Plant Molecular Biology **27**: 1197–1204, 1995.

Short communication

Structure and induction pattern of a novel proteinase inhibitor class II gene of tobacco

Teresa Balandin, Chris van der Does [1], José-Maria Bellés Albert [2], John F. Bol and Huub J. M. Linthorst *

Institute of Molecular Plant Sciences, Gorlaeus Laboratories, P.O. Box 9500, 2300 RA Leiden, The Netherlands (author for correspondence); Present addresses: [1] Department of Molecular Microbiology, Groningen University, P.O. Box 14, 9750 AA Haren, The Netherlands; [2] Dep. Biotecnología, E.T.S.I. Agrónomos, Univ. Politécnica, Camino de Vera, 14, 46022-Valencia, Spain*

Received 21 April 1994; accepted in revised form 19 January 1995

Key words: pathogenesis-related proteins, proteinase inhibitor, signal transduction, wounding

Abstract

A cDNA and a corresponding genomic clone encoding a protein with partial identity to type II proteinase inhibitors from potato, tomato and *Nicotiana alata*, were isolated from tobacco libraries. The protein of 197 amino acids contains a putative signal peptide of 24 residues and three homologous domains, each with a different reactive site. The tobacco *PI-II* gene is not expressed in leaves of healthy plants, but is locally induced in leaves subjected to different types of stress (TMV infection, wounding, UV irradiation) and upon ethephon treatment. As opposed to the analogous *PI-II* genes of potato and tomato, the tobacco gene is not systemically induced by wounding or pathogenic infection. A far-upstream region in the *PI-II* promoter, containing various direct and indirect repeats, shares considerable sequence similarity to a similar region in the stress-inducible Cu/Zn-superoxide dismutase gene of *N. plumbaginifolia*.

Plants reacting to environmental stress conditions change their gene expression patterns to adapt to the new situation. This usually results in the accumulation of numerous defense proteins of which the serine proteinase inhibitors (PI) are widely represented in the plant kingdom. Two non-homologous PIs, inhibitor I and II, are by far the best characterized. They are present in storage organs, in vegetative cells and in reproductive organs. In potato tubers PI-I and PI-II represent approximately 2% and 5% of the soluble proteins, respectively. PI-I and PI-II are powerful inhibitors of serine endopeptidases of animals and microorganisms [for review see 21]. PI-I and PI-II are synthesized in leaves of tomato and potato plants in response to wounding [8]. They accumulate not only in the wounded leaf (local) but also in distant, undamaged tissues (systemic).

The nucleotide sequence data reported will appear in the EMBL Nucleotide Sequence Library under the accession number Z29537 (gPI2-1).

The accumulation of PIs after wounding is thought to be a defensive response that interferes with the digestive processes of attacking pests (bacteria, fungi) and insects.

PIs of type II contain two reactive sites, one of which inhibits chymotrypsin and the other trypsin [4, 19, 1]. Tomato PI-II is also a strong inhibitor of the bacterial subtilisins [19]. The nucleotide sequence of cDNA or genomic clones corresponding to PI-II of tomato [6], potato [25, 12], and *Nicotiana alata* [1] have been reported. The tomato and potato sequences show that the encoded *PI-II* proproteins contain two homologous domains, whereas the *N. alata* precursor consists of six repeated domains.

After the screening of a λZAP cDNA library from TMV-infected tobacco leaves [14], with a ^{32}P-labelled fragment corresponding to a *PI-II* gene of tomato [6], a hybridizing cDNA clone (cPI2-2) was isolated. The cDNA insert was ca. 800 bp long and contained a nucleotide sequence similar to PI-II sequences reported from other plants (see below). The subsequent screening of a tobacco genomic library using cPI2-2 as a probe resulted in the isolation of three independent genomic clones (gPI2-1, gPI2-2 and gPI2-13). Southern blot analysis of *Hind* III-digested DNA from these clones with a cPI2-2 probe revealed a common 2.8 kb hybridizing band, suggesting that the three genomic clones probably contain the same gene. However, an extra hybridizing band (with a size of ca. 7 kb) was found in gPI2-13, suggesting the existence of a second gene (data not shown).

DNA blots from tobacco genomic DNA digested with *Eco* RI and *Hind* III and hybridized at high stringency with cPI2-2 insert as probe, revealed the presence of 4 hybridizing fragments in each digest. This indicates that tobacco contains a limited number of *PI-II* genes (data not shown). The 2.8 kb *Hind* III fragment of clone gPI2-1 was subcloned and selected for characterization of the tobacco *PI-II* gene.

The complete nucleotide sequence of the cDNA insert of clone cPI2-2 and the 2.8 kb *Hind* III fragment of gPI2-1 was elucidated. Figure 1 shows the nucleotide sequence of clone gPI2-1. The genomic fragment is 2768 bp long and completely overlapped the 778 bp (excluding the poly(A) stretch at the 3′ end) of the cDNA insert of clone cPI2-2 (the 5′ residue of cPI2-2 is located at position 1716). Both nucleotide sequences were identical, indicating that the gene present in clone gPI2-1 is expressed. The sequence of the cDNA is interrupted in the genomic clone by an intron of 203 bp. The relative position of this intron, located in the open reading frame, is conserved in the genes for potato PI-II [12, 25]. The location of the transcription start site (indicated by the first bold residue in Fig. 1) was determined by primer extension dideoxy sequencing on poly(A) RNA from TMV-infected plants. The primer extension resulted in the elucidation of more than 50 5′-terminal nucleotides and did not show heterogeneity (data not shown). TATAAA and CAAT boxes (underlined) are present at −31 and −83, respectively, upstream of the transcription start site. The poly(A) tail in cDNA clone cPI2-2 is preceded by the putative polyadenylation signal AATATT (underlined).

The open reading frame encodes a protein of 197 amino acid residues (Fig. 1). The highly hydrophobic, N-terminal region of 24 residues is expected to function as a signal peptide for subcellular targeting, similar to tomato and potato PI-II. Cleavage between Ala-24 and Lys-25 would result in a peptide of 173 amino acids. The amino acid sequence encoded by the tobacco gene, shares a considerable homology with those of the *PI-II* genes of other plants from the Solanaceae family. There is 72% identity with potato PI-II [25] and 69% identity with tomato PI-II [6]. However, the polypeptide encoded by the open reading frame is considerably longer than those of the tomato and potato *PI-II* genes, which each contain two catalytic domains with similar amino acid sequences [6, 22]. Yet, the tobacco PI-II protein is shorter than the protein encoded by a cDNA clone from *N. alata* stigma tissue, containing six repeated domains with very high sequence similarity [1]. In agreement with its size, the protein encoded by the tobacco *PI-II* gene is composed of a repeat of three domains. A tomato

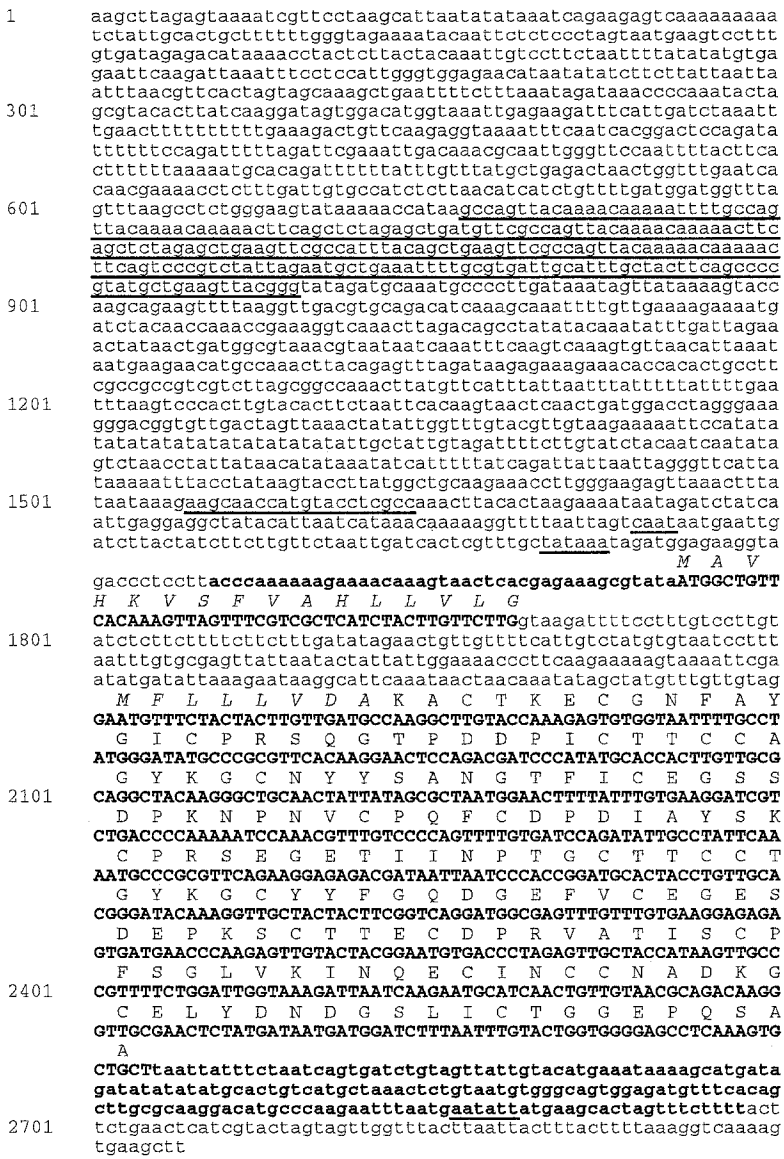

Fig. 1. Nucleotide sequence and deduced amino acid sequence of clone gPI2-1, encoding tobacco PI-II. The coding region is given in capitals. Putative CAAT and TATAAA boxes in the upstream region and a polyadenylation signal in the downstream region are underlined. Also underlined are upstream regions with sequence similarity to promoter sequences of other genes (see text). The sequence transcribed into mRNA is indicated in bold. The amino acid sequence of the encoded protein is given in one-letter code above the nucleotide sequence. The hydrophobic N-terminal region, probably functioning as signal peptide, is given in italics.

PI-II-type gene which is auxin-inducible in roots and containing three highly similar domains was recently characterized [24]. Figure 2 shows a comparison of the putative tobacco PI-II (tob PI-II) domains (excluding the signal peptide) with those of PI-II domains deduced from the tomato (tom PI-II), potato (pot PI-II) and *N. alata* (N.a. PI-II) genes. Also included in the comparison are PI-II-like proteins from tobacco (TTI [17]), eggplant (EP [20]) and potato (PTI, PCI [9]), which were sequenced at the amino acid level. These peptides are each composed of a single catalytic domain. It is evident that the various PI-II domains are homologous. In particular,

Fig. 2. Similarity of tobacco PI-II to other PI-II-type inhibitors. The amino acid sequence of the protein encoded by cDNA clone cPI2-2, minus the 24 residues of the N-terminal signal peptide, is given (tob PI-II) and compared to the corresponding sequences of PI-II proteins encoded in cDNA or genomic clones from potato (pot PI-5 [22]) and tomato (tom PI-II [6]), *Nicotiana alata* (N.a. PI [1]). These deduced sequences were arranged such as to obtain a maximum alignment with the sequences of tobacco trypsin inhibitor (TTI [17], eggplant trypsin inhibitor (EP [20]), potato trypsin inhibitor (PTI [9]) and potato chymotrypsin inhibitor (PCI [9]). The three domains with putative reactive sites present in the tobacco PI-II protein are indicated. Highly conserved residues in the sequences are indicated by bars. The consensus sequence for PI-II type inhibitor domains is shown above. Absolutely conserved residues are indicated in capitals. The active site residue is indicated (▼).

the positions of the eight Cys residues, involved in intramolecular folding [7], are highly conserved. A PI-II protein preparation isolated from *N. alata* stigmas was shown to be a mixture of peptides, each encompassing one PI-II domain, presumably processed from the 6-domain precursor encoded by the cDNA clone. Also, the TTI inhibitor from wounded tobacco leaves is a mixture of six low-molecular-weight peptides. Two of these peptides have very high (>95% identical) sequence similarity to the *Nicotiana alata* domains [17]. None of the putative domains of the tobacco *PI-II* characterized in the present study is more than 66% identical to the sequenced TTI inhibitors.

Each of the three domains of the tobacco PI-II contains a putative reactive site, supposed to be at the same position as in the other members of the family [20]. The basic residue (Lys) at the active site (P1 position) of domain 1 suggests that this domain could possess anti-trypsin activity. The P1 position in the putative active sites of domain 2 (Gln) and domain 3 (Thr) have not been found in PI-II proteins in plants, but are present in the active site of the third domain of several avian ovomucoid serine proteinase inhibitors [13]. Viral serine proteinases like the potyviral NIa proteinase require a Gln residue at the P1 position and Val at the P4 position [5]. In the second domain of the tobacco PI-II protein these residues are present at the P1 and P4 position, respectively. Whether domain 2 or the other putative tobacco PI-II domains would be capable to render plant (viral) proteinases inactive by locking the specific substrate-binding pocket remains to be resolved.

Expression of tobacco PI-II *genes after different treatments*

Defence genes of plants are known to be inducible by diverse stimuli such as wounding, UV irradiation, pathogen infection and ethylene. TMV infection is a strong inducer of *PI-I* genes in tobacco [15, 10]. We have tested the expression of *PI-II* genes in tobacco leaves after applying different types of stress.

Seven weeks old tobacco plants (*Nicotiana tabacum* var. Samsun NN) were grown in a growth chamber under controlled conditions (70% relative humidity, 23 °C, 16 h light/8 h dark). At different times after applying the stress, leaves were taken for RNA extraction. In the cases in which systemic expression of *PI-II* genes was tested, leaves directly above the stressed ones were sampled. The sampled leaves were immediately frozen in liquid nitrogen and stored at −80 °C.

Figure 3 shows the results of hybridization of cPI2-2 probe to total RNA isolated from tobacco leaves. The left panel of Fig. 3 shows that the *PI-II* gene(s) are not expressed in leaves of healthy plants (lane H), but that PI-II mRNA accumulation is high in TMV-infected leaves at three days after inoculation (lane T). Similar to what we found for PI-I [15] and most basic PR proteins [3], TMV-induced *PI-II* gene expression in tobacco coincides with local lesion formation

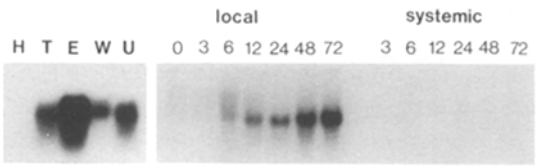

Fig. 3. Expression of the tobacco *PI-II* gene upon stress. Total RNA was isolated from tobacco leaves, electrophoresed, blotted and hybridized to PI-II cDNA probe. The left panel shows PI-II mRNA accumulation in tobacco leaf 3 days after inoculation with water (lane H) or TMV (lane T), 1 day after ethephon treatment (lane E), 2 days after incision wounding (lane W), or 1 day after UV irradiation (lane U). The right panel shows PI-II mRNA accumulation at the indicated times (h) after incision-wounding in the wounded leaves (local) and the unwounded leaves directly above the wounded leaves (systemic).

and is only apparent in the infected leaves. Any detectable induction of the genes in the uninfected, systemic leaves is absent (data not shown). The genes are also highly induced 1 day after spraying with 10 mM ethephon (lane E), 2 days after wounding by making 30 incisions parallel to the lateral veins (lane W), and 1 day after UV irradiation (lane U).

In tomato and potato plants, *PI* genes quickly respond to wounding, not only in the affected leaves but in the unwounded leaves as well [8, 2]. PI mRNA can be detected in the unwounded leaves of potato within 20 min of local wounding [18]. The right panel of Fig. 3 shows a time course of tobacco *PI-II* gene expression upon incision wounding. The first increase in PI-II mRNA could be measured at 6 h after wounding and accumulation increased for at least 3 days (Fig. 3; local, 0–72). However, in the upper, non-wounded leaves there was no increased PI-II mRNA synthesis measurable during the same time course (Fig. 3; systemic, 0–72). A similar absence of a systemic wound response was obtained with very young tobacco plants (3 weeks old) and with plants subjected to different types of wounding, including incision, pricking, squeezing and scorching (data not shown).

In tobacco plants transformed with a potato *PI-II* gene containing the coding region and 5′ upstream and 3′ downstream sequences, wound-induced expression of the *PI-II* gene followed the same pattern as in potato plants [23]. Based on these observations, it is evident that tobacco plants contain the cellular elements necessary to respond systemically to a wound stress. Apparently, the tobacco PI-II gene identified here is devoid of the proper *cis*-acting sequences necessary for responding to systemic signalling. In fact, neither the 1690 bp of the 5′ promoter region, nor the 242 bp of the 3′ region of the tobacco *PI-II* gene, show extended regions with high similarity to the analogous regions of the inhibitors of tomato [6] or potato [25]. The 3′-flanking regions of the tomato and potato genes, share a short palindromic sequence, 30 to 40 bp upstream of the polyadenylation signal, which has been proposed to be important in the systemic response to

		systemic	elicitor	
potato	-165	AAGCGTAA--GTACCTTGCC		-148
tobacco	-182	AAGCaaccatGTACCTcGCC		-163

Fig. 4. Conserved elicitor-responsive element in the tobacco *PI-II* promoter. The left part of the sequence of the potato *PI-II* gene involved in systemic induction (systemic) and the right part involved in elicitor-induced expression (elicitor) are given [16]. Nucleotides of the tobacco PI-II promoter identical to the potato sequence are in capitals. Numbering is relative to the transcription start sites.

wounding [25]. This sequence is not present in the tobacco gene.

In the potato *PI-II* promoter region, a sequence of 10 bp located between −165 and −156 has been implicated in the systemic response to wounding. Immediately adjacent to this sequence lies a conserved sequence that has been identified as an elicitor- and light-inducible motif in a number of stress-inducible genes [16, and references therein]. A stretch with sequence similarity to this region is present in the promoter of the tobacco *PI-II* gene (position 1509 to 1528 in Fig. 1, or −182 to −163, relative to the transcription start site). The elicitor motif of the potato gene is almost perfectly conserved in the tobacco promoter, whereas the 'systemic response domain' is only partly conserved (Fig. 4).

Interestingly, the far-upstream region from 633 to 858 (−1057 to −832, relative to the transcription start site, underlined in Fig. 1) contains a stretch with high sequence similarity to a comparable region in the promoter of the stress-inducible Cu/Zn-superoxide dismutase (SOD) gene of *N. plumbaginifolia* [11]. Figure 5 shows a comparison of the respective promoter regions. Several direct repeats are present in the region of the *SOD* promoter (Fig. 5, repeats C and D). Surprisingly, in spite of the sequence similarity between the two regions, the repeats of the *SOD* gene are not conserved in the *PI-II* gene. However, other, more extended direct repeats are present in the *PI-II* promoter (Fig. 5, repeats A and B). Because of a palindromic sequence, also an inverted repeat is present, superimposed on part of the direct repeats (Fig. 5, open arrows). The conservation of the primary sequence and the presence of various direct and indirect repeats in the far-upstream regions of the promoters of two stress-induced genes from different plant species hints at a function for these regions in induced gene expression. Analyses of the promoter se-

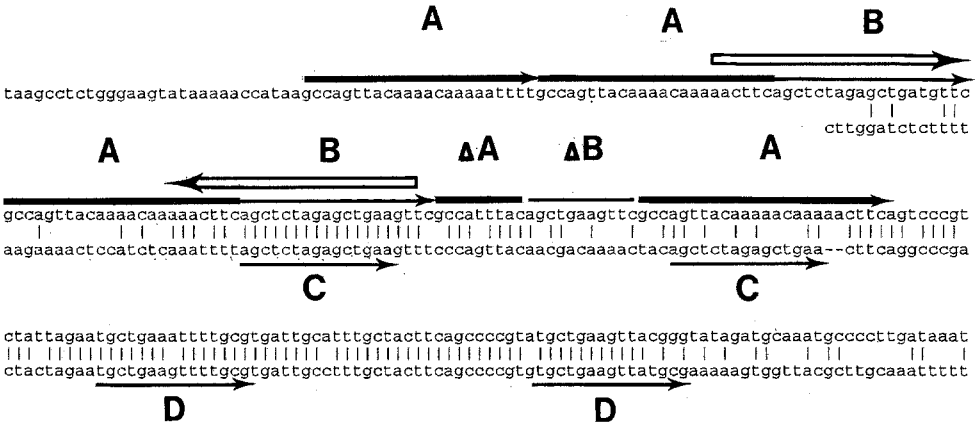

Fig. 5. Comparison of far-upstream promoter regions of genes encoding tobacco PI-II and *N. plumbaginifolia* Cu/Zn-SOD. The sequence in the top lines shows nucleotides 604 to 885 (−1087 to −866, relative to the transcription start site) of clone gPI2-1, while the bottom lines show the sequence of the promoter of the SOD gene from −741 to −542, relative to the transcription start site [11]. The two sequences are aligned to show identical residues, indicated by bars. Direct repeats A and B in the *PI-II* gene are indicated by thick and thin arrows/overlining the *PI-II* sequence, respectively. ΔA and ΔB indicate partial repeats containing the 5′ part of repeat A and the 3′ part of repeat B, respectively. The open arrows indicate inverted repeats superimposed on the direct repeats in the *PI-II* gene. Repeats C and D in the *SOD* gene are indicated by arrows underneath the *SOD* sequence.

quences responsible for induction by wounding via expression of reporter gene fusions in transgenic plants are in progress.

Acknowledgements

We are grateful to Dr C. A. Ryan for providing us with the tomato cDNA clone pT2-47. T. B. was a recipient of post-doctoral fellowships from the Ministerio de Educación y Ciencia (Spain) and from the European Communities. J. M. B. was a recipient of a fellowship from the Generalitat Valenciana (Spain).

References

1. Atkinson AH, Heath RL, Simpson RJ, Clarke AE, Anderson MA: Proteinase inhibitors in *Nicotiana alata* stigmas are derived from a precursor protein which is processed into 5 homologous inhibitors. Plant Cell 5: 203–213 (1993).
2. Bowles DJ: Defense-related proteins in higher plants. Annu Rev Biochem 59: 873–907 (1990).
3. Brederode FT, Linthorst HJM, Bol JF: Differential induction of acquired resistance and PR gene expression in tobacco by virus infection, ethephon treatment, UV light and wounding. Plant Mol Biol 17: 1117–1125 (1991).
4. Bryant J, Green TR, Gurusaddaiah T, Ryan CA: Proteinase inhibitor II from potatoes: isolation and characterization of its protomer components. Biochemistry 15: 3418–3424 (1976).
5. Dougherty WG, Semler BL: Expression of virus-encoded proteinases: functional and structural similarities with cellular enzymes. Microbiol Rev 57: 781–822 (1993).
6. Graham JS, Pearce G, Merryweather J, Titani K, Ericsson LH, Ryan CA: Wound-induced proteinase inhibitors from tomato leaves. II. The cDNA-deduced primary structure of pre-inhibitor II. J Biol Chem 260: 6561–6564 (1985).
7. Greenblatt HM, Ryan CA, James MNG: Structure of the complex of *Streptomyces griseus* proteinase b and polypeptide chymotrypsin inhibitor-I from Russet Burbank potato tubers at 2.1 Å resolution. J Mol Biol 205: 201–228 (1989).
8. Green TR, Ryan CA: Wound-induced proteinase inhibitors in plant leaves: a possible defense mechanism against insects. Science 175: 776–777 (1972).
9. Hass GM, Hermodson MA, Ryan CA, Gentry L: Primary structures of two low molecular weight proteinase inhibitors from potatoes. Biochemistry 21: 752–756 (1982).
10. Heitz T, Geoffroy P, Stintzi A, Fritig B, Legrand M: cDNA cloning and gene expression analysis of the microbial proteinase inhibitor of tobacco. J Biol Chem 268: 16987–16992 (1993).
11. Herouart DM, Van Montagu MM, Inze DM: Redox-activated expression of the cytosolic copper/zinc superoxide dismutase gene in Nicotiana. Proc Natl Acad Sci USA 90: 3108–3112 (1993).
12. Keil M, Sanchez-Serrano J, Schell J, Willmitzer L: Primary structure of a proteinase inhibitor II gene from potato (*Solanum tuberosum*). Nucl Acids Res 14: 5641–5650 (1986).
13. Laskowski M Jr, Kato I, Ardelt W, Cook J, Denton A, Empie MW, Kohr WJ, Park SJ, Parks K, *et al.*: Ovomucoid third domains from 100 avian species: isolation, sequences, and hypervariability of enzyme-inhibitor contact residues. Biochemistry 26: 202–221 (1987).
14. Linthorst HJM, van Loon LC, van Rossum CMA, Mayer A, Bol JF, van Roekel JSC, Meulenhoff EJS, Cornelissen BJC: Analysis of acidic and basic chitinases from tobacco and petunia and their constitutive expression in transgenic tobacco. Mol Plant-Microbe Interact 3: 252–258 (1990).
15. Linthorst HJM, Brederode FT, van der Does C, Bol JF: Tobacco proteinase inhibitor-I genes are locally, but not systemically induced by stress. Plant Mol Biol 21: 985–992 (1993).
16. Palm CJ, Costa MA, An G, Ryan CA: Wound-inducible nuclear protein binds DNA fragments that regulate a proteinase inhibitor II gene from potato. Proc Natl Acad Sci USA 87: 603–607 (1990).
17. Pearce G, Johnson S, Ryan CA: Purification and characterization from tobacco (*Nicotiana tabacum*) leaves of 6 small, wound-inducible, proteinase isoinhibitors of the potato inhibitor-II family. Plant Physiol 102: 639–644 (1993).
18. Pena-Cortes H, Sánchez-Serrano J, Rocha-Sosa M, Willmitzer L: Systemic induction of proteinase inhibitor-II gene expression in potato plants by wounding. Planta 174: 84–89 (1988).
19. Plunkett G, Senear DF, Zuroske G, Ryan CA: Proteinase inhibitor I and II from leaves of wounded tomato plants: purification and properties. Arch Biochem Biophys 213: 463–472 (1982).
20. Richardson M: The complete amino acid sequence and the trypsin reactive (inhibitory) site of the major proteinase inhibitor from the fruits of aubergine (*Solanum melongena* L.). FEBS Lett 104: 322–326 (1979).
21. Ryan CA: Protease inhibitors in plants: genes for improving defenses against insects and pathogens. Annu Rev Phytopath 28: 425–449 (1990).
22. Sánchez-Serrano J, Schmidt R, Schell J, Willmitzer L: Nucleotide sequence of proteinase inhibitor II encoding cDNA of potato (*Solanum tuberosum*) and its mode of expression. Mol Gen Genet 203: 15–20 (1986).
23. Sánchez-Serrano JJ, Keil M, O'Connor A, Schell J,

Willmitzer L: Wound-induced expression of a potato proteinase inhibitor II gene in transgenic tobacco plants. EMBO J 6: 303–306 (1987).

24. Taylor BH, Young RJ, Scheuring CF: Induction of a proteinase inhibitor II-class gene by auxin in tomato roots. Plant Mol Biol 23: 1005–1014 (1993).
25. Thornburg RW, An G, Cleveland TE, Johnson R, Ryan CA: Wound-inducible expression of a potato inhibitor II-chloramphenicol acetyltransferase gene fusion in transgenic tobacco plants. Proc Natl Acad Sci USA 84: 744–748 (1987).
26. Wildon DC, Thain JF, Minchin PEH, Gubb IR, Reilly AJ, Skipper YD, Doherty HM, O'Donnell PJ, Bowles DJ: Electrical signalling and systemic proteinase inhibitor induction in the wounded plant. Nature 360: 62–65 (1992).