

G. candidum was not transformed with any of the plasmids tested. Several possible explanations such as DNA degradation, DNA uptake problems, nonintegration of the plasmids, or even lethal homologous integration may be advanced. Degradation prior to uptake of transforming DNA was ruled out because intact plasmids could be visualized on gels after incubation with *G. candidum* protoplasts. Generally, fungal transformation systems employing protoplasts are much the same with only minor variations, but it is possible that *G. candidum* may require some as yet unidentified factor for successful transformation to occur.

Acknowledgments. The first author was supported in part by a USDA National Needs Graduate Fellowship Program Grant No. 84-GRAD-9-0035 and by a University of California, Riverside Graduate Council Fellowship.

References

- Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (1987) Current protocols in molecular biology. Greene Publishing Associates and Wiley-Interscience, New York
- Ballance DI (1986) Sequences important for gene expression in filamentous fungi. *Yeast* 2:229-236
- Beach D, Nurse P (1981) High-frequency transformation of the fission yeast *Schizosaccharomyces pombe*. *Nature* 290:140-142
- Butler L, Fogle D (1988) *Galactomyces citri-aurantii* a newly found teleomorph of *Geotrichum citri-aurantii* the cause of sour rot of citrus fruit. *Mycotaxon* 33:197-212
- Crouse GF, Frischauf A, Lehrach H (1983) An integrated and simplified approach to cloning into plasmids and single-stranded phages. *Methods Enzymol* 101:78-89
- Davies LC, Flach W (1977) Differential binding of methyl benzimidazol-2-yl carbamate to fungal tubulin as a mechanism of resistance to this antimitotic agent in mutant strains of *Aspergillus nidulans*. *J Cell Biol* 72:174-193
- Devereux J, Haeblerli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res* 12:387-395
- Elwood HJ, Olson GJ, Sogin ML (1985) The small subunit ribosomal RNA gene structure from the hypotrichous ciliates *Ovtrichia nora* and *Stylomastix pustulata*. *Mol Biol Evol* 2:399-410
- Frischaut A-M, Lehrach H, Poustka A, Murray N (1983) Lambda replacement vectors carrying polylinker sequences. *J Mol Biol* 170:827-842
- Gallwitz D, Halfter H, Mertins P (1987) Splicing of mRNA precursors in yeast. In: Kinghorn JR (ed) Gene structure in eukaryotic microbes. Special Publications of the Society for General Microbiology, Oxford, Washington DC, pp 27-40
- Gurr SJ, Unkles SE, Kinghorn JR (1987) The structure and organization of nuclear genes of filamentous fungi. In: Kinghorn JR (ed) Gene structure in eukaryotic microbes. Special Publications of the Society for General Microbiology, Oxford, Washington DC, pp 93-139
- Harper JF, Mages W (1988) Organization and structure of *Vulvox* β -tubulin genes. *Mol Gen Genet* 213:315-324
- Henikoff S (1984) Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* 28:351-359
- Hirakawa Y, Toda T, Yanagida M (1984) The *ND43* gene of fission yeast encodes β -tubulin: a cold-sensitive *nda3* mutation reversibly blocks spindle formation and chromosome movement in mitosis. *Cell* 39:349-358
- Kaster KR, Burgett SG, Ingolia TD (1984) Hygromycin B resistance as dominant selectable marker in yeast. *Curr Genet* 8:353-358
- Keen NT, Tamaki S, Kobayashi D, Trottlinger D (1988) Improved broad-host-range plasmids for cloning in Gram-negative bacteria. *Gene* 70:191-197
- Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
- May GS, Gambino J, Weatherbee JA, Morris NR (1985) Identification and functional analysis of beta-tubulin genes by site specific integrative transformation in *Aspergillus nidulans*. *J Cell Biol* 101:712-719
- May GS, Tsang L-S, Smith H, Fidel S, Morris NR (1987) *Aspergillus nidulans* β -tubulin genes are unusually divergent. *Gene* 55:231-243
- May GS, Waring RB, Morris NR (1990) Increasing *tubC* β -tubulin synthesis by placing it under the control of a *benA* β -tubulin upstream sequence causes a reduction in *benA* β -tubulin level but has no effect on microtubule function. *Cell Motil Cytoskeleton* 16:214-220
- Mitchison JM (1970) Physiological and cytological methods for *Schizosaccharomyces pombe*. *Methods Cell Physiol* 4:131-165
- Neff NF, Thomas JH, Grisafi P, Botstein D (1983) Isolation of the β -tubulin gene from yeast and demonstration of its essential function in vivo. *Cell* 33:211-219
- Orbach MJ, Porro EB, Yanofsky C (1986) Cloning and characterization of the gene for β -tubulin from a benomyl-resistant mutant of *Neurospora crassa* and its use as a dominant selectable marker. *Mol Cell Biol* 6:2452-2461
- Panaccione DG, Hanan RM (1990) Characterization of two divergent β -tubulin genes from *Colletotrichum graminicola*. *Gene* 86:163-170
- Pastell J, Kafatos FC (1984) A convenient and adaptable package of computer programs for DNA and protein sequence management, analysis and homology determination. *Nucleic Acids Res* 12:643-655
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74:5463-5467
- Sharp PM, Li W-H (1986) An evolutionary perspective on synonymous codon usage in unicellular organisms. *J Mol Evol* 24:28-38
- Sharp PM, Tuohy TMF, Mosurski KR (1986) Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. *Nucleic Acids Res* 14:5125-5143
- Sherwood JE, Sommerville SC (1989) Molecular characterization of a β -tubulin gene from *Erysiphe graminis*. *Phytopathology* 79:1205 Abstract
- Smith HA, Gorman JW, Koltin Y, Gorman JA (1990) Functional expression of the *Candida albicans* β -tubulin gene in *Saccharomyces cerevisiae*. *Gene* 90:115-123
- Sogin ML, Miotto K, Miller L (1986) Primary structure of the *Neurospora crassa* small subunit ribosomal RNA coding region. *Nucleic Acids Res* 14:9540
- Sorensen MA, Kurland CG, Pederson S (1989) Codon usage determines translation rate in *Escherichia coli*. *J Mol Biol* 207:365-377
- Tamaki SJ, Gold S, Robeson M, Manulis S, Keen NT (1988) Structure and organization of the *pel* genes from *Erwinia chrysanthemi* EC16. *J Bacteriol* 170:3468-3478
- Turgeon BG, Garber RC, Yoder OC (1987) Development of a fungal transformation system based on selection of sequences with promoter activity. *Mol Cell Biol* 7:3297-3305
- Vieira J, Messing J (1987) Production of single-stranded plasmid DNA. *Methods Enzymol* 153:3-11
- Vollmer SJ, Yanofsky C (1986) Efficient cloning of genes of *Neurospora crassa*. *Proc Natl Acad Sci USA* 83:4869-4873
- Weatherbee JA, Morris NR (1984) *Aspergillus* contains multiple tubulin genes. *J Biol Chem* 259:15452-15459
- Yoder OC (1988) *Cochliobolus heterostrophus*, cause of corn leaf blight. *Adv Plant Pathol* 6:93-112

Communicated by C. van den Hondel

Pathogenesis-related protein 4 is structurally homologous to the carboxy-terminal domains of hevein, Win-1 and Win-2

Leslie Friedrich, Mary Moyer, Eric Ward, and John Ryals

Agricultural Biotechnology Research Unit, CIBA-GEIGY Corporation, P.O. Box 12257 Research Triangle Park, NC, 27709-2257, USA

Received May 21, 1991

Summary. The extracellular, acidic pathogenesis-related protein, PR-4, was purified to homogeneity from leaves of *Nicotiana tabacum* infected with tobacco mosaic virus (TMV) and characterized by partial amino acid sequencing. Complementary DNA clones encoding PR-4 were isolated using an oligonucleotide probe based on the sequence of one of the peptides. The deduced PR-4 protein sequence was found to be related to a family of proteins including hevein and Win-1, which have an amino-terminal lectin domain and a carboxy-terminal domain of unknown function. PR-4 is homologous to the carboxy-terminus of these proteins but does not contain the lectin domain. Thus, the organization of the PR-4 family of proteins is similar to that of the plant chitinase family, in that both contain structural subclasses characterized by the presence or absence of an amino-terminal lectin domain. This observation is consistent with the proposal that the DNA encoding the lectin domain may be capable of transposing to form new genes encoding proteins of more complex, multi-domain structure. The expression of PR-4 mRNA was found to increase dramatically in response to TMV infection and the time course of RNA accumulation was similar to that of other PR proteins.

Key words: cDNA cloning Tobacco Defense genes Lectins

Introduction

When phytopathogenic bacteria, fungi or viruses infect plants, the plant may respond by forming a necrotic lesion surrounding the infection site. The biochemical basis of this hypersensitive reaction (HR) is unknown but it functions to limit the systemic spread of the pathogen (Fritig et al. 1987; Van Loon et al. 1990). Along with the HR, other cellular processes occur that may reduce primary and secondary pathogen infections, in-

cluding the production of antibiotic metabolites (e.g. phytoalexins) and the synthesis of extracellular and vacuolar proteins (Lamb et al. 1989). Among the newly synthesized proteins are enzymes such as chitinase and β -1,3-glucanase that may act directly by hydrolyzing components of the fungal cell wall (Boller 1988), and other enzymes such as peroxidase that may act indirectly by strengthening the plant cell wall, making the tissue less susceptible to infection (Bowles 1990). Other newly synthesized proteins, such as pathogenesis-related (PR) proteins also accumulate in the extracellular spaces of the infected tissue to levels approaching 1% of the total cellular protein (Van Loon 1985; Bol et al. 1990; Carr and Klessig 1990).

The PR-proteins were first described as acidic, extracellular proteins of relatively low molecular weight that accumulate in tobacco leaves reacting hypersensitively to tobacco mosaic virus (TMV) infection (Van Loon and Van Kammen 1970; Gianinazzi et al. 1970). In tobacco, 10 PR-proteins, originally designated PR-1a, PR-1b, PR-1c, PR-2, PR-3, PR-4, PR-5, PR-6, PR-7, PR-8, PR-9, and PR-10, have been described, which migrate with decreasing mobility on non-denaturing polyacrylamide gels (Van Loon et al. 1983). Recently, a change in nomenclature has been suggested based on the serological relationships between the proteins (Van Loon et al. 1987). The designations PR-1a, PR-1b and PR-1c remain the same; PR-2, PR-3, PR-4, PR-5, PR-6, PR-7, PR-8, PR-9, and PR-10 are now termed PR-2a, PR-2b and PR-2c, respectively; PR-P and PR-Q become PR-3a and PR-3b, respectively; PR-R becomes PR-4; and PR-S becomes PR-5. In this report we will refer to the proteins using this new nomenclature. While no activity has been described for either the PR-1 or PR-4 proteins, it has been demonstrated that the PR-2 family of proteins have in vitro β -1,3-glucanase activity (Kauffmann et al. 1987). PR-3 proteins have chitinase activity (Legrand et al. 1987) and the PR-5 proteins are structurally similar to an alpha-amylase/trypsin inhibitor from maize (Cornelissen et al. 1986; Pierpoint et al. 1987; Richardson et al. 1987).

In recent years, cDNA and genomic clones encoding PR-1, PR-2, PR-3 and PR-5 have been isolated, which has facilitated studies of the molecular biology of the

Offprint requests to: J. Ryals

31951

MGG
© Springer-Verlag 1991

proteins they encode (Bol et al. 1990; Carr and Klessig 1990 for recent reviews). However, neither cDNAs nor genomic clones encoding the PR-4 proteins have been described. Here, we report isolation of the PR-4 protein and its characterization by partial amino acid sequencing. The structure of the PR-4 proteins, deduced from cDNA clones encoding two isoforms, was found to be homologous to the carboxy-terminal end of both hevein, a major protein constituent of rubber tree latex (Broek-aert et al. 1990) and Win-1 and Win-2, two wound-inducible proteins from potato tubers (Stanford et al. 1989). These proteins have an amino-terminal, cysteine-rich lectin domain that is missing in PR-4, supporting the suggestion that DNA encoding the lectin domain is capable of recombining into target genes to form multi-functional proteins (Shinshi et al. 1990).

Materials and methods

Plant materials and virus infection. *Nicotiana tabacum* cv. Xanthi.nc plants were grown in a greenhouse and inoculated with the common strain (U1) of tobacco mosaic virus (TMV) or mock-inoculated as previously described (Payne et al. 1990a).

Protein purification and amino acid sequencing. Intercellular fluid (ICF) from TMV-infected plants was isolated as described by Parent and Asselin (1984). Concentrated ICF was fractionated on a Vydac phenyl reverse phase column with a linear 20–80% gradient of acetonitrile: isopropanol (1:1) in 0.1% trifluoroacetic acid (TFA). PR-4 co-eluted from this column with PR-1a, -1b, and -1c as determined by electrophoresis on native polyacrylamide gels. PR-4 was separated from PR-1a, -1b and -1c using a Brownlee Labs AX-300 anion exchange column equilibrated in 10 mM ammonium acetate, pH 7.0, and eluted with a linear gradient of ammonium acetate (pH 7.0) from 10 to 350 mM. Proteins were analyzed by electrophoresis under non-denaturing conditions using a 17–27% mini-Septagel (Integrated Separation Systems) in TRIS-glycine buffer, pH 8.3, and under denaturing conditions using a 10–20% Novex mini-gel in TRIS-Tricine buffer. Novex wide range protein standards were used for the polyacrylamide gel electrophoresis (SDS-PAGE) and proteins were visualized by Coomassie Brilliant Blue R (Sigma) staining.

Purified PR-4 was digested with either trypsin or endoprotease Asp-N (Boehringer-Mannheim); peptides were separated on a Brownlee Labs RP-300 reverse phase column with a linear gradient of 0–60% acetonitrile: isopropanol (1:1) in 0.1% TFA. Automated Edman degradations were performed using an Applied Biosystems 470A sequencer. Phenylthiohydantoin (PTH) amino acids were identified with an on-line Applied Biosystems 120A PTH analyzer.

Isolation of cDNA clones encoding PR-4. A cDNA library was prepared from RNA extracted from TMV-infected tobacco leaves as described by Payne et al. (1990a). Putative cDNA clones were identified by plat-

ing the cDNA library and screening duplicate filters with a mixed pool of oligonucleotides having 32-fold degeneracy and the following design: 5' G G Y T T R T C N G C R T C C C A 3' (where Y represents a pyrimidine, R represents a purine and N represents a mixture of all four bases). Oligonucleotides were synthesized using β -cyanoethylphosphoramidite chemistry on the Applied Biosystems 380A DNA synthesizer and 5' end-labeled with T4 polynucleotide kinase and [^{32} P]ATP (>6000 Ci/mmol). The filters were hybridized and washed under low stringency conditions (125 mM NaCl, 1% SDS, 40 mM sodium phosphate, pH 7.2, 1 mM EDTA at 50°C) (Church and Gilbert 1984).

Inserts from the recombinant phage were amplified via PCR using the GeneAmp kit from Perkin-Elmer/Cetus as recommended by the manufacturer and subcloned into the pBluescript plasmid (Stratagene). The DNA sequence of the inserts was determined by dideoxy sequencing of double-stranded templates (Hattori and Sakaki 1986).

RNA and DNA hybridization. Total RNA was prepared as described (Lagrimini et al. 1987) from TMV-infected leaves or from leaves acropetal to the inoculated leaves (secondary leaves). Total DNA was extracted from tobacco leaves according to Ausubel et al. (1987). RNA and DNA blot analyses were performed as described (Ausubel et al. 1987) using the PR-4a cDNA insert as probe.

Primer extension mapping. Primer extension reactions were carried out on total RNA as described by Metraux et al. (1989) using an oligonucleotide with the following sequence: 5' GACGAGGCTGTCTGGCATC-3'.

Results

Purification and characterization of PR-4 protein

A total protein extract of the intercellular fluid from TMV-infected tobacco leaves was separated by reverse phase HPLC resulting in a single peak (Fig. 1A) that contained the PR-1 and PR-4 proteins as determined by their relative mobility in native polyacrylamide gels. This peak was collected and PR-4 was further separated using ion-exchange chromatography. The PR-4 protein eluted as a single peak at 5 min (Fig. 1B) and was well resolved from the other proteins in the mixture. The PR-4 peak was determined to be at least 90% pure by electrophoresis on native and SDS polyacrylamide gels (Fig. 1C, D). However, in another experiment, in which a preparation of PR-4 protein was chromatographed on a reverse phase phenyl column, we obtained evidence for the presence of either two isoforms of PR-4 or a second contaminating protein virtually co-eluting with PR-4 (data not shown).

The purified protein was found to be blocked at the amino terminus and did not yield an amino-terminal sequence even after enzymatically deblocking the protein with pyroglutamate aminopeptidase. Therefore, purified

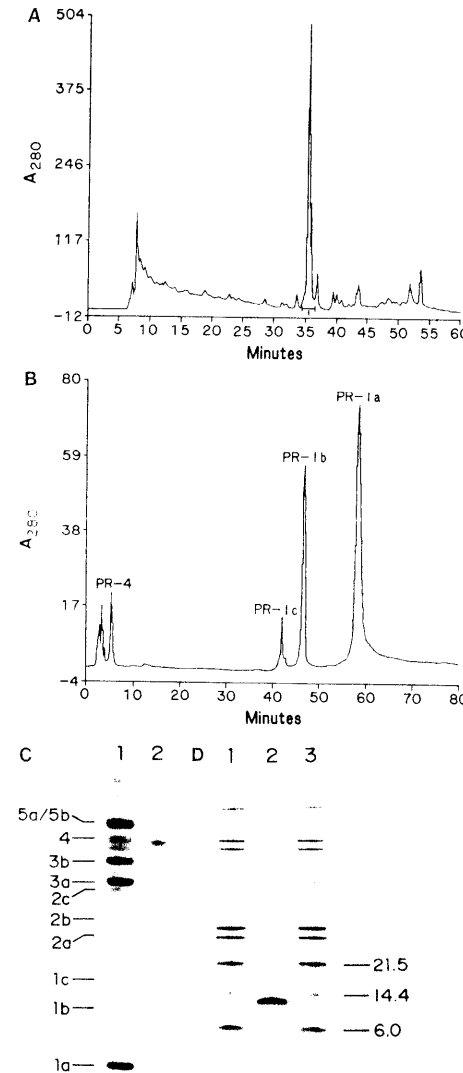


Fig. 1A–D. Purification of PR-4 protein. Intracellular fluid (ICF) from TMV-infected tobacco leaves was fractionated using reverse-phase HPLC on a phenyl column (Panel A). The peak containing PR-4 and PR-1a, -1b, and -1c (indicated by the bar under the peak in panel A at 35 min) was collected and PR-4 was separated from the PR-1 proteins by anion exchange HPLC (Panel B). Purified PR-4 was electrophoresed on native (Panel C) and denaturing (Panel D) polyacrylamide gels. Panel C: Lane 1, ICF from TMV-infected tobacco leaves; Lane 2, the purified PR-4 fraction. Panel D: Lanes 1 and 3, molecular weight markers (sizes in kDa); Lane 2, purified PR-4.

PR-4 protein was digested with the endoproteases Asp-N and trypsin and the resulting peptides were separated using reverse phase HPLC. Amino acid sequences were obtained for six of the purified peptides and these data are shown as the underlined residues in Fig. 2. In early attempts at sequencing, reduction and pyridylethylation were carried out to modify cysteine residues. However, much of the protein was lost in subsequent desalting, and this step was omitted when the peptides were prepared. Therefore, cysteine residues were not detected in the peptide sequences.

Isolation of cDNA clones encoding PR-4

A 32-fold degenerate mixture of 17-base oligonucleotides was used as a probe to screen a cDNA library prepared using RNA isolated from TMV-infected tobacco leaves (Payne et al. 1990a). The oligonucleotide mixture was designed to be complementary to mRNA that would encode the amino acid sequence WDADKP, derived from a portion of the Tryp48.8 peptide. Approximately 300,000 plaques were screened in duplicate and twenty-four positively hybridizing plaques were isolated. Seven of these phage were purified and the cDNA inserts were subcloned and sequenced. Among the seven clones, two types of cDNA sequences, designated PR-4a (3 of 7) and PR-4b (4 of 7) were found. The sequences of these two types of cDNAs compiled from the complete analysis of all 7 clones are shown in Fig. 2. PR-4a and PR-4b differed at only six nucleotide positions in the predicted coding region. The amino acid sequences of all of the peptides were found in the deduced amino acid sequence, confirming that the cDNA encodes the PR-4 protein. The results of primer extension mapping indicated that the longest cDNA extends to within approximately 15–20 bases of the 5' end of the mRNA (data not shown).

The cDNAs encoded proteins with 147 amino acid residues, having calculated molecular masses of 16222 for PR-4a and 16236 for PR-4b. The molecular mass of the mature PR-4 protein determined from SDS-PAGE (Fig. 1D) is about 13500. So, apparently the PR-4 protein is synthesized as a pre-protein which includes a signal peptide of about 25 residues. The first 25 amino acids of the deduced protein have structural characteristics similar to the signal peptides of other PR-proteins (Metraux et al. 1989; Payne et al. 1988, 1990a, b; Pierpoint et al. 1987; Ward et al. 1991), and by analogy cleavage of the putative signal should occur between the alanine at position 25 and the glutamine at position 26. If this is the case, a pyroglutamate modification could protect the protein from N-terminal Edman degradation, which would be consistent with our failure to obtain N-terminal protein sequences. However, it is not clear why attempts at enzymatic deblocking were unsuccessful. Assuming cleavage of a signal peptide, both mature proteins would have a calculated isoelectric point of 6.28 and molecular masses of 13483 (PR-4a) or 13479 (PR-4b) which are in close agreement with the size estimated from SDS-PAGE.

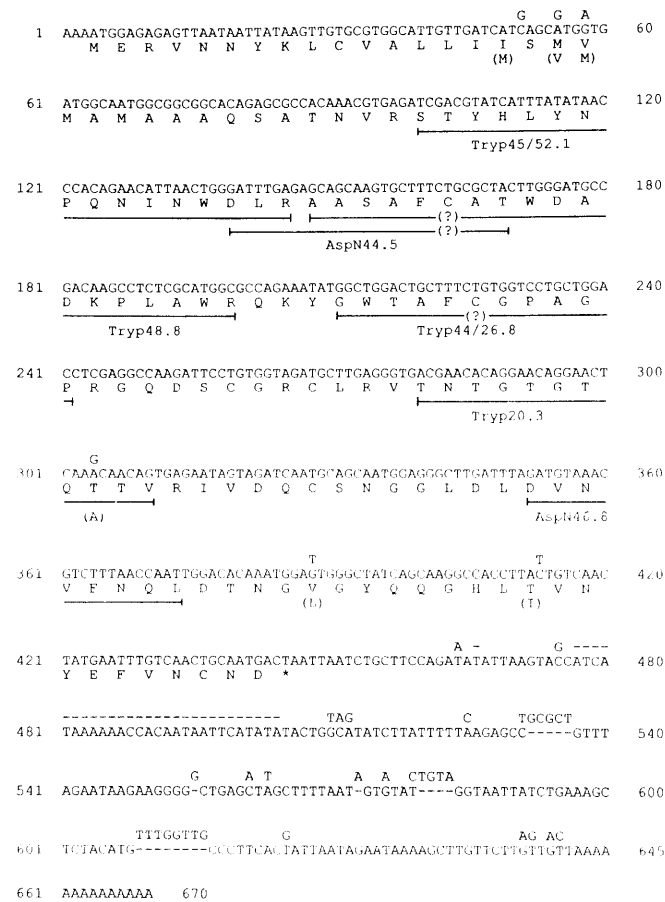


Fig. 2. Nucleotide and deduced amino acid sequences of PR-4 cDNA clones. The DNA sequence was compiled from 3 independent isolates and represents the PR-4a clone. The deduced amino acid sequence of the PR-4 protein is given below the nucleotide sequence. The DNA sequence for PR-4b was compiled from 4 independent isolates. Differences between PR-4b and PR-4a are indicated in parentheses above (nucleotides) and below (amino acids) the PR-4a sequences. Numbers refer to the nucleotide sequence. The stop codon at the end of the coding region is marked with an asterisk. Tryptic and Asp-N peptides isolated from digests of the PR-4 protein are indicated by the solid lines below the amino acid sequence. These sequences have been deposited with GenBank.

Three of the nucleotide differences between PR-4a and PR-4b result in amino acid changes in the putative signal peptide and the other three differences result in amino acid changes in the mature protein sequence. The peptide Tryp20.3 contained a threonine, indicating that it was derived from the PR-4a protein; however, none of the other peptides were from parts of the protein that would allow a discrimination between the two forms. Therefore, there is no positive proof that the PR-4b protein co-purified with PR-4a.

The deduced amino acid sequence of PR-4 was found to share approximately 75% identity (Fig. 3) with those of two wound-induced proteins (Win-1 and Win-2) from potato (Stanford et al. 1989) and with hevein from rubber tree latex (Broekaert et al. 1990). However, these proteins contain additional sequences that comprise an

amino-terminal "lectin domain", which PR-4 lacks. The lectin domain is a highly conserved, cysteine-rich region of approximately 40 amino acids that is found in other proteins including wheat germ agglutinin (Wright et al. 1984) and the basic, class I chitinase from tobacco (Shinshi et al. 1990). Figure 4 schematically illustrates the domain arrangements of several of these proteins. The enzymatic activities of the PR-4 protein and of the carboxy-terminal domains of hevein and the Win proteins are not known.

Southern blot analysis of tobacco genomic DNA using PR-4 cDNA as a probe was carried out to determine the complexity of this gene family. Enzymes that do not restrict PR-4 cDNA were used in the analysis. The data suggest that the PR-4 gene family is relatively small, comprised of possibly 3-4 members (data not shown).

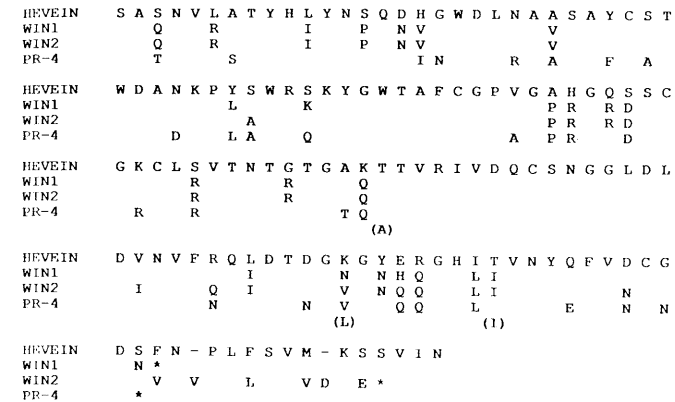


Fig. 3. Comparison of the amino acid sequence of PR-4 to related proteins. The amino acid sequences are derived from hevein, residues 69-204 (Broekaert et al. 1990), Win-1, residues 78-200 and Win-2, residues 78-211 (Stanford et al. 1989). The three amino acid differences found in PR-4b are shown in parentheses. Only amino acids differing from the hevein sequence are shown. Asterisks denote the end of the proteins.

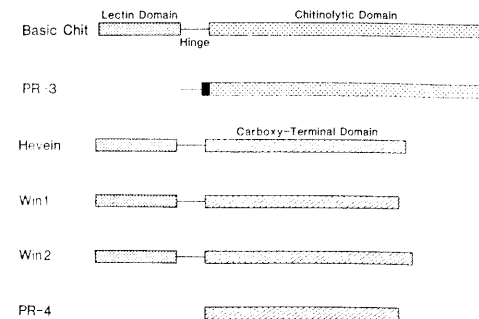


Fig. 4. Schematic representation of the domains present in various proteins: tobacco basic chitinase (Shinshi et al. 1990), hevein (Broekaert et al. 1990), Win-1 and Win-2 (Stanford et al. 1989) and PR-4. The lectin domain is stippled, the hinge region is shown as a thin line between the two domains, the chitinolytic domain is dotted and the domain with homology to PR-4 is hatched.

This result is consistent with those for other tobacco PR-proteins, which have been shown to be encoded by gene families of similar size (Pfitzner and Goodman 1987; Payne et al. 1990b; Ward et al. 1991; Bol et al. 1990 and references therein).

Expression of PR-4 mRNA

The expression profile of PR-4 mRNA was examined by Northern blot analysis of RNA prepared from TMV-infected tobacco plants. On day 0, a group of plants was inoculated on several leaves with a suspension of TMV. At various times after inoculation, the treated leaves (primary) and leaves immediately acropetal (secondary) to the TMV-inoculated leaves were harvested and RNA was isolated from samples pooled from 3 plants. The PR-4 mRNA began to accumulate by day 2 in the primary leaves and reached its highest level

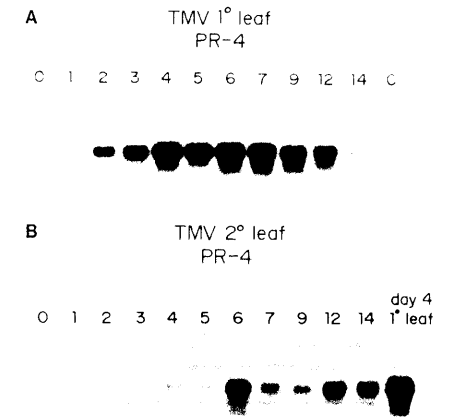


Fig. 5A, B. Time course of the expression of PR-4 mRNA in TMV-infected (1°, A) and uninfected (2°, B) tobacco leaves. Each lane contains 10 µg of total RNA extracted at various times post-inoculation. The RNA blots were probed with PR-4a cDNA. A day seven control (C) of mock-inoculated plants is included on the blot of 1° tissue. The numbers refer to days after infection.

6-7 days after infection (Fig. 5A). The accumulation of mRNA in the secondary leaves occurred more slowly and reached levels lower than those attained in the primary leaves (Fig. 5B). Previous experiments have shown similar patterns of expression for other PR-proteins (Ward et al. 1991; Ryals et al. 1991).

Discussion

In this report, we have detailed the purification and partial amino acid sequencing of the PR-4 protein and the

isolation of cDNA clones encoding the protein. Two forms of the PR-4 cDNA, PR-4a and PR-4b, which were 98.6% identical in nucleotide sequence within the coding sequence and 89.5% identical in overall structure, were isolated in roughly equivalent numbers. The occurrence of two isoforms of a particular protein in tobacco is common because its genome is stably amphidiploid. It has been suggested that the ancestral tobacco genome was derived by sexual hybridization between *N. tomentosiformis* and *N. sylvestris* followed by chromosomal duplication (Goodspeed and Clausen 1928). Thus, each of the subgenomes may account for one form of the PR-4 protein in tobacco. The PR-4 family is similar to the other tobacco PR-protein families, each of which is comprised of at least two members (Bol et al. 1990; Carr and Klessig 1990). Interestingly, the calculated isoelectric point for PR-4 is about 6.3 and at pH 7.0 its calculated net charge is about -1. This low net acidic charge and low charge-to-mass ratio probably accounts for the observation that the PR-4 protein routinely migrates as a diffuse band of low mobility in non-denaturing PAGE.

Plant chitinases have been grouped into at least three distinct classes based on their protein structures. The class I chitinase is a complex protein consisting of an amino-terminal, cysteine-rich lectin domain and a carboxy-terminal catalytic domain (Shinshi et al. 1990). The class II chitinase is a smaller protein that contains the catalytic portion of the protein but is missing the amino-terminal lectin domain (Payne et al. 1990a). It has been suggested that the more complex class I chitinase evolved from a proto-chitinase by insertion of the DNA encoding the lectin domain into the proto-chitinase gene, thereby adding a new carbohydrate-binding domain to the enzyme (Shinshi et al. 1990). The class II chitinase was proposed to have then evolved from the class I gene by excision of that DNA, leaving a 9 bp direct repeat in the class II gene (Shinshi et al. 1990). It was further proposed that related lectins containing a homologous, cysteine-rich region could have evolved by a similar transposition mechanism (Shinshi et al. 1990). The data presented here are consistent with this model, in that the PR-4 gene family including Win-1, Win-2 and hevein have related carboxy-terminal domains with or without associated amino-terminal lectin domains. The PR-4 family is, therefore, a structurally distinct example of the phenomenon previously observed in chitinases. Presumably other gene families encoding still different activities, will be found which also have members distinguished by the presence or absence of the lectin domain.

Acknowledgements. We would like to thank our colleagues Drs. Mary-Dell Chilton, Laura Privalle, Jim Ligon and Tom Luntz for their comments on the manuscript.

References

- Ausubel F, Brent R, Kingston R, Moore D, Seidman J, Smith J, Struhl K (1987) Current protocols in molecular biology. Wiley and Sons, New York
- Bol JF, Linthorst HJM, Cornelissen BJC (1990) Pathogenesis-related proteins induced by virus infection. *Annu Rev Phytopathol* 28:113-138
- Boller T (1988) Ethylene and the regulation of antifungal hydrolases in plants. In: Milfin BJ (ed) Oxford surveys of plant molecular biology, vol 5. Wiley and Sons, New York, pp 145-175
- Bowles D (1990) Defense-related proteins in higher plants. *Annu Rev Biochem* 59:873-907
- Broekaert W, Lee HJ, Kush A, Chua N-H, Raikhel N (1990) Wound-induced accumulation of mRNA containing a hevein sequence in laticifers of rubber tree (*Hevea brasiliensis*). *Proc Natl Acad Sci USA* 87:7633-7637
- Carr J, Klessig D (1990) The pathogenesis-related proteins of plants. In: Setlow J (ed) Genetic engineering of plants: Principles and methods, vol 11. Plenum Press, New York, pp 65-109
- Church G, Gilbert W (1984) Genomic sequencing. *Proc Natl Acad Sci USA* 81:1991-1995
- Cornelissen BJC, Hooft van Huijsduijnen RAM, Bol JF (1986) A tobacco mosaic virus-induced protein is homologous to the sweet-tasting protein thaumatin. *Nature* 321:531-532
- Cornelissen B, Horowitz J, Van Kan J, Goldberg R, Bol J (1987) Structure of tobacco genes encoding pathogenesis-related proteins from the PR-1 group. *Nucleic Acids Res* 15:6799-6811
- Fritig B, Kauffmann S, Dumas B, Geoffroy P, Kopp M, Legrand M (1987) Mechanism of the hypersensitive reaction of plants. In: Evered D, Harnett S (eds) Plant resistance to viruses. CIBA Foundation Symposia, vol. 133. Wiley and Sons, Chichester, pp 92-108
- Graninazzi S, Martin C, Vallee JC (1970) Hypersensibilité aux virus, températures et protéines solubles chez le *Nicotiana glauca*. Apparition de nouvelles macromolécules lors de la répression de la synthèse virale. *C R Acad Sci Paris D* 270:2283-2286
- Goodspeed TH, Clausen RE (1928) Interspecific hybridization in *Nicotiana*. VIII. The *syvestris-tomentosa-tabacum* hybrid triangle and its bearing on the origin of tobacco. *Univ Calif Publ Bot* 11:245-256
- Hattori M, Sakaki Y (1986) Dideoxy sequencing method using denatured plasmid templates. *Anal Biochem* 152:232-238
- Kauffmann S, Legrand M, Geoffroy P, Fritig B (1987) Biological function of "pathogenesis-related" proteins. Four PR-proteins of tobacco have 1,3- β -glucanase activity. *JMBIO J* 6:3209-3212
- Lagrimini M, Burkhardt W, Moyer M, Rothstein S (1987) Molecular cloning of complementary DNA encoding the lignin-forming peroxidase from tobacco: Molecular analysis and tissue-specific expression. *Proc Natl Acad Sci USA* 84:7542-7546
- Lamb CJ, Lawton MA, Dron M, Dixon RA (1989) Signals and transduction mechanisms for activation of plant defenses against microbial attack. *Cell* 56:215-224
- Legrand M, Kauffmann S, Geoffroy P, Fritig B (1987) Biological function of "pathogenesis-related" proteins: four tobacco PR-proteins are chitinases. *Proc Natl Acad Sci USA* 84:6750-6754
- Metraux JP, Burkhardt W, Moyer M, Dincher S, Middlesteadt W, Williams S, Payne G, Carnes M, Ryals J (1989) Isolation of a complementary DNA encoding a chitinase with structural homology to a bifunctional lysozymic chitinase. *Proc Natl Acad Sci USA* 86:896-900
- Parent JG, Asselin A (1984) Detection of pathogenesis-related proteins (PR or b) and of other proteins in the intercellular fluid of hypersensitive plants infected with tobacco mosaic virus. *Can J Bot* 62:564-569
- Payne G, Parks TD, Burkhardt W, Dincher S, Ahl P, Metraux JP, Ryals J (1988) Isolation of the genomic clone for pathogenesis-related protein 1a from *Nicotiana tabacum* cv. Xanthi nc. *Plant Molec Biol* 11:89-94
- Payne G, Ahl P, Moyer M, Harper A, Beck J, Meins F, Ryals J (1990a) Isolation of complementary DNA clones encoding pathogenesis-related proteins P and Q, two acidic chitinases from tobacco. *Proc Natl Acad Sci USA* 87:98-102
- Payne G, Ward F, Gaffney T, Goy P, Moyer M, Harper A, Meins F, Ryals J (1990b) Evidence for a third structural class of β -1,3-glucanase in tobacco. *Plant Mol Biol* 15:797-808
- Pfützner UM, Goodman HM (1987) Isolation and characterization of cDNA clones encoding pathogenesis-related proteins from tobacco mosaic virus-infected tobacco plants. *Nucleic Acids Res* 15:4449-4465
- Pierpoint WS, Tatham AS, Pappin DJC (1987) Identification of the virus-induced protein of tobacco leaves that resembles the sweet protein thaumatin. *Physiol Molec Plant Pathol* 31:291-298
- Richardson M, Valdes-Rodriguez S, Blanco-Labra A (1987) A possible function for thaumatin and a TMV-induced protein suggested by homology to a maize inhibitor. *Nature* 327:432-434
- Ryals J, Ward E, Metraux JP (1991) Systemic acquired resistance: an inducible defense mechanism in plants. In: Biochemistry and molecular biology of inducible enzymes and proteins in higher plants. Cambridge University Press, London, in press
- Shinshi H, Neuhaus JM, Ryals J, Meins F (1990) Structure of a tobacco endochitinase gene: evidence that different chitinase genes can arise by transposition of sequences encoding a cysteine-rich domain. *Plant Mol Biol* 14:357-368
- Stanford A, Bevan M, Northcot D (1989) Differential expression within a family of novel wound-induced genes in potato. *Mol Gen Genet* 215:200-208
- Van Loon LC (1985) Pathogenesis-related proteins. *Plant Molec Biol* 4:111-116
- Van Loon LC, Gianinazzi S, White RF, Abu-Jawdah Y, Ahl P, Antoniw J, Boller T, Camacho-Henriquez A, Conejero V, Coussirat JC, Goodman RN, Maiss E, Redolfi P, Wilson TM (1983) Electrophoretic and serological comparisons of pathogenesis-related (b) proteins from different plant species. *Neth J Plant Path* 89:292-303
- Van Loon LC (1989) Stress proteins in infected plants. In: Kosuge T, Nester EW (eds) Plant-Microbe Interactions: Molecular and Genetic perspectives. Macmillan Press, New York, pp 198-237
- Van Loon LC, Gerritsen YAM, Ritter CE (1987) Identification, purification, and characterization of pathogenesis-related proteins from virus-infected Samsun NN tobacco leaves. *Plant Molec Biol* 9:593-609
- Van Loon LC, Van Kammen A (1970) Polyacrylamide disk electrophoresis of the soluble leaf proteins from *Nicotiana tabacum* var. "Samsun" and "Samsun NN". II Changes in protein constitution after infection with tobacco mosaic virus. *Virology* 40:199-211
- Ward E, Payne G, Moyer M, Williams S, Dincher S, Sharkey K, Beck J, Taylor H, Ahl-Goy P, Meins F, Ryals J (1991) Differential regulation of β -1,3-glucanase mRNAs in response to pathogen infection. *Plant Physiol* 96:390-397
- Wright CS, Gavilanes F, Peterson DL (1984) Primary structure of wheat germ agglutinin isolectin 2. Peptide order deduced from X-ray structure. *Biochemistry* 23:280-287

Communicated by J. Schell