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

Teresa Balandin, Chris van der Does¹, José-Maria Bellés Albert², 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: ¹Department of Molecular Microbiology, Groningen University, P.O. Box 14, 9750 AA Haren, The Netherlands; ²Dep. Biotecnologia, 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 (gPIf2-1).
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 32P-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