Journal of Plant Physiology 167 (2010) 1101–1108

Contents lists available at ScienceDirect

Journal of Plant Physiology

journal homepage: www.elsevier.de/jplph

A distinct subfamily of papain-like cystein proteinases regulated by senescence and stresses in *Glycine max*

Belén Esteban-García, José Antonio Garrido-Cárdenas, Diego López Alonso, Federico García-Maroto*

Grupo de Biotecnología de Productos Naturales (BIO-279), Facultad de Ciencias Experimentales, Universidad de Almería, 04120 Almería, Spain

ARTICLE INFO

Article history:
Received 22 January 2010
Received in revised form 3 March 2010
Accepted 3 March 2010

Keywords:
Nonsense-mediated mRNA decay
Papain-like protease
Plant stress
Senescence
Soybean

ABSTRACT

GMCP3 encodes a cystein proteinase of *Glycine max* belonging to the papain-like family (C1A in MEROPS database) that was previously found to be involved in the mobilization of protein reserves during seed germination. Here, we report that *GMCP3* is induced by senescence and diverse stresses in non-seed tissues, thus indicating a more general function in plants. Cladistic analysis of papain-like proteins of plants indicated that GMCP3, along with related proteases of other species, belongs to a distinct new group within the C1A family, which can also be distinguished by the four-exon structure of the gene. We also describe the genomic organization of *GMCP3* revealing the presence of two closely related copies that are transcriptionally regulated in a similar way, although only one appears to be functional.

Introduction

Proteinase enzymes catalyze either limited site-specific or complete hydrolysis on their substrates, acting in coordination with other post-translational modifications to regulate turnover, direction and activity of proteins. Among these, cysteine (sulfhydryl) proteinases (CPs) are currently receiving much attention due to their involvement in very diverse aspects of plant physiology and development. They play a role in protein storage and mobilization, senescence, programmed cell death, and are regulated by various types of environmental stresses (revised in Grudkowska and Zagdanska, 2004; Schaller, 2004). Involvement of proteolytic enzymes, including CPs, in the ubiquitin (Ub)/26S proteasome proteolytic pathway is shedding light on the connection of these components to many aspects of plant biology including senescence, cell-cycle, embryogenesis, flower morphogenesis, hormone signaling, and defense (Avci et al., 2008; Vierstra, 2009).

The most investigated CPs are members of the papain-like group (C1A subfamily in the MEROPS catalogue http://merops.sanger.ac.uk). Along with legumains (C13 family), papain-like enzymes are the major CP in plants. C1A proteases are synthesized as relatively inactive precursors (40–50 kDa) containing an N-terminal signal peptide responsible for endoplasmic reticulum (ER) translocation, a variable pro-sequence with regulatory functions in folding and activity, and a 22–35 kDa mature enzyme. Protease activation takes place through limited intra- or inter-molecular proteolysis (Wiederanders, 2003). The structure of proteins from the C1 family exhibits the typical papain-like fold (Drenth et al., 1968), composed of two domains, an α-helix-rich (L) domain and a β-barrel-like (R) domain, separated by a groove containing the catalytic dyad formed by residues Cys25 and His159 (papain numbering), each one on each domain. Two other residues also seem to be important Gln19 aids the formation of the 'oxyanion hole' and Asn175 is important for proper orientation of the His imidazolium ring (Beers et al., 2004).

A number of genes encoding C1A peptidases have been identified in diverse plant species. In *Arabidopsis thaliana*, for instance, the papain-like family comprises 30 genes that have been classified into eight distinct groups (C1A-1 to C1A-8) according to their gene structures and phylogenetic relationships (Beers et al., 2004). Although information about the roles of individual genes is still fragmentary in this plant, the available data indicate that they are regulated by senescence (SAG2 and SAG12, Hensel et al., 1993; Gan and Amasino, 1995), both drought stress and senescence (RD19 and RD21, Koizumi et al., 1993), programmed cell death (CEP genes, Gietl and Schmid, 2001), and during xylogenesis (XCP genes, Funk et al., 2002). More recently, involvement of RD19 in pathogen resistance response has been demonstrated (Bernoux et al., 2008).

GMCP3 is a papain-like protease previously identified in *Glycine max* (Nong et al., 1995). Preliminary gene expression studies showed that the *GMCP3* transcript is stored in the dry seeds, and fur-

Abbreviations: 4-MU, 4-methylumbelliferone; ABA, abscisic acid; CP, cystein proteinase; CSPD, Disodium 3-(4-methoxyspiro {1,2-dioxetane-3,2′-(5′-chloro)tricyclo [3.3.1.13,7]decan}-4-yl)phenyl phosphate; DIG, digoxigenin; ER, endoplasmic reticulum; EtBr, ethidium bromide; IPCR, inverse PCR; MeJa, methyl-jasmonate; NMD, nonsense-mediated mRNA decay; nts, nucleotides; PTC, premature termination codon.

* Corresponding author. Tel.: +34 950015033; fax: +34 950015008.
E-mail address: fgmaroto@ual.es (F. García-Maroto).

0176-1617/$ – see front matter © 2010 Elsevier GmbH. All rights reserved.
doi:10.1016/j.jplph.2010.03.012

ther degraded very early during germination. A possible role in the initiation of storage protein mobilization has been suggested (Nong et al., 1995). Here, we report further characterization of *GMCP3*, including data about the genomic organization and gene structure, as well as relationships to other papain-like proteases. Expression analysis results also showed regulation of *GMCP3* by senescence and different stresses, thus indicating a more general role in the plant.

Materials and methods

Biological material and plant treatments

The Osumi variety of *Glycine max* was used in all the experiments. Seedlings (2–4 true leaves stage) were grown at 25 °C under the controlled conditions of culture cabinets with a 16 h light/8 h dark photoperiod and 70% relative humidity. Leaf material from etiolated seedlings was used as a source of genomic DNA. The different organs analyzed were sampled from mature plants cultivated under greenhouse conditions. The plant treatments were performed using plantlets grown in soil up to the stage of two to four developed true leaves. For dehydration experiments, plantlets were laid on dry filter paper for the indicated times. Wounding was performed using a dented forceps applied to the main veins of the leaves. Treatments with abscisic acid (ABA) and methyl-jasmonate (MeJa) (both from Sigma–Aldrich) were accomplished by pulverization of a 100 μM solution in water. At the end of each treatment, leaves were collected, pooling three plants for each experimental point, frozen in liquid N_2 and stored at −70 °C until RNA extraction.

Experiments performed with transgenic tobacco plants containing the *GMCP3Ψ* promoter fused to the *GUS* gene were performed with plants developed to 8–10 visible leaves stage, except for the assay of senescent leaves, which was carried out on adult plants. Leaf discs, pooled from three plants, were collected and analyzed for GUS activity as indicated below. Stress treatments were performed as in soybean. The results shown correspond to a representative transgenic line, but the same regulation was confirmed in at least two other lines, depending on the treatment.

Cloning of the GMCP3 and GMCP3Ψ genes

The genomic sequence comprising the whole coding region of *GMCP3* was assembled from two overlapping clones obtained by PCR on genomic DNA. The two DNA fragments of 1.5 kb and 1.3 kb were generated using the oligonucleotide pairs OCEP-1 (5′-ATGGAGGCAAAGCGAGGCCATGCCCTCAT-3′)/OCEP-2 (5′-TTGTTGTCACAGTCAAGGAGCTGTTGTTCAC-3′), and OCEP-3 (5′-GTGCGAGATAGGTAGAAGGAACAGAGTG-3′)/OCEP-4 (5′-CAGAACTAAGGAAAACCAAAGAGGAGATGC-3′), respectively. PCR amplification was conducted using a proofreading polymerase (Pfx®, Invitrogen) and a program consisting of a denaturation step of 2 min at 94 °C, followed by 35 cycles of 15 s at 94 °C, 45 s at 60 °C and 2 min at 68 °C, ending with a 7 min step at 68 °C. The resulting fragments were cloned in the vector pGEM-T-Easy® (Promega) and several clones were fully sequenced. A full cDNA clone was also obtained for *GMCP3*. Briefly, a cDNA was synthesized, as described below, using total RNA obtained from developing soybean seeds, and RT-PCR amplification was performed using the flanking oligonucleotide primers OCEP-1 and OCEP-4. The sequence of the 3′-untranslated region was completed by 3′-RACE as indicated below.

Cloning of *GMCP3Ψ* was accomplished by obtaining of a 1.8 kb partial fragment by genomic PCR using primers OCEP-5 (5′-CAAGCAAGCACAGAGATGGAGGCAAAG-3′) and OCEP-6 (5′-GACCTCCATTACACCCATTGTCACATG-3′), followed by bi-directional genomic walking through inverse PCR (IPCR). The method by Ochman et al. (1990) was employed with minor modifications (Garcia-Maroto et al., 2007). Thus, from the partial sequence of *GMCP3Ψ*, two nested upstream primers OCEP-7 (5′-GAGCTAAGCACGTGATGGAATGGTCTC-3′) and OCEP-8 (5′-AGCAGAGGATAGGGTTAGAGCGAAGAG-3′), and two nested downstream primers OCEP-9 (5′-TCTGCGCACGGAGAAGAAGTTCAAGGT-3′) and OCEP-10 (5′-CATGGAGAATTACGGGAGGAGCTACTC-3′), were designed to perform the IPCR. Either EcoRI or VspI restriction enzymes were used to cut genomic DNA. Released fragments were circularized with T4 DNA ligase and subjected to two nested rounds of PCR amplification, allowing us to obtain 1.6 kb (from EcoRI digestion) and 2.5 kb (VspI digestion) -long products that were useful for the walking process. These fragments were analyzed and sequenced as described above, allowing the assemblage of the whole coding sequence for *GMCP3Ψ*, including 1.4 kb of the 5′-flanking region.

The sequences have been deposited in the GenBank under the accession numbers GU452503 (*GMCP3* cDNA), GU452501 (*GMCP3* gene) and GU452502 (*GMCP3Ψ* gene).

RT-PCR and 3′-RACE analysis

The total RNA, employed for RT-PCR and RACE analysis, was purified from different soybean tissues (leaves and developing seeds) using the RNeasy Plant Mini Kit (QUIAGEN). Elimination of contaminating genomic DNA was performed using DNase-I (RQ1 RNase-Free DNase, Promega). The cDNA for RT-PCR was synthesized from 1 to 8 μg of RNA using oligo(dT) and employing the SuperScript First-Strand Synthesis System for RT-PCR kit (Invitrogen), following manufacturer's instructions.

Semi-quantitative RT-PCR was performed using the oligonucleotide primers UP-N (5′-ATGGAGGCAAAGCGAGGCCATGCCCTCAT-3′) and UP-Ψ (5′-ATGGAGGCAAAGCGAGACCATTCCATCAC-3′) as specific upstream primers for *GMCP3* and *GMCP3Ψ*, respectively, located on the first exon, in combination with the C13 primer (5′-CTTAACTTCGGTGACAGCTCCTTTCTCTC-3′) common to both genes and located on the second exon (see Fig. 4A). The PCR reactions were performed in 25 μL containing cDNA derived from 0.5 μg of RNA from developing seeds, using a hot-start Taq polymerase (Taq Platinum®, Invitrogen) and a program consisting of a denaturation step of 2 min at 94 °C, followed by 25–36 cycles of 15 s at 94 °C, 30 s at 58 °C and 1 min at 72 °C, ending with a 5 min step at 72 °C. Identity of the respective amplification products was confirmed by direct sequencing of the fragments.

Similarly, relative expression of *GMPCP3* and *GMCP3Ψ* in plantlets subjected to dehydration (experiment in Fig. 5C) was estimated by RT-PCR using either UP-N or UP-Ψ as the upstream specific primers in combination with C11 primer (5′-GAGAGAAAGGAGCTGTCACCGAAGTTAAG-3′), common to both genes and located at the end of the first exon. PCR conditions were as described above, using 32 cycles. As a loading control, amplification of the actin-2 was carried out in parallel on equivalent samples using the oligonucleotide primers ACT2-F (5′-TCCCTCAATCTCATCTTCTTCC-3′) and ACT2-R (5′-GACCTGCCTCATCAATCTTCTTCC-3′).

3′-RACE was performed using the SMART RACE® kit (Clontech), following manufacturer's protocols. 1–2 μg of total RNA from developing seeds (Fig. 4C and Suppl. Fig. 2), or leaves from normal and dehydrated seedlings (Suppl. Fig. 2) were used for the synthesis of the oligo(dT) primed cDNA linked to an appropriate adaptor. Either the UP-N primer specific for *GMCP3*, UP-Ψ and C17 (5′-GAGAAGCTCTACACCGGCTCCGTCGAC-3′) for *GMCP3Ψ*, or the C6 (5′-TCTGCGCACGGAGAAGAAGTTCAAGGT-3′) primer, common to both *GMCP3* and *GMCP3Ψ*, were used as sense gene-specific primers for PCR amplification, as indicated in the experiments. The