A Genomic Cluster Containing Four Differentially Regulated Subtilisin-like Processing Protease Genes Is in Tomato Plants*

Lucia Jordáš, Alberto Coego, Vicente Conejero, and Pablo Vera§

From the Instituto de Biología Molecular y Celular de Plantas (IBMCP), Universidad Politécnica-Consejo Superior de Investigaciones Científicas, Camino de Vera s/n, 46022-Valencia, Spain

Screening of a genomic library from tomato plants (Lycopersicon esculentum) with a cDNA probe encoding a subtilisin-like protease (PR-P69) that is induced at the transcriptional level following pathogen attack (Torrent, P., Conejero, V., and Vera, P. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 6332–6337) resulted in the isolation of a cluster of genomic clones that comprise a tandem of four different subtilisin-like protease genes (P69A, P69B, P69C, and P69D). Sequence analyses and comparison of the encoded proteins revealed that all are closely related (79 to 88% identity), suggesting that all are derived from a common ancestral gene. mRNA expression analysis as well as studies of transgenic plants transformed with promoter-β-glucuronidase fusions for each of these genes revealed that the four genes exhibit differential transcriptional regulation and expression patterns. P69A and P69D are expressed constitutively, but with different expression profiles during development, whereas the P69B and P69C genes show expression following infection with Pseudomonas syringae and are also up-regulated by salicylic acid. We propose that these four P69-like proteases, as members of a complex gene family of plant subtilisin-like proteases, may be involved in a number of specific proteolytic events that occur in the plant during development and/or pathogenesis.

Proteolysis is fundamental for the normal functioning of multicellular organisms and plays key roles in a variety of processes such as development, physiology, defense and stress responses, and adaptation to the changing environment. In plants, despite the importance of all these processes and involvement of different classes of proteinases (Refs. 1–5, and references therein), it still remains to be defined more precisely what components and molecular mechanisms are responsible for regulating specific aspects of protein degradation/processing. A major task for research will be to determine which proteolytic enzymes in higher organisms. They can be grouped in six clans, of which one of the largest is the subtilisin-like clan (EC 3.4.21.14) that includes over 200 different members. Despite this wealth of knowledge, very little is know about subtilisin-like proteases in plants. Recently, we and others have shown the existence of members of this clan in plants, including Arabidopsis (6), tomato (7, 8), melon (9), and Lilium plants (10). According to a recent classification (11), the subtilisin-like proteases from plants can be grouped within the Pyrolysin subfamily, which is highly related to the Kexin subfamily of proteases involved in the posttranslational processing of peptide hormones (12, 13). Comparative molecular, biochemical, and cellular studies indicate that the subgroups of plant subtilisin-like enzymes are characterized by the presence of a large polypeptide sequence insertion preceding the reactive Ser residue and/or long C-terminal extensions relative to all other subtilisin-like proteases. Furthermore, they were found to be glycosylated and to be secreted to the plant extracellular matrix (ECM) where they accumulated and presumably exert their biochemical function(s) by recognizing and processing pericellular substrates (7–11).

We here describe the isolation and characterization of a genomic cluster comprised of four genes encoding different, but highly related, members of subtilisin-like proteases (named as P69A, P69B, P69C, and P69D). While the four clustered proteases exhibit a high degree of amino acid sequence identity, we show that they are differentially regulated at the transcriptional level, each showing a different expression pattern, either during normal plant development or following pathogenic attack.

EXPERIMENTAL PROCEDURES

**Plant Material, Growth Conditions, Treatments—Lycopersicon esculentum cv. Rutgers and Arabidopsis thaliana (Col-0) plants were grown at 22 °C in growth chambers programmed for a 14-h light and 10-h dark cycle. Fully expanded leaves or rosette leaves were sprayed with SA (0.5 mM) or buffer alone (50 mM phosphate buffer, pH 7.2), and samples were taken for analysis after 48 h. Suspensions of Pseudomonas syringae strains (0.09 O.D.) were infiltrated locally in one part of the leaf. Control (mock) plants were injected similarly with the solution containing no bacterium. Samples were analyzed 24–48 h post-inoculation.**

**Library Screening and DNA Sequence Analysis—**A tomato genomic DNA library constructed in λ-EMBL3 was screened at 65 °C as described (14) with the radiolabeled p26 cDNA encoding the prepro sequence of the PR-P69 protein described previously (7). The positive clones were isolated and characterized with the routine described previously (14).

**RT-PCR—**cDNA synthesis, quantification of the products, and reverse transcriptase-mediated PCR were conducted as described (8). The oligonucleotide primer pairs (50 pmol each), a1 + a2 (ATGGGATT CTTGAAAATCCTT), b1 + b2 (ATGGGATT CTTGAAAATCCTT + CCTAGGCCAGACCAACTG)

* This work was supported in part by the Spanish Ministry of Science and Education. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1724 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL and DDBJ Data Banks with accession number(s) Y17275 (P69A), Y17276 (P69B), Y17277 (P69C), and Y17278 (P69D).

‡ Supported by a Fellowship from the Conselleria de Educación y Ciencia de la Generalitat de Valencia.

§ To whom correspondence should be addressed. Tel.: 34-96-3877864; Fax: 34-96-3877869; E-mail: vera@ibmcp.upv.es.

1 The abbreviations used are: PCR, polymerase chain reaction; RT, reverse transcription; Pst, Pseudomonas syringae pv. tomato; SA, salicylic acid; GUS, β-glucuronidase; aa, amino acid; HR, hypersensitive response; aa, amino acids.
CAAT), c1 + c2 (ATGGGAT TCTTGGAAATCCTT + TCAATC ATG TCTCTCTCAAGAG) and d1 + d2 (ATGGGATTCTTGGAAAT + TTA TATCACGAGACTCTAAGTCG), specific for the amplification of P69A, P69B, P69C, and P69D. Hybrids of the PCR products were sequenced using specific primers. The constructs were verified by sequencing the amplicons from the plasmid DNA containing each of the four P69 ORFs as template.

**RESULTS**

Characterization of a Genomic Cluster Containing Four Genes Encoding Highly Related P69 Proteases—A DNA fragment encoding the signal peptide and propeptide for the previously identified P69 protease was obtained from plasmid p26 (7) and used as a radiolabeled probe to screen a tomato genomic library constructed in plasmid DNA containing each of the three P69 genomic clones. The resulting transcriptional fusions were verified by nucleotide sequence analysis using specific primers. The constructs were introduced into Arabidopsis plants by Agrobacterium tumefaciens-mediated transformation. Transformants were selected on MS agar medium containing kanamycin, transferred to soil, and allowed to self. The transgenic lines were assayed for GUS activity by a fluorimetric assay or by an in situ assay using the colorigenic substrate X-gluc (16).

**Fig. 1. P69 genomic cluster.** Four P69-like open reading frames sequences (boxes) are arranged in a tandem array (named as P69A, P69D, P69C, and P69D), that share the same orientation. The EcoRI (E) restriction sites are indicated in each clone. The distances are only approximate. Arrows indicate direction of transcription.

**Fig. 2. Amino acid sequence alignment for the predicted P69 gene products.** The sequence of P69A is shown in full and compared with predicted open reading frames for P69B, P69A, and P69D. Dashes represent sequence identity. The catalytically important Asp, His, Asn, and Ser residues are shown in bold with an asterisk. The propeptide domains are dashed. Dots were introduced to maximize alignment. Amino acid residues of each protease are numbered from the precursor sequence.
conserved Thr-115, identified also by comparison with other plant subtilisin-like proteases (6–11). The predicted mature enzymes thus contain 631, 631, 552, and 632 amino acids for the P69A, P69B, P69C, and P69D isoforms, respectively. Within the four mature proteins, the amino acid residues Asp-146, His-203, and Ser-532 (or Ser-531 for P69B and P69C), corresponding to the catalytic site (catalytic triad) essential for the enzymatic activity of subtilisin-like members to function as proteases, were identified (Fig. 2). Also the proteases of the four P69s in the cluster have an Asn residue (Asn-306, or Asn-305 for P69A, P69B, and P69C) that has been found to be highly conserved in this position and that is catalytically important in the subtilisins (12, 13). However, sequences close to this Asn-306 are highly variable within the four P69s (Fig. 2). In all cases there is also an insertion of a long sequence (226 amino acids) between the stabilizing Asn-306 and the reactive Ser-532, relative to all other subtilisin-like proteases (11), in which these two residues are separated by much shorter distances. This displacement has also been observed in the three other subtilisin-like proteinases recently identified from plants (6, 9, 10) and could represent a characteristic signature of the subtilisin enzymes from plants (Fig. 3).

Expression Analysis of P69 Genes—The differential expression pattern of the four P69 genes was initially determined by gene-specific RT-PCR reactions combined with Southern blot analyses. In vitro synthesized single-stranded cDNAs from mRNA samples of fully grown leaves from healthy and Pseudomonas syringae pv. syringae-infected tomato plants were assayed by PCR using sets of primers which were specific for each P69 gene member. Primer specificity was demonstrated in pilot experiments using each of the individually cloned gene as template in the PCR reaction (Fig. 4). Whereas P69A mRNAs accumulate to detectable levels in fully expanded leaves of healthy tomato plants, the P69D mRNAs accumulate in marginal amounts in the same leaf samples. Neither P69A nor P69D were found to be induced over basal levels in leaves that were infected with P. syringae. Conversely, similar RT-PCR analysis with primers specific for the P69B gene and the P69C gene indicated that while the corresponding mRNAs were nearly undetectable in leaves from healthy control plants, a dramatic accumulation of the transcripts occurred in P. syringae-infected leaves (Fig. 4).

Developmental and Tissue-specific Regulation of the Different P69 Promoters in Transgenic Plants—To investigate in detail the spatial pattern of expression of the four P69 genes, each of the 5′ flanking promoter regions was fused to the β-glucuronidase (GUS) reporter gene in plasmid pBI101.1 to generate constructs pP69A::GUS, pP69B::GUS, pP69C::GUS, and pP69D::GUS (Fig. 5). These constructs were introduced separately into Arabidopsis plants by transformation with A. tumefaciens, and a minimum of ten independent kanamycin-resistant transformants were generated for each construct. GUS activity was initially analyzed in plants grown under normal growth conditions and also after inoculation with the avirulent bacteria P. syringae DC3000 (AvrRpm1) by a fluorimetric assay (16) (data not shown). These assays revealed constitutive expression of GUS activity in the transgenic plants generated with construct pP69A::GUS and P69D::GUS, whereas those harboring the P69B::GUS or the P69C::GUS cassette expressed GUS activity only upon bacterial infection (see below). These results were, to some extent, coincident with the RT-PCR studies shown above (Fig. 4).

To study comparatively the distribution of GUS activity in planta, the initial transgenic plants generated for each construct were selfed, and the kanamycin-resistant progeny was analyzed in situ using the chromogenic substrate X-Gluc (Fig. 6). Expression of GUS activity driven by the P69A promoter was detected in the seedlings as well as in fully grown plants. As deduced from the tissue staining pattern, it seems as if the P69A gene is expressed in a general fashion in all organs of the plant, except in roots and in flower organs, where no GUS activity could be detected in any plant.

Likewise, transgenic plants in which GUS expression was driven by the P69D promoter revealed that this is active predominantly in expanding cotyledons and leaves (Fig. 6). This expression was transient because it disappeared once the leaves or cotyledons had enlarged and matured. Interestingly, the P69D gene is also observed to be expressed in inflorescences, and more particularly, in stigmas. We still do not know whether or not this expression pattern in flowers is also transient.

Conversely, no constitutive expression of the GUS gene
driven by the P69B or P69C promoters could be detected in any of the transgenic plants generated (Fig. 6). However, in some of the P69C::GUS transgenic lines, we could detect GUS activity in discrete groups of cells (islands) that appeared sporadically, and with an unpredictable location, either in stems, roots or leaves (not shown).

Induced expression was also analyzed in rosette leaves from transgenic plants before and after inoculation with P. syringae pv. tomato (Pst) strain DC3000 carrying the avrRpm1 avirulence gene, which is recognized by the corresponding resistance gene in A. thaliana ecotype Col-0. Pst DC3000(avrRpm1) and causes a macroscopic hypersensitive response (HR) at the inoculation site. These studies revealed that GUS expression driven by either the P69B and P69C promoters was induced in the infected leaves. Induction occurred throughout the inoculated leaves and was not restricted to the site of inoculation where the HR became apparent (Fig. 7). Conversely, in transgenic plants carrying P69A::GUS or P69D::GUS constructs, GUS expression was not induced during the course of infection.

Because salicylic acid (SA) has been demonstrated to be a master regulatory molecule mediating most of the plant defense responses to challenging pathogens (19), we tested whether or not SA could act as inducer of any of these genes. Spraying of healthy transgenic Arabidopsis plants with a 0.5 mM solution of SA resulted in the induction of GUS activity only in plants harboring the P69B::GUS and P69C::GUS transgenes, and in a manner similar to that observed following bacteria inoculation (Fig. 7). These results further reinforce our consideration that the P69B and P69C gene pair is involved in pathogenic response of the plant to challenging pathogens, whereas the P69A and P69D pair is more related to development or to a basic biochemical function.

**DISCUSSION**

In this work, we provide structural and functional information on a genomic cluster comprising four different members of a family of plant genes, which on the basis of amino acid sequence conservation and structural organization are related to the subtilisin-like protease clan (EC 3.4.21.14) (11). The genomic cluster was identified by screening of a genomic DNA library from tomato plants with a partial cDNA probe for the previously identified pathogenesis-related PR-P69 protease (7) (here renamed as P69A). The predicted primary structure of the four P69 proteases designated P69A, P69B, P69C, and P69D, indicate that all of them are synthesized as precursor proteins (preproenzyme) composed of three distinct domains: a 22-amino acid signal peptide, a 92-amino acid propolypeptide, and a mature polypeptide of variable length for each protease in the cluster. Within the mature polypeptides, the amino acid sequences surrounding Asp-146, His-203, and Ser-531 are the most salient features of all these proteases, which are identical to those of the catalytic sites (catalytic triad) of subtilisins (11). Interesting also is the conservation of an insertion of a long sequence (226 amino acids) between the conserved Asn residue and the reactive Ser residues of the catalytic triad, relative to...
all other subtilisin-like proteases. The meaning of such a conserved displacement remains unknown, but its conservation suggests it may subserve important functions in regulating the properties of this subgroup of subtilases in plants.

Comparative studies of the mode of expression by RT-PCR and by analysis of transgenic plants harboring independent promoter-GUS fusions for the four P69 genes indicate that they are regulated differently. P69A is transcribed at all stages of plant growth and in all organs except roots and flowers where the expression is absent. Likewise, P69D is also transcribed under resting conditions in emerging leaves, but its expression is transient. This transient expression pattern, which has been observed also in transgenic tobacco plants (data not shown), is presumably associated with the elongation processes in these organs because transcription is repressed once the leaves cease elongation. Although P69D is not expressed in roots, transcription is specifically recovered in flowers, and in the stigma in particular. Neither P69A nor P69D gene expression is induced over basal levels during pathogenesis.

Conversely, although neither P69B or P69C show constitutive expression in any organ of plants grown under resting conditions, both are transcriptionally activated upon infection with avirulent bacteria, either in transgenic Arabidopsis plants or in tomato plants. This latter result is consistent with our previous observations that one member of this family, now identified as P69B on the basis of sequence identity, is coordinately induced with a set of pathogenesis-related (PR) proteins associated with the defense response in tomato plants (7, 20, 21).

Interestingly, the observation that the induced transcriptional activation of P69B and P69C is not restricted to the point of inoculation with the avirulent bacteria (Fig. 7), where the HR cell death and the induction of protective genes occur (22), but rather is extended throughout the afflicted leaf blade suggests that their induction is mediated by a long distance signaling process. Because the benzoic acid-derivative SA is one of the candidate signal molecules mediating long distance activation of plant defense reactions (19), we tested the ability of SA to induce expression of these two pathogen-inducible genes. The result presented in Fig. 7 demonstrates that local application of SA to leaves promotes the transcription of both genes de novo. Thus, these results reconcile with the idea that proteases, and in particular members of the subtilisin-like family, are components of the general plant response to attacking pathogens.

The differential expression profile found for these four P69 genes indicate that different regulatory mechanisms have evolved to control the expression of these genes either during

---

**Fig. 7.** GUS staining patterns in rosette leaves of transgenic Arabidopsis (Col-0) plants carrying the P69A::GUS, P69B::GUS, P69C::GUS, and P69D::GUS transgenes. Left panel, GUS staining pattern in leaves from plants that have been sprayed with a 0.5 mM salicylic acid solution; middle panel, GUS staining pattern in leaves from plants that have been inoculated with P. syringae pv. tomato DC3000 carrying the avirulent avrRpm1 gene. The site of inoculation and the area undergoing the HR are indicated with an arrow in each case. Right panel, GUS staining pattern in full expanded leaves from healthy mock-inoculated control plants.
normal growth or under pathological situations. It remains to be demonstrated whether or not these differential expression patterns also imply different roles for the protein products. In this regard, it is tempting to speculate that the permanent expression of the P69A gene at the different stages of growth of the plant may suggest a housekeeping function for the P69A protein and that this function may be advantageous enhanced during processes of plant cell elongation and/or pathogenesis by supplementing with the other isoforms that may serve backup functions to the principal protease. It has been demonstrated in animal systems, where natural substrates for subtilisin-like proteases have been identified (12, 13), that when the different protease members of a family are removed from their biological context and assayed in vitro, many of these proteases are able to process the same substrates. This observation again raises the question of whether such a functional redundancy exists among the family members in vivo and how this might be regulated. One insight into defining functionality has been provided by examining the expression of the individual protease members (23). From such analysis, it has been shown that protease substrate specificity in vivo is influenced by restricting expression to particular tissues and also by compartmentalization of the individual enzymes to specific intracellular locations (24). Thus, it may be the case that a similar regulation also applies for the differentially regulated P69 protease isoforms under consideration as a likely explanation for delimiting redundant functions in vivo. Alternatively, we cannot disregard the possibility that each member functions separately by recognizing different substrate(s) and thus implies that both function and gene expression patterns evolved coordinately, but separately, for each P69 protease.

Whatever the meaning of such complexity is, the diversification of either the regulation of gene expression or function for each member of the P69 clan would be in agreement with the polymorphism found for other unrelated gene families in higher plants, which arise from gene duplication events of a common ancestral gene (25, 26).

So far, only two protein substrates have been identified for these plant subtilisin-like proteases. One is systemin, the traveling peptide hormone mediating signaling processes during wound response in plants (27), the other is LRP (28), an extracellular matrix associated leucine-rich repeat (LRR) protein that is part of a family of proteins that mediate molecular recognition and/or protein interaction processes (29). Subtilisin-like enzymes have been shown to be secreted to the plant extracellular matrix (ECM) (30). Thus, the consideration that the P69s may be mediating in pericellular processing/degradation events may indicate a role, similar to those in some animal systems (31, 32), in modulating the interaction of the plant cell surface with the extracellular environment.

Acknowledgments—We thank Jeff Dangl and Pablo Tornero for providing the Pst DC3000 strain.

REFERENCES

1. Glazer, A. N., and Smith, E. L. (1971) in The Enzymes (Boyer, P. D., ed) Vol. 3, 3rd Ed., pp. 561–546, Academic Press, New York.
2. Ryan, C. A. (1981) in The Biochemistry of Plants, Vol. 6, pp. 321–350, Academic Press, New York.
3. Schaller, A., and Ryan, C. A. (1996) Plant Mol. Biol. 31, 1073–1077.
4. Pautot, V., Holzer, F. M. Reisch, B., and Walling, L. L. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 9906–9910.
5. Kinoshita, T., Fukuzawa, H., Shimada, T., Saito, T., and Matsuda, Y. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 4693–4697.
6. Ribeiro, A., Akkermans, A. D. L., van Kammen, A., Bisseling, T., and Pawlowski, K. (1995) Plant Cell 7, 789–794.
7. Tornero, P., Conejero, V., and Vera, P. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 6332–6337.
8. Tornero, P., Conejero, V., and Vera, P. (1997) J. Biol. Chem. 272, 14412–14419.
9. Yamagata, H., Masuzawa, T., Nagahara, Y., Ohsaki, T., and Iwasaki, T. (1994) J. Biol. Chem. 269, 27225–27231.
10. Taylor, A. A., Horsch, R. A., Raczczak, A., Hasenkampf, C. A., and Daniel Riggs, C. (1997) Plant J. 12, 1261–1271.
11. Siezen, R. J., and Leunissen, J. A. M. (1997) Protein Sci. 6, 501–523.
12. Barr, P. J. (1991) Cell 66, 1–3.
13. Steiner, D. F., Smeenk, S. P., Ohagi, S., and Chan, S. J. (1992) J. Biol. Chem. 267, 23435–23438.
14. Tornero, P., Conejero, V., and Vera, P. (1996) Plant J. 9, 639–648.
15. Kunkel, T. A., Roberts, J. D., and Zakour, R. A. (1987) Methods Enzymol. 154, 367–381.
16. Jefferson, R. A. (1987) Plant Mol. Biol. Rep. 5, 387–405.
17. Von Heijne, G. (1986) Nucleic Acids Res. 14, 4683–4690.
18. Zhou, A., Paquet, L., and Mains, E. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 4683–4690.
19. Enyedi, A. J., Yalpani, N., Silverman, P., and Raskin, Y. (1992) Cell 70, 879–886.
20. Vera, P., and Conejero, V. (1988) Plant Physiol. 87, 58–63.
21. Vera, P., and Conejero, V. (1989) Physiol. Mol. Plant Pathol. 34, 323–334.
22. Dixon, R. A., and Lamb, C. J. (1990) Annu. Rev. Plant Physiol. Plant Mol. Biol. 41, 339–367.
23. Molloy, S. S., Thomas, L., VanSlyke, J. K., Stenberg, P. E., and Thomas, G. (1994) EMBO J. 13, 18–33.
24. Seidah, N. G., Day, R., Hamelin, J., Gaspar, A., Collard, M. W., and Chretien, M. (1992) Mol. Endocrinol. 6, 1559–1570.
25. Sharrock, R. A., and Quail, P. H. (1989) Genes Dev. 3, 1745–1757.
26. Liang, X., Abel, S., Keller, J. A., Shen, N. F., and Theologis, A. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 11046–11050.
27. Schaller, A., and Ryan, C. A. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 11802–11806.
28. Tornero, P., Mayda, M. E., Gómez, M. D., Cuénas, L., Conejero, V., and Vera, P. (1999) Plant Cell 11, 315–330.
29. Kobe, B., and Deisenhofer, J. (1994) Trends Biochem. Sci. 19, 415–421.
30. Vera, P., Hernandez-Yago, J., and Conejero, V. (1989) Plant Physiol. 91, 119–123.
31. Alexander, C. M., and Werb, Z. (1991) in Cell Biology of Extracellular Matrix (Hay, E., ed) pp. 255–302, Plenum Press, New York.
32. Chen, W-T. (1992) Curr. Opin. Cell Biol. 4, 802–809.
