A Soybean Vacuolar Protein (P34) Related to Thiol Proteases Is Synthesized as a Glycoprotein Precursor during Seed Maturation*

(Received for publication, January 16, 1992)

Andrew Kalinski, Diane L. Melroy*, Radhey Shyam Dwivedi§, and Eliot M. Herman†

From the Plant Molecular Biology Laboratory, United States Department of Agriculture, Agricultural Research Service, Beltsville, Maryland 20705

We have examined the synthesis, posttranslational processing, and localization of soybean P34, a member of the papain superfamily. P34 has been identified as a constituent of oil storage organelles or oil bodies isolated from seed lysates and has been assumed to be one of the oil body proteins. Electron microscopic immunocytochemistry with a monoclonal antibody demonstrated that P34 is localized in the protein storage vacuoles but not in the oil bodies. Immunocytochemical observations of partially disrupted seed cells showed that the association of P34 with oil bodies appears to occur as a consequence of cell lysis. In vitro synthesis of P34 results in the formation of a 46-kDa polypeptide that increases to 47 kDa due to core glycosylation by canine microsomes. In vitro synthesis studies in the presence and absence of tunicamycin, an inhibitor of N-linked glycosylation, indicate that pro-P34 is 47 kDa. Since the cDNA sequence of prepro-P34 contains a single putative glycosylation site in the precursor domain, we conclude that P34, like a few other vacuolar proteins, is synthesized as a glycoprotein precursor. Pulse-chase experiments showed that the processing of pro-P34 to mature P34 occurs in a single step and that this posttranslational cleavage occurs on the carboxyl side of an Asn, which is typical of seed vacuolar proteins. Pro-P34 (47 kDa) is detected in immunoblots of maturing seeds. Analysis of RNA indicates that the P34 genes are expressed only during seed maturation and that the P34 mRNA is related to other thiol protease mRNAs detectable in other organs and plants. Unlike other seed thiol proteases that are synthesized only after seed germination, P34 accumulates during seed maturation.

Thiol proteases belong to the papain superfamily (EC 3.4.22), which are expressed in diverse eukaryotic species (see Bond and Butler (1987) for review). The sequences of cDNA and genomic clones have been determined for many members of this superfamily, including the cathepsins of animal cells (Tahio et al., 1983; Chan et al., 1986; Ishidoh et al., 1987, 1989), thiol proteases expressed in slime molds (Williams et al., 1985; Pears et al., 1985; Datta and Firtel, 1987), plant cell proteases including papaya fruit papain (Cohen et al., 1986), barley (Rogers et al., 1985; Koehler and Ho, 1988; Koehler and Ho, 1990), pea (Guerrero et al., 1990), rice (Watanabe et al., 1991), mung bean (Mitsuhashi et al., 1986; Akasofu et al., 1989), and kiwi fruit actinidin (Prenekelt et al., 1988; Podivinsky et al., 1989). Sequence comparisons between the protein and deduced sequences show considerable homology in several domains, which indicates that this gene family has been highly conserved. The sequence homology among thiol proteases extends from the 5' to 3' ends of the open reading frames, including not only the mature polypeptide but also the signal sequence and precursor domains (North, 1986).

Although the sequence and structure of members of the thiol protease family is well characterized, the elucidation of the physiological roles of these proteins, especially in higher plants, is still not well understood. The expression of plant thiol protease genes has been shown to be correlated with the onset of physical stress (Schaffer and Fischer, 1988; Guerrero et al., 1990) and the mobilization of vacuolar reserve storage proteins, such as in legume (Baumgartner and Chrispeels, 1977; Baumgartner et al., 1978), rice (Watanabe et al., 1991), barley seeds (Koehler and Ho, 1988; Holwerda et al., 1990), and brassica seedlings (Dietrich et al., 1989). In other instances, the presence of abundant thiol protease(s) does not appear to be well correlated with an obvious physiological function, such as in the case of bromelin, a thiol protease of pineapple plant stems.

Where there is a good correlation between synthesis of thiol proteases and physiological function there still may be considerable complexities involved in thiol protease gene expression. For example barley aleurones express at least three different thiol proteases, termed cysteine proteases EP-A and EP-B (Koehler and Ho, 1990) and aulurain (Rogers et al., 1985), during germination and in response to the phytohormone gibberellin. Two of these thiol proteases (EP-A and EP-B) are secreted from the cell, whereas the other (aulurain) is sequestered within the aleurone cells. The intracellular thiol protease appears to be localized in a vacuole-like subcellular compartment that is distinct from the protein storage vacuoles, termed aleurone grains, that contain seed storage proteins. Aulurain has not yet been shown to possess proteolytic activity, and the functional role and exact characterization of the aulurain-containing organelle remains to be determined.

We have previously shown that isolated soybean oil bodies contain polypeptides of 34, 24, 18, and 17 kDa, which we have termed P34, P24, P18, and P17, respectively (Herman 1987; Herman et al., 1990). Using monoclonal antibodies directed against P34, we observed that it undergoes developmentally regulated processing to a 32-kDa protein during seedling...
growth as the consequence of the removal of an amino-terminal decapeptide (Herman et al., 1990). We isolated a cDNA clone of P34 and found it was highly homologous to the papain superfamily of thiol proteases (Kalinski et al., 1990). The apparent expression of P34 during seed maturation rather than seed germination sets it apart from other described seed thiol proteases. The apparent localization of a thiol protease-related gene product in the oil bodies was surprising, and as a consequence, we undertook a series of studies to examine the developmental regulation of the synthesis, processing, and localization of P34 in order to better define its potential physiological roles.

**EXPERIMENTAL PROCEDURES**

**Materials**—Soybean plants (Glycine max Merr. L. cv. Century) used as a source of developmentally staged seeds were maintained as previously described (Herman et al., 1990). Germinating seeds were grown by imbibing dry seeds overnight and sowing the seeds on wet paper towels. The seedlings were maintained in the dark at room temperature as a source of staged cotyledons. Monoantibodies P3E1 and P4B5, which are directed against two different epitopes of P34, were previously described in Herman et al. (1990). A cDNA clone encoding a papain superfamily of thiol protease-related gene product in the oil bodies was surprising, and as a consequence, we undertook a series of studies to examine the developmental regulation of the synthesis, processing, and localization of P34 in order to better define its potential physiological roles.

**RNA and DNA Isolation**—Total RNA was isolated as previously described (Kalisz et al., 1990). Poly(A)^+ RNA was fractionated from total RNA on Poly(A) Qwik column (Stratagene Cloning Systems). Phagemid DNA containing a P34 cDNA insert was isolated from recombinant bacteria using a boiling method (Hols and Geigley, 1981). The P34 cDNA insert was released from pBluescript SK(–) phagemid DNA by EcoRI and BamHI digestions, separated from the vector DNA by electrophoresis on 0.8% agarose gel, and then extracted from the gel (Samboek et al., 1989). RNA and DNA were quantitated spectrophotometrically.

**In Vitro Translation, in Vivo Labeling Immuno precipitation, Oil Body Protein Isolation, SDS-PAGE**—Immunoblots—mAb P3E1- and P4B5-secreting cell lines were expanded in vivo as ascites tumors. Ascites fluid (0.33 ml) was diluted 1:1 with 25 mM Tris-HCl, pH 7.4, 150 mM NaCl (TBS) and clarified by centrifugation in a Brinkmann microcentrifuge for 5 min. The supernatant was passed over a protein G-Sepharose 4B column (Pharmacia-LKB, Uppsala, Sweden) and washed with TBS. The bound IgGs were eluted with 0.025 M glycine HCl, pH 2.8, containing 0.1 M sodium borate/boric acid buffer, pH 8.6, and dialyzed against 4 liters of borate-buffered saline (10 mM sodium borate, 150 mM NaCl, pH 8.6) overnight at 7°C. Equal amounts of mAbs P3E1 and P4B5 IgGs were combined on CNBr-Sepharose 4B with excess IgG in order to obtain maximum amounts of bound antibody, according to the manufacturer's directions. In vitro translation of midmaturation (150 mg, fresh weight) soybean seed poly(A)^+ mRNA was accomplished using a reticulocyte translation kit (in vitro express, Stratagene) with [35S]methionine, according to the supplied protocol. Canine microsomes (Fromega Biotech) (1.8 or 3.6 eq) were added to some of the translation reactions in order to assess cotranslational processing of newly synthesized polypeptides.

**In vivo labeling was accomplished by incubation of thin slices (1 mm thick) in 25 μCi of carrier-free ([35S]methionine. Samples were initially pulsed for 1 h and, in some instances, chased by transferring the tissue slices to 4% sucrose, 1% asparagine, 1 mM methionine for varying lengths of time. In order to examine the glycosylation status of the newly synthesized proteins, tissue samples were incubated in 0.05 mg/ml tunicamycin for 30 min, then 25 μCi of ([35S]methionine was added. Tissue samples were then washed with 1X PBS, 12069, and centrifuged at 15000 × g for 1 h. Parallel sets of thin slices were incubated in the same solutions without tunicamycin. The in vivo labeled samples were placed in microcentrifuge tubes and then frozen with liquid nitrogen and extracted by sonication in a bath sonicator for 15 min after the addition of 50 μl of 6 M urea, 5% (w/v) Triton X-100, 0.1 M Tris-HCl, pH 6.8, 5% (v/v) β-mercaptoethanol (SDS-urea sample buffer). The resulting lysates were removed to a new microcentrifuge tube, diluted with TBS containing 2% v/v Triton X-100, and then clarified by centrifugation in a Brinkmann microcentrifuge. The labeled polypeptides were recovered by adding 50 μl of mAb P3E1/P4B5 Sepharose 4B and tumbling the samples overnight at room temperature. The Sepharose-antibody-P34 complexes were washed by washing the beads 3 times with TBS. Following the final wash, the buffer was removed and the Sepharose-antibody-P34 complex was denatured by the addition of 100 μl of 8 M urea, 5% (v/v) SDS, 50 mM Tris-HCl, pH 6.8, with 5% (v/v) β-mercaptoethanol.

**Electron Microscopy and Immunocytochemistry**—Developmentally staged soybean seed cotyledons were cut into approximately 1-mm cubes and fixed with 4% formaldehyde and 2% glutaraldehyde in 0.1 M sodium cacodylate buffer, pH 7.2, and embedded in 10% fetal bovine serum/TBST, pH 7.4, overnight at 4°C. The grids were then washed with TBST and indirectly labeled with 10 nm colloidal gold particles coupled to rabbit antimouse 1 the abbreviations used are: SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; mAb, monoclonal antibody.
Soybean Vacuolar Protein (P34)

RESULTS

P34 Is Found in Purified Oil Bodies and Can Be Selectively Removed from the Oil Body Fraction by Carbonate Extraction—Oil bodies were purified by centrifugal flotation of seed lysates. The SDS-PAGE of proteins extracted from crude soybean oil bodies isolated from whole seeds, total cotyledon, and axis homogenates are shown in Fig. 1 (lanes 1–3). As previously shown, four major polypeptides, 34, 24 (oleosin), 18, and 17 kDa, are enriched as a consequence of purifying soybean oil bodies (Herman, 1987; Murphy and Cummins, 1989; Tzen et al., 1990). Additional high molecular weight proteins corresponding to the molecular weight of soybean vacuolar storage proteins are also observed. In order to examine whether P34 is embedded in the oil body membrane, isolated oil bodies were treated with 0.1 M sodium carbonate for 1 h on ice. The carbonate treatment of the oil bodies resulted in the selective extraction of P34, whereas the 24- and 18-kDa oleosins remained firmly embedded in the oil body membrane (Fig. 1, lane 4). The selective removal of P34 by carbonate may indicate that P34 is not embedded in the oil body membrane. This would be consistent with the deduced sequence of the P34 (Kalinski et al., 1990), which does not contain any obvious domain for membrane insertion.

Immunological Detection of pro-P34 in Maturing Soybean Seed—The open reading frame of P34 indicated that there is a potential coding region encoding 122 amino acids 5' of the experimentally determined amino terminus (Kalinski et al., 1990; Herman et al., 1990), which is homologous to the prepro regions of the deduced sequences of other plant thiol proteases. In order to examine whether there is an accumulation of high molecular weight putative precursors of P34, samples of soybean seed cotyledons at different developmental stages were extracted with SDS-sample buffer. The samples were fractionated by SDS-PAGE and blotted onto nitrocellulose membranes. Replicate membranes were stained with Amido Black to visualize total proteins or immunolabeled with P3E1 mAb (Fig. 2). As previously shown, the 34-kDa P34 is accumulated during the course of seed development (Herman et al., 1990). A putative pro-P34 polypeptide is immunolabeled with an apparent molecular mass of 47 kDa (Fig. 2B). The size of the putative pro-P34 observed on the immunoblots is similar to the predicted size of the pro-P34 based on its primary structure. This high molecular weight polypeptide is not observed during germination and seedling growth (data not shown), which is consistent with the accumulation of P34 mRNA and its translation to P34 polypeptide during seed maturation.

P34 Is Localized in Protein Storage Vacuoles—Biochemical fractionation of soybean seeds and analysis of the proteins found in the oil bodies recovered by centrifugal flotation have shown the presence of four abundant polypeptides of 34, 24, 18, and 17 kDa (Herman, 1987; Murphy and Cummins, 1989; Herman et al. 1990; Tzen et al., 1990). The 24- and 18-kDa polypeptides are members of the oleosin family of integral membrane proteins of the oil bodies. The primary structure of the 34-kDa protein indicates that it belongs to the papain superfamily of thiol proteases, which in plants includes members that are localized in the vacuole or are secreted. The 34-kDa protein does not have a membrane insertion region, which suggests that if it is associated with the oil bodies, it is either bound to the oil body surface or to other proteins such as the oleosins. The biochemical characteristics of P34, synthesis as a prepropolypeptide, and its glycosylation, are consistent with the synthesis of a secretory protein destined for secretion or accumulation in the vacuole. This suggests that P34 may not be an authentic oil body protein. In order to resolve the question of the localization of P34, electron microscopic immunocytochemical analysis was performed with immunoaffinity-purified monoclonal antibodies and indirect labeling with colloidal gold second antibody. The purified mAb P4B5 densely labeled the protein storage vacuoles or protein bodies in the storage parenchyma cells (Fig. 3). Segments of endoplasmic reticulum in the cytoplasm surrounding the protein storage vacuoles were also labeled with gold particles, which is consistent with the synthesis of P34 as an N-linked glycoprotein precursor. Oil bodies in the surrounding cytoplasm were not labeled by mAb P4B5 and the colloidal gold second antibody (Fig. 3), which indicates that P34 is not a genuine constituent of the oil bodies. The localization of P34 within the protein storage vacuoles is consistent with the...
biosynthesis and posttranslational data presented in the next sections, which show P34 to be posttranslationally processed by mechanisms typical for vacuolar proteins.

The immunocytochemical observations suggested that P34 may become associated with the oil bodies as a consequence of cellular lysis and/or fractionation. We also conducted immunogold examination of partially disrupted cells on the edge of the tissue blocks. Disruption of these cells is a consequence of mincing the tissue as a preparation for fixation and embedding for electron microscopy. In these cells, the protein storage vacuoles were observed to be disrupted, releasing their contents into the cytoplasm. Immunogold localization of P34 in the disrupted cells with mAb P4B5 shows extensive labeling of the oil bodies in these cells (Fig. 4). Therefore, based on our immunocytochemical observations, we consider the association of P34 with the oil bodies to be the consequence of affinity of this protein for the oil body surface.

**P34 Remains Associated with the Protein Storage Vacuole during Mobilization of Storage Proteins**—During seed ontogeny, the protein storage vacuoles undergo developmentally regulated and temporal differentiation or subdivision of the central vacuole into numerous small protein storage vacuoles, which are filled with storage proteins and other enzymes, lectins, and defense proteins. After germination, the storage proteins are progressively digested, and the small protein storage vacuoles fuse to reform the central vacuole (for a review, see Herman 1992). The immunocytochemical localization of P34 was examined during seed germination and seedling growth. This laboratory has isolated P34 and observed its apparent association with oil bodies in imbibed and germinated seeds (Herman et al., 1990). P34 was localized in the protein storage vacuoles during the developmental stage.
of storage protein mobilization that is shown in Fig. 6. Very little gold label was apparent in the other cellular constituents, including the oil bodies. However, in cells that contained disrupted protein vacuoles, P34 was redistributed and observed to have a very specific association with the oil bodies (data not shown).

**P34 Gene Expression Occurs in Midmaturation Soybean Seeds**—The expression of mRNAs encoding P34 during the course of soybean seed maturation and seedling growth was examined with slot blot hybridization assays. Identical amounts of total RNA derived from cotyledons (15-, 45-, and 150-mg fresh weight seeds) and 3, 7, and 12 days of seedling growth were blotted onto nitrocellulose membranes. The blot was washed under high stringency conditions in order to assess the abundance of messages specifically encoding P34 and not other members of the papain superfamily of thiol proteases. In order to determine whether similar mRNAs present in other organs and plants would hybridize to P34 cDNA under high stringency conditions, total RNA from soybean, tomato, rice, and barley leaves was blotted onto adjacent slots of the nitrocellulose membrane. The resulting hybridization clearly shows that P34 mRNA is accumulated in midmaturation soybean seed cotyledons and is not found prior to the initiation of storage substance accumulation in 15-mg fresh weight seeds or during seedling growth (Fig. 7A). Further, P34 mRNA was not observed in soybean leaves nor was any very closely related mRNA observed in tomato, rice, or barley leaves.

**P34 Message Is Related to Other mRNAs Encoding Thiol Proteases**—The thiol proteases are widely distributed among eukaryotic cells. Southern blot analysis of several plant species has shown, including in soybean (Kalinski et al., 1990), that these genomes contain several thiol protease genes. In order to assess whether similar thiol protease genes are expressed in germinating soybean seeds and mature leaves, Northern blots were probed with P34 cDNA (Fig. 7B). Low and moderate stringency washes (42 and 58 °C, respectively) indicate that there is a similar thiol protease mRNA synthesized during seedling growth in soybean cotyledons. The mRNA for this thiol protease is 1.7 kilobases, slightly larger than that of P34, 1.5 kilobases, indicating that it is encoded by a distinct gene. This protease is likely to be similar to other described proteases, which are expressed during seedling growth that probably mediates the mobilization of storage proteins. The P34 cDNA strongly cross-hybridizes with a 1.7-kilobase mRNA of mature soybean leaves at low stringency. However the P34 cDNA cross-hybridizes to the leaf thiol protease mRNA poorly at moderate stringency (58 °C). This indicates that the thiol protease mRNAs observed during
seedling growth and in mature leaves are different mRNAs, although they are similar in size. The P34 cDNA cross-
hybridizes to thiol protease mRNAs present in tomato, rice, and barley leaves at low and moderate stringency washes (Fig. 7B). These results indicate that the P34 gene appears to be
expressed only during soybean seed maturation. Further, these results demonstrate that P34 is a member of the papain family of thiol proteases expressed in soybean plants.

The tissue-specific distribution of P34 in soybean tissues and organs was also examined by immunoblot analysis. Samples obtained from a soybean plant in the reproductive stage, including late maturation cotyledon, axis, aleurone, seed coat, and pod, as well as stem, roots, and mature leaves, were fractionated with SDS-PAGE and transferred to nitrocellulose membranes. Replicate blots were stained with Amido Black to visualize total protein or probed with immunoaffinity-purified mAb P4B5. A 34 kDa band was immunolabeled only in the samples derived from cotyledons, axis, and aleurone, indicating that expression of P34 is restricted to embryonic tissues (data not shown). The relative intensity of the immunolabeled P34 band in the axis lane was much reduced, as compared with the cotyledon lane, indicating that there appears to be differential amount of P34 accumulation among embryonic tissues.

In Vitro and in Vivo Synthesis of Prepro and Pro-P34 and Its Posttranslational Processing—In order to assess the molecular size of newly synthesized P34 and to examine whether it is posttranslationally processed, a series of in vitro and in vivo synthesis studies were conducted. Poly(A)⁺ mRNA isolated from midmaturation soybean seeds was translated in rabbit reticulocyte lysate. Immunoprecipitation with mAbs P3E1/P4B5 coupled to Sepharose 4B beads and size analysis by SDS-PAGE fluorography resulted in the identification of a 46 kDa band (Fig. 8A, lane 1). The size of the isolated polypeptide is similar to the estimated 43-kDa polypeptide, which would result from translation of the open reading frame of the P34 cDNA clone. Two additional labeled bands (39 and 34 kDa) were immunoprecipitated by P3E1/P4B5 mAb-Sepharose 4B beads (Fig. 8A, lane 1). These additional polypeptide bands may result from translation initiation at methionine codons corresponding to amino acid residue numbers 78 and 120 of the P34 open reading frame (Fig. 9). In vitro translation reactions supplemented with canine pancreatic microsomes were undertaken in order to estimate the size of the putative signal sequence identified in the open reading frame of the P34 cDNA clone. However, translation in the presence of microsomes resulted in a net increase in molecular mass of the immunoprecipitated translation product. Fig. 8A, lanes 2 and 3, show the effect of the addition of two different concentrations of canine microsomes supplementing the translation mixture. Note that in lane 2 both 46- and 47-kDa polypeptides were recovered, whereas reactions with twice the microsome concentration of lane 2 resulted in a nearly complete conversion of 46-kDa prepro-P34 to 47-

kb pro-P34. These results are consistent with core N-linked glycosylation at the consensus site (Kornfeld and Kornfeld, 1985) in the precursor segment (Fig. 9). The additional minor polypeptides at 39 and 34 kDa were also observed in translation reactions in the presence of canine microsomes. The 39- and 34-kDa secondary translation products do not shift in molecular weight in response to added canine microsomes (Fig. 8A, lanes 1–3, asterisks). This is consistent with analysis of the primary structure of the P34 cDNA, which shows a single putative asparagine-linked glycosylation site at amino acid residue 70 (Fig. 9, boxed), preceding the codons for methionines 2 and 3. Therefore, these truncated secondary translation products lacking the consensus glycosylation site would not be expected to undergo molecular weight shift by core glycosylation.

This conclusion is supported by in vitro transcription/translation of a P34 cDNA clone (Kalinski et al., 1990) in which two labeled immunoprecipitated polypeptides of 39 and 34 kDa were isolated (data not shown). Supplementing the translation reaction with canine microsomes did not alter the molecular weight of either of these two bands, indicating the absence of a cleavable signal sequence or a core glycosylation site on the truncated translation product (data not shown). The capped synthetic mRNA did not appear to direct the synthesis of the entire open reading frame, which we hypothesize to be the consequence of a short 5' -untranslated leader of 54 base pairs, including the T7 polymerization promoter. The synthetic mRNA was experimentally determined to correspond in size to the full length of the cDNA. Therefore, although the P34 cDNA apparently encodes the entire open
reading frame, the in vitro translation of the synthetic RNA appears be initiated at the second and third methionines corresponding to residues 78 and 120 in the precursor region.

In order to examine the molecular size of in vivo synthesized pro-P34 and its posttranslational product, thin slices of immature soybean cotyledons were preincubated in the core glycosylation inhibitor tunicamycin. After the 2-h preincubation in the inhibitor, the incubation medium was supplemented with 25 μCi of (14C)methionine. Radioactive pro-P34 was recovered from the seed lysates by immunoaffinity with mAb P3E1/P4B5 beads and analyzed by SDS-PAGE and fluorography. The result of this experiment is shown in Fig. 8B. Control samples of pro-P34 had an apparent mass of 47 kDa (lane −), whereas pro-P34 synthesized in the presence of tunicamycin is about 2 kDa smaller (lane +). SDS-PAGE fluorography of mixed control and inhibitor samples shows a doublet separated by 2 kDa (lane mix). This result is consistent with the mass of a single high mannose glycan side chain attached to pro-P34.

**DISCUSSION**

Herman et al. (Herman, 1987; Herman et al., 1990) and others (Murphy and Cummins, 1989; Tzen et al., 1990) have identified P34 as a constituent of the seed oil storage organelle or oil bodies. Remarkably, immunolocalization reveals that the biochemically deduced localization is an artifact of cell lysis and fractionation used in the oil body preparation procedure. Why is P34 purified by the simple expedient of isolating oil bodies? Based on the information we have detailed in this paper, we envision that the affinity of P34 for the oil bodies results from solubilization with the storage proteins and its binding to the oil bodies. Kalinski et al. (1991) have
determined the DNA sequence of isoforms of 24-kDa soybean oleosin, the major oil body membrane protein. The sequence has not revealed any structural features that are obvious candidates for P34's binding to oil bodies. Soybean oil bodies also contain a minor oleosin (18 kDa) of unknown sequence. We cannot exclude the possibility that P34 interacts with the minor, rather than the major, soybean oleosin. Whether the binding of P34 to the oil body membranes is merely fortuitous or whether it is indicative of some aspect of its potential function(s) will require additional investigation.

The posttranslational processing of pro-P34 to mature P34 apparently involves the cleavage on the carboxyl side of Asn at residue 122 (Kalinski et al. 1990) of P34, producing the experimentally determined amino terminus (Herman et al. 1990). Asn-specific processing has been demonstrated in a wide variety of plant seed vacuolar proteins, such as the 11s superfamily of storage proteins. Although processing on the carboxyl side of Asn is a common form of processing of seed vacuole proteins, processing in many other sites is also known. Except for papain, processing after Asn for maturation of the pro form is not apparently characteristic of other plant thiol proteases (Mitsuhashi and Minamikawa, 1988; Koehler and Ho, 1990). The seed vacuolar-specific mode of processing of P34 provides indirect biochemical evidence in support of the immunochemical localization. The processing of other described plant cell thiol proteases has indicated that there may be one or more intermediate products that give rise to the final mature product. In this regard, P34 is quite different in that there is apparently only a single processing site of the precursor during seed maturation. The amino-terminal sequence of P34 exhibits a second processing during seedling growth. Removal of a decapeptide from the amino terminus produces a 32-kDa polypeptide (Herman et al., 1990) whose amino terminus aligns with the amino terminus of other plant and animal thiol proteases.

In this paper, we have shown that P34 is synthesized as a glycoprotein in vivo and in vitro. The N-glycosylation of pro-P34 likely occurs at the consensus site in the precursor domain. The glycosylation of precursor segments has been shown for a limited number of plant vacuolar proteins. A glycoprotein precursor was first shown for the jack bean lectin concanavalin A (Herman et al., 1985) and has subsequently been shown in wheat and rice lectins (Mansfield et al., 1988), as well as &beta;-glucanase (Shinshi et al., 1988). The glycopeptide excised from pro-concanavalin A is also processed on the carboxyl side of Asn residues (Faye and Chrispeels, 1987). Analysis of the deduced sequences available for plant thiol proteases shows that putative glycosylation sites are present in some of the precursor domains of plant thiol protease genes, such as papain (Cohen et al., 1986), actinin (Podivinsky et al., 1989), barley cysteine protease (Koehler and Ho, 1990), and &gamma;-oryzain (Watanabe et al., 1991), but this is not a highly conserved feature, since consensus glycosylation sites are absent in the precursor domain of aleurain (Rogers et al., 1985), mung bean (Mitsuhashi and Minamikawa, 1989), and &alpha; and &beta;-oryzains (Watanabe et al., 1991). Some of these proteins, such as barley aleurain, possess glycosylation sites, but these are located in the mature domain (Holwerda et al., 1990). Those genes that possess glycosylation sites in the precursor domain are not in aligned positions. Why might glycosylation of the precursor be an essential feature for some thiol proteases? Vernet et al. (1990) have shown that papain secretion from insect cells is inhibited in the translation of mutated genes that possess deleted consensus glycosylation sites. One interpretation of this observation is that lack of glycosylation in the precursor domain results in misfolding with consequent recognition and disposal by the protein quality control mechanisms of animal cells in the pre-Golgi region of the endomembrane system (for a review, see Pelham (1989)). Cotranslational core glycosylation may be essential for the cooperative interactions of other enzymes, for example binding protein and protein disulfide isomerase (for a review, see Pelham (1989)), which mediate the folding of the precursor protein into a correct tertiary structure.

P34 is a moderately abundant protein of soybean embryonic tissues; however, the physiological function of P34 remains unknown. Many vacuolar proteins undergo endoproteolytic cleavage upon deposition in the vacuole. A primary processing site is on the carboxyl side of exposed Asn residues. Such cleavage sites are highly conserved in the 11s storage protein family and are present in legume lectins such as concanavalin A (Bowles et al., 1986) and P34 (this paper). Scott et al. (1992) have isolated the 11s-processing protease from soybean seeds. Although the processing protease has been inferred to be a thiol protease, the protein Scott et al. (1992) have isolated and P34 have different molecular weights, indicating that P34 is not the processing protease. We have considered the possibility that P34 is the protease responsible for degradation of storage proteins during seedling growth. The removal of a decapeptide from the amino terminus of P34 during seedling growth coincides with the initiation of storage protein mobilization. The activation of thiol proteases by removal of small protein segments has been documented in cathespins (for a review, see Bond and Butler (1987) and Erickson et al. (1991)). However, we still have not shown that the 32-kDa processed form of P34 possesses proteolytic activity. We have also shown in this paper that soybean seeds appear to synthesize a new mRNA related to P34 during seedling growth. This mRNA is likely to encode a protease that functions to mobilize storage proteins. Recent measurements of proteolytic activity in germinating soybean seeds also clearly show that the protease is either activated or newly synthesized after several days of seedling growth (Wilson et al., 1988; Papastoïtis and Wilson, 1991). Therefore, it appears that P34 is neither a storage protein processing enzyme nor the endoprotease induced during seedling growth. We do not yet know the physiological function of this protein. Without an obvious functional role, and in its developmental regulation, P34 appears to be different from the other described seed thiol proteases.

REFERENCES

Aksoy, H., Yamauchi, D., Mitsuhashi, W., and Minamikawa, T. (1989) Nucleic Acids Res. 17, 6733
Baumgartner, B., and Chrispeels, M. J. (1977) Eur. J. Biochem. 77, 233-233
Baumgartner, B., Tokuyasu, K. T., and Chrispeels, M. J. (1978) J. Cell Biol. 79, 10-19
Bond, J. S., and Butler, P. E. (1987) Annu. Rev. Biochem. 56, 333-364
Bowles, D. J., Marcus, S. E., Pappin, D. J. C., Findlay, J. B. C., Eliopoulos, E., Maycox, P. R., and Burgess, J. (1986) J. Cell Biol. 102, 1284-1297
Chan, S. J., San Segundo, B., McCormick, M. B., and Steiner, D. F. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 7721-7725
Cohen, L. W., Coghlan, V. M., and Dihel, L. C. (1986) Gene (Amst.) 48, 219-227
Datta, S., and Firtel, R. A. (1987) Mol. Cell. Biol. 7, 149-159
Dietrich, R. A., Maslyer, D. J., Heupel, R. C., and Harada, J. J. (1989) Plant Cell 1, 73-80
Erickson, A. H. (1989) J. Cell. Biol. 109, 31-41
Faye, L., and Chrispeels, M. J. (1987) Plant. Sci. 170, 217-224
Feinberg, A. P., and Vogelstein, B. (1983) Anal. Biochem. 132, 6-13
Fujiki, Y., Hubbard, A. L., Fowler, S., and Lazarow, P. B. (1982) J. Cell Biol. 93, 97-102
Guerrero, F. D., Jones, J. T., and Mullet, J. E. (1990) Plant Mol. Biol. 15, 11-26
