

# Sequence of a *Nicotiana plumbaginifolia* $\beta(1,3)$ -glucanase gene encoding a vacuolar isoform

Godelieve Gheysen, Dirk Inzé, Piet Soetaert<sup>1</sup>, Marc Van Montagu\* and Carmen Castresana  
Laboratorium voor Genetica, Rijksuniversiteit Gent, K.L.Ledeganckstraat 35, B-9000 Gent and <sup>1</sup>Plant Genetic Systems N.V., Plateaustraat 22, B-9000 Gent, Belgium

Submitted October 18, 1990

EMBL accession no. X54742

$\beta(1,3)$ -glucanases are a group of hormonally and developmentally regulated hydrolytic plant enzymes, which are also induced upon pathogen infection. A cDNA and genomic clone of the *Nicotiana plumbaginifolia*  $\beta(1,3)$ -glucanase *gn1* gene have been analyzed previously (1, 2). Here, we report the nucleotide and deduced amino acid sequences of a different *N. plumbaginifolia*  $\beta(1,3)$ -glucanase gene designated *gn2*.

The coding region of *gn2* is contained within two exons and yields a precursor protein of 365 amino acids. In the absence of the corresponding cDNA, the length of the single intron present in *gn2* was determined by primer extension and sequencing of the transcribed mRNA by the dideoxy procedure (3). A synthetic oligonucleotide complementary to nucleotides 16 to 56 downstream from the 5' end of the second exon was utilized for this experiment. Similarly, primer extension analysis using the same oligonucleotide defined the cap site (arrow in the figure) of the transcript derived from *gn2* at 14 base pairs (bp) upstream from the first ATG of the coding region. The sequences TATAAAT and CCAAT (underlined in the figure) likely corresponding to the *gn2* TATA and CAAT boxes, reside at positions -29 to -35 and -78 to -82 from the cap site, respectively. The N-terminal amino acid sequence of the GN2 protein has previously been determined (4) and the start of the mature GN2 was identified (boxed in figure). The nucleotide sequence data revealed that *gn2* encodes a precursor protein containing a 32-amino acid N-terminal signal peptide for translocation to the endoplasmic reticulum.

Partial amino acid sequence analysis identified two distinct classes of  $\beta(1,3)$ -glucanases in tobacco plants: the extracellular acidic isoforms and the basic isoforms localized primarily in the central vacuole of the cell (5). The *gn2* gene product shows more than 95% identity to the tobacco vacuolar  $\beta(1,3)$ -glucanases, whereas it is only 50–65% homologous to the extracellular isoforms. In addition, *gn2* shows 97.5% identity to the partial cDNA sequence of an intracellular isoform from tobacco (6). These results indicate that GN2 most likely represents a vacuolar  $\beta(1,3)$ -glucanase.

## ACKNOWLEDGEMENTS

This work has been supported by a grant from the Services of the Prime Minister (U.I.A.P. 12OC0187). C.C. is indebted to the Commission of European Communities for a postdoctoral fellowship.

```

-783 GGGAAACATTTAATAGCGAAACGGGACCGCAATTCCTAAACAGCATCGCCGGCTGCTAC -724
-723 ACTCTCGCTGCCATATTCACATAGCGCTCTGTGGTACGCAAAATAGTAAATTAATTC -664
-663 GGGATAGAATTTGAGATTCATTAATTCAGCTTTAATTAATTAAGTATTAGCTAAATTAAT -604
-603 AGCTAATCCGATAGCCAGCCGCGGGAATCGACTTCCTCAATAAATGACTAGCAATAT -544
-543 TAGAAACTTTTCATTAATAATCTAGCATCGCAATGCAAAATGAAAGCTCATATGAACTTTCA -484
-483 AATCTCGAATGCTATATGCGAATTCATTAATGCGAATGCTATTAATGCTACTTATC -424
-423 GTTAAATAATATTTTATTTAATCTGCAATGAAATATTAATTAATGCTTGGTTTAAAT -364
-363 ATAAAGATAGATATCGCTTACTATTAACCCAGAACATTTAATTAAGCTTAAGCATGT -304
-303 AAACAGCAATTCCTGCTGACGAGGATGAAAGAACTAATCACTGATATAATTTCTTTT -244
-243 GGAAGTCAAGATAGTACGGCCTTATCATTTATGCGAATAATTAATAAATTTAATCATCAT -184
-183 CAACTCATCTCTTTTTCGCAAAATGACAGCTTGCATTAATAACGAGCTCAGCCG -124
-123 GCGCGCTCTTAATCTTCAATAGCTAGCAATAACAACTGCAATATATTCGACTCTA -64

-63 ATATATATTTACAGAGGCTATCTCAATATAATAGCTCTCTGCTGATCTTAATTTGCC -4
-3 CAATAGCTCTTCATCATGCTGACTCATATAAACAATAACTCTGCAATGGCTGCTA 57
I H S T L H K H N T P Q H A A
58 TCAGACTCTAGGATTAAGTCTGCTGCGACGAGCAATTCAGATAGCACTTTCTGCTCAA 117
I T L L G L L L L V A S S I E I A
118 ATATTTCAACTTCGAGGCAAAAATTTATTTGCTATAAATTTCTGCTGCAAAATTTAA 177
178 TTTTGATAATTTCTCTTCAAAATGCTCTGCTTTGCTCATATACTTAAGTACAGGCT 237
238 ATTCATTAGCGATAGCTGCTTCTGCTATTTTCCGGTAACTGACTCAACAGCTTTTCA 297
298 ATTTACGGGGGAGCGCGTAAATCAAGATTTGAAATTTAGCTTTAGCTAATTCGAT 357
358 TTGATTTATGGCTAGAGCTCTATTTTGTATAGTAAATCTTTAAGCATAGACCGAT 417
418 CTGAATCAAGGCTACTATCTGCTATTAAGCCATAGAGATGCTGCTGCTGCTAAATTT 477
478 ATGTACACACTTCTTATAAATTTGCTCCAAAGAGAAATGCGATTTTCAATTTAGTATC 537
538 ACTTTACCTTTAACTTTTATTTTAACTAATAGAGCACTTTTAACTAGCATCAATAA 597
598 TTATGGAATAGATAAGCAGCAATTTGCAAAATTTGCAAAATTTAATGCTATATAATTTAGCA 657
658 TTTTAAATAGCAACTTTCAAAATTTCTTCTTCCCTTAAACAAATTTATGCTATATAATTTAGCA 717
718 TTAATTTAGAGGGGTTATAGTACATGATCAATGAGCGGCTAGTGGAGGGGAGCTACTGATGT 777
778 TTTTAACTTTGATTTGATTAAGAGCTGCTTTGCGACATAAATTTAAT 837
838 TCTTTTTCGATAAGCAAAATTTATTTTCTTAAAGAGGACTTTTCAAAAATTTTCA 897
898 ATTTCACTCATTTAAATTTTAAAAATTTGCAAAATTTTAAATGCTATATAATTTCAAAAT 957
958 TTCAGTCAAAATCATTACAGCACTGCAAAAGAGCGCAAAATGCTTAAACAAAATTTT 1017
1018 TCAGCACTGTTAGAAATTTTATTCGCAAGCGGCACTGACTGATTTTAACTATATTTGA 1077
1078 GCAAGCGGCTTATGATAGCTTTTCTTATGAAATGCTGAGCAAACTGCGAAATCAAT
G A A S I G V C Y G H L G N N L P N H
1138 TGGGAAGTTATAGAGCTCTCAAGCAATCAAGAAATAGCAAGAGCTGAGGCTTTATGATCCA 1197
W E V I Q L Y K S R N I G R L R L Y D P
1198 AATCATGAGCTTTACAGCATTAAGAGGCTCAACATAAATGCTGATGAGAGCTGCC 1257
N H G L Q A I K C S H I E V H G L P
1258 AATTCAGATGCAAGCAGTCTGCTCCGAGTGGAGCTAGTACTGAGTCTGGTACAGAAA 1317
N S D V K H I A S G H E A R W V W V Q K
1318 AATGTTAAGATTTTGGCCAGATTAAGATTAAGTATTTGCTGCTGGAAATGAACT 1377
N V K P F W P D K E V I A G C H E I
1378 AGCATTTGCTGAGATAGCTTATGATTAAGAGCTGCTTTGCGACATAAATTTAAT 1437
S P V T G T S Y L T S F L T P A M V N I
1438 TAGAAGCAATTTGCTGAGCTCTTGGAAACCAATGCAAGAGCTGAGGCTTTATGATGAC 1497
Y K A I G E A C L C N N I K V S T S V D
1498 ATGAGCTTTGATGCAAGCTTTATCAGCAGTACAGGCTGCTTTGAGAGGATGCTAGG 1557
H T L I C G S D P S G F S R N D A R
1558 TGCTTTGATCGGATTTGCTGCTTCTTGAAGGACACAGCTGACCTTACTGCTTAA 1617
W F V D P I V C F L R D T L R A P L L V N
1618 ATTTAGCCATTTAGAGTATTTGCTGATTAAGGCGAGATTTCTGCTGCTGATTTCTTT 1677
I Y P F S W P S G F P C Q I S P Y S L
1678 TTTAGAGCAAAATGCTGACTAGCAAGCTGCTCAGCGCAATAGCAATTTGATTTGAT 1737
F T A P N V V V Q D G S R Q Y R N L F D
1738 GCAATCTGATGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1797
A H L D S V Y A A I E R S G C A S V G I
1798 GTTCTGCGAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1857
V V S E S W P S G F S R N D A R
1858 GCAACTTCTTGAAGCAATTAATCAACCGCTAAGAGGCTGACCCCAAGAGGCTGAG 1917
A T Y L K N L I Q H A K E G S P R K P R
1918 CCTATTGAGCCTATATTTGCCATGTTTATGAGAGCAACCAAGCCTGAGCTGAG 1977
F I E F Y I F A H F D E N N K N P E L E
1978 AAGACTTTGATTTGCTGCGACAGAGCCGCAATTAATTAATGCTGCTGCTGCTGCTGCTG 2037
K H F C L F S P N K Q P K Y N L N F G V
2038 TCTGCTCAGTGAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2097
S G S V E T N A T A S L I S E I
2098 ACAGCTGAAATCTCTTACATAAGTATTAACCTAGTAAAGCAGAGACTTTTTTTTT 2157
2158 TTTTAAATGAGAGCAATAGATAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2217
2218 GCGAAGCATGCAACTCTGAG

```

## REFERENCES

- De Loose, M., Alliotte, T., Gheysen, G., Genetello, C., Gielen, J., Soetaert, P., Van Montagu, M. and Inzé, D. (1988) *Gene* **70**, 13–23.
- Castresana, C., de Carvalho, F., Gheysen, G., Habets, M., Inzé, D. and Van Montagu, M. (1990) *The Plant Cell* **2**, in press.
- Bertholet, C., Van Meir, E., ten Heggeler-Bordier, B. and Wittek, R. (1987) *Cell* **50**, 153–160.
- Bauw, G., De Loose, M., Inzé, D., Van Montagu, M. and Vandekerckhove, J. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 4806–4810.
- Van den Bulcke, M., Bauw, G., Castresana, C., Van Montagu, M. and Vandekerckhove, J. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 2673–2677.
- Shinshi, H., Wenzler, H., Neuhaus, J.-M., Felix, G., Hofsteenge, J. and Meins, F.Jr. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5541–5545.

\* To whom correspondence should be addressed