Tobacco and Tomato PR Proteins Homologous to *win* and Pro-Hevein Lack the “Hevein” Domain

Huub J. M. Linthorst1, Nadia Danhash2, Frans T. Brederode1, Jan A. L. Van Kan2, Pierre J. G. M. De Wit2, and John F. Bol1

1Department of Biochemistry, Leiden University, The Netherlands, and 2Department of Phytopathology, Wageningen Agricultural University, The Netherlands.
Received 10 June 1991. Accepted 22 August 1991.

Clones corresponding to tobacco pathogenesis-related (PR) proteins PR-4 and tomato PR protein P2 were isolated from phage cDNA libraries of tobacco infected with tobacco mosaic virus and tomato infected with *Cladosporium fulvum*, respectively. The probe used in these screenings was a polymerase chain reaction product, synthesized on phage DNA from the tobacco cDNA library, using a synthetic oligonucleotide primer whose sequence corresponded to the partial amino acid sequence available for P2. The different cDNA sequences from the tobacco and tomato clones contained open reading frames for small proteins with 80–90% amino acid sequence identity. Both tobacco PR-4 and tomato P2 are synthesized as precursor proteins, with an N-terminal signal peptide involved in extracellular targeting. The proteins are highly similar to putative wound-induced proteins of potato (*win*) and to the precursor protein of hevein. However, in contrast to the hevein pro-protein and *win* proteins, PR-4 and P2 do not contain N-terminal, chitin-binding “hevein” domains. The tobacco and tomato genomes contain a limited number of genes corresponding to PR-4 or P2, whose expression is induced upon infection with the above-mentioned pathogens.

Additional keyword: acquired resistance.

The hypersensitive response of plants to infection with certain pathogens is accompanied by the induced synthesis of a large number of proteins, among which are the so-called pathogenesis-related (PR) proteins. In Samsun NN tobacco, at least five groups of acidic PR proteins are induced by infection with tobacco mosaic virus (TMV); inducible basic isofoms of four of these groups (PRs 1, 2, 3, and 5) have been identified. In parenchymatic tissue and xylem elements of leaves, the acidic proteins accumulate in the extracellular space, whereas the basic proteins are targeted to the vacuoles (for recent reviews see Bol et al. 1990; Linthorst 1991). Several of these proteins are enzymes, capable of hydrolyzing β-1,3-glucans and chitin. Because these polysaccharides are major structural components of fungal cell walls, it has been proposed that glucanases (PR-2) and chitinases (PR-3) may be involved in acquired resistance to fungi, a phenomenon that frequently coincides with the hypersensitive response. In parallel with this view, it is also likely that the other PR proteins, for which no enzymatic activity or function has yet been found, may play a role in resistance to pathogens.

In recent years, cDNA and genomic sequences corresponding to PR proteins of tobacco (groups 1, 2, 3, and 5) have been obtained and it was established that extensive similarities exist with proteins from other plants, induced by hormone stress, pathogenesis, wounding, or other stresses (see Linthorst 1991; Linthorst et al. 1990a). PR proteins may also be involved in tissue-specific and developmentally regulated processes, since a number of PR and PR-like proteins have been found to accumulate in roots, older leaves, and flower organs during normal plant development (Fraser 1981; Lotan et al. 1989). Whether they have a function in defense in these organs is unknown.

In this paper, we present sequence data on cDNA clones corresponding to PR-4 proteins of tobacco and of the serologically related PR protein P2 of tomato (Joosten et al. 1990). This work reveals similarities and differences with putative wound-induced proteins of potato (*win* proteins, Stanford et al. 1989) and with pro-hevein, the precursor of the small, chitin-binding and wound-induced lectin of *Hevea brasiliensis* (Willd. ex ADR. Juss.) Muell. Arg. (Brockaert et al. 1990).

MATERIALS AND METHODS

Protein purification and sequencing. Protein P2 was purified from the intercellular fluid isolated from tomato cultivar CF4, inoculated with the avirulent race 5 of *Cladosporium fulvum* Cooke (Joosten et al. 1990). After reduction, pyridylethylolation, and digestion with trypsin, P2 fragments were isolated by high-performance liquid chromatography (HPLC) and sequenced by Edman degradation.

Polymerase chain reaction (PCR). PCRs were set up in which one primer was complementary to the T7 promoter, downstream of the 3′ end of the cDNA insert in the lambda ZAP cloning vector (Stratagene, La Jolla, CA). The degenerate second primer mixture consisted of oligo-
omers with sequences based on amino acid sequence data obtained from P2. Lambda ZAP DNA was isolated from an amplified library by polyethylene glycol precipitation of phage particles, followed by phenol extraction and ethanol precipitation. DNA amplification was performed during 40 cycles of sequential incubations at 94°C for 0.5 min, 52°C for 5 min, and 74°C for 1 min, in a 50-μl reaction mixture, containing 1 μg of DNA, isolated from the tobacco lambda ZAP cDNA library, 0.2 μM T7 promoter primer (Promega, Madison, WI), 1 μM degenerate primer mixture, and 2.5 units of Thermus aquaticus DNA polymerase. The reaction products were electrophoresed in 1.5% agarose gels and excised from the gel for labeling.

**Synthesis and analysis of cDNA libraries.** Isolation of poly(A)+ RNA from TMV-infected tobacco and tomato infected with *C. fulvum*, and the construction of lambda ZAP cDNA libraries were as described (Linhorst et al. 1990b; Van Kan et al., unpublished). Recombinant phage were screened by using a 32P-labeled PCR product as a probe. cDNA inserts were subcloned in M13 derivatives and sequenced, or were directly sequenced from denatured plasmid DNA, using T3, T7, or SP6 promoter primes (Promega; Sanger et al. 1977). Apart from the M13 sequencing primer, the primer mixture LS13 (see Results section) was also used for sequencing on single-stranded DNA.

**RNA blot and DNA blot analysis.** Total RNA from noninfected tobacco, from tobacco 4 days after inoculation with TMV (strain U1), or spraying with 5 mM sodium salicylate, was extracted from frozen leaf tissue by homogenization in extraction buffer (1 M Tris-HCl, 0.1 M LiCl, 10 mM EDTA, 1% sodium dodecyl sulphate [SDS], pH 9.0). The homogenate was extracted with phenol and chloroform and the RNA was precipitated with 2 M LiCl. The RNA was electrophoresed in agarose gels after glyoxalization and blotted to Genescreen (New England Nuclear) membranes. Tobacco genomic DNA was isolated and digested, electrophoresed, and blotted as described (Cornelissen et al. 1987). Hybridization of tobacco DNA was assayed with the 32P-labeled cDNA insert of clone cPR4-6 in 5× SSC (1× SSC is 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0, 2% SDS, 50% formamide, 100 μg/ml -1 herring sperm DNA at 42°C. The blots were finally washed in 0.2× SSC, 0.1% SDS, at 55°C and autoradiographed. Tomato and *C. fulvum* tissue were homogenized in guanidine-HCl buffer (8 M guanidine-HCl, 20 mM morpholinoethanesulfonic acid, 20 mM EDTA, 50 mM β-mercaptoethanol, pH 7.0). RNA was precipitated with 2 M LiCl and poly(A)+ RNA was obtained by chromatography on oligo(dT)-cellulose. After electrophoresis in formamide-containing gels the RNA was blotted onto Hybond N+ membranes (Amerham, UK). Tomato genomic DNA was isolated according to Rogers and Bendich (1988), digested, depurinated, denatured, electrophoresed, and blotted. The northern and Southern blots containing the tomato nucleic acids were hybridized with the 32P-labeled ZapP2-1 cDNA insert in 5× SSC, 5× Denhardt’s, 0.5% SDS, 100 μg/ml -1 salmon sperm DNA at 65°C, and finally washed in 0.5× SSC, 0.5% SDS, at 65°C.

**RESULTS**

**Partial amino acid sequence analysis of P2.** N-terminal sequencing on tomato P2 yielded no information, after digestion, suggesting that the N-terminus is blocked. After digestion, sequencing of some of the separated fragments gave

---

Vol. 4, No. 6, 1991 587
the following results: (T) ASVYCATWDADKPLE for fragment 1, and ATYHYLNPQMNWDLR for fragment 2. A third peptide sample gave a mixture of sequences: Y/G N/D/G/L N/D/T A/V F/N/D Y/G/E/P/F A/V G/N P X G Q (X: unidentifiable amino acid).

**PCR amplification of PR-4 sequences and screening of cDNA libraries.** Because there was no information on the positions of the tryptic fragments in P2, complementary oligonucleotide mixtures were synthesized with nucleotide sequences based on the amino acid sequences of fragments 1 and 2. The oligonucleotide mixtures corresponding to fragment 1 were LS13 (co-linear with the cDNA) and LS14 (complementary to the cDNA). Oligonucleotides LS15 (co-linear with the cDNA) and LS16 (complementary to the cDNA) corresponding to fragment 2 were also synthesized. The primers had the following sequences (I = inosine):

- **LS13:** CAGGATCCGGT/CTTG/ATCIGCG/ATCCCAIITGCG/ACAG/ATA
- **LS14:** CAGGATCTTC/TTGC/TGCIACITGGGAC/TGGCG/AA/GC
- **LS15:** CAGGATCCAC/TC/TTITAC/TAAC/TCCICAG/AAAC/TATAAC/TTGGGA
- **LS16:** CAGGATCCCTCAG/ATTIATG/ATTGC/TTTGGG/ATTG/ATAIAG/AG/ATG

All oligos contained BamHI restriction sites at their 5' termini, initially intended to be used for cloning of the PCR fragments. However, PCR's performed with the appropriate combinations of primers on genomic DNA from tomato or tobacco, lambda ZAP library cDNA, or randomly primed cDNA synthesized on RNA from tobacco infected with TMV, did not produce distinct fragments. This indicated that fragments 1 and 2 were possibly located in close proximity in the P2 and PR-4 protein. Therefore, another strategy was chosen in which PCR's were performed on the tobacco cDNA library, with combinations of the T7 primer, which was complementary to the T7 promoter located downstream of the 3' end of the cDNA in the lambda ZAP cloning vector and the LS13 or LS15 primers, respectively. This resulted in the amplification of DNA fragments, which, after labeling, were used as probes to screen the cDNA libraries. The screenings resulted in the isolation and further analysis of eight and four positive clones from the tobacco and tomato cDNA libraries, respectively.

**Analysis of cDNA clones and encoded proteins.** Although different in size, five of the tobacco cDNA clones contained the same sequence, while the other three clones had a different, highly similar nucleotide sequence. All four tomato clones had identical cDNA sequences. All clones contained poly(A) tails of varying sizes. There was no heterogeneity in polyadenylation sites between different clones of each group. The nucleotide sequence and the deduced amino acid sequence of the largest open reading frame of the longest clone of each group is shown in Figure 1. Located upstream of the first ATG codon in clone Zap2:1 is an in-frame TAA stop codon, indicating that this ATG is the initiation start site. Primer extension analysis on poly(A) RNA from tobacco infected with TMV suggested that clones cPR-4-5 and cPR-4-6 are nearly full length. Two extension products were found, demonstrating that the corresponding mRNAs have seven and 22 additional 5' terminal residues, which were not present in the largest clone (cPR-4-6, results not shown).

Protein purification and electrophoretic analyses of tomato PR proteins have indicated the presence of only one PR-P2 isomer in tomato leaf tissue infected with C. fulvum. The amino acid sequences of PR-2 puttyc fragments 1 and 2 completely corresponded to the sequence deduced from the cDNA clones, whereas two sequences deduced from the peptide mixture could also be mapped (underline in Fig. 1). As is shown in Figure 2, the tomato PR-P2 and tobacco PR-4 proteins are very similar (>80% amino acid sequence identity). The N-terminal region of all PR proteins is highly hydrophobic and is expected to function as a signal peptide for translocation through the endoplasmic reticulum. The expected cleavage site in all proteins would be between the first alanine and glutamine residues, analogous to the cleavage sites of acidic chitinases, glucanases, hevein, and other proteins (Shinshi et al. 1988; Brockaert et al. 1990; Linthorst et al. 1990a,b; Payne et al. 1990). This would result in mature proteins with molecular weights of 13,485, 13,468, and 13,480 for PR-P2, PR-4a, and PR-4b, respectively. The putative mature tobacco proteins PR-4a and PR-4b, encoded by cPR-4-6 and cPR-4-5, respectively, differ by only four amino acids, one of which is an acidic residue. This will probably result in a small difference in net charge between the two proteins, as...
observed during electrophoretic analyses of tobacco PR proteins (Joosten et al. 1990; Van Loon et al. 1987; Kauffmann et al. 1990).

**Similarity of PR-P2 and PR-4 to other stress proteins.** Comparison of the amino acid sequences of PR-4 and PR-P2 with sequences stored in the SwissProt protein sequence database (rel. 16, November 1990) revealed a large similarity to potato wound-induced proteins win1 and win2 (Stanford et al. 1989). Both win proteins contain a so-called "hevein" domain, immediately following a hydrophobic N-terminal signal peptide region. The tobacco and tomato PR proteins do not, however, accommodate such a "hevein" domain. Recently, it was shown that hevein, the small lectin from *H. brasiliensis* (bold underline in Fig. 2), is cleaved from a much larger precursor protein, which is approximately 65% identical to win1 and win2 (Broekaert et al. 1990). Figure 2 shows the similarity of the PR-4 proteins and PR-P2 to win1, win2, and pro-hevein. The PR proteins are approximately 50% identical to the win proteins and pro-hevein.

**Southern blot hybridization.** The cDNA inserts from clones cPR4-6 and ZapP2-1 were used to probe blots containing genomic tobacco and tomato DNA, respectively. Figure 3A shows that after autoradiography, two to five bands were visible in the lanes containing digested tobacco DNA, demonstrating the presence of a limited number of genes (probably two to four). Tomato DNA digested with EcoRI and EcoRV (Fig. 3B) showed only one band hybridizing with the probe, whereas the HindIII digest contained four hybridizing fragments (the 9-kb band is a doublet, Fig. 3B). Because the PR-P2 cDNAs have an internal HindIII site, these results indicate that there are probably only two very similar PR-P2 genes in the diploid tomato genome.

**Accumulation of PR-4 and PR-P2 mRNA upon infection.** Figure 4 shows a northern blot containing total RNA extracted from an uninfected Samsun NN tobacco leaf (lane H) and from leaves sprayed with salicylate (lane S) or inoculated with TMV (lane T), 4 days before sampling. It is evident that expression of the PR-4 genes is highly induced by both treatments. Similar induction patterns have been found for the other tobacco PR genes (Linthorst 1991).

The northern blot shown in Figure 5 indicates that

---

**Fig. 3.** Southern blots of digested genomic tobacco and tomato DNA. A, Tobacco DNA was digested with HindIII (H) or EcoRI (E), electrophoresed, blotted, and hybridized to labeled cPR4-6 cDNA. B, Tomato DNA was digested with EcoRI (E), EcoRV (V), or HindIII (H), electrophoresed, blotted, and hybridized with the labeled cDNA insert from clone ZapP2-1. The sizes of the hybridizing fragments (kb) are indicated on the left and right.
tomato PR-P2 gene(s) are induced by fungal infection with *C. fulvum*. Two types of interaction were studied. Tomato genotype Cf4 infected by *C. fulvum* race 5 reacts hypersensitively, leading to very limited local necrosis without visible symptoms, whereas tomato genotype Cf5 is completely colonized by *C. fulvum* race 5, resulting in massive sporulation of the fungus. PR-P2 mRNA is not detected in uninoculated leaves nor in RNA isolated from *C. fulvum* grown in vitro. In the incompatible interaction, PR-P2 mRNA accumulation is detected 4 days after inoculation, and this increases to reach a maximum at 8 days after inoculation. In the compatible interaction, PR-P2 mRNA can only be detected 6 days after inoculation, but reaches a higher level during later stages of the infection. Identical induction patterns have been observed for the genes encoding tomato PR proteins of groups 1 and 2 (P14-like proteins and β-1,3-glucanases, respectively; Van Kan et al., unpublished).

**DISCUSSION**

Two to four PR proteins with molecular weights in the range of 13,000 to 14,500 have been identified in tobacco.

![Graph](image)

Fig. 4. Induced accumulation of tobacco PR-4 mRNA. Total RNA was isolated from uninoculated (H), salicylic acid-sprayed (S), or tobacco mosaic virus-infected (T) tobacco leaf, electrophoresed, and blotted. The blot was hybridized with a labeled cDNA fragment from clone cPR4-6.

![Graph](image)

Fig. 5. Time course of accumulation of tomato PR-P2 mRNA. Poly(A)⁺ RNA was isolated from healthy Cf5 plants (P), *Cladosporium fulvum* (race 5) grown in vitro (F), from an incompatible interaction (between tomato genotype Cf4 and fungal race 5), and from a compatible interaction (between tomato genotype Cf5 and fungal race 5), at 4, 6, 8, 10, 12, and 14 days after inoculation, as indicated. After electrophoresis, the blot was hybridized with the labeled cDNA insert from clone ZapP2-1.

infected with TMV (Van Loon et al. 1987; Kauffmann et al. 1990; Pierpoint 1986). Provisionally, these have been placed in PR protein group 4 (Van Loon et al. 1987). In tomato infected by *C. fulvum*, PR-P2 was detected in a group of three small PR proteins. Two of these proteins were serologically related and similar to the tobacco PR-1 proteins (Van Kan et al., unpublished). The third, PR-P2, was serologically distinct, but PR-P2 antisera cross-reacted with tobacco PR-4 proteins (Joosten et al. 1990). This relationship was used to isolate the cDNA clones described above. The encoded tobacco proteins, after removal of the putative signal peptide, contain approximately equal amounts of acidic and basic residues, which is in support of the virtually neutral character of the PR-4 proteins (Van Loon et al. 1987; Kauffmann et al. 1990). The mature protein encoded by ZapP2-1 contains more basic amino acids, which corresponds to the high pI of PR-P2 (Joosten et al. 1990). The presence of a hydrophobic N-terminal domain, reminiscent of a signal peptide, is in agreement with the extracellular location of PR-4 and PR-P2.

PR-4 and PR-P2 proteins appeared to be very similar to the wound-induced *win* proteins from potato and to the recently characterized pro-hevein protein (Stanford et al. 1989; Broekaert et al. 1990). The *win* proteins have not yet been identified and were only characterized from genomic and cDNA clones. Their function is still unknown.

Hevein, a small chitin-binding lectin from *H. brasiliensis*, is a protein consisting of 43 amino acids, of which a large proportion are cysteine and glycine residues. Hevein is processed by proteolytic cleavage from a much larger precursor protein, which contains the hevein domain between an N-terminal signal peptide and a C-terminal region of 144 amino acids. The protein, which is present in large quantities in the latex expelled from wound sites, has a potent antifungal activity in vitro (Van Parijs et al. 1991). The similarity of the *win* proteins to the hevein precursor suggests that they too may be processed to release the “hevein” domain. The finding that in tobacco and tomato *win*-like proteins without a “hevein” domain are synthesized suggests that the C-terminal domains of *win* and pro-hevein have a yet unknown function.

It also seems likely that in tobacco and tomato, “hevein” domain-containing *win*-like proteins exist, because potato...
win DNA probe hybridized to restriction fragments of tobacco and tomato genomic DNA (Stanford et al. 1989). Cross hybridization of the win probe to genes corresponding to the PR-4 or PR-P2 proteins described here is unlikely, because of the low nucleotide sequence similarity between these genes and the win genes (42-50%). Furthermore, similarly digested tobacco and tomato DNA resulted in differently sized fragments hybridizing with the win and PR-4/PR-P2 probes.

Figure 2 shows, that in comparison with the extracellular PR proteins, pro-hevein and win2 have an extended C-terminal region. In agreement with current ideas, such C-terminal extensions could be involved in vacuolar targeting (Shinshi et al. 1988; Bednarek et al. 1990). All tobacco PR proteins characterized thus far can be classified in groups of either acidic, extracellular proteins or basic, vacuolar proteins. Fitting this scheme, win-like or pro-hevein-like tobacco proteins could be the vacuolar counterparts of the PR-4 proteins described here.

The similarity of extracellular PR-4 to vacuolar pro-hevein compares with that of the tomato extracellular PR-3 chitinases and their vacuolar counterparts. Also, the extracellular chitinases lack the "hevein" domain present at the N-terminus of the mature vacuolar chitinases (Linthorst et al. 1990b; Payne et al. 1990; Shinshi et al. 1987). In addition, the vacuolar chitinases contain a C-terminal extension not present in the extracellular proteins. These chitinases are not further processed to release the "hevein" domain, which explains their high substrate affinity. Furthermore, the basic tobacco chitinases, like pro-hevein and win proteins, are highly wound inducible, in contrast to the extracellular PR-3 and PR-4 proteins (Brederode et al., 1991).

Tomato PR-P2 is one of a set of extracellular PR proteins that accumulate almost simultaneously in tomato-C. fulvum interactions (Joosten and De Wit 1989). In an incompatible interaction (where the plant resistance reaction is effective) accumulation of these proteins is observed at 6 days after inoculation, which is 2-4 days earlier than in a compatible interaction (where plant defense is ineffective). The data presented in Figure 5 indicate that the mRNA for tomato PR-P2 is already detectable at 4 days after inoculation in an incompatible interaction, but not in a compatible interaction. The timing of induction of mRNAs encoding other extracellular tomato PR proteins, such as two P14 isomers and β-1,3-glucanases, is similar to the induction of PR-P2 mRNA (Van Kan et al., unpublished). C. fulvum is a biotrophic pathogen that enters the stomata around the third day after inoculation and colonizes the intercellular space of the leaves without penetrating host cells. In an incompatible interaction fungal growth is inhibited around the fourth day after inoculation (Lazarovitz and Higgins 1976; De Wit and Van der Meer 1986). Thus, the correlation between the inhibition of fungal growth and the early induction of mRNAs for PR-P2 and other extracellular PR proteins suggests the involvement of one or several of these proteins in a successful defense reaction against C. fulvum.

The interaction of TMV with tobacco can be considered as analogous to the incompatible tomato-C. fulvum interaction. In both cases the replication/growth of the pathogen is inhibited early after inoculation and this inhibition is correlated with the induced expression of genes for PR proteins as can be seen in Figures 4 and 5. Infection of tobacco with alfalfa mosaic virus does not result in a hypersensitive response or the induced expression of PR genes, and the virus replicates to high titers (results not shown). Contrary to the compatible plant-virus interaction, the compatible tomato-C. fulvum interaction does result in (significantly delayed) PR gene expression. This late expression of PR-P2 genes may be due to secondary effects caused by the rapidly growing fungus. In this case, it could be argued that P2 or other tomato PR proteins with antifungal activity are produced too late to overcome the massively growing fungus.

ACKNOWLEDGMENTS

We are grateful to Matthieu Joosten for purification of PR-P2. D. Shah at Monsanto, St. Louis, MO, is gratefully acknowledged for performing the protein sequence analysis. We would like to thank Leo Meichers and Ben Cornelissen at Mogen Int. N.V., Leiden, the Netherlands, for providing synthetic primers. This work was sponsored in part by the Netherlands Technology Foundation (STW) and was coordinated by the Foundation of Biological Research (BION).

LITERATURE CITED


