (Clone C4) were purchased from ICN Biomedical. Monoclonal antibodies to BiP, an ER-localized protein, were purchased from StressGen Biotechnologies Corp. Anti-mannosidase I antibodies were used as a Golgi marker (Nebenfuhr *et al.*, 1999; Yu-Bing *et al.*, 2002). Glycinin antibodies were provided by Niels Nielsen (Purdue University).

Protein extraction and electrophoresis

Crude protein was extracted by grinding developing soybean seed in liquid N₂. The protein was extracted in buffer containing 25 mM Tris-HCl pH 7.2, 4% w/v SDS, 10% v/v glycerol, 5 mM EDTA, 1 mM PMSF, 1 μ M pepstatin, and 1 μ M leupeptin. The supernatants were clarified by centrifuging at 20 000 $\times g$ for 2 min, and protein concentration was determined with the BCA protein assay (Pierce Chemical). Aliquots of protein were precipitated with a final concentration of 80% acetone (-20°C) and 5% v/v TCA. Samples were incubated at -20°C for 20 min. The protein was pelleted at 10 000 $\times g$ for 15 min. The protein pellet was washed twice in 80% acetone (-20°C), and one time in 100% acetone (-20°C). The protein pellet was then dried under vacuum and stored at -80°C .

Samples analyzed by 2D gel electrophoresis were solubilized to 10 $\mu\text{g}/\mu\text{l}$ final concentration in UTC buffer containing 7 M urea, 2 M thiourea, 4% w/v CHAPS, 50 mM DTT, and 1% pH 3–10 ampholytes (Bio-Rad). Cysteine residues were alkylated with 55 mM iodoacetamide prior to running the IEF gel.

IEF gels (1.5 mm) contained 5% acrylamide, 0.13% bis-acrylamide, 8 M urea, 2% Triton X-100 (v/v), 6.66% pH 3–10 ampholytes (Bio-Rad), and 3.33% pH 5–8 ampholytes (Bio-Rad). Gels were overlaid with 8 M urea, and pre-focused for 30 min at 500 V using degassed 20 mM NaOH as the upper buffer and 10 mM H₃PO₄ as the lower buffer. Samples containing 100 μg of protein were loaded, and focusing was carried out at 750 V for 3 h. IEF gels were extruded and equilibrated in buffer containing 40 mM Tris-HCl, pH 6.8, 3% SDS w/v SDS, 10% glycerol, 2 mM DTT, and 0.0001% v/v bromophenol. SDS-PAGE was run according to Laemmli (1970) using gels consisting of a 4% stacking gel and a 10% separating gel. Gels were fixed in 10% v/v methanol and 7% v/v acetic acid prior to staining with Sypro Ruby (Molecular Probes) according to the manufacturer's instructions. For quantification, Sypro Ruby-stained gels were scanned on a Molecular Dynamics Storm 860 and the data were analyzed using ImageQuant 4.1 software (Molecular Dynamics). The gels were photographed with Polaroid 667 film and the Sypro S-6656 filter (Molecular Probes).

Samples to be analyzed by 1D gel electrophoresis were diluted from 10 $\mu\text{g}/\mu\text{l}$ in UTC buffer to 5 $\mu\text{g}/\mu\text{l}$ in 2 $\times$ SDS-PAGE sample buffer, and run according to Laemmli (1970) with gels consisting of a 4% stacking gel and a 10% separating gel. Gels were either stained with 0.1% Coomassie brilliant blue R-250 or transferred onto nitrocellulose according to Towbin *et al.* (1979).

Immunoblots

Monoclonal antibodies required membranes blocked with 0.25% w/v gelatin, 0.1% w/v ovalbumin in citrate-buffered saline (CBS) containing 25 mM sodium citrate pH 5.5, 375 mM NaCl, and 0.05% v/v Tween 20. Monoclonal BiP antibodies and actin antibodies were diluted 1:5000 in the same blocking buffer. Primary monoclonal antibodies were detected using Protein G-HRP conjugated secondary diluted 1:3000 in gelatin/ovalbumin blocking buffer. Membranes incubated with the other antibodies were blocked with 5% Carnation non-fat milk (NFM) in Tris-buffered saline containing 20 mM Tris-HCl pH 7.4, 375 mM NaCl, and 0.2% v/v Tween 20 (TBST375). Affinity purified SBP-CDA and anti-SBP-CTP antibodies were diluted to 0.1 $\mu\text{g}/\text{ml}$ in 1% NFM/TBST375. Serum containing antibodies to glycinin was diluted 1:20 000 in 1% NFM/TBST375.

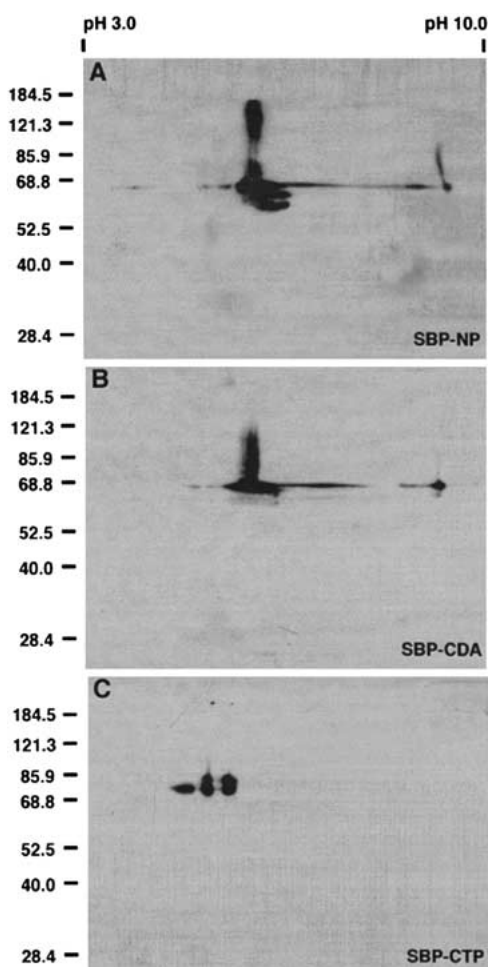


Figure 1. Immunoblot analysis of total soybean seed protein separated by 2D gel electrophoresis. A. Anti-SBP-NP antibody diluted 1:30 000 labels a cluster of 12–15 proteins. B. Anti-SBP-CDA antibody diluted at 1 μ g of antibody in 10 ml of blocking buffer labels a group of 4 proteins. C. Anti-SBP-CTP antibody diluted at 1 μ g of antibody in 10 ml of blocking buffer identifies the α/α' subunits of β -conglycinin. All three antibodies were incubated for 1 h. The blots were incubated in goat anti-rabbit HRP-conjugate for 1 h and developed using SuperSignal West Pico ECL and Hyperfilm from Amersham Pharmacia Biotech.

Secondary goat anti-rabbit HRP-conjugated antibodies were diluted 1:4000 in 1% NFM/TBST375. Blots were incubated with either SuperSignal West Pico or SuperSignal Dura (Pierce Chemical) and exposed to Hyperfilm MP (Amersham Pharmacia).

Silver enhanced immunogold labeling

Cotyledon tissue from developing soybean seed was cut into 1 mm slices while submerged in 50 mM potassium phosphate pH 7.4. The tissue was fixed

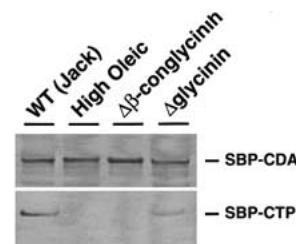


Figure 2. Expression of GmSBP and β -conglycinin α/α' subunits in wild type (WT) soybean, high oleic acid soybean (High Ole), β -conglycinin α/α' null soybean ($\Delta\beta$ -conglycinin), and glycinin null soybean (Δ glycinin). The anti-SBP-CDA (0.1 μ g/ml) antibody demonstrated unaltered expression of GmSBP, while the anti-SBP-CTP (0.1 μ g/ml) antibody identified the α/α' subunits of β -conglycinin.

in 1% v/v paraformaldehyde and 2% v/v glutaraldehyde. Thick sections (700 nm) were heat fixed to gelatin-coated slides. Sections were washed once in TBS containing 500 mM NaCl and 0.2% Tween 20 (TBST500). Sections were then incubated in blocking buffer (2% BSA and 0.1% PVP in TBST500) for 1 h. Primary antibody (SBP-CDA) was diluted to 1 μ g/150 μ l in TBST500 plus 1% BSA, and incubated on the sections for 1.5 h. Sections were washed twice in TBST500 plus 1% BSA. Goat anti-rabbit secondary (15 nm gold particle) was diluted 1:50 in TBST500 plus 1% BSA and incubated for 1 h. The sections were washed twice in TBST500, and twice in dsH_2O . Silver enhancement was carried out with the IntenSE M Silver Enhancement Kit (Amersham Life Science). The tissue was post-stained with 1% (w/v) aqueous safranin. Pictures were taken on Kodak ASA200 slide film with a 35 mm camera and a POL cube (Leitz, Wetzlar, Germany).

Electron microscopy

Cotyledon tissue from developing soybean seed was cut into 1 mm slices while submerged in 50 mM cacodylate pH 7.4. The tissue was fixed following the procedures of Hoh *et al.* (1995), except the $K_4Fe[CN]_6 \cdot 3H_2O$ was omitted from the osmium fixative. Light gold thin sections were picked up on Formvar-coated 200-mesh nickel grids. Sections were incubated in 0.1% $NaBH_4$ for 15 min, washed twice in dsH_2O , and once in TBST500. The tissue was incubated in blocking buffer (2% BSA and 0.1% PVP in TBST500) for 1 h and primary antibodies (SBP-CDA at 1 μ g/150 μ l; anti-SBP-CTP at 1 μ g/100 μ l; glycinin at 1:200 in TBST500 plus 1% BSA) for 1.5 h. Sections were washed twice in TBST500 plus 1%

BSA. Goat anti-rabbit secondary (15 nm gold particle) was diluted 1:50 in TBST500 plus 1% BSA and incubated for 1 h. Sections were washed twice in TBST500 plus 1% BSA, twice in TBST500, and twice in dsH_2O . Prior to drying, the grids were gently dipped in dsH_2O 30 times. Sections used for immunogold labeling were stained with 4% uranyl acetate. Sections not used for immunogold labeling were stained with 3% uranyl acetate and 0.25% KMnO_4 .

¹⁴C-sucrose uptake

The uptake of sucrose into yeast expressing GmSBP1, GmSBP2, or GmSUT1 was measured according to the procedure set forth in Overvoorde *et al.* (1996). Time points were taken at 0, 1, 2, 3, and 5 min. All measurements were performed in triplicate and used to calculate standard deviations.

Protein storage vacuole isolation

Mature soybean seed (5 g) were ground to a fine powder in liquid N_2 . The powder was transferred to 20 ml of glycerol, and ground until thoroughly mixed. Glycerol slurry was filtered through four layers of cheesecloth. Protein storage vacuoles, or protein bodies, were isolated as described in Hara-Nishimura *et al.* (1982).

Results

Antibody development and characterization

The disparity between GmSBP-mediated sucrose transport across the plasma membrane and the sequence homology to seed storage proteins required a critical re-evaluation of the biological function of GmSBP. Two additional antibodies were developed to confirm the expression and immunolocalization data collected using the anti-SBP-NP antibody. The C-terminal domain antibody (anti-SBP-CDA) included the C-terminal 165 amino acids of GmSBP1 and the C-terminal peptide antibody (anti-SBP-CTP) was a synthetic peptide, N'-CELPREERGRRADA-C, at the extreme C-terminus of GmSBP1. These new GmSBP antibodies as well as the other antibodies used in this paper are listed in Table 1. The specificity of the various GmSBP antibodies was compared using 2D gel electrophoresis. Immunoblot analysis with the anti-SBP-NP antibody identified 12–15 protein spots (Figure 1A). The number of protein spots identified by this

antibody indicated that it was not a specific probe for GmSBP. Immunoblot analysis with affinity-purified anti-SBP-CDA antibodies identified four protein spots near the predicted molecular mass for GmSBP1 and with isoelectric points ranging from 5.6 to 6.1. (Figure 1B). The four proteins identified with the anti-SBP-CDA antibody may be explained by a GmSBP gene family or by modifications (post-translational or carbamylation) affecting the pI of GmSBP1. The anti-SBP-CTP antibody should have produced the most specific GmSBP labeling. However, the five protein spots identified with the anti-SBP-CTP antibody had molecular masses and isoelectric points matching the α/α' subunits of β -conglycinin (Figure 1C). The cross reactivity with the α/α' subunits of β -conglycinin was unexpected, especially since the anti-SBP-CTP peptide sequence cannot be identified within the α/α' sequences. However, we believe this antibody recognizes a structural epitope.

Soybean lines lacking the expression of glycinin or β -conglycinin were used to confirm the anti-SBP-CDA antibody and anti-SBP-CTP antibody specificity. Protein labeled by the anti-SBP-CDA antibody did not change among the soybean lines tested, indicating that there was no background labeling due to antibody cross-reactivity with any of the glycinins or β -conglycinins (Figure 2). Examination of these soybean lines with the anti-SBP-CTP antibody identified protein in the wild-type (Jack) and glycinin null soybean lines, but not in the β -conglycinin null or high-oleic acid soybean lines (Figure 2). This confirmed that this antibody does label the α/α' subunits of β -conglycinin.

Visualization of protein storage vacuoles during seed filling

Questions raised about the specificity of the anti-SBP-NP antibody compelled our lab to re-evaluate the immunolocalization of GmSBP with the new anti-SBP-CDA antibody: this demonstrated a higher specificity in 2D gel immunoblot analysis. During seed development, filling, and maturation, cells within the storage parenchyma undergo a series of morphological changes that correlate with the onset of storage protein expression (Hara-Nishimura *et al.*, 1993; Hoh *et al.*, 1995; Kinney *et al.*, 2001). In young developing cotyledons, 4–5 mm, the cells within the storage parenchyma are highly vacuolated (Figure 3A) and have very light labeling as indicated by the silver-enhanced immunogold labeling (data not shown). Around 6 mm,

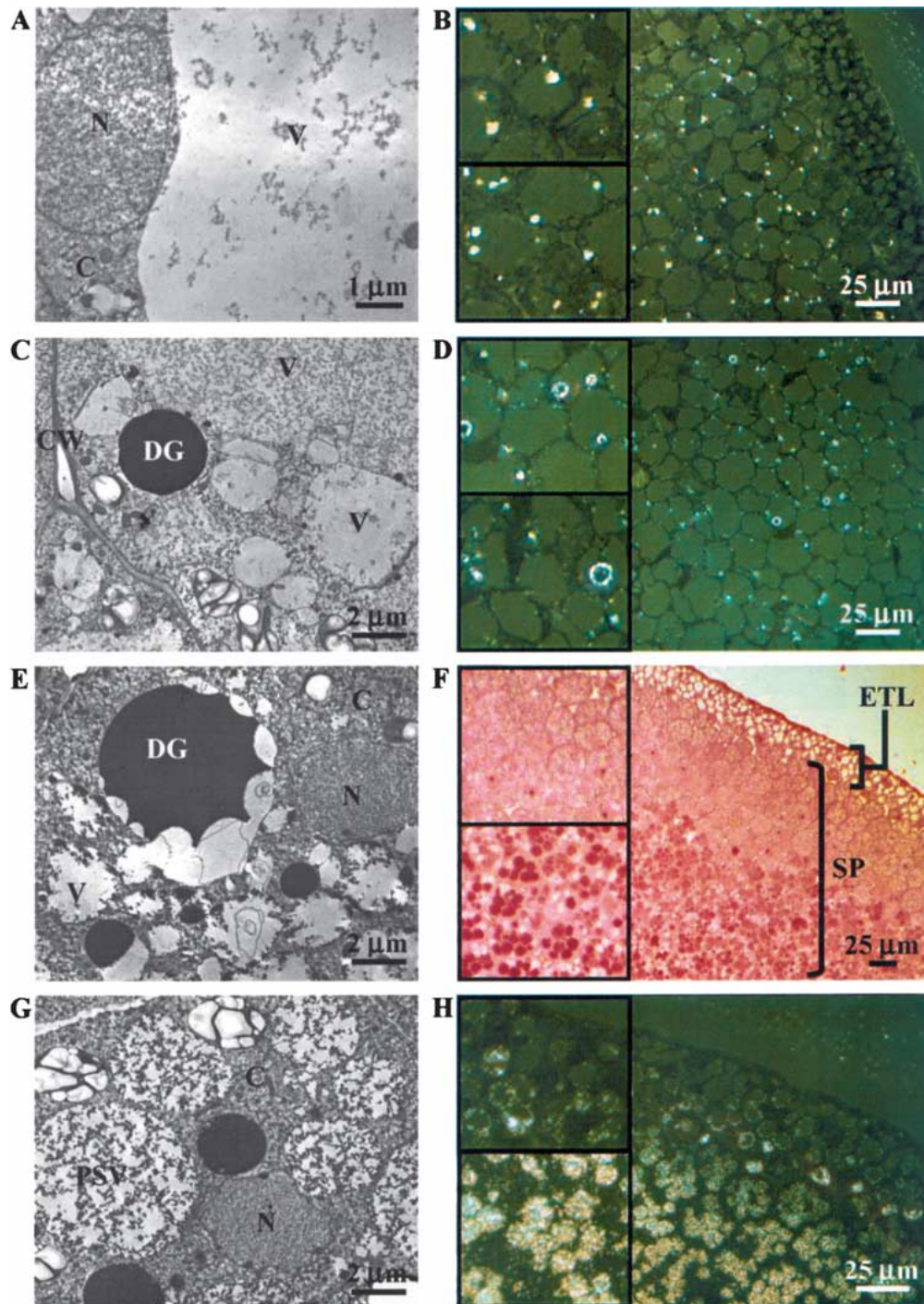


Figure 3. Developmental changes in vacuolar morphology and GmSBP expression in soybean cotyledon storage parenchyma. A. Low magnification (6000 $\times$) transmission electron microscopy of a storage parenchyma cell from 5 mm cotyledon showing the large central vacuole and oppressed cytoplasm. B. Epipolarization of silver-enhanced immunogold labeling of storage parenchyma from a 6 mm cotyledon with increased labeling by the anti-SBP-CDA antibody. C. Low magnification (4000 $\times$) transmission electron microscopy of a storage parenchyma cell from a 7 mm cotyledon with multiple vacuolar compartments and dense globules. D. Epipolarization of silver enhanced immunogold labeling of storage parenchyma from an 8 mm cotyledon with anti-SBP-CDA antibody. Ring shaped labeling is visible on the large dense globules. E. Low magnification (4000 $\times$) transmission electron microscopy of a storage parenchyma cell from a 7.9 mm cotyledon with multiple vacuolar compartments (V) and large dense globules (DG). F. Safranin stained 10 mm soybean cotyledon storage parenchyma (SP) and adaxial epidermal transfer layer (ETL). G. Low magnification (4000 $\times$) transmission electron microscopy of a storage parenchyma cell from a 9 mm cotyledon with mature protein storage vacuoles (PSV). H. Epipolarization of silver enhanced immunogold labeling of storage parenchyma from a 10 mm cotyledon with anti-SBP-CDA antibody. The progression of seed filling toward the adaxial epidermis is clearly visible in the PSV labeled with the anti-SBP-CDA antibody. N, nucleus; C, cytoplasm; V, vacuole; CW, cell wall; DG, dense globule; PSV, protein storage vacuole.

dense globules begin to appear as punctate spots in silver-enhanced immunogold labeling (Figure 3B). As the cotyledons enter seed filling, around 7 mm, the dense globules are present in both the cytoplasm and the large central vacuole (Figure 3C). As the seed develops, there is a transition from a single vacuole to multiple protein storage vacuoles. The protein storage vacuoles appear via by *de novo* synthesis (Hoh *et al.*, 1995). At 8.0 mm, large dense globules appear as punctate structures or ring shaped structures with silver-enhanced immunogold labeling (Figure 3D). These large dense globules are present within the cytoplasm and the vacuolar compartments (Figure 3E). Seed filling progresses toward the adaxial epidermal transfer layer (ETL) shown in Figure 3F. During the late stages of seed filling, 11 mm, the dense globules within the protein storage vacuoles (PSV) are processed into their mature forms (Figure 3G). Labeling of the mature protein storage vacuoles by the anti-SBP-CDA antibody is readily apparent late in seed development (Figure 3H).

The localization of GmSBP to the mature protein storage vacuole was also investigated using enriched protein storage vacuoles isolated according to Hara-Nishimura *et al.* (1982). Immunoblot analysis was used to identify glycinin, β -conglycinin, and GmSBP within the enriched protein storage vacuole (Figure 4). This supports the data obtained by immunolocalization, and demonstrates that GmSBP is ultimately transported to and stored within the mature protein storage vacuole.

In addition to the localization to the mature protein storage vacuoles demonstrated by silver-enhanced immunogold labeling, immunogold labeling with β -conglycinin (anti-SBP-CTP), glycinin, and GmSBP (anti-SBP-CDA) antibodies identified large dense globules present in both the cytoplasm and the vacuole (Figure 5B, 5C, and 5D). While these large dense globules in the cytoplasm probably represent ER-derived precursor accumulating vesicles (Hara-Nishimura *et al.*, 1998), they might also represent small protein bodies formed by budding off the protein storage vacuoles, larger equivalents of the pea dense vesicles (Hinz *et al.*, 1999), or a cross-section of a lobed protein storage vacuole after sectioning. Unlike immunolocalization with the anti-SBP-NP antibodies (Warmbrodt *et al.*, 1989; Warmbrodt *et al.*, 1991; Grimes *et al.*, 1992; Tegeder *et al.*, 1999), the anti-SBP-CDA antibody does not label protein at the plasma membrane. In addition, immunolocalization with the anti-SBP-NP antibodies included labeling of

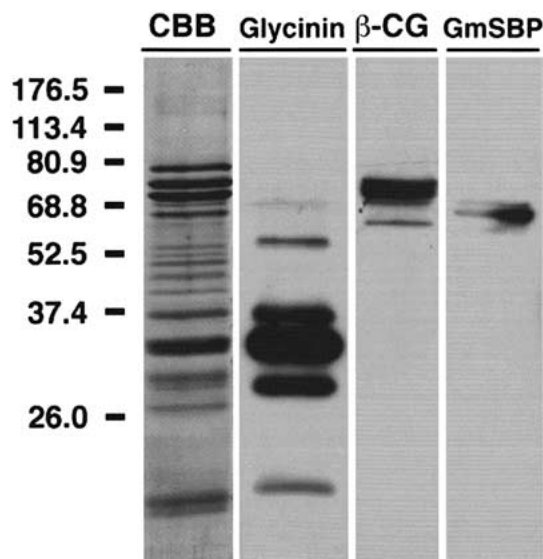


Figure 4. Colocalization of GmSBP, glycinin, and β -conglycinin in the protein storage vacuoles (PSV) from mature soybean seed. The PSVs were isolated according to Hara-Nishimura *et al.* (1982). Immunoblots were developed using affinity purified antibodies to GmSBP (anti-SBP-CDA at 0.1 μ g/ml), glycinin (1:20 000), β -conglycinin (anti-SBP-CTP at 1 μ g/ml). The glycinin antibody indicated mature as well as processed forms of the protein were present.

the protein storage vacuoles as well as the plasma membrane (data not shown).

Identification and characterization of a GmSBP gene family

Southern blot analysis of *Eco*RI digested DNA was used to determine if *GmSBP* was a unique protein or if it belonged to a gene family. The appearance of bands at 8.0 kb, 6.0 kb, and 1.7 kb after hybridization with a ^{32} P-labeled *GmSBP1* probe indicated two or more *GmSBP* genes present in soybean (data not shown). The second soybean sucrose-binding protein, *GmSBP2*, was cloned from a cDNA library using primers designed to the *GmSBP1* sequence (See Materials and Methods). An additional soybean sucrose-binding protein, *GmSBP64*, was cloned by Pedra *et al.* (2000) using an antibody raised to a soybean seed crude membrane fraction. Based on the southern blot data and additional proteins cloned from soybean, we concluded that there is a small *GmSBP* gene family in soybean with at least three members.

Sequence analysis indicated that GmSBP1 and GmSBP2 are 91% identical in amino acid composition, while GmSBP2 and GmSBP64 are 99% identical

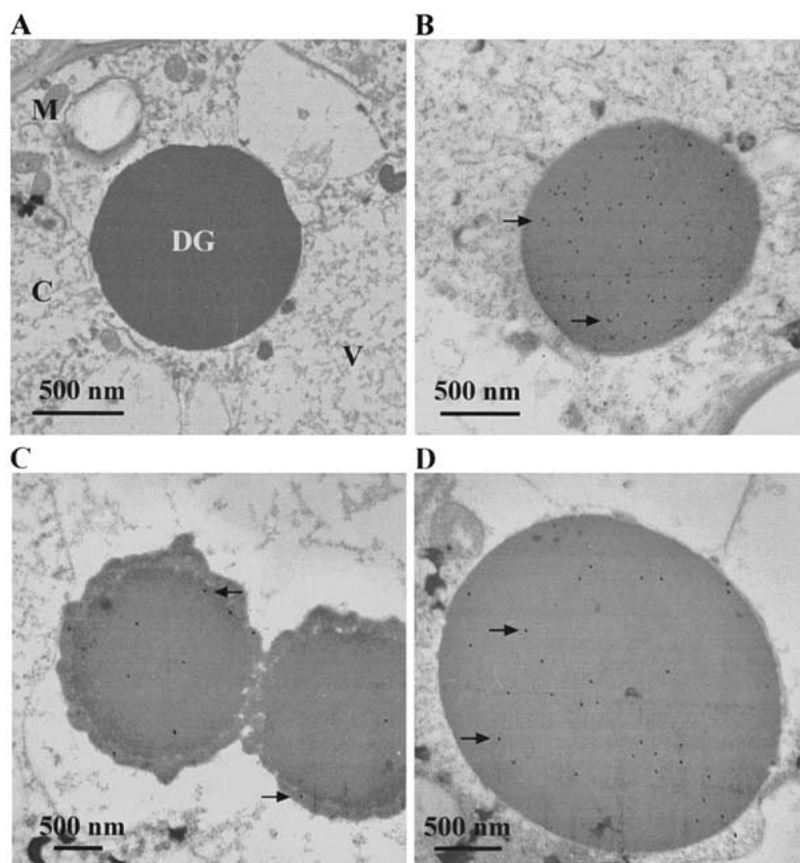


Figure 5. Immunogold labeling (10 nm gold particle) of dense globules from storage parenchyma cells in 9 mm soybean cotyledons. A. Tissue was incubated with goat anti-rabbit secondary antibody (10 nm gold particle). B. Tissue labeled with glycinin antibody (1:200). C. Tissue labeled with anti-SBP-CDA antibody (1 μ g/150 μ l). D. Tissue labeled with anti-SBP-CTP antibody (1 μ g/150 μ l). All three antibodies identify large dense globules within the cytoplasm or vacuole.

in amino acid composition. The predicted molecular mass and isoelectric point (pI) of GmSBP1 (60 465 Da and 6.42) differ slightly from GmSBP2 (55 775 Da and 6.12) and GmSBP64 (55 834 Da 6.32). Differences between GmSBP2 and GmSBP64 are limited to four amino acid substitutions that may be accounted for by allelic or varietal differences. A comparison between GmSBP1 and GmSBP2 includes both amino acid substitutions and gaps in the GmSBP2 sequence (Figure 6).

Structural modeling has been used to establish GmSBP as a vicilin-like protein that oligomerizes into trimeric molecules (Ko *et al.*, 1993; Lawrence *et al.*, 1994; Maruyama *et al.*, 2001). Combining the crystallographic studies with multiple sequence alignments of vicilin-like proteins identified 28 strictly conserved amino acids (Figure 6; solid circles; Lawrence *et al.*, 1994) and 26 structurally significant amino acids (Fig-

ure 6; open circles; Ko *et al.*, 1993). Six of these amino acids were common to both the strictly conserved and structurally conserved amino acids (Figure 6; asterisks). Multiple sequence alignment of phaseolin, canavalin, GmSBP1, GmSBP2, and GmSBP64 was used to compare the 48 strictly or structurally conserved amino acids. This analysis identified 35 conserved amino acids and 13 non-conserved amino acids within the GmSBP2 and GmSBP64 sequences (Figure 6). The only difference in non-conserved structural residues among GmSBPs occurred at Q289 of GmSBP1, resulting in 36 conserved amino acids and 12 non-conserved amino acids. Within the 13 non-conserved residues, there are two key amino acid changes, P157A and P361I (Overvoorde *et al.*, 1997). In canavalin, phaseolin, and β -conglycinin β -subunit, these proline residues aid in forming β -bulges that participate in protein-protein interactions between sub-

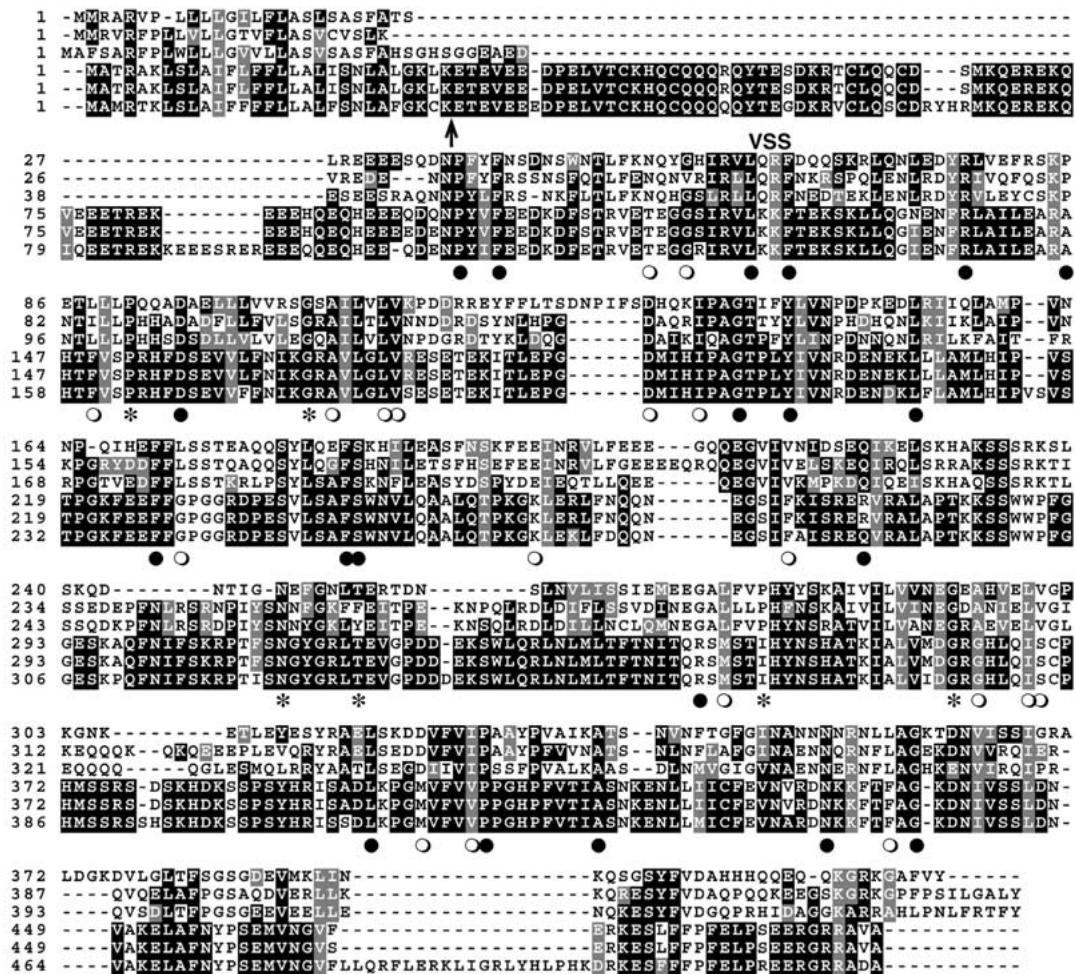


Figure 6. CLUSTAL W multiple sequence alignment with GmSBP1, GmSBP2, GmSBP64, canavalin, soybean β -conglycinin β -subunit (Gm β CG β), and phaseolin. Identical residues are shaded black with white text. Similar residues are shaded gray with white text. The vertical arrow represents the position of the cleaved N-terminal signal sequence. Open circles represent structurally conserved residues, closed circles represent strictly conserved residues, and asterisks represent residues that were common to both the strictly conserved residues and the structurally conserved residues.

units within a trimeric oligomer (Ko *et al.*, 1993; Lawrence *et al.*, 1994; Maruyama *et al.*, 2001). The structural impact of these amino acid substitutions is uncertain, but previous work from our lab established that GmSBP assembles into a trimeric structure *in planta* (Overvoorde *et al.*, 1997) or in a heterologous yeast expression system (Grimes and Overvoorde, 1996).

Both GmSBP1 and GmSBP2 were used to search the non-redundant database (SDSC) by BLASTP (Molecular Workbench). An unrooted dendrogram of 38 proteins from BLASTP shows the inferred relationship among vicilin, globulin, and GmSBP-like proteins (Figure 7). Previous sequence analyses linked

GmSBP1 with vicilin-like storage proteins from gymnosperms and cycads (Newton *et al.*, 1992; Bäumlein *et al.*, 1995; Braun *et al.*, 1996). Within this dendrogram, GmSBP-like proteins cluster with herbaceous and monocot globulins (Figure 7). The other legume seed storage proteins (β -conglycinins, canavalins, convicilins, and vicilins) form a separate branch within the dendrogram. This established a clear divergence between legume seed storage proteins and GmSBP as well as suggesting a possible evolutionary link to monocot seed storage proteins. In this regard, it is noteworthy that the C-terminus of GmSBP2 contains the tripeptide sequence AVA, which might serve as the putative vacuolar sorting sequence (VSS) sim-

ilar to the C-terminal, hydrophobic, VSS documented by Frigerio *et al.* (1998). In contrast, the C-terminus of GmSBP1 contains the tripeptide sequence ADA, which has a charged residue (D) in place of a hydrophobic residue (V). The insertion of a charged residue would likely result in a non-functional VSS, and may also indicate a different sorting mechanism for GmSBP1.

Previous work in our lab has established that GmSBP1 transports sucrose in a heterologous yeast expression system (Overvoorde *et al.*, 1996). The unexpected localization of GmSPB to the protein storage vacuole necessitated the re-evaluation of the yeast sucrose uptake analysis. A comparison of ^{14}C -sucrose uptake revealed the rate of uptake by yeast expressing either GmSBP1 or GmSBP2 measured 146.10 pmol per mg cells after 1 min and 151.99 pmol per mg cells after 3 min, respectively. In contrast, ^{14}C -sucrose uptake by GmSUT1, a H^{+} -dependent sucrose transporter from soybean cotyledon, measured 2447.02 pmol per mg cells after 3 min. The rate of sucrose transport mediated by GmSBP1 and GmSBP2 was 16-fold less than the rate of sucrose transport mediated by GmSUT1. While consistent and reproducible, the uptake of sucrose by GmSBP1 and GmSBP2 in the yeast expression system was measurable, but may not be biologically significant.

The 5'-flanking sequences of *GmSBP1* and *GmSBP2* were cloned from a Genome Walker 'library' using a *GmSBP1* gene-specific primer. Analysis of the sequences obtained from the library identified one sequence that contained nucleotides overlapping the *GmSBP1* open reading frame and a second sequence belonging to *GmSBP2*. Further cloning, using either *GmSBP1* or *GmSBP2* gene-specific primers, produced 0.8 kb and 0.9 kb of 5'-flanking sequence for *GmSBP1* and *GmSBP2*, respectively. These 5'-flanking sequences will be referred to as promoter regions. The first 200 nucleotides upstream of the translation start site displayed more conservation than the sequences further upstream. This suggests that *GmSBP1* and *GmSBP2* may be physically linked and may have originated from a single ancestral gene (Lee *et al.*, 1986; Stockhaus *et al.*, 1994).

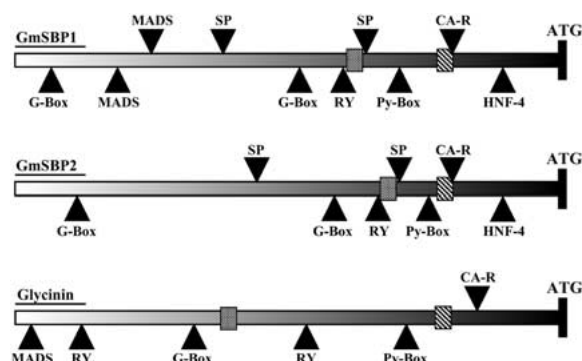


Figure 8. Diagrammatic representation of the 5'-flanking regions, or promoter regions, from *GmSBP1*, *GmSBP2*, and glycinin (*Gy3*). The position of the TATA boxes are indicated by a hatched rectangle, the position of the CCAAT box by a checked rectangle, and the position of *cis*-acting regulatory elements by a black triangle. Each of the *cis*-acting regulatory elements is proportionally positioned along a 5'-flanking sequence of 325 nucleotides.

Sequence analysis of the promoter regions showed that *GmSBP1* and *GmSBP2* have a unique organization of transcriptional consensus sequences and putative *cis*-acting regulatory elements (*cis*-RE; Figure 8). The TATA box was located -72 nt upstream of the translation start site for *GmSBP1* and -65 nt upstream of the translation start site for *GmSBP2*, compared to -79 for glycinin (Figure 8). The CAATT site was located -126 nt upstream of the translation start site for *GmSBP1* and -106 upstream of the translation start site for *GmSBP2*, compared with glycinin which has a CAATT box in the reverse orientation at -193 nt (Figure 8). Putative *cis*-RE present in the *GmSBP* promoter regions include CA-rich region (Kawagoe and Murai, 1992; Vellanoweth and Okita, 1993; Ezcurra *et al.*, 1999), pyrimidine-rich sequence (Daraselia *et al.*, 1996; Kitamoto *et al.*, 1998), MADS-like sequences (West and Sharrocks, 1997; Tilly *et al.*, 1998), seed specific sequences (Itoh *et al.*, 1993; Zhou and Fan, 1993; Shasany and Koundal, 2000), G-boxes (Loake *et al.*, 1992; Fiedler *et al.*, 1993; Kawagoe *et al.*, 1994), and a RY sequence (Bäumlein *et al.*, 1992; Fujiwara and Beachy, 1994; Bobb *et al.*, 1997; Sakata *et al.*, 1997; Figure 8). The G-boxes and RY-sequences present in the glycinin promoter were clearly different than those present within the *GmSBP* promoter region with respect to both the order and spacing. The biological activity of the seed specific *cis*-RE is supported by the seed specific expression of β -glucuronidase driven by *VfSBPL*, a *GmSBP* homologue, promoter sequence (Heim *et al.*, 2001). Analysis the *VfSBPL* promoter also identified putative sugar responsive ele-

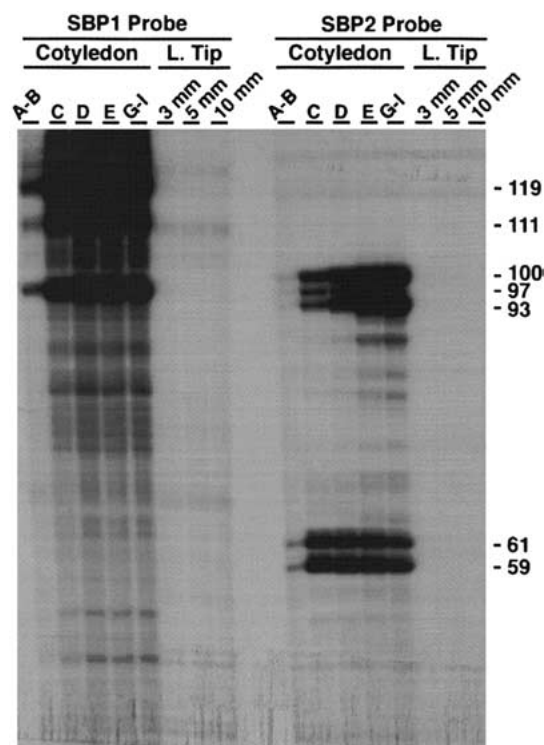


Figure 9. RNase protection assay using ^{32}P -labeled gene specific probes to characterize the seed specific expression, transcriptional control of mRNA expression, and the length of the 5'-untranslated region for each gene. *GmSBP1* and *GmSBP2* demonstrate differential expression patterns during seed development. Cotyledon samples are categorized by developmental stage: stage AB <4.0 mm, stage C = 5.0 mm, stage D = 7.0 mm, stage E = 8.0 – 9.0 mm, stage GI = 1.1 – 1.2 mm, stage S = initiation of desiccation. Developmental leaf series includes 3 mm, 5 mm, and 10 mm leaves. The size of the fragments was measures in nucleotides.

ments (SURE; Heim *et al.*, 2001). Treating excised soybean cotyledons with 100 mM or 200 mM sucrose produced no change in *GmSBP1* or *GmSBP2* mRNA (data not shown). Thus, any remnants of SURE *cis*-RE present within the *GmSBP1* and *GmSBP2* promoters are inactive in excised soybean cotyledons.

RNase protection assay (RPA) was used to characterize the 5'-untranslated region (5'-UTR) and examine the expression of *GmSBP1* and *GmSBP2* transcripts (Figure 9).

Probes were constructed from either sense strand (antisense probe) or antisense strand (sense probe) using short genomic clones containing the 5'-flanking region of *GmSBP1* or *GmSBP2*. Sense probes did not produce any protected fragments. The *GmSBP1* antisense probe protected three major fragments (119 , 111 , and 97 nt), confirming that three different transcription

start sites were used (Figure 9). Although the antisense probe for *GmSBP2* protected five fragments, only the 100, 97, and 93 nt fragments are believed to represent transcriptional start sites for the full-length *GmSBP2* protein (Figure 9). From these data, it was possible to determine the longest 5'-UTR was 47 nt in *GmSBP1* and 28 nt in *GmSBP2* (Figure 9). The 61 and 59 nt fragments from *GmSBP2* would place the transcriptional start sites downstream of the translational start site. Three possibilities could explain these fragments: there is a third less conserved gene, the internal transcript start site led to the production of a truncated *GmSBP2* gene, or there is an allelic mutation.

Gene specific expression of *GmSBP1* and *GmSBP2* mRNA was also examined by a RNase protection assay (RPA). A developmental cotyledon series was hybridized with a probe designed to the antisense strand indicated that *GmSBP1* mRNA peaked at stage D (7 mm) and was maintained through stage G-I (11–12 mm; Figure 9). A developmental cotyledon series was hybridized with antisense *GmSBP2* probe indicated that *GmSBP2* mRNA levels increase gradually through development and peaked at stage G-I (11–12 mm; Figure 9). Seed specific expression indicated by putative *cis*-RE present in the promoter region (Figure 7) was supported by the lack of *GmSBP1* or *GmSBP2* transcript in developing leaf tissue (Figure 9).

Expression of GmSBP1 and GmSBP2 during seed development and germination

Seed specific expression and mRNA expression patterns, demonstrated in Figure 9, prompted us to examine the accumulation of GmSBP protein during seed development and germination. Protein extracted from a finely graded series of developing cotyledons was probed with anti-GmSBP (anti-SBP-CDA), anti- β -conglycinin (anti-SBP-CTP), anti-actin (Clone C4), and anti-glycinin antibodies (Figure 10). GmSBP first appeared in 5.5 mm cotyledons, increased throughout seed filling, and then decreased as the seed approached maturity (Figure 10). The glycinin antibody identified bands at 50 and 31 kDa (Figure 10). The 31 kDa acidic subunit first appears in 6.0 mm cotyledons, then increases throughout seed filling and maturation. The α/α' -subunits of β -conglycinin first appear around 6.0–7.0 mm, they remain expressed at high levels through 12.0 mm, and then decrease as the seed matures (Figure 10). The apparent decrease in GmSBP and β -conglycinin protein late in seed development

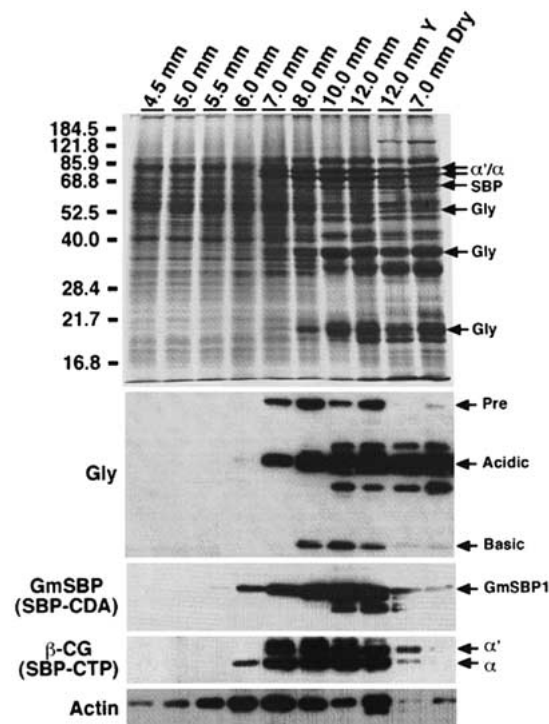


Figure 10. Expression of GmSBP protein in developing soybean seed was compared to glycinin, β -conglycinin, and actin. Total protein extracted from developing soybean seed was separated on three SDS-PAGE gels, and blotted to nitrocellulose. Immunolabeling of GmSBP (anti-SBP-CDA antibody diluted at 1 μ g of antibody per 10 mm of 1% NFM blocking buffer), glycinin (1:30 000 in 1% NFM blocking buffer), and β -conglycinin (anti-SBP-CTP antibody diluted at 1 μ g of antibody per 10 ml of 1% NFM blocking buffer) shows expression started around 6–7 mm and increased during seed maturation. The apparent decrease in β -conglycinin and GmSBP was due to inefficient transfer commonly encountered with 7S seed storage proteins.

may be attributed to difficulties in transfer efficiency. The actin antibody was used as a control for relative protein levels, and showed an increase in actin through 7.0 mm. After 7.0 mm, coinciding with the onset of storage protein accumulation, the actin levels vary and almost disappear in the mature seed (Figure 10).

Although GmSBP appears to decrease as the seed approaches maturity, we were able to detect GmSBP protein in soybean cotyledons after imbibition. GmSBP protein levels decreased gradually after imbibition between 24 to 72 h, with ca. 26 kDa proteolytic fragments appearing at 72 h (Figure 11). The decrease in GmSBP protein is more dramatic after 96 h (Figure 11). This rate of degradation appears similar to β -conglycinin (Bryant *et al.*, 1995). The gradual decrease in GmSBP after imbibition may represent the loss of protein due to the absence of mRNA ex-

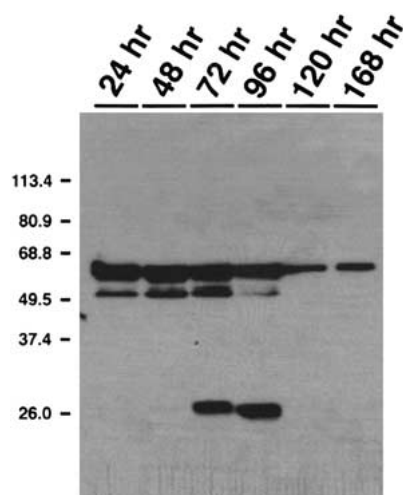


Figure 11. Changes in GmSBP during germination. Total protein was extracted from cotyledonary tissue at 24 h intervals after imbibition and separated on a 10% gel by SDS-PAGE. Protein was detected using affinity purified anti-SBP-CDA antibody (1 μ g/10ml). A gradual decrease in protein was observed, with 26 kDa fragments appearing at 72 h after imbibition.

pression during imbibition previously established by Overvoorde *et al.* (1997).

Semi-quantitative comparison between GmSBP, Glycinin, and β -conglycinin

The accumulating evidence based on sequence similarity, immunolocalization, seed-specific expression, and protein expression patterns during seed development, and protein degradation during germination indicated that GmSBP shares many characteristics with seed storage proteins. The critical question that remained was the quantitative relationship between GmSBP, glycinin, and β -conglycinin. In essence, does GmSBP accumulate to levels high enough for it to play a role in nitrogen storage? Sypro Ruby-stained gels (Steinberg *et al.*, 2000), scanned with a Molecular Dynamics Storm 860, were used to determine the relative contribution of GmSBP1 and GmSBP2 to total seed protein. β -conglycinin α/α' subunits, corresponding to spots 1–5, make-up 23% of the total seed protein (Figure 12). Glycinins, the most abundant legume seed proteins, are comprised of both acidic and basic subunits. The acidic subunits, corresponding to spots 6–11, represent 22.4% of the total seed protein. The four spots identified by the anti-SBP-CDA antibody, spots 12–15, accumulate to 2.4% of the total seed protein (Figure 12). In comparison to β -conglycinin α/α' subunits and glycinin, the GmSBPs would contribute

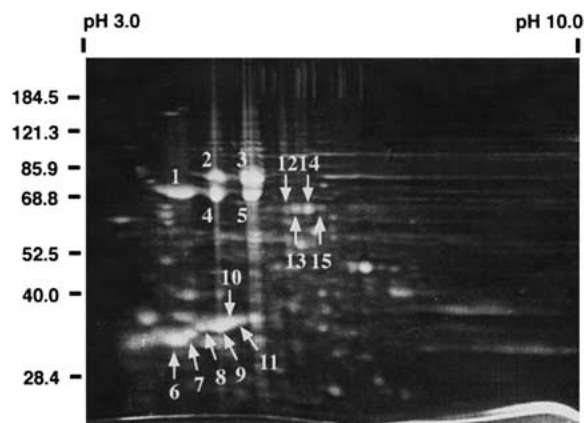


Figure 12. Semi-quantitative analysis of GmSBP, glycinin, and β -conglycinin in 10 mm developing soybean seed. Total seed protein (100 μ g) was focused on 1.5 mm diameter tube gels for 3 h. The second dimension separated protein on a 10% gel using SDS-PAGE. The gel was stained with Sypro Ruby (Molecular Probes), scanned with a Molecular Dynamics Storm 860, and photographed with on Polaroid 667 film. β -conglycinin (spots 1-5), glycinin acidic subunits (spots 6-11), and GmSBP (spots 12-15) were identified by immunoblot prior to semi-quantitative analysis.

19-fold less to the stored nitrogen than glycinin and β -conglycinin combined.

Discussion

This paper presents a detailed characterization of the cell biology and storage protein-like characteristics presented by GmSBP. Previous studies, in our lab and others, have implicated GmSBP and GmSBP-like proteins as sucrose transporters or markers for sucrose transport (Ripp *et al.*, 1988; Warmbrodt *et al.*, 1989, 1991; Grimes *et al.*, 1992; Overvoorde and Grimes, 1994; Overvoorde *et al.*, 1996; Tegeder *et al.*, 1999; Pedra *et al.*, 2000). However, contrasting with the biochemical and immunological data supporting the sucrose transporter-like characteristic is the sequence homology and putative structural homology to globulin-like and vicilin-like seed storage proteins (Bäumlein *et al.*, 1995; Shutov *et al.*, 1995; Braun *et al.*, 1996; Overvoorde *et al.*, 1997).

One major finding presented here with new antibodies (the anti-SBP-CDA antibody) prepared against bacterially expressed and purified GmSBP1 is the localization to protein storage vacuoles. Additional immunolocalization with the anti-SBP-CDA antibody identified GmSBP within large dense globules that are may represent ER-derived precursor accumulating vesicles (Hara-Nishimura *et al.*, 1998), larger equivalent-

ents of the pea dense vesicles (Hinz *et al.*, 1999), or a cross section of a lobed protein storage vacuole. Transmission electron microscopy shows that the soybean large dense globules often exceed 500 nm in diameter and are often located in the cytoplasm, separate from the large central vacuole. These dense globules appear to be processed and eventually form a mature protein storage vacuole.

Light microscopy, and epipolarization, was used to follow the formation of protein storage vacuoles during seed filling using anti-SBP-CDA antibodies. Deposition of GmSBP-like proteins within the protein storage vacuole was also demonstrated for VfSBPL, a GmSBP homologue from *Vicia faba* (Heim *et al.*, 2001). Together, these data suggest that GmSBP and its homologues would be present in protein storage vacuoles of developing soybean cotyledons, consistent with the immunolocalization observed with both VfSBPL and GmSBP.

Since the present study localizes GmSBP into PSVs, the question arises as to why the original polyclonal antibody recognized epitopes present in the plasma membrane and related structures. One explanation may be that the original antibody used in Ripp *et al.* (1988) was developed to partially-purified, ¹²⁵I-6'-HABS-labeled protein from a crude soybean cotyledon membrane preparation that included GmSBP1 and one or more highly antigenic plasma membrane proteins. The plasma membrane labeling may be due to the presence of antibodies that recognize these plasma membrane epitopes with high avidity. Alternatively, sequence comparisons suggest that GmSBP1 shares immunogenic properties with the peanut allergen (AraH1), a vicilin-like protein known for eliciting strong and some times fatal immune responses (Maleki *et al.*, 2000; Rabjohn *et al.*, 2002). The apparent antigenicity of the GmSBP-like proteins was illustrated by the cloning of GmSBP64 using a polyclonal antibody developed to proteins from a crude membrane preparation (Pedra *et al.*, 2000).

If the GmSBPs were localized to the protein storage vacuoles, they would be expected to traverse the secretory system similarly to other legume storage proteins. Indeed, the oligomeric organization of GmSBP (Grime and Overvoorde, 1996; Overvoorde *et al.*, 1996) resembles the trimeric form of mature vicilin-like proteins such as phaseolin, canavalin, and the β -subunit of β -conglycinin. The formation of these trimeric structures requires entry into the ER via an N-terminal signal sequence that is cleaved prior to folding and oligomerization (Coleman *et al.*, 1995;

Jung *et al.*, 1997). Both GmSBP1 and the vicilin-like proteins contain N-terminal signal sequences and N-terminal CXXXC motifs (Kriz, 1989; McHenry and Fritz, 1992; Heck *et al.*, 1993; Braun *et al.*, 1996).

Despite the similarities between GmSBP and seed storage proteins, divergence of GmSBP from the other legume seed storage proteins was evident with a close examination of the 5'-flanking regions of *GmSBP1* and *GmSBP2*. Sequence analysis identified several *cis*-RE related to seed-specific expression and seed storage protein expression. However, the organization and spacing of the *cis*-RE was not consistent with other seed storage protein promoters such as glycinin (Itoh *et al.*, 1993, 1994). The divergence in the promoter region may account for quantitative differences between GmSBP, glycinin, and β -conglycinin. In fact, semi-quantitative analysis of storage protein levels suggested that GmSBP contributed significantly less to nitrogen storage than glycinin and β -conglycinin.

GmSBP1, GmSBP2, and GmSBP64 are members of the cupin superfamily by virtue of their sequence similarity and the structural conservation of the modular Swiss-roll β -barrel domain. The GmSBPs, and VfSBPL, have a pseudo-dyad primary structure in which the N-terminal and C-terminal domains arose from gene duplication (Argos *et al.*, 1985; Lawrence *et al.*, 1994; Shutov *et al.*, 1995). Our sequence analysis of the GmSBPs, with an expanded set of globulin-like proteins, concluded that GmSBP clustered with monocot and herbaceous globulins, rather than cycad and gymnosperm vicilin-like seed storage proteins as reported by Bäumlein *et al.* (1995).

The sequence and putative structural homology of GmSBP to the cupin superfamily raises the possibility of additional biological function. The cupin superfamily contains prokaryotic proteins such as phosphomannose isomerase and polyketide synthase, fungal proteins such as spherulins and various dioxygenases, as well as plant proteins such as dioxygenases, oxalate decarboxylase, seed storage proteins, and GmSBP (Dunwell *et al.*, 2000). The cupin superfamily links GmSBP, and related proteins, with a functionally diverse set of proteins spanning at five of the six kingdoms. The biological function of the SBP-like proteins remains uncertain, but the cell biology and subcellular localization may represent an evolutionary link to the globulin-like ancestor that gave rise to the 7S family of seed storage proteins.

Acknowledgements

We would like to thank Anthony Kinney for his assistance in testing our antibodies on various soybean lines. We would also like to thank Vince Franceschi and Valerie Lynch-Holm for their assistance with the immunolocalization as well as John Rogers, John Browse, and Kathleen Postle for their critical reading of this manuscript. This work was supported in part by NSF grant 0078279.

References

- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25: 3389–3402.
- Argos, P., Narayana, S.V.L. and Nielsen, N.C. 1985. Structural similarity between legumin and vicilin storage proteins from legumes. *EMBO J.* 4: 1111–1117.
- Bäumlein, H., Nagy, I., Villarroel, R., Inze, D. and Wobus, U. 1992. *Cis*-analysis of a seed protein gene promoter: the conservative RY repeat CATGCATG within the legumin box is essential for tissue-specific expression of a legumin gene. *Plant J.* 2: 233–239.
- Bäumlein H., Braun H., Kakhovskaya I.A. and Shutov A.D. 1995. Seed storage proteins of spermatophytes share a common ancestor with desiccation proteins of fungi. *J. Mol. Evol.* 41: 1070–1075.
- Bobb, A.J., Chern, M.S. and Bustos, M.M., 1997. Conserved RY-repeats mediate transactivation of seed-specific promoters by the developmental regulator PvALF. *Nucl. Acids Res.* 25: 641–647.
- Braun, H., Czihal, A., Shutov, A.D. and Bäumlein, H. 1996. A vicilin-like seed protein of cycads: similarity to sucrose-binding proteins. *Plant Mol. Biol.* 31: 35–44.
- Bryant, J.L., Hosangadi, P. and Beachy, R.N. 1995. The β -conglycinins of soybean remain assembled in a 7S conformation while undergoing proteolytic cleavage during germination and early seedling growth. *Plant Cell Physiol.* 36: 1059–1065.
- Chao, W.S., Liu, V., Thompson, W.W., Platt, K. and Walling, L.L. 1995. The impact of chlorophyll-retention mutants, d1d2 and cyt-G1, during embryogeny in soybean. *Plant Physiol.* 107: 253–262.
- Coleman, C.E., Lopes, M.A., Gillikein, J.W., Boston, R.S. and Larkins, B.A. 1995. A defective signal peptide in the maize high-lysine mutant fleury-2. *Proc. Natl. Acad. Sci. USA.* 92: 6828–6831.
- Daraselia, N.D., Tarchevskaya, S. and Narita, J.O. 1996. The promoter for tomato 3-hydroxy-3-methylglutaryl coenzyme A reductase gene 2 has unusual regulatory elements that direct high-level expression. *Plant Physiol.* 112: 727–733.
- Dunwell, J.M., Khuri, S. and Gane, P.J. 2000. Microbial relatives of the seed storage proteins of higher plants: conservation of structure and diversification of function during evolution of the cupin superfamily. *Microbiol. Mol. Biol. Rev.* 64: 153–179.
- Ezcurra, I., Ellerstrom, M., Wycliffe, P., Stalberg, K. and Rask, L. 1999. Interaction between composite elements in the *napA* promoter: both the B-box ABA-responsive complex and the RY/G complex are necessary for seed-specific expression. *Plant Mol. Biol.* 40: 699–709.
- Felsenstein, J. 1989. PHYLIP: Phylogeny Inference Package (Version 3.2). *Cladistics.* 5: 164–166.
- Fiedler, U., Filistein, R., Wobus, U. and Bäumlein, H. 1993. A complex ensemble of *cis*-regulatory elements controls the expression of a *Vicia faba* non-storage seed protein gene. *Plant Mol. Biol.* 22: 669–679.
- Fujiwara, T. and Beachy, R.N. 1994. Tissue-specific and temporal regulation of a β -conglycinin gene: roles of the RY repeat and other *cis*-acting elements. *Plant Mol. Biol.* 24: 261–272.
- Grimes, H.D. and Overvoorde, P.J. 1996. Functional characterization of sucrose binding protein-mediated sucrose uptake in yeast. *J. Exp. Bot.* 47: 1217–1222.
- Grimes, H.D., Overvoorde, P.J., Ripp, K., Franceschi, V.R. and Hitz, W.D. 1992. A 62-kD sucrose binding protein is expressed and localized in tissues actively engaged in sucrose transport. *Plant Cell.* 4: 1561–1574.
- Hara-Nishimura, I., Nishimura, M., Matsubara, H. and Akazawa, T. 1982. Suborganellar localization of proteinase catalyzing the limited hydrolysis of pumpkin globulin. *Plant Physiol.* 70: 370–382.
- Hara-Hishimura, I., Takeuchi, Y., Inoue, K. and Nishimura, M. 1993. Vesicle transport and processing of the precursor to 2S albumin in pumpkin. *Plant J.* 4: 793–800.
- Hara-Nishimura, I., Shimada, T., Hatano, K., Takeuchi, Y., Nishimura, M. 1998. Transport of storage proteins to protein storage vacuoles is mediated by large precursor-accumulating vesicles. *Plant Cell.* 10: 825–836.
- Heck, G.R., Chamberlain, A.K. and Ho, T.H. 1993. Barley embryo globulin 1 gene, *Beg1*: characterization of cDNA, chromosome mapping and regulation of expression. *Mol. Gen. Genet.* 239: 209–218.
- Heim, U., Wang, Q., Kurz, T., Borisjuk, L., Golombek, S., Neubohn, B., Adler, K., Gahrz, M., Sauer, N., Weber, H. and Wob, U. 2001. Expression patterns and subcellular localization of a 52 kDa sucrose-binding protein homologue of *Vicia faba* (VfSBPL) suggest different functions during development. *Plant Mol. Biol.* 47: 461–474.
- Hinz, G., Hillmer, S., Baumer, M. and Hohl, I.I. 1999. Vacuolar storage proteins and the putative vacuolar sorting receptor BP-80 exit the Golgi apparatus of developing pea cotyledons in different transport vesicles. *Plant Cell.* 11: 1509–1524.
- Hoh, B., Hinz, G., Jeong, B.K. and Robinson, D.G. 1995. Protein storage vacuoles form *de novo* during pea cotyledon development. *J. Cell Sci.* 108 (Pt 1): 299–310.
- Itoh, Y., Kitamura, Y., Arahira, M. and Fukazawa, C. 1993. *Cis*-acting regulatory regions of the soybean seed storage 11S globulin gene and their interactions with seed embryo factors. *Plant Mol. Biol.* 21: 973–984.
- Itoh, Y., Kitamura, Y. and Fukazawa, C. 1994. The glycinin box: a soybean embryo factor binding motif within the quantitative regulatory region of the 11S seed storage globulin promoter. *Mol. Gen. Genet.* 243: 353–357.
- Jung, R., Nam, Y.-W., Saalbach, I., and Nielsen, N.C. 1997. Role of sulfhydryl redox state and disulfide bonds in processing and assembly of 11S seed globulins. *Plant Cell.* 9: 2037–2050.
- Kawagoe Y., Campbell, B.R. and Murai, N. 1994. Synergism between CACGTG (G-box) and CACCTG *cis*-elements is required for activation of the bean seed storage protein β -phaseolin gene. *Plant J.* 5: 885–890.
- Kawagoe, Y. and Murai, N. 1992. Four distinct nuclear proteins recognize in vitro the proximal promoter of the bean seed storage protein beta-phaseolin gene conferring spatial and temporal control. *Plant J.* 2: 927–936.

- Kinney A.J., Jung, R. and Herman, E.M. 2001. Cosuppression of the α subunits of β -conglycinin in transgenic soybean seeds induces the formation of endoplasmic reticulum-derived protein bodies. *Plant Cell*. 13: 1165–1178.
- Kitamoto, N., Matsui, J., Kawai, Y., Kato, A., Yoshino, S., Ohmiya, K. and Tsukagoshi, N. 1998. Utilization of the TEF1- α gene (TEF1) promoter for expression of polygalacturonase genes, *pgaA* and *pgaB*, in *Aspergillus oryzae*. *Appl. Microbiol. Biotechnol.* 50: 85–92.
- Ko, T.P., Ng, J.D. and McPherson, A. 1993. The three-dimensional structure of canvalin from jack bean (*Canavalia ensiformis*). *Plant Physiol.* 101: 729–744.
- Kriz, A.L. 1989. Characterization of embryo globulins encoded by the maize Glb genes. *Biochem. Genet.* 27: 239–251.
- Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. 227: 680–685.
- Lawrence, M.C., Izard, T., Beuchat, M., Blagrove, R.J. and Colman, P.M. 1994. Structure of phaseolin at 2.2 Å resolution. Implications for a common vicilin/legumin structure and the genetic engineering of seed storage proteins. *J. Mol. Biol.* 238: 748–766.
- Lee, J.S., Brown, W.E., Graham, J.S., Pearce, G., Fox, E.A., Dreher, T.W., Ahern, K.G., Pearson, G.D. and Ryan, C.A. 1986. Molecular characterization and phylogenetic studies of a wound-inducible proteinase inhibitor I gene in *Lycopersicon* species. *Proc Natl Acad Sci USA*. 83: 7277–7281.
- Loake, G.J., Faktor, O., Lamb, C.J. and Dixon, R.A. 1992. Combination of H-box [CCTACC(N)7CT] and G-box (CACGTG) *cis* elements is necessary for feed-forward stimulation of a chalcone synthase promoter by the phenylpropanoid-pathway intermediate *p*-coumaric acid. *Proc Natl Acad Sci USA*. 89: 9230–9234.
- Maleki, S.J., Kopper, R.A., Shin, D.S., Park, C.W., Compadre, C.M., Sampson, H., Burks, A.W., and Bannon, G.A. 2000. Structure of the major peanut allergen Ara h 1 may protect IgE-binding epitopes from degradation. *J. Immunol.* 164: 5844–5849.
- Maruyama, N., Adachi, M., Takahashi, K., Yagasaki, K., Kohno, M., Takenaka, Y., Okuda, E., Nakagawa, S., Mikami, B. and Utsumi, S. 2001. Crystal structures of recombinant and native soybean β -conglycinin β homotrimers. *Eur. J. Biochem.* 268: 3595–3604.
- McHenry, L. and Fritz, P.J. 1992. Comparison of the structure and nucleotide sequences of vicilin genes of cocoa and cotton raise questions about vicilin evolution. *Plant Mol. Biol.* 18: 1173–1176.
- Nebenfuhr, A., Gallagher, L.A., Dunahay, T.G., Frohlick, J.A., Mazurkiewicz, A.M., Meehl, J.B. and Staehelin, L.A. 1999. Stop-and-go movements of plant Golgi stacks are mediated by the acto-myosin system. *Plant Physiol.* 121: 1127–1142.
- Newton, C.H., Flinn, B.S. and Sutton, B.C. 1992. Vicilin-like seed storage proteins in the gymnosperm interior spruce (*Picea glauca engelmannii*). *Plant Mol. Biol.* 20: 315–322.
- Overvoorde, P.J. and Grimes, H. 1994. Topographical analysis of the plasma membrane-associated sucrose binding protein from soybean. *J. Biol. Chem.* 269: 15154–15161.
- Overvoorde, P.J., Frommer, W.B. and Grimes, H.D. 1996. A soybean sucrose binding protein independently mediates nonsaturable sucrose uptake in yeast. *Plant Cell*. 8: 271–280.
- Overvoorde, P.J., Chao, W.S. and Grimes, H.D. 1997. A plasma membrane sucrose-binding protein that mediates sucrose uptake shares structural and sequence similarity with seed storage proteins but remains functionally distinct. *J. Biol. Chem.* 272: 15898–15904.
- Pedra, J.H.F., Delu-Filho, N., Pirovani, C.P., Contim, L.A.S., Dewey, R.E., Otoni, W.C. and Fontes, E.P.B. 2000. Antisense and sense expression of a sucrose binding protein homologue gene from soybean in transgenic tobacco affects plant growth and carbohydrate partitioning in leaves. *Plant Sci.* 152: 87–99.
- Rabjohn P., West, C.M., Connaughton, C., Sampson, H.A., Helm, R.M., Burks, A.W., Bannon, G.A. 2002. Modification of Peanut Allergen Ara h 3: Effects on IgE Binding and T Cell Stimulation. *Int. Arch. Allergy Immunol.* 128: 15–23.
- Ripp, K.G., Vitanen, P.V., Hitz, W.D. and Fransceschi, V.R. 1988. Identification of a membrane protein associated with sucrose transport into cells of developing soybean cotyledons. *Plant Physiol.* 88: 1435–1445.
- Sakata, Y., Chiba, Y., Fukushima, H., Matsubara, N., Habu, Y., Naito, S. and Ohno, T. 1997. The RY sequence is necessary but not sufficient for the transcription activation of a winged bean chymotrypsin inhibitor gene in developing seeds. *Plant Mol Biol.* 34: 191–197.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. Molecular cloning: A Laboratory Manual. 2nd ed. Cold Spring Harbor Laboratory Press, Plainview, NY.
- Shasany, A.K. and Koundal, K.R. 2000. Isolation and characterisation of legumin promoter sequence from chickpea (*Cicer arietinum* L.). *Indian J Exp Biol.* 38: 363–372.
- Shutov, A.D., Kakhovskaya, I.A., Braun, H., Bäumlein, H. and Müntz, K. 1995. Legumin-like and vicilin-like seed storage proteins: evidence for a common single-domain ancestral gene. *J. Mol. Evol.* 41: 1057–1069.
- Steinberg, T.H., Chernokalskaya, E., Berggren, K., Lopez, M.F., Diwu, Z., Haugland, R.P., Patton, W.F. 2000. Ultrasensitive fluorescence protein detection in isoelectric focusing gels using a ruthenium metal chelate stain. *Electrophoresis* 21: 486–496.
- Stockhaus, J., Poetsch, W., Steinmuller, K. and Westhoff, P. 1994. Evolution of the C4 phosphoenolpyruvate carboxylase promoter of the C4 dicot *Flaveria trinervia*: an expression analysis in the C3 plant tobacco. *Mol Gen Genet.* 245: 286–293.
- Tegeder, M., Wang, X.D., Frommer, W.B., Offler, C.E. and Patrick, J.W. 1999. Sucrose transport into developing seeds of *Pisium sativum* L. *Plant J.* 18: 151–161.
- Thompson, J.D., Higgins, D.G. and Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22: 4673–4680.
- Tilly, J.J., Allen, D.W. and Jack, T. 1998. The CArG boxes in the promoter of the *Arabidopsis* floral organ identity gene *APETALA3* mediate diverse regulatory effects. *Development*. 125: 1647–1657.
- Towbin, H., Staelin, T. and Gordon, J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc. Natl. Acad. Sci. USA*. 76: 4350–4354.
- Vellanoweth, R.L., and Okita, T.W. 1993. Analysis of nuclear proteins interacting with a wheat α/β -gliadin seed storage protein gene. *Plant Mol Biol.* 22: 25–41.
- Warmbrodt, R.D., Buckhout, T.J. and Hitz, W.D. 1989. Localization of a protein, immunologically similar to a sucrose-binding protein from developing soybean cotyledons, on the plasma membrane of sieve-tube members of spinach leaves. *Planta*. 180: 105–115.
- Warmbrodt, R.D., Vanderwoude, W.J. and Hitz, W.D. 1991. Studies on the localization of a protein, immunologically similar to a 62-kD sucrose-binding protein isolated from developing soybean cotyledons, in the shoot and root of spinach. *New Phytol.* 118: 501–512.

- West A.G. and Sharrocks, A.D. 1997. The role of DNA-bending in MADS-box transcription factor function. *Biochem Soc Trans.* 25: S639.
- Yu-Bing, L., Rogers, S.W., Tse, Y.C., Lo, S.W., Sun, S.S.M., Juah, G.-Y. and Jiang, L. 2002. BP-80 and homologs are concentrated in post-Golgi, lytic prevacuolar compartments. *Plant Cell Physiol.* in press.
- Zhou, X. and Fan, Y.L. 1993. The endosperm-specific expression of a rice prolamin chimeric gene in transgenic tobacco plants. *Transgenic Res.* 2: 141–146.