

# Analysis of gene families encoding acidic and basic $\beta$ -1,3-glucanases of tobacco

(pathogenesis-related proteins/virus-induced proteins/plant defense genes/salicylic acid)

HUUB J. M. LINTHORST\*<sup>†</sup>, LEO S. MELCHERS<sup>‡</sup>, ANDREAS MAYER\*, JEROEN S. C. VAN ROEKEL<sup>‡</sup>, BEN J. C. CORNELISSEN<sup>‡</sup>, AND JOHN F. BOL\*

\*Department of Biochemistry, Gorlaeus Laboratories, Leiden University, Einsteinweg 5, 2333 CC Leiden, The Netherlands; and <sup>†</sup>Mogen Int. N.V., Einsteinweg 97, 2333 CB Leiden, The Netherlands

Communicated by Clarence A. Ryan, July 27, 1990 (received for review May 15, 1990)

**ABSTRACT** Healthy tobacco plants accumulate  $\beta$ -1,3-glucanases (glucan endo-1,3- $\beta$ -glucosidase; EC 3.2.1.39) in their roots and in specific parts of the flowers. After infection with tobacco mosaic virus, acidic and basic  $\beta$ -1,3-glucanases are induced in the inoculated and virus-free leaves of the plant. An analysis of cDNA clones demonstrated that at least five genes for acidic  $\beta$ -1,3-glucanases are induced after tobacco mosaic virus infection. Southern blot analysis indicated that the tobacco genome contains approximately eight genes for acidic  $\beta$ -1,3-glucanases and a smaller number of genes encoding basic  $\beta$ -1,3-glucanases. Genes from both gene families were cloned and sequenced. The basic isozymes contain a C-terminal extension that is cleaved off during their targeting to the vacuoles. This extension is absent in the acidic isozymes, which accumulate extracellularly. Northern blot hybridization showed that genes encoding acidic and basic  $\beta$ -1,3-glucanases are strongly induced after tobacco mosaic virus infection or salicylate treatment of tobacco. The cloning of these genes is a first step toward the identification of regulatory elements involved in their coordinate induction.

The hypersensitive response of plants to pathogens triggers the *de novo* synthesis of a large number of proteins with a putative role in defense mechanisms. A subset of these proteins is known as the "pathogenesis-related (PR) proteins" (for recent reviews, see refs. 1 and 2). The PR proteins induced in tobacco after infection with tobacco mosaic virus (TMV) have been classified into five groups (3). Groups 1, 2, 3, and 5 contain acidic isoforms that are secreted into the intercellular space of the leaf as well as basic isoforms that accumulate in the vacuoles of leaf cells. In group 4 only acidic extracellular proteins have been identified so far. Genes for the basic proteins in groups 1, 2, and 3 are constitutively expressed in the roots of healthy plants and their synthesis is coordinately induced by the plant hormone ethylene (4). Moreover, the accumulation of proteins from groups 1, 2, and 3 in specific parts of tobacco flowers indicates that these putative defense proteins may fulfill highly specialized functions in healthy plants as well (5).

So far, an enzymatic function has been identified only for the tobacco PR proteins in groups 2 and 3. Within the PR-2 group four acidic proteins (PR proteins 2, N, O, and Q') and one basic protein were shown to have  $\beta$ -1,3-glucanase (glucan endo-1,3- $\beta$ -glucosidase; EC 3.2.1.39) activity (6, 7). Group 3 contains two acidic proteins (PR-P and PR-Q) and two basic proteins (Ch.32 and Ch.34) with endochitinase activity (8). Because a mixture of  $\beta$ -1,3-glucanases and chitinases was found to inhibit fungal growth on agar plates,

it has been suggested that these hydrolases play a role in the induced resistance of plants to fungal infection (9).

To obtain a better understanding of the function of these proteins in the healthy and diseased plants we have recently characterized cDNA and genomic clones encoding acidic and basic chitinases of tobacco (10). The present study describes an analysis of the genes encoding  $\beta$ -1,3-glucanases that are expressed after TMV infection of Samsun NN tobacco.<sup>§</sup>

## MATERIALS AND METHODS

**Synthesis and Analysis of cDNA Libraries.** Isolation of poly(A) RNA from TMV-infected tobacco and the construction of  $\lambda$ gt11 and  $\lambda$ ZAP (Stratagene) cDNA libraries were as described (10). Recombinant phages were screened using a polyclonal antiserum to PR proteins 2 and N (11) or with [<sup>32</sup>P]cDNA inserts with glucanase sequences. The antiserum was a gift from L. C. van Loon (Wageningen, The Netherlands). cDNA inserts were subcloned in M13 derivatives for sequence analysis (10).

**Analysis of Genomic Clones.** A genomic DNA library of Samsun NN tobacco in Charon 35 was screened for the presence of glucanase sequences using [<sup>32</sup>P]cDNA inserts (12). DNA from positive clones was isolated and restriction fragments containing glucanase sequences were subcloned for sequence analysis.

**DNA Blot Analysis.** Nuclear genomic DNA, isolated from Samsun NN tobacco leaf, was digested with *Eco*RI or *Hind*III, electrophoresed on 0.8% agarose gels, blotted, and hybridized to <sup>32</sup>P-labeled subfragments of cDNA clones corresponding to acidic and basic  $\beta$ -1,3-glucanases (12). The subfragments were devoid of internal *Eco*RI or *Hind*III sites. After hybridization the blots were washed at 65°C in 30 mM NaCl/3 mM sodium citrate, pH 7.5, containing 0.1% sodium dodecyl sulphate.

**RNA Blot Analysis.** Total RNA from uninfected, TMV-infected, or salicylate-sprayed tobacco leaves was electrophoresed in agarose gels, blotted, and hybridized to specific glucanase cDNA probes as described (10). RNA was isolated 4 days after TMV inoculation or salicylate treatment (13).

## RESULTS

**Characterization of cDNA Clones Encoding  $\beta$ -1,3-Glucanases.** Screening of a  $\lambda$ gt11 cDNA expression library of TMV-infected Samsun NN tobacco leaf with an antiserum to PR proteins 2 and N resulted in the isolation of several

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: PR, pathogenesis-related; TMV, tobacco mosaic virus.

<sup>†</sup>To whom reprint requests should be addressed.

<sup>§</sup>The sequences reported in this paper have been deposited in the GenBank data base [accession nos. M34087 for the acidic  $\beta$ -1,3-glucanase gene and M34086 for the basic  $\beta$ -1,3-glucanase gene].

positively reacting clones. Sequence analysis of one of the inserts showed that the encoded protein was similar to the sequence of a basic  $\beta$ -1,3-glucanase protein encoded in cDNA clones derived from tobacco callus incubated on hormone-free medium (14), indicating that it could correspond to an acidic  $\beta$ -1,3-glucanase indeed. Rescreening of the library with the  $^{32}$ P-labeled insert of this clone yielded 30 independent clones that subsequently were isolated and characterized by sequence analysis or restriction enzyme mapping. All clones were  $\approx$ 600 base pairs (bp) long and terminated at an internal *Eco*RI site in the cDNA. Subsequent screening of a  $\lambda$ ZAP library resulted in several additional clones containing longer inserts. Five different  $\beta$ -1,3-glucanase cDNA sequences were found (Fig. 1), none of which corresponds to the cDNA sequences encoding basic  $\beta$ -1,3-glucanase (14). In total, 34 clones were characterized, with 22 clones belonging to the groups represented by cI101 and cI125. Of 6 clones that were (partially) sequenced, 3 belonged to the cI101 group and 3 were identical to the cI125 sequence. These two cDNA sequences were very similar, with only three differences in the sequenced regions (Fig. 1). Groups cI30, cI32, and cI37 were represented by 3, 2, and 7 clones, respectively. Subclones used for sequencing of

cDNA clone cI32 differed in the length of the (AT)<sub>n</sub> repeat in the 3'-terminal noncoding region (underlined in Fig. 1). The value of *n* was found to vary between 10 and >60, whereas in clone cI44 of this group *n* appeared to be 13. We assume that this variance is a cloning artifact caused by the (AT) repeat rather than the result of the presence of different genes. The various clones of groups cI101 and cI125 were found to have three different poly(A) sites (indicated in Fig. 1).

**Characterization of Genomic Clones Encoding Acidic and Basic  $\beta$ -1,3-Glucanases.** Screening of a genomic library of tobacco with labeled  $\beta$ -1,3-glucanase cDNA from one of the clones of group cI125 resulted in six genomic clones with different restriction enzyme maps. Furthermore, the screening of a tobacco genomic DNA library with a cDNA probe corresponding to a basic glucanase (clone pCNT11, 4) resulted in two groups of clones with overlapping inserts, each representing a unique gene. The sequences of the coding and adjacent regions of clone gI9, corresponding to an acidic isozyme, and of clone ggIb50, corresponding to a basic isozyme, are shown in Figs. 2 and 3, respectively. The sequence of the gene of clone gI9 matches exactly the cDNA sequence of the clones of group cI125, whereas the ggIb50 sequence was identical to the sequence of basic glucanase

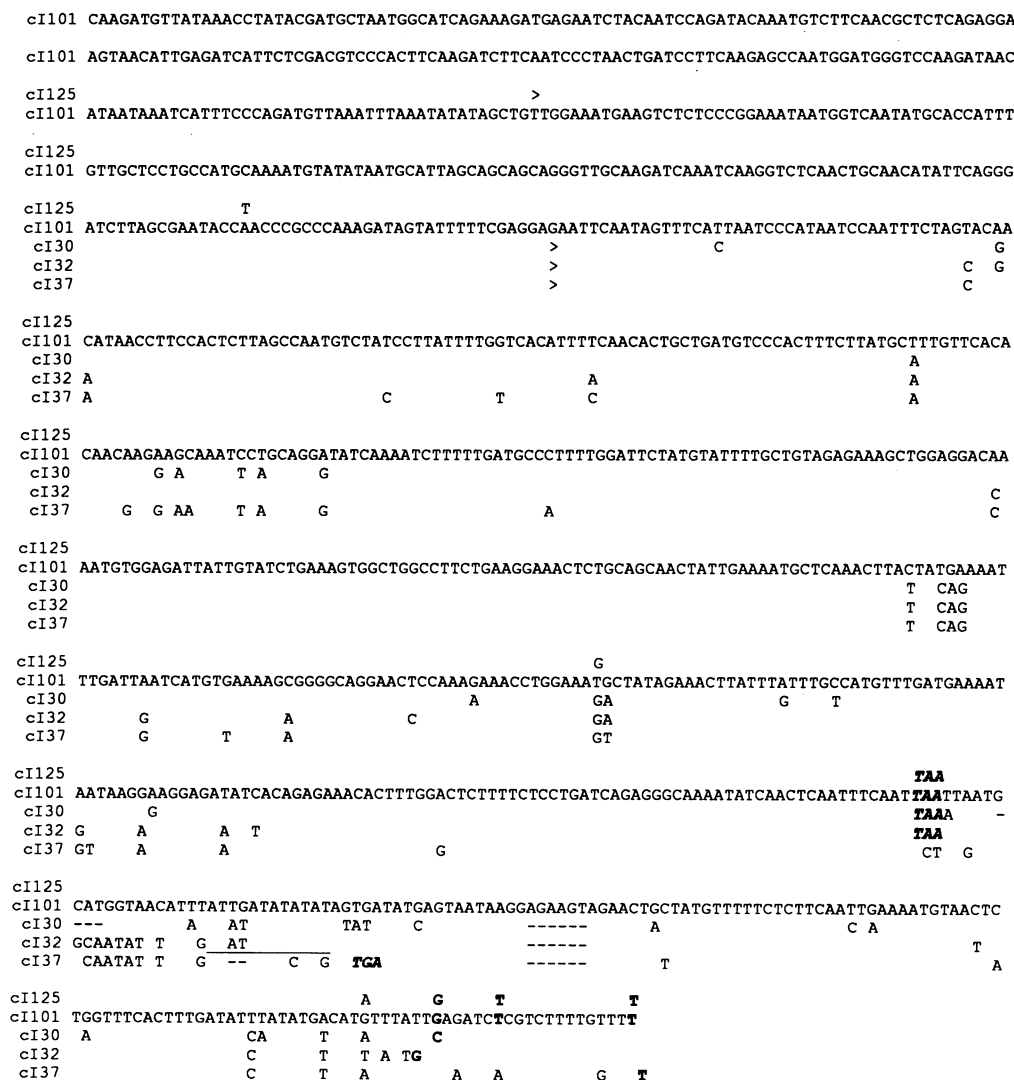


FIG. 1. Nucleotide sequences of different groups of cDNA clones corresponding to tobacco  $\beta$ -1,3-glucanases. The sequence of the longest clone (cI101) is given. Above and below are shown the nucleotide differences of cDNA sequences represented by clones cI125, cI30, cI32, and cI37. Gaps (-) are introduced for optimal alignment. The 5'-terminal residue of the last four sequences (>), the stop codons in the different clones (bold italic), and poly(A) sites present in different clones (bold) are indicated. For alignment, six (AT) repeats of clones of group cI32 are given (underlined), although in different subclones the repeat value varied from 10 to >60 (see text).

```

1 CCTTCTCGAATATTCGACCATTTTTTCTCATAAATCTATGCTCTCCATCACAACATCAGATTGAACCTTTCACATTTTTTTTTTTC
91 TTCACCTGTGACAAAAATACACTGTTTTAAATCCATGACGTGCTTACTGGGTGTGACTTTGGCTTTTTAAAGCTTTTGTGGGCATCTT
181 CTGGCTGTCTATAAAGAACGAGCTAAATGGTTCCTTAAACACACAAATTCAGCTCAAGTGTTCCTTACTCTCAATTTCCATTTAGCT
271 ATGACTTTATGCATAAAAATGGCTTTCTTGCAGCTGCCCTTGTACTTTGTGGGCTGTAATTTGCAGATTCACAAATGATAGGCTCTCT
361 CTCTCACACACACACTTTCTCTCATGATACATGACATGACCTTGTATGATGCGGATCAACTTATGTACACTAATAGCGTAATAAT
451 TTTTACAAATATATATAGGATTAATATATTTAACATGTTGTGTCAGGTAATCTACCTTATTTATAGTCACTTATATGAATAGTTAC
541 TTATAGTACTTCTGGGTGACCCGACACTATAATGTGGCTAGAGAAGAACTAAATAGAGAATCATGGTTAGTGAGAATATTCATTTAT
631 TCGACACCAACTTATTTGGGGACTGAAACTCTTTTGAATATACTTTTTTTCACAAATCCAGGGCACAATCTATTGGAGTATGCTATG
721 GAAACACTGCACAAATTTACCAATCAGACCAAGATGTTATAAACCTATACAATGCTAATGGCATCAGAAAGATGAGAACTCAATCCAG
811 ATACAAATGCTTCAACGCTCTCAGAGGAAGTAACATTTAGACTTCTCGACGCTCCCACTCAAGATCTCAATCCCTCAATGCTATG
901 CAAGAGCAATGGATGGGTCCAGATAAATAAATCAATTTCCAGATGTTAAATTTAAATATATAGCTGTGGAAATGAAGTCTCTC
991 CCGAAATTAATGGTCAATATGCACCATTGTGCTCCTGCCATGCCAAATGTATATAATGCATTAGCAGCAGCAGGCTTCAAGATCAAA
1081 TCAAGGCTCACTGCAACATATTCAGGGATCTTAGCGAATACCTACCCGCAAGATAGTATTTTTCGAGGAGAAATCAATAGTTTCA
1171 TAAATCCCAATAATCCAAATTTCTAGTACAACAATCCCTTCCACTCTTAGCCAATGCTATCTTATTTTGGTCAATTTTCAACACTGCTG
1261 ATGCTCCACTTTCTTATGCTTTGTTACACAACAAGAAAGCAATCTCTGAGGATATCAAAATCTTTTTGATGCCCTTTTGGATTCTATGT
1351 ATTTGCTGTAGAGAAGCTGGAGCAAAATGTGGAGATTTGTATCTGAAAGTGGCTGGCCCTTCTGAGGAAACTCTCGACCAACTG
1441 TTGAAATGCTCAAACTACTATGAAAATTTGATTAATCAATGTTGAAAAGCGGGCAGGAACTCCAAAGAACTTGGAAAGCTTATAGAAA
1531 CTTATTTTATTTGCCATGTTTGTATGAAAATAAAGGAAGGAGATATCACAGAGAACACTTTGGACTCTTTTCTCTGATCAGAGGGCAA
1621 AAATCAACTCAATTTCAATTAATTAATGCATGGTAACATTTATGATATATATAGTATATAGTAAATAGGAGAAGTAGAATGCTAT
1711 GTTTTCTTCAATTTGAAAATGTAACCTCTGGTTTCACTTTGATATTTATATGACATATTTATGAGATCTCTGCTTTTTTAAATTC
1801 TTGGCTTCTATTTGGCAAAATCTGCGCTAATTTTCAATTTGTTTAAAAAATTAAGCTCAAAGAGTGACTACCAATATATTTGATTT
1891 ATTAATATCCCGTGTCTGGGGACCGGGTGGGGGGTGGGGGGATGACGAAAAAAGTAAATGAAAACCGGTTGCTATGATGAT
1981 GCTCTTTTAACTCCCAAAATATGATGGTTTGTGCTTGGAGAGTGTAAAGCTACTTCTCTCAGAAATTTCTTGGTCACTTCT
2071 TAACGTAATGCTTTTAAATTTCTTAATATCGGTAACCCCTCGAAACAAAAGGAAATTAAGCTAGGAGATGACTCGTATTCATAATGTT
2161 TTACCTTGGATCAACCCGCTTTATATTTTACATACGATGTTTTTTCATTGAAAGGATGAACATAGACTTTTTTATCATGCCCAAGG
2251 AAATCAAAA

```

Fig. 2. Nucleotide sequence of genomic clone g19, encoding tobacco acidic  $\beta$ -1,3-glucanase. Coding regions, separated by an intron, are indicated (bold).

cDNA clone pGL36 (14), indicating that both genes are probably expressed.

The coding region of clone g19 contains an intron of 341 bp, located in the penultimate codon of the putative glucanase signal peptide. A putative TATAA box is located at position 192 but a distinct CAAT box is lacking. The coding region of the basic glucanase gene in clone g1b50 is divided over two

exons separated by an intron of 787 bp. Analogous to the gene in clone g19 this intron is located in the glycine codon preceding the last codon of the signal peptide. The gene contains putative TATAAA and CAAAT boxes at positions 1507 and 1463, respectively.

Partial sequence analysis of genomic clone g1x33 (data not shown), from the second group of basic glucanase genes,

```

1 CTCGAGGTCGGGCTGGATGCCACAATAACCCATAAGTCAAGTCCCGATGTTGGCTGTATACAATATGATTAATAAAGTCCACCAAG
91 ATTTAGCTCCTTAGGGCCAAAGGGCAAACCTTTTGTAAATGGGAAAAAATATAGACCAAGAGTTTGTAAATAGCTCAATTTCA
181 ATTTCAAAGGGGAAAATTTAACTGTTTTGGCCCTTATATCTTTTGTGCCGCAACATAAAATATCCCAATCCGAAATCCAAATGGTCCAT
271 TATCGGCAAGTAGCTTTCTTTTATATAGTATAGTTGACAAAACACTATCGAGATATCATAGTATAATAAATACTTCAAAGTCCATCTTA
361 CTTAGCTGCTCCTCATTAGAGCCGCCATAAATAAGACCGATCAAAATAAAGCCGCCATAAATAAATGAATTTTAGSACTCAAT
451 GTCAGCTAAGTCCCAAAACTCTTCCAACTTTGCTGCAACTTGGGGCTGTAGCTTCTGAGCTTCTGGGATATTTCTATGTTTATCTC
541 TTAATTAATCTCACTAATATCAAGAAATTAACACAGTGCAGCAATATAAATTTCTCTAAAGAAAGAAATGACTCCCGTTACTG
631 ATTCATGGCCCTTTCAGAGTCTGCGTGCATATCTCACTAATGGTCTGTTGGTCAAGAAATATGATAAATTTGAAATGAAATTTG
721 GGATTTCAATTTTCAATTAATTAATAAATATAGCTAATTTGCAAAATAAATTTACATTAATAATAGTAAATCAACTATCTCATATA
811 TAAGTGGAAATAGCTAATCTCATAGCCACTCAATAGAAATCCAGTTTCTCTAATAAATGACAGGAAATATAGAAAGCTTTTCAATA
901 TCAATTCATATAAATTTAAATACTAGACATGGAATAAAGAAACGATTCGAGACTGTTATGGAAGGCGTGGCTTCCATGTAGATTTCCAT
991 CCATTTGCTTCTGCAATAGCAATATGACATCTTATCTTGAAGTACTTTTAAATGAAAGTCAATAGAACTTTAAATACTCTCAACT
1081 GTTTAAGGGAAATCCAAATACGACCAATATTTACTTACTTATGATAAATTTAAATAATATGATAATTTTATCTGAAATGGATGA
1171 AAATTTAAATTTACTTGGTTTAAATATAAACAATAGATATCGTTAAGTATTTACCAACAACATTTTATAGTTGTAACAGATGATTAAGCAG
1261 AAATTCCTCTGGTTGTGAGGATGAAAGAACTAATCAATTTCTTTTGTAAAGTCAAGATAGTACGGCACTTATGAGGAACTT
1351 AAATAACTTTACATCACTCAACTCATCTTTTTTCCACAATAATGATCAAGTTGACATGTTAATAGCCAGTCCAGGGGGCGCTCTT
1441 AACTTCATTAGCTACGAAATAACAATCCAAATATATATTTACACAAGGCTATATATATCTCAATATAAATAGCTCGTTGTTTCACT
1531 AATTCCTCCAAACAGTCTTCCCACTATGCTACCTCACATAAACAATAAATCTCCCAAAATGGGCTGCTATCACACTTACTCT
1621 TGTTGCCAGCAGCATTGACATAGCAGGTTTCTGGTCAAAATTTGAACTTCCAGCCAAAATAATGTTCTTATAAATTTTGGTGGCCAAA
1711 ATTTTAAATTTAGTTGATAGTTATTTGCTTATTTTTCTTTTCAAATTTGCTTGGTTTTTTTCTCAAATTAATTCGACCGTATCTATTAG
1801 CGATAGTATTTGCTCTATTTTGTGTAACACTCACTCAAACTTTTCAATTTAGGGGGAGGACAGTGAATCTAAGATGAAATTTATGA
1891 GTTAAATAGACTAATTTCCATTTGATTTATGGCTAGAAGTCAATTTTGCATAGTGTGCTTTTAAACACAGATTTGAGTTAAGG
2981 TACTAGTTCGTATTAACCCATAACATATACACTCTGTTCTAATTTCTTGGACATTTTTGTAGTTTGTTCAAAAGGACGACAT
2071 ATTTGATATTTGAAATCTTTACCTTAACCTTAATAGAAATTTTTTATGACATCACAATATATAGGAATAATACGACCAATAATTTCA
2161 AATCTTATAGTCGTACAAAATATATAGACTGTTTAAATACCACAACCTTTCAAATTTCTTTTCTTAAACAAAATATGTCACATA
2251 ATTAATAAGAGGAAATATACATCAATCAGCCCTAGTGGAGGGGACCTACTGTAAGTTTTAAGTTTTCAAGAATTCAGTAAATGA
2341 TTAGAGCCCGCTGACATAAATAAATAAATTTCTTTTCCAAAATAATGCCCATAAATTTCTAACACTTTTTGTAATCTTTATGAG
2431 CAGGGCTCAATCGATAGGTTTGGCTATGGAATGCTAGGCAACAACCTGCCAAATCAATGGGAAGTTATACAGCTCTACAGTCAAGAA
2521 ACATAGGAAGACTGAGGCTTTATGATCCAAATCAAGGACTTTACAGCAATTAAGAGCTCAATATTGAAGTTATGTTAGGATTTCCCA
2611 ATTCAGATGGAAGCAATTTGCTTCCGGAATGGAACATGCAAGTATGGTGGTACAGAAAAATTTAAAGATTTTGGCCAGATGTTAAGA
2701 TATAGTATATTTGCTGTTGGGAATGAAATCAGCCCTGCTCAGGCACTTACCTTAACCTCATTTCTACTCTGCTATGCTGAAATATTT
2791 CAAAAGCAATTTGGTGAAGCTGGTTTGGGAAACAATCAAGGCTCAACTTCTGTAGACATGACCTTTGATGGAACTCTTATCCACCAT
2881 CACAGGTTCTGTTAGGACGATGCTAGGTTGTTGTTGATGCCATTTGTTGGCTTCTTAAAGGACACAGCTGCACCTTTACTCGTAAACA
3971 TTTACCCTATTTTCAATTTCTGGTAAATCCAGCCAGATTTCTCTCCCTATTTCTTTTTTACAGCACAATATGGTGGTACAAGATG
3061 GTTCCCGCAATATAGGAATTTATGATGCAATGCTGGATTTCTGTATGCTGCCCTCGAGCGATCAGGAGGGGCACTCTGAGGGATTTG
3151 TTGTTCCGAGATGGCTGCCATCTGCTGGTGCATTTGGAGCCACATATGCAATTCAGCAACTTACTTTAGGAACTTAAATCAACACG
3241 CTAAAGGGGTAGCCCAAGAAAGCTTGGACCTATTTAGACTATATATTTGCCATTTTGTAGAGAACCAAGAACTTCACTGACGCA
3331 AACATTTTGGATTTTCCCCCAACAGCAGCCAAATATAATATCAACTTTGGGGTCTCTGGTGGAGTTTGGGACAGTTCAGTTGAAA
3421 CTAATGCTACTGCTTCTCTGTAAGTGAATGTAGCTATGAGACACTTGAATCTCTTACATACGTATTTCTTGGATGGAAACCTA
3511 GTAAAACCAAGAGAAATTTTTCTTTATGCAAGATAAATAACATTTGCATGCTCTGTAAGTCTCATGGATTTGTTCCAGTACGCA
3601 TGCAACTCTGAGTGGTTTTAAATTCCTTTCTTTGATATTTGGTAATTTGGCAAGAACTTTCTGTAAGTTTGGAAATTTCAATGATCA
3691 TAAATTAACATCAGTTCCATGTTGATCAGATGGGATTTGGTAACCTCAATGTTAGTATAAATAATAGTGTCTTTATCATGACTATC
3781 AAATTAATCTTATTTGGCAAGGCTTGAATATTTAGTACTCTTAGGATTTGCAAGCAACTGATCTTCTTTTACCCGTTTCTGGCT
3871 TAAACCTATTAAGAAATATATTAATGTCACCTACTCTGTTGTTAAGACATTTCCCTTACATTAAGGATTTTCAAGTCTGATCAGT
4961 CGAAAAAATAATGGTACGCTTCTTATCACAATAATCTCTCAACTTCTAGACCAATGAAATCTTGTCTCCCAATAGGATTTGCTTTT
4051 ACTGTATGTTTCTCTATCAATTCAGGCTCAAGCCATCAAAATCGCTGGTATTTCTCGCTCCCATTTAACCAATCGAGCTTTAGCTC
4141 TGCTAAAAGTCTCAATTTCTAGCTTCTCAAAACCTCAGAGCTTCTCAACCAAAAAATAGAAAAAATCTTTGATCCCATGTAATAACCC
4231 CAGAAATGTCATTCGAGTGAAGGAAGTGTGGAAACCGGCATCAAAACCGGATTTGCCAGCTGATTTGTCATAGGAAAAAGAAAGTTG
4321 ATCAGATGTTCTGCGGGGATTTTCCACTCTGCTTTGATGACCTATATATTTGCCATTTTGTAGAGAACCAAGAACTTCACTGACG
4411 GAATCCATAAATCTAAGAGCATGCCACAACAACCGGATTAATCTCAGCTAAGAGATCTATTTGATCTTACCCCAACATAGTCCGG
4501 GTGAATTTGCTATTTCACTCAACTCAAAGCGAAAGAAAGACTCTATTACAGAGCAGCAAAAGAGCTTACTTTGCTTCCCTAGTGGGCTC
4591 TAAAGCAATTTGTTCCCGGATGATTTTGGCTTGTATGTTGTCAAAGGAACACATGATGAACAAGTCCATGGATAGTGTACATAG
4681 GGCATGTCATATCAACAAAAGGCAATGCACCTGCAAGGATACGTATAAGTGGTTATATCTGATAATCCCTGAACGAAAAGCTT

```

Fig. 3. Nucleotide sequence of genomic clone g1b50, encoding basic  $\beta$ -1,3-glucanase of tobacco. Coding regions, separated by an intron, are indicated (bold).

showed that it is 85% similar at the nucleotide level to the gene of *gglb50*. Clone *ggIx33* contained two intervening sequences of which the first one is located in the same relative position as the intron in *gglb50*. However, the two introns do not seem to have correct splice junctions, indicating that the *ggIx33* gene may be a pseudogene.

A comparison of the genomic clones *gglb50* and *gI9* indicated a similarity between the nucleotide sequences of the coding regions of 57%. No striking similarity in upstream, intron, or downstream sequences was observed (data not shown).

**Comparison of Different  $\beta$ -1,3-Glucanase Proteins.** Fig. 4 shows that the proteins encoded in clones *gI9* and *gglb50* share extensive sequence similarity (49% identity). By analogy to the basic tobacco glucanase (14), the hydrophobic N-terminal region of the protein encoded in *gI9* is assumed to be a signal peptide. The glucanase encoded in this clone terminates exactly at the position where in clone *gglb50* the C-terminal peptide begins, which upon glycosylation is cleaved off (14). Apparently, the  $\beta$ -1,3-glucanase encoded in clone *gI9* is not posttranslationally processed at the C terminus.

Depending on which methionine codon is used, the signal peptide encoded in the glucanase clone *gglb50* is 21 or 32 amino acid residues long. Also in *Nicotiana plumbaginifolia* (15) two methionine codons are present, and the sequence of the putative long *N. plumbaginifolia* signal peptide is highly conserved in this tobacco protein. The mature tobacco glucanase encoded in clone *gglb50* has a molecular weight of 37,150 and is a neutral, rather than a basic, protein with a calculated isoelectric point of 7.71. However, the protein encoded in *gI9*, upon removal of the signal peptide, has a molecular weight of 34,763 and is highly acidic (isoelectric point, 4.89). The partial amino acid sequences of the proteins encoded in the cDNA clones of groups *cI101*, *cI30*, *cI32*, and *cI37* were very similar (>90% identical) to the one encoded in clone *gI9*. Probably, these clones encode other acidic  $\beta$ -1,3-glucanase isoforms. The protein encoded in *cI37* is 13 amino acid residues longer at the C terminus than the other putative acidic glucanases. However, this region does not show any similarity to the C-terminal peptide of the basic glucanase, nor does it contain an asparagine glycosylation site (NXS/T).

**Genomic DNA Blot Analysis.** The cDNA insert of a clone from the *cI125* group was used to probe a blot containing electrophoretically separated digested genomic DNA. Upon

autoradiography, the blot showed at least 8 distinct bands in both lanes containing *EcoRI*- and *HindIII*-digested DNA (Fig. 5A). The band that gave the highest signal corresponds to a fragment containing genomic sequences, homologous to the probe, as was shown by the characterization of genomic clone *gI9*, which contained similarly sized restriction fragments (results not shown). The other bands may correspond to related genes for acidic  $\beta$ -1,3-glucanases, including genes that express proteins corresponding to cDNA clones *cI101*, *cI30*, *cI32*, and *cI37*. When a cDNA insert of a clone corresponding to basic glucanase was used as probe, five or six bands were found to give a hybridization signal (Fig. 5B). These bands did not comigrate with the fragments containing acidic glucanase gene sequences, indicating that no cross-hybridization occurred.

**RNA Blot Analysis.** Fig. 6 shows the results of hybridization of acidic and basic glucanase cDNA probes to RNA from noninduced tobacco leaf and from leaves infected with TMV or treated with salicylic acid. It is evident that the expression of both types of glucanase genes is highly induced by TMV infection. Furthermore, salicylic acid also induces the genes, although the induction of basic glucanase genes is relatively low.

**DISCUSSION**

The analysis of cDNA clones revealed that at least five genes of the approximately eight genes for acidic  $\beta$ -1,3-glucanases of tobacco are expressed after TMV infection. Recently, we have transformed tobacco with the *gI9* gene fused to the cauliflower mosaic virus 35S promoter. The protein encoded by the chimeric gene accumulated in the extracellular space in the leaves of the transgenic plants and comigrated in nondenaturing gels with the acidic  $\beta$ -1,3-glucanase PR-N (unpublished results). The relationship of cDNA clones *cI101*, *cI30*, *cI32*, and *cI37* to the known TMV-induced acidic  $\beta$ -1,3-glucanases PR-2, -O, and -Q' is not yet clear. Three extracellular  $\beta$ -1,3-glucanases, induced after infection with *Pseudomonas syringae* or salicylate treatment of tobacco, have been partially sequenced (16). These were named PR(35), PR(36), and PR(37). The protein encoded in our clone *gI9* (or in the highly similar cDNA clone *cI101*) corresponds to the partial PR(36) sequence with only one mismatch. The deduced amino acid sequences of the other clones do not correspond to either PR(35) or PR(37). This could indicate



FIG. 4. Comparison of the amino acid sequences of  $\beta$ -1,3-glucanases of tobacco. The amino acid sequences are given of basic and acidic glucanase, encoded in genomic clones *gglb50* and *gI9*, respectively. Gaps (-) were introduced for alignment and identical residues in the two sequences are connected. Signal peptides (underlined) and C-terminal residues (\*) are indicated. Differences in the partial sequences encoded in cDNA clones *cI101*, *cI30*, *cI32*, and *cI37* [5'-terminal encoded residue indicated (>)] with the protein sequence of clone *gI9* are given.

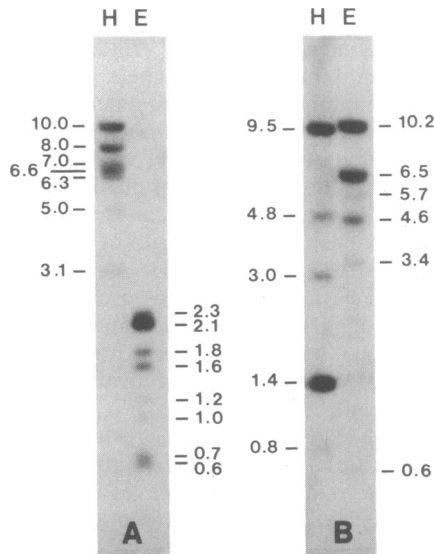


FIG. 5. Hybridization analysis of genomic tobacco DNA. *Hind*III- (H) or *Eco*RI- (E) digested genomic DNA was electrophoresed in agarose gels and blotted to filter membranes. The blots were hybridized to cDNA probes corresponding to acidic (A) or basic (B) tobacco glucanases. The size (kilobases) of the hybridizing fragments, based on the migration of a set of marker DNAs, is given.

that after infection of tobacco more than five acidic  $\beta$ -1,3-glucanases are expressed in the leaves.

Recently, two glycosylated extracellular acidic  $\beta$ -1,3-glucanases have been identified in the styles of tobacco flowers (R. Fluhr, Weizmann Institute, Israel; personal communication). These enzymes are not induced in TMV-infected leaves; their sequence similarity to TMV-induced acidic and basic  $\beta$ -1,3-glucanases is 80% and 49%, respectively. Determination of the possibility that some of the minor bands in the genomic blot of Fig. 5A correspond to genes for stylar-specific glucanases awaits further investigation.

The basic glucanases studied by Shinshi *et al.* (14) are known to accumulate in the vacuole of stressed tobacco cells (16). Three incomplete cDNA sequences have been characterized, indicating that at least three genes encoding basic glucanase are expressed (14). The observation that the sequence of the putative pseudogene in clone ggIx33 is similar but not identical to these cDNA sequences demonstrates that the tobacco gene family of basic  $\beta$ -1,3-glucanases contains at least four members. This is in agreement with the genomic blot shown in Fig. 5B. A preliminary characterization of a hormonally regulated glucanase gene from *N. plumbaginifolia* (15) indicated that the exon/intron organization of this

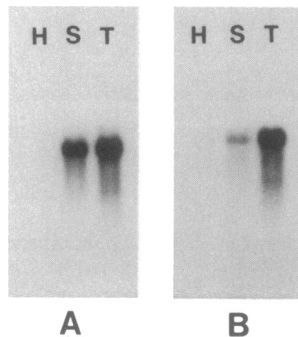


FIG. 6. Hybridization analysis of tobacco RNA. Total RNA was isolated from uninfected (H), TMV-infected (T), or salicylic acid-sprayed (S) tobacco leaf, electrophoresed, and blotted. Identical blots were hybridized to cDNA probes corresponding to acidic (A) and basic (B) tobacco glucanases.

gene is similar to that of the *Nicotiana tabacum*  $\beta$ -1,3-glucanase genes.

After removal of the N-terminal signal peptide, the C-terminal sequence of 22 residues of the basic tobacco  $\beta$ -1,3-glucanase becomes glycosylated and is subsequently cleaved off during vacuolar targeting of the protein (14). The protein encoded by clone gI9, which was definitively identified as an extracellular acidic  $\beta$ -1,3-glucanase, lacks this C-terminal extension. A possible role of such a C-terminal extension in vacuolar targeting can be studied by the expression of modified proteins in transgenic plants.

A comparison of the upstream sequences of PR genes from groups 1 and 5 revealed little sequence similarities (17). Recently, we reported that in the PR-1a promoter TMV- and salicylate-responsive elements are probably located between nucleotides 625 and 689 upstream of the transcription start site (18). Preliminary results indicate that in the PR-S promoter (group 5) TMV-responsive elements are located >1 kilobase upstream of the gene (H. Albrecht and J.F.B., unpublished). The cloning of several group 2 genes reported in this study permits a further investigation on the mechanism of the induction of these putative defense genes.

The antiserum against PR-2 and -N was a gift of Dr. L. C. Van Loon and cDNA clone pCNT11 was kindly donated by Dr. Johan Memelink. We gratefully acknowledge the assistance of Dr. Marianne Huisman, Luc van de Seyp, and Menno Knetsch. The continuous interest of Dr. Peter van den Elzen is highly appreciated. This work was sponsored in part by the Foundation for Technical Sciences, with financial support from the Netherlands Organization for Scientific Research.

- Carr, J. P. & Klessig, D. F. (1989) in *Genetic Engineering: Principles and Methods*, ed. Setlow, J. K. (Plenum, New York), Vol. 11, pp. 65-109.
- Bol, J. F., Linthorst, H. J. M. & Cornelissen, B. J. C. (1990) *Annu. Rev. Phytopathol.* **28**, 113-138.
- Van Loon, L. C., Gerritsen, Y. A. M. & Ritter, C. E. (1987) *Plant Mol. Biol.* **9**, 593-609.
- Memelink, J., Linthorst, H. J. M., Schilperoort, R. A. & Hoge, H. C. (1990) *Plant Mol. Biol.* **14**, 119-126.
- Lotan, T., Ori, N. & Fluhr, R. (1989) *Plant Cell* **1**, 881-887.
- Kauffmann, S., Legrand, M., Geoffroy, P. & Fritig, B. (1987) *EMBO J.* **6**, 3209-3212.
- Fritig, B., Rouster, J., Kauffmann, S., Stinzi, A., Geoffroy, P., Kopp, M. & Legrand, M. (1989) in *Signal Molecules in Plants and Plant-Microbe Interactions*, ed. Lugtenburg, B. J. J. (Springer, Berlin), pp. 161-168.
- Legrand, M., Kauffmann, S., Geoffroy, P. & Fritig, B. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 6750-6754.
- Mauch, T., Mauch-Mani, B. & Boller, T. (1988) *Plant Physiol.* **88**, 936-942.
- Linthorst, H. J. M., Van Loon, L. C., Van Rossum, C. M. A., Mayer, A., Bol, J. F., Van Roekel, J. S. C., Meulenhoff, E. J. S. & Cornelissen, B. J. C. (1990) *Mol. Plant-Microbe Interact.*, 252-258.
- Young, R. A. & Davis, R. W. (1983) *Science* **222**, 280-287.
- Cornelissen, B. J. C., Horowitz, J., Van Kan, J. A. L., Goldberg, R. B. & Bol, J. F. (1987) *Nucleic Acids Res.* **15**, 6799-6811.
- Hoof van Huijsduijnen, R. A. M., Alblas, S. W., De Rijk, R. H. & Bol, J. F. (1986) *J. Gen. Virol.* **67**, 2135-2143.
- Shinshi, H., Wenzler, H., Neuhaus, J.-M., Felix, G., Hofsteenge, J. & Meins, F. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5541-5545.
- De Loose, M., Alliotte, T., Gheysen, G., Genetello, C., Gielen, J., Soetaert, P., Van Montagu, M. & Inzé, D. (1988) *Gene* **70**, 13-23.
- Van den Bulcke, M., Bauw, G., Castresana, C., Van Montagu, M. & Vandekerckhove, J. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 2673-2676.
- Van Kan, J. A. L., Van de Rhee, M. D., Zuidema, D., Cornelissen, B. J. C. & Bol, J. F. (1989) *Plant Mol. Biol.* **12**, 153-155.
- Van de Rhee, M. D., Van Kan, J. A. L., González-Jaén, M. T. & Bol, J. F. (1990) *Plant Cell* **2**, 357-366.