RESEARCH PAPER

A multifunctional bicupin serves as precursor for a chromosomal protein of Pisum sativum seeds

Josefa Castillo*, Ainhoa Genovés†, Luis Franco‡ and M. Isabel Rodrigo

Department of Biochemistry and Molecular Biology, University of Valencia, E-46100, Burjassot, Valencia, Spain

Received 27 July 2005; Accepted 1 September 2005

Abstract

The fact that the psp54 gene codes for p16, a seed chromatin protein of Pisum sativum, has been described previously. In the present paper it is shown that p54, the p16 precursor, also exists as a free polypeptide in pea and that it also yields p38, a second polypeptide from the N-terminal region of p54, which is co-localized at a subcellular level with p16. By using antibodies against pea p16 and p38, it was found that these proteins are present in the members of the tribe Viciae examined. Sequence analysis and 3D modelling indicates that p54 proteins belong to the cupin superfamily, and that they are related to sucrose binding proteins and, to a lesser extent, to vicilin-type seed storage proteins. Nevertheless, several distinctive characteristics of psp54 expression have been found: (i) the gene is differentially induced by ABA and several stress situations, in accordance with the presence of putative separate ABA and stress responsive elements in its promoter; (ii) the proteins are present in pods and seed coats, tissues of maternal origin; and (iii) p54 mRNA accumulates in the dry seeds. In view of both the functional properties of p54-derived proteins and the features of the psp54 gene expression, it is concluded that p54 represents a novel class within the cupin superfamily.

Key words: Abscisic acid, ABA, cupins, pea seed germination, processing of precursor peptides, protein modelling, stress-induced gene expression.

Introduction

The cupin superfamily of proteins, named after cupa, the Latin word for barrel (Dunwell, 1998), is characterized by the presence of a conserved β-barrel (Gane et al., 1998). The existence of this superfamily was first suspected when Lane et al. (1991) found that the ‘germin box’, a nonapeptide present in the wheat protein germin, was shared by spherulin, a stress-induced protein of the slime mould Physarum polycephalum. Posterior studies based on sequence analysis showed that this similarity also extended to other proteins, such as the seed storage proteins vicilin and legumin (Bäumlein et al., 1995). The availability of the three-dimensional structure of some of these proteins allowed Dunwell and Gane (1998) to refine the search for other similar proteins on the grounds of a structure-based alignment. The search eventually ended with the proposal of the existence of the widespread cupin superfamily, characterized by the presence of the cupin domains, marked by two conserved motifs, each comprising two β-strands, intervened by a loop of variable length and sequence (Dunwell, 1998; Woo et al., 2000).

From a structural point of view, cupins may be grouped in monocupins, bicupins, and multicupins, according to the number of cupin domains in their structure (for reviews see Dunwell et al., 2001, 2004). The functional classification is not so simple, as the diversity of roles assigned to the members of the cupin superfamily is continuously increasing, and at least 18 functional classes are now accepted. Apart from those of the founding members of the superfamily, these classes comprise, among others, enzymes, of which the dioxygenases are the most representative ones, auxin-binding proteins, and some nuclear or DNA-binding...
proteins of both prokaryotic and eukaryotic origin (for a review see Dunwell et al., 2004)). The existence of divergent roles in the structurally-related members of the superfamily poses some interesting questions. For instance, a soybean bicupin that was originally described as a sucrose-binding protein (SBP) (Grimes et al., 1992), was shown to share similarity with the functionally unrelated seed storage proteins (Overvoorde et al., 1997).

It has previously been described that p16, a small protein partitioned between the chromatin and protein bodies of pea (Pisum sativum) seeds, is encoded by psp54, a large gene whose mRNA would be able to yield a polypeptide of a 54.4 kDa and which has been named p54 (Castillo et al., 2000). Although proof of the actual presence of p54 in peas has not been obtained, it was assumed that p16 resulted from its post-translational processing. Based on the identity (60%) between SBP and p54 (Castillo et al., 2000), Heim et al. (2001) classified p54 as an SBP-like protein. These authors described a 52 kDa polypeptide of Vicia faba, which they assumed to be homologous to SBP and termed accordingly VfSBPL.

In the present paper it is shown that p54 actually exists in pea and that it is processed to yield p16 and a larger polypeptide, p38, corresponding to the N-terminal region of p54, which coincides with p16 in the subcellular localization. The sequence analysis and structural modelling of p54 indicate that p54 is a bicupin in agreement with the previous suggestion of Dunwell et al. (2000). The cloning of the psp54 promoter, which contains an ABA-responsive element as well as a stress-responsive element, is described further. This agrees with the finding that the psp54 gene is regulated by abscisic acid as well as by different stress conditions. All these properties of the gene and of its products suggest that the multifunctional bicupin p54 and its analogues present in several members (Vicia, Pisum, and Lens) of the tribe Vicieae cannot be properly classified as an SBP-like protein. These results add a novel function to a member of the cupin superfamily, namely the role of a precursor of multifunctional proteins present in nuclei.

**Materials and methods**

**Plant materials and treatments**

The culture of peas (Pisum sativum, cv. Lincoln), obtaining all plant materials, treatment with ABA, and desiccation of plant tissues were carried out as previously described by Castillo et al. (2000).

**Preparation of His-tagged p38 and antibody raising**

cDNA, coding for amino acids 1–371 of p54, was obtained by PCR amplification using psp54 cDNA (Castillo et al., 2000) as the template. The primers used were: 5’-CCCCCTCGAGATGGCCGTATTAAACCAAG (forward) and 5’-CCCCCTCGAGTCAGTTTGAGCATAGGATGACCTTGAGTGCC (reverse primer). The PCR product was digested at the Xhol site, located at the end of the primers, and cloned in phase in the Xhol site of plasmid pRSETA. In this study, no expression of the recombinant plasmid took place, irrespective of the E. coli strain and culture conditions used. An attempt was made to eliminate the N-terminal region containing the leader peptide. To do this, the recombinant plasmid was digested with BamHI to eliminate the first 93 nucleotides and then with Xhol. The resulting fragment was then cloned into the site BamHI-Xhol of plasmid pRSETB. The construct was used to transform E. coli BL21 (DE3)pLysS. After inducing with IPTG, his$_R$p38 (ps) was purified by affinity chromatography on Ni-NTA (Qiagen, CA, USA), eluted with 300 mM imidazole. His$_R$tagged p38 was dialysed against PBS buffer and used to immunize rabbits as described by Castillo et al. (2000).

**Cloning of the 5’ flank of the psp54 gene**

The 5’-flanking region of the psp54 gene was cloned by genome walking, using the Universal Genome Walker kit (Clontech Laboratories, Palo Alto, CA, USA) and following the instructions of the manufacturer. Genomic DNA was prepared from pea leaves (Michaels et al., 1994) and digested with DraI. The blunt-end fragments generated were then ligated to a Genome Walker adaptor to construct a Genome Walker library. The c. 1 kb fragment putatively containing the psp54 promoter was isolated by two successive PCR-based DNA walkings in the Genomic Walker library. The primary PCR was performed with a gene-specific primer (5’-CACACCTGCT CATATCCGTTGTC-3’) and the outer adaptor primer (AP1), using the GeneAmp PCR system 9700 (Perkin Elmer-Applied Biosystems). The amplification started at 94 °C for 5 min, followed by 30 cycles of 94 °C for 1 min, 60 °C for 1 min, and 72 °C for 4 min. The final extension was done at 72 °C for 7 min. The diluted primary products served as template for the secondary PCR with a nested gene-specific primer (5’-CTGCCCTAGGGCAGATTGTAAGACGAT-3’) and the nested adaptor primer AP2. The secondary PCR products were analysed by agarose gel electrophoresis. The major band was extracted and purified from the gel with Qiaquick gel extraction kit (Qiagen), cloned into pGEMT vector (Promega, Madison, WI, USA) and sequenced. The search for putative cis elements in the obtained sequence was carried out with PLACE (http://www.dna.afric.go.jp/PLACE/signalscan.html).

**Other biochemical methods**

Total protein extractions were done according to Rivin and Gruth (1991). Other biochemical and molecular biology methods, including the detection of in vivo cross-linked proteins via disulphide bridges, were carried out as previously described (Castillo et al., 2000). For western blotting, the p16 and p38 antisera were used at dilutions of 1:500 and 1:1000, respectively; the membranes were probed with a secondary, phosphatase-conjugated, goat anti-rabbit IgG.

**Informatic analysis of p54 sequence and prediction of its tertiary structure**

Alignment of the p54 sequence with those of VfSBPL, a homologue from Vicia faba (Heim et al., 2001), SBP and the storage proteins β-conglycinin and phaseolin was done with the DNAAMAN programme (Lynnon BioSoft). The 3D structure of the p54 protein was modelled using SwissModel and the SwissPDBViewer protein modelling programme version 3.7 (Guex and Peitsch, 1997), which are both available from the ExPaSy website (http://www.expasy.ch/spdwr).

**Results**

**Products of p54 processing**

In a previous paper (Castillo et al., 2000), it was suggested that, probably after removal of its leader peptide, the putative precursor p54 would be processed to yield p16.
Fig. 1. (A) Total proteins from pea embryonic axes were separated by SDS-PAGE and blotted. Equivalent blots were probed with p38 antibody (lane 1), with p16 antibody (lane 2) or with peroxidase-conjugated concanavalin-A to detect glycosylated proteins (lane 3). (B) Western blots of total proteins from ungerminated pea embryonic axes, with p16 antibody (lane 2) or with peroxidase-conjugated p38 antibody (lane 1), and pea (lane 3) and from whole ungerminated seeds of soybean (lane 2). The blots were probed with p38 and p16 antibodies. The mobility of molecular weight markers is given between both panels.

and another polypeptide, of about 340 residues, from the N-terminal region of p54. The question was asked whether this polypeptide, which will henceforth be referred to as p38 according to its calculated molecular weight, actually existed in pea embryos. To address this issue, the recombinant p38 and the antibodies against it have been obtained, as described in the Materials and methods. It should be mentioned that all these procedures were only possible when they were started from a cDNA in which the nucleotides corresponding to the first 31 amino acids (containing the leader sequence) had been removed. Any other attempt to produce recombinant p38 was unsuccessful in all the E. coli strains and culture conditions assayed.

The specificity of p38 antibody was first assayed with recombinant p54, p38 and p16. The antibody reacted with p38 and p54, but not with p16. On the other hand, p38 was not recognized by the p16 antibody, which was able to react to some extent with the precursor p54 (results not shown). Then the p38 antibody was used to probe a western blot containing total proteins from ungerminated pea embryonic axes. The results given in Fig. 1A show that a clear signal in the region of p38 and p54 was developed, while, when a similar blot was probed to the p16 antibody, the band of p16 was clearly seen and, to a lesser extent due to the reactivity of this antibody, the p54 band also appeared. These results indicate that p54 is actually present in pea and that its processing yields not only p16, but also p38. It should be noted that the mobility displayed by p16, p38, and p54 in SDS polyacrylamide gels correspond to molecular weights higher than the actual ones (Fig. 1A). To study the subcellular localization of p38 and p54, a subcellular fractionation experiment was carried out as previously described for p16 (Castillo et al., 2002) and it was found that both the precursor and the mature p38 were present in the same fractions as p16, i.e. the nuclear, the mitochondrial, and the microsomal fractions. No trace of p38 or p54 was found in the 100 000 g soluble fraction.

It has previously been suggested that p54 is processed by an asparagine-specific proteinase to yield, finally, p16 and p38. In the case of legumins and of some lectins, proteolytic cleavage occurs at a specific Asn residue lying between two cysteinyl residues that form a disulphide bridge, so the mature protein retains both fragments of the limited proteolysis as two chains linked by the disulphide bridge (for a review see Müntz and Shutov, 2002). It may be queried whether a similar situation occurs in p54 processing. If it were, p16 and p38 would only represent two types of polypeptide chains present in the mature protein and their presence in pea tissues would then be an artefact resulting from the reduction of the disulphide bridge. Previous results (see Fig. 7 of Castillo et al., 2000) seem to discard this possibility. Nevertheless, to explore this question further an additional experiment was carried out. A protein extract from nuclei purified from pea embryos was prepared in the absence of reducing agents and in the presence of iodoacetamide to block thiol groups. Under these conditions, disulphide bridges existing in vivo are preserved and no additional bridges can be artefactually formed during the preparation of the extract. The latter was submitted to two-dimensional electrophoresis. The first dimension was run in the absence of β-mercaptoethanol. The gel was then incubated with this reducing reagent, which was also added to the second dimension buffer. The gel was finally treated sequentially with antibodies against p16 and p38 (Fig. 2A). Spots outside the diagonal ought to correspond to p16 and/or p38 that participate in disulphide bridges, and their mobility in the second dimension, together with the antibody specificity, allowed the discrimination between p16 and p38. It is obvious that p16, but not p38, participates in disulphide bridges in vivo. The products identified were those previously reported, i.e. p16 homodimers and p16-H3 heterodimers. When a protein extract in which iodoacetamide was omitted, was treated in the same way, some p38-containing products were obtained, but they are obviously artefacts due to the oxidation of cysteine thiols during manipulation. Moreover, p16 does not participate in these products (Fig. 2B). Therefore, it can be concluded that p16 and p38 are not linked in vivo and that they are present in nuclei as individual polypeptides rather than as chains of a larger protein.

Once it had been ascertained that p38 and p54 actually exist in embryo axes, both polypeptides were recovered by preparative electrophoresis and an attempt was made to determine, by microsequencing, whether the leader peptide was present in them, but no sequence was obtained in any case. It is not yet known whether this negative result means that both polypeptides were blocked at their N-terminal ends.
Fig. 2. Two-dimensional electrophoretical detection of nuclear proteins cross-linked to p16 and p38 by disulphide bridges. (A) Nuclei from pea embryonic axes were isolated in the absence of reducing agents in a iodoacetamide-containing buffer to preserve the disulphide bridges existing in vivo. The nuclear proteins were then resolved by SDS-polyacrylamide gel electrophoresis in the absence of β-mercaptoethanol (first dimension, I). The gel strip was incubated in β-mercaptoethanol and the second dimension (II) was run in the same electrophoresis system. The gel was then western-blotted and developed sequentially with antisera against p16 and p38. The spots outside the diagonal (1 and 2), whose mobility in the second dimension coincides with that of p16, correspond to in vivo crosslinked p16-containing products. They were identified, respectively, with heterodimers p16-H3 and homodimers of p16 (Castillo et al., 2000). (B) Experiment similar to that of (A), in which iodoacetamide was omitted from the isolating buffer. Apart from spots 1 and 2, some p38-containing products, marked by stars, are now seen. In both panels the mobility of p54, p38, and p16 in the second dimension is indicated on the right margin, and spots 3 and 4 correspond to in vivo unmodified free p16 and p38.

It has previously been mentioned that, among the species analysed, only lentil and broad bean seeds possess a protein similar in size to p16 and able to cross-react with its antibody. On the contrary, no soybean protein cross-reacted with p16 antibody in spite of the similarity in mobility, V. faba embryos not only contain a ∼54 kDa polypeptide immunologically related to P. sativum p54, but also that this polypeptide is processed in a manner similar to that of the pea precursor. What the causes may be for the discrepancy between this study’s results and those of Heim et al. (2001) are not known. On the other hand, no related polypeptide or processing products are found in soybean seed proteins, although all the lanes were loaded with equivalent amounts of protein.

Analysis of p54 structure

It has previously been shown that psp54, the gene encoding p16, shares similarity with those of the vicilins (Castillo et al., 2000) and, based on sequence analysis, Dunwell et al. (2000) classified p54 as a cupin. Figure 3 shows the alignment of sequences from the precursors of p54, VfSBPL, SBP, and two vicilins, namely Glycine max β-conglycinin and Phaseolus vulgaris phaseolin, whose 3D structures are known (Lawrence et al., 1994; Maruyama et al., 2001). There is a clear homology that extends over the two cupin domains of the known bicupins, showing that p54 may also be considered as a bicupin. The p54 N-terminal cupin domain contains the two conserved cupin motifs. Motif 1 (P-X-13–14-G), present in the region of strands C and D of conglycinin, is strictly conserved in all the proteins shown in Fig. 3, whereas in motif 2 (GX,IPXG), which corresponds to conglycinin strands G and H, a valine residue substitutes for the isoleucine in p54 and in VfSBPL. In the C-terminal cupin domain, the proline residue of motif 1, which is well conserved among the vicilins (Shutov et al., 1995), is replaced by an isoleucine in p54 (residue 339), as well as in VfSBPL and SBP. Apart from this divergence, some other differences make the C-terminal cupin domain less conserved than the N-terminal one when the alignment includes the vicilins, but the homology is clear if only p54, VfSBPL, and SBP are compared (Fig. 3). It is noteworthy that the overall homology between p54 and VfSBPL is as much as 85%, and that this figure rises to 94% when the conservative residue changes were considered. VfSBPL shares only 58% overall homology with Glycine max SBP, so the Vicia faba protein is closer to pea p54 than to soybean SBP.

Figure 3 shows that p54, VfSBPL, and SBP possess a glycosylation consensus motif (Asn330–Ile-Thr). To check whether the motif is functional in p54, as it is in some storage proteins, a third lane of the blot used for the western analysis of Fig. 1A was treated with peroxidase-conjugated concanavalin A, a procedure commonly used to detect glycosylated proteins. Although some glycosylated proteins were detected in the pea embryo extract (see lane 3 of Fig. 1A), neither of them corresponded to p54, p38, or p16. The bands in the upper part of the gel move faster than p38, and the band above p16 does not react with the
antibody raised against this protein, which, in addition, does not contain any glycosylation consensus sequence. Therefore, it can be concluded that neither of the proteins p54, p38, and p16 is glycosylated.

It was next asked whether the sequence homology results in a possible 3D structural similarity. To address this issue, the structure of p54 was modelled on the basis of the known structure of *Glycine max* β-conglycinin (Fig. 4A). The model can be built for the residues 87–454 of p54 (Fig. 4B), therefore comprising the two cupin domains. The first p54 residues cannot be obviously modelled because they correspond to a hydrophilic N-terminal extension absent from the model β-conglycinin 3D structure. The structural modelling resulted in a good fit, in which no severe errors affecting the polypeptide backbone were found and only minor errors affecting side chains were detected, especially in the helix-containing loops flanking the two cupin domains (Fig. 4B). Mature p16 starts at serine 372 (Castillo et al., 2000). The site Asn 371-Ser372 is conserved in VfSBPL, which explains the results found here on its processing (Fig. 1B). Interestingly, in SBP a histidyl residue substitutes for Asn 371 (Fig. 3) and this would agree with the fact that processing of SBP had not been observed (see Fig. 4 of Overvoorde et al., 1997).

**Time-course and organ-specificity of psp54 expression**

The structural analyses depicted in Figs 2 and 3 show that p54 is homologous to vicilins, as well as to SBP. On the other hand, previous reports from this laboratory clearly indicate that, at least p16, is a nuclear, chromatin-bound protein (Castillo et al., 2000, 2002). As the expression patterns of the genes coding for pea vicilins and for SBP show some characteristic features, some properties of the expression of psp54 were examined in an attempt to obtain some additional clue as to the function of p54 and/or its processing products p16 and p38.

First, the time-course of gene expression in the embryonic organs, axes and cotyledons, was studied. Figure 5 shows that ungerminated embryonic axes contain p54 mRNA, which is actually translated as the three gene products, namely p54, p38, and p16, are present.

**Fig. 3.** Alignment of p54 with related plant proteins. The primary structure of the precursors for *Vicia faba* VfSBPL (sbpVF, accession No. AI292221), *Glycine max* SBP (sbpGm, accession No. Q04672), the β-chain of *Glycine max* β-conglycinin (accession No. P25974), and *Phaseolus vulgaris* phaseolin (accession No. P02853) are shown. The residues identical in all five proteins are boxed and those identical at least in two of the first three proteins are shown over a grey background. Below the sequences, the secondary structure elements of β-conglycinin are shown as arrows for β-strand and cylinders for α-helix. The β-strands are designated after the usual convention for cupin domains. The curved arrow points to the processing site of p54.
mRNA disappears soon after imbibition, but the proteins are relatively stable and they remain at a high concentration until the end of the germination period, when the radicle protrudes through the testa tissues. At that time, the levels of the proteins begin to decline, although p16 seems to be somewhat more stable. The decay of p54 mRNA in cotyledons is slower than in the axes and, although somewhat degraded, it is clearly observed at 12 HAI (hours after imbibition). Curiously enough, the increase of mRNA in cotyledons, detected by northern blot from 48 HAI to the end of the experiment (Fig. 5A), shows a peculiar re-expression of the psp54 gene in these organs, which does not occur in vicilin and SBP genes. The western blot of Fig. 5B shows that the level of proteins remains high in the cotyledons throughout the whole period and, in the case of p16, even increases, although some degradation occurs as shown by the appearance of fast moving bands. It is also worth mentioning that mature p16 was also found in pods and seed coats (Fig. 5B), two tissues of maternal origin that have a role in embryo protection and nutrition, but p54 and p38 are present to a much lesser extent. The senescence of these tissues starts at the end of embryogenesis, and the stability of the RNAs is severely reduced. This may be the cause by which intact RNA could not be recovered and the northern analysis was not carried out. The above results point to some functional peculiarities in the expression of psp54 when compared to other related genes. First, the mRNAs of pea vicilins (Chandler et al., 1984) and of SBP (Overvoorde et al., 1997) do not accumulate in the dry seeds. On the other hand, the expression of pea vicilin genes does not result in the presence of the proteins in maternal tissues.

It has been suggested that the expression of the psp54 gene is regulated by the hydric state of the plant (Castillo et al., 2000). In this context it would seem reasonable to think that the re-expression of psp54 in cotyledons (Fig. 5A) results from some hydric stress during germination. To check whether this assumption was right, the water content of embryonic axes and cotyledons was determined. During imbibition at 4 °C, the water gained resulted in a 60% increase in the weight of the seed (Fig. 6). The seeds were then transferred to 28 °C to start germination (HAI=0). Between 24 HAI, namely when the radicle emerged at the end of germination, and 36 HAI embryonic axes gained a further 30% of water, but the cotyledons did not lose water, whose content remained constant to the end of the experiment (Fig. 6). Therefore, hydric stress cannot be the cause of the resumption of psp54 expression in cotyledons.

Another possible cause for psp54 re-expression may be the deterioration of cotyledon tissues during senescence. To check this possibility, it was next examined whether the gene may be induced by wounding. Wounds were inflicted by puncturing the plant material three times with a hypodermic needle. This experiment was carried out with 18 HAI embryonic axes, which do not naturally express the gene (Fig. 4), and found that an 1.4 kb mRNA corresponding to psp54 readily accumulates as detected by northern analysis. In many instances signal transduction mechanisms after wounding involve ethylene, but in this study, treatment of 18 HAI embryonic axes with 0.14 ppm ethylene for 24 h did not result in psp54 expression (data not shown).

ABA and hydric stress are capable of inducing psp54 during germination (Castillo et al., 2000) and the question was next asked whether this induction also took place in later phases of plant development. Figure 7A shows that

---

Fig. 4. Modelling of the 3D structure of p54. (A) Structure of the β-chain of Glycine max β-conglycinin (PDB 1IPJ), corresponding to residues 31–414 of the sequence given in Fig. 2. (B) Modelled structure for the residues 87–454 of p54. The drawings have been made with the SwissPDBViewer protein modelling programme version 3.7 (Guex and Peitsch, 1997).

Fig. 5. Temporal changes in the expression of psp54 gene in different pea embryo and seedling tissues. (A) Northern blot analysis of the 1.4 kb psp54 mRNA. The upper panel shows the major rRNA species as loading control. (B) Western blot analysis of the mature proteins p16 and p38 and of the precursor p54, probed with the antibodies raised against p16 and p38. In the case of seed pods (P) and coats (SC) only the western blots are shown. The time scale is given in h after imbibition (HAI).
treatment of post-germination 72 HAI seedlings with ABA eventually resulted in the appearance of the 1.4 kb p54 mRNA, although to a lesser extent than in the germinating 18 HAI embryo axes. In the absence of ABA the gene typically is not expressed in either stages (Fig. 4). On the other hand, psp54 is absolutely insensitive to ABA treatment in adult organs (Fig. 7B), so it is obvious that the responsiveness to ABA diminishes with age. By contrast, desiccation induces the gene not only in seedlings (Fig. 7A) but also in adult organs (Fig. 7B). Moreover, the response to desiccation in 72 HAI seedlings is even faster than in 18 HAI embryonic axes (Fig. 7A). A possible explanation for the latter result may be that the testa tissues, still intact in 18 HAI seedlings, protect the inner tissues from water loss, while in 72 HAI seedlings, after the protrusion of the radicle, the rupture of the testa leaves both the epicotyl and hypocotyl more exposed to changes in ambient humidity.

Other aspects of p54 mRNA translation

The psp54 gene may be induced by hydric stress in the post-germination stages of plant development. However, in western blot experiments neither the primary polypeptide p54 nor the mature proteins (p16 and p38) were detected in any of the plant materials examined (72 HAI seedlings and adult organs, either during the stress, or when the plant was recovering from it). This suggests the existence of a distinct regulatory process at the translational level.

One can wonder whether translation is also dissociated from transcription when the psp54 gene is induced in situations in which it is not naturally expressed. To address this issue, 18 HAI embryonic axes, which do not express the psp54 gene but still have a high level of the mature proteins p16 and p38 (Fig. 5), were treated with ABA and analysed for the presence of p54 mRNA and of mature p16. Figure 8 shows that treatment with ABA results in the presence of psp54 mRNA as early as 6 h after applying the fitohormone and that, after 48 h, the steady-state level of psp54 mRNA is still high. Western blot analysis revealed that, while p16 diminishes in a time-dependent manner in the control seedlings, the ABA treatment results in an increase in the persistence of the protein. The levels of p16 do not show a clear relationship with those of mRNA, so it remains to be determined whether the prolonged presence of the protein is solely due to the transcriptional activation or whether ABA treatment also has some effect on protein stability. It has systematically been observed that in seedlings of different ages the level of p16 was never increased after the ABA-induced increase in the steady-state level of p54 mRNA. All these facts point to the possibility that ABA had some influence on p16 stability as well as on the transcription of the gene.

Presence of ABA and stress responsive elements in the psp54 promoter

1127 bp of the 5’ flank of the psp54 gene have been cloned, which contains the promoter of the gene (GenBank accession No. AY822015). Several cis elements were found upon analysing the sequence. First, a putative TATA box (sequence TATAAAA) is present at −78 relative to the initiation codon. Putative ABRE and STRE sequences were also found at −146 and −125, respectively. At −120 the sequence TGATGC corresponds to a consensus RY motif (CATGCA in the complementary strand) often found in seed-specific genes (Ezcurra et al., 2000). The putative ABRE, GACACGTA, exactly matches the sequence of a functional element of response to ABA found in rice (Hattori et al., 1995). The putative STRE has the sequence CCCCT, which has been characterized in yeast as a functional element of response to stress (Martinez-Pastor et al., 1996; Garay-Arroyo and Covarrubias, 1999) and it has also been found in some plant promoters (see, for instance, Perez-Espinosa et al., 2001). It is not yet known whether
these three elements are functional in the psp54 promoter, but if they were, the simultaneous presence of these separate regulatory elements would explain why the gene maintains its capacity to respond to dehydration along the whole life cycle of the plant, while the sensitivity to ABA is restricted to the seed and early developmental stages.

Discussion

p54 is processed to give two polypeptides

It has already been shown that p16, a chromatin-associated protein of pea seeds, is coded by the 3' region of psp54, which is transcribed to a single mRNA and presumably translated to give a large precursor of 54.4 kDa (Castillo et al., 2000). The results presented in this paper demonstrate that the precursor polypeptide of 54 kDa, and a second processing product, p38, corresponding to the N-terminal region of the precursor, are also present in peas. The data of Fig. 2 show that p16 and p38 are not subunits of a larger protein, linked by disulphide bridges, as occurs in some lectins and legumins (Münzt and Shutov, 2002). It is interesting to note that some pea genes encoding vicilin precursors are also known to yield several protein products (Gatehouse et al., 1983; Lycett et al., 1983). Thus, the pea psp54 gene adds to the list of plant genes that encode two or more proteins, although the appearance of two gene products does not mean that psp54 represents a dicistronic locus, whose existence has been suspected, but not proved in plants (Blumenthal, 1998).

p54 is a bicupin

The results presented in this paper strongly support the idea that p54 is a bicupin. Although the C-terminal cupin domain is less conserved than the N-terminal one, both sequence alignment (Fig. 3) and 3D structure modelling (Fig. 4) support the above assumption. The relative peculiarity of the C-terminal domain may explain the failure to model this region (Fig. 4). It is worth noting that, in spite of the similarity of both domains at a 3D level, the antibodies raised against p16 and p38 do not cross-react with each other. Obviously, the structural similarity between the two cupin domains of p54 does not imply that they share the same antigenic determinants.

The precursor p54 is obviously related to vicilins, as pointed out some time ago by Castillo et al. (2000), but a particular similarity among p54, VfSBPL, and SBP is obvious from Fig. 3. There is a need to emphasize that the Vicia faba protein shares more identity with p54 than with Glycine max SBP, so the 52 kDa protein of V. faba, would be aptly described as p54-like, rather than as SBP-like. In this connection, it can be emphasized that polypeptides related to p16, p38, and to their precursor p54, both in size and immunological properties, do actually exist in V. faba extracts (Fig. 1B). There is every reason to think that p54 and VfSBPL actually are homologous proteins. This reasoning is not only based on structural similarity but, in particular, on functional considerations. For instance, both p54 and VfSBPL do possess nuclear localization signals, while SBP does not. On the other hand, no property of either p54 or its processing products is related to sucrose binding and/or transport. In this context, it should be mentioned that Tegeder et al. (1999) have used full-length cDNA from G. max SBP as a probe to analyse P. sativum RNA and found a positive signal of a 1.7 kb transcript in cotyledon tissues enriched in epidermal transfer cells, but not in seed coats. Accordingly, the antibody raised against G. max SBP (Grimes et al., 1992) only cross-reacted with a protein of M, 62 000 present in the same cotyledon tissues (Tegeder et al., 1999) but, again, no signal was detected in seed coats. In view of the data given in this paper, it is thought that the protein detected by Tegeder et al. (1999) represents a bona fide pea SBP, clearly different from our p54.

Despite the above statements it is obvious that p54 proteins from P. sativum and from V. faba are related to SBP proteins, as they are, albeit to a lesser extent, to seed storage vicilins. The unrooted dendogram proposed by Elmer et al. (2003) places p54 and VfSBPL in a common twig within the SBP branch, and well apart from the vicilin branches. Nevertheless, there are several interesting functional differences among all these bicupins, namely p54, SBPs, and vicilins. For instance, as far as is known, the expression of pea vicilins in maternal organs of seeds has not been reported to date, while p54 and its processing products are found in both seed pods and coats (Fig. 5), as is the

Fig. 8. Effect of prolonged treatments with abscisic acid on the expression of psp54 gene in 18 HAI pea seedlings. (A) Northern analysis of the 1.4 kb mRNA. The loading control is given as in previous figures. (B) Western blot analysis of mature p16. The duration of treatment is given over each lane.
VfSBPL protein (Heim et al., 2001). The patterns of psp54 expression, which will be discussed below, add more reasons in support of the statement that p54 proteins represent a functionally unique group of proteins within the cupin superfamily. This functional approach to assess the homology among structurally-related cupins has already been used by other authors (see, for instance, Elmer et al., 2003).

The temporal appearance of p16, p38, and of psp54 mRNA

The transcription of the psp54 gene, as shown by northern analysis, is limited to the tissues containing p16 and p38 and vice versa. Therefore, processing of p54 seems to occur from newly translated precursor rather than from stored p54. Nevertheless, the gene may be expressed under certain non-physiological conditions (see below), without mRNA translation.

It has previously been found that mRNA is abundant in imbibed, non-germinated embryonic axes and it is no longer present at 6 HAI, whereas the level of p16 in chromatin steadily decreases until 18 HAI, and is no longer found after that time (Castillo et al., 2000). The data of Fig. 5 indicate that total p16 remains visible for a longer period. As the protein is also found in other organelles (Castillo et al., 2002), the present results show that p16 is more stable in these organelles than in nuclei. The time-course of whole p16 in cotyledons is more difficult to analyse due to the resumption of psp54 expression (Fig. 5). In axes, the half-life of whole p38 seems to be shorter than that of whole p16 (Fig. 5), and comparable to that of nuclear p16.

Regulation of psp54 expression

The data presented here reinforce the previous idea that expression of the psp54 gene responds to ABA, and that the gene is also induced in several situations of stress. For instance, hydric stress seems an important factor, because the expression of the gene ceases in rehydration periods. Moreover, desiccation also induces the gene in seedlings as well as in adult organs (Fig. 7). Sorbitol-induced osmotic stress causes the expression of psp54 in seedlings (Castillo et al., 2000) as does wounding. This latter effect occurs through an ethylene-independent mechanism and it has to be noted that the injury-induced expression of a family of proteinase inhibitor genes from potato and tomato also occurs in response to ABA and jasmonic acid in the absence of ethylene mediation (Peña-Cortés and Willmitzer, 1995). Some other stress situations, for instance, exposure to extreme temperature, may cause the induction of the gene in 18 HAI seedlings, but, compared with the heat-shock genes, the induction only takes place after 24 h (J Castillo and MI Rodrigo, unpublished results).

The analysis of the promoter sequence offers an explanation to the above considerations. First, the existence of an ABRE exactly coinciding with a functional one in rice (Hattori et al., 1995), would explain the induction by ABA. The RY motif may also be functional as a coupling element, which may allow the physical interaction of the transcription factors equivalent to Arabidopsis AB13 and AB15, resulting in a synergistic activation of the transcription of many ABA-inducible genes (Nambara and Marion-Poll, 2003). In this line, failure of ABA to induce the psp54 gene during post-germinative periods of plant development may result either from the absence of these transcriptional factors after germination or from changes in the chromatin structure that render the cis elements inaccessible to the trans factors. It should be noted that early data from this laboratory showed that subtle yet general changes occur in pea chromatin structure during germination (Ull and Franco, 1986). On the other hand, the putative presence of a STRE in the psp54 promoter may explain why the gene is induced in response to desiccation (Fig. 7).

The data given in Fig. 8 point to the possibility that ABA has an effect in the mRNA translation and/or the stability of p16. An effect of ABA in altering the stability of transcriptional factors has already been described (Lopez-Molina et al., 2001, 2003). In summary, the available data indicate that the regulation of psp54 expression is very complex, and it may occur both at the transcriptional and translational levels depending on tissues, developmental state, and hydric stress.

Is p54 the founder member of a novel subclass of cupins?

The mature products of p54 processing are present under physiological conditions in cotyledons and embryonic axes, but the proteins are not found in adult plant organs. This is a feature shared by storage proteins, but the latter, in contrast to p16 and p38, are not usually found in tissues of maternal origin such as the testa and pods (Evans et al., 1984). Curiously enough, p16 and p38 are localized to the same subcellular compartments, i.e., nuclei and other organelles present in the mitochondrial and/or microsomal fractions, probably protein bodies. The presence of p16 and p38 in nuclei may be topogenically related to the existence of nuclear localization signals in the sequence of both mature proteins. The precursor p54 also contains a leader peptide (Castillo et al., 2000), which may explain why their processing products are also found in other organelles, probably protein bodies.

Proof as to the association of p16 to chromatin has already been obtained, but it is not yet known whether nuclear p38 is also bound to chromatin. Some time ago, Chiatante et al. (1995) described a pea nuclear protein, QP47, with an apparent Mr identical to that of p38, i.e. 47 000. The expression pattern of QP47 coincides with that of p38 as does the subcellular localization. Moreover, a protein of size similar to QP47 and immunologically
related to it was found in *Vicia faba* (Chiatante et al., 1995). All these coincidences suggest that QP47 and p38 may be the same protein. It is proposed that one of the roles of p16 may be the protection of seed chromatin against desiccation. By contrast, QP47, though nuclear, is not a chromosomal protein, but it may also have a protective role against desiccation (Chiatante *et al*., 1995). If p38 and QP47 actually are the same protein, then p16 and p38 would possess complementary roles in the protection of nuclear components against desiccation and this would explain why they are synthesized from a single precursor. The proteolytic processing of p54 would, then, be of functional significance because it yields equivalent amounts of the two final products.

It may be that the mature proteins also play a secondary role as storage proteins, but taking into account their expression pattern (see above), they do not seem typical seed storage proteins. At any rate, p16 is a multifunctional protein and the data on p38 point to the same conclusion. The expression pattern of the *psp54* gene points, in turn, to the idea that the function played by p16 in chromatin is related to protecting the chromatin against stress injuries.

All the above considerations suggest that p54 possesses a hitherto undescribed role within the cupin superfamily, namely, to serve as a precursor of two stress-related multifunctional seed proteins, which, among other functions, play a function in nuclear processes.

Acknowledgements

This work was supported by grant BMC2001-2868 from the Ministerio de Ciencia y Tecnología, Spain. We are very indebted to Dr J Salgado for his invaluable aid in protein modelling and to Professor JV Castell for facilitating the raising of p38 antibodies.

References


Lawrence MC, Izard T, Beuchat M, Blagrove RJ, Colman PM. 1994. Structure of phaeolin at 2.2 Å resolution. Implications for a


