



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵ : C12N 15/29, C07K 13/00 C12N 15/70, 15/82, C12Q 1/68	A1	(11) International Publication Number: WO 93/11241 (43) International Publication Date: 10 June 1993 (10.06.93)
(21) International Application Number: PCT/IT92/00158 (22) International Filing Date: 4 December 1992 (04.12.92) (30) Priority data: RM91A000915 6 December 1991 (06.12.91) IT (71) Applicants (for all designated States except US): CONSIGLIO NAZIONALE DELLE RICERCHE [IT/IT]; P. le Aldo Moro, 7, I-00185 Rome (IT). UNIVERSITY OF GEORGIA RESEARCH FOUNDATION INC. [US/US]; Body Graduate Studies Research Center, Athens, GA 30602-7411 (US). (72) Inventors; and (75) Inventors/Applicants (for US only) : CERVONE, Felice [IT/IT]; DE LORENZO, Giulia [IT/IT]; SALVI, Giovanni [IT/IT]; Dip. di Biologia vegetale, Università degli Studi di Roma "La Sapienza", P. le Aldo Moro, 5, I-00185 Rome (IT). ALBERSHEIM, Peter [US/US]; 1575 Morton Road, Athens, GA 30605 (US). DARVILL, Alan [GB/US]; 440 Sandstone Drive, Athens, GA 30605 (US). BERGMANN, Carl [US/US]; 124 Crest Way, Athens, GA 30605 (US).		(74) Agents: BANCHETTI, Marina et al.; Ing. Barzano' & Zanardo Roma S.p.A., Via Piemonte, 26, I-00187 Rome (IT). (81) Designated States: AT, AU, BB, BG, BR, CA, CH, CS, DE, DK, ES, FI, GB, HU, JP, KP, KR, LK, LU, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, SN, TD, TG). Published <i>With international search report.</i>
(54) Title: NUCLEOTIDE SEQUENCES CODING AN ENDOPOLYGALACTURONASE INHIBITOR (57) Abstract The gene coding a fungine endopolygalacturonase inhibitor (PGIP) is cloned and sequenced from both <i>Phaseolus vulgaris</i> genomic and cDNA libraries. The PGIP amino acid sequence, vectors comprising the PGIP gene or parts thereof, cells transformed by said vectors are also disclosed. The isolated sequences are useful to transform plants, which are sensitive to fungi or microorganism pathogen activities, into plants producing the PGIP protein.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NL	Netherlands
BE	Belgium	GN	Guinea	NO	Norway
BF	Burkina Faso	GR	Greece	NZ	New Zealand
BG	Bulgaria	HU	Hungary	PL	Poland
BJ	Benin	IE	Ireland	PT	Portugal
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SK	Slovak Republic
CI	Côte d'Ivoire	LJ	Liechtenstein	SN	Senegal
CM	Cameroon	LK	Sri Lanka	SU	Soviet Union
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	MC	Monaco	TG	Togo
DE	Germany	MG	Madagascar	UA	Ukraine
DK	Denmark	ML	Mali	US	United States of America
ES	Spain	MN	Mongolia	VN	Viet Nam
FI	Finland				

NUCLEOTIDE SEQUENCES CODING AN ENDOPOLYGALACTURONASE
INHIBITOR

This invention relates to nucleotide sequences
5 coding a endopolygalacturonase inhibitor.

More particularly, this invention relates to a
gene coding a new protein having an inhibiting activity
against the fungine endopolygalacturonase enzyme, or
fractions thereof, and to the protein itself or
10 fractions thereof. The invention relates also to
recombinant vectors comprising such sequences, to cells
and plants transformed by means of such recombinant
vectors and to the use of such vectors and cells to
produce the protein.

15 References are enclosed at the end of this
specification as they are orderly quoted.

The fungine endo alpha-1,4-D-polygalactorunase
enzyme (EC 3.21.15) is an important pathogenic factor
for plants (1,2). Actually, the enzyme catalizes the
20 scission reaction of polygalacturonic acid and induces
the subsequent solubilization of plant cell wall
homo-alpha-1,4-D-galactorunanes, thus facilitating the
fungi to penetrate into plant tissues and then
supplying the same fungi with nutritional substrates.

25 Moreover, the activity of such an enzyme could,
if controlled, produce non-virulence factors, i.e.

factors able to activate some plant defence mechanisms by converting plant cell wall homogalacturonanes into oligogalacturonides, which are known to stimulate phytoalexine synthesis and other defence mechanisms as well (2).

An inhibiting activity of fungine endopolygalacturonase, associated to the cell wall, named PGIP, has been detected in dicot plants.

The authors of the present invention have already detected such activity in different plant tissues of Phaseolus vulgaris L. (3), showing that the same is specific for fungine endopolygalacturonases, with no activity with either bacterial or vegetable enzymes, or other pectic enzymes of microbial origin (4). Moreover, the authors found that the PGIP is able to stabilize in vitro an oligogalacturonide mixture which stimulates phytoalexines (5).

Therefore it is evident the need to purify and isolate the molecule responsible for such activity by means of cloning coding PGIP sequences. Said sequences may be inserted in suitable vectors to transform plants that are sensitive to fungi or other microorganism pathogens, into plants producing the PGIP protein on large scale. The protein could be then used with many advantages in the food industry.

Till now attempts to isolate the PGIP gene have been unsuccessful, essentially due to low levels of the protein in vegetable tissues.

The authors of the invention have now isolated
5 and sequenced the PGIP protein coding DNA, both from a genomic and cDNA library, and have also identified the amino acid sequence of the PGIP protein.

It is an object of the present invention a nucleotide sequence comprising a sequence coding a PGIP
10 protein having an inhibiting activity of the fungine endo-alpha-1,4-D-polygalacturonase enzyme (PG), or parts thereof. Preferably said coding sequence is of vegetable origin, more preferably isolated either from plants, or parts of plants, or in vitro cultured
15 vegetable cells of the Phaseolus genus, most preferably of the Phaseolus vulgaris species.

According to a preferred embodiment said coding sequence codes a PGIP protein having the amino acid sequence of SEQ ID N2, or parts thereof. Alternatively
20 said amino acid sequence SEQ ID N2 lacks or is substituted of one or more amino acids, preferably is a variant of a PGIP protein, more preferably said variant of a PGIP protein comprises the amino acid sequence SEQ ID N4, or parts thereof.

25 According to another embodiment said nucleotide sequence comprises the nucleotide sequence of SEQ ID N1, consisting of:

- a coding region from nucleotide 1 to nucleotide 1026;
- a 3'end untranslated sequence from nucleotide 1027 to nucleotide 1116.

Alternatively said nucleotide sequence lacks or
5 is substituted of one or more nucleotides, preferably
comprising the nucleotide sequence of SEQ ID N3.

It is another object of the invention a
nucleotide sequence hybridizing to at least one of the
sequences above described, preferably complementary to
10 them.

It is another object of the invention a
nucleotide sequence comprising an allele derivative of
either SEQ ID N1 or SEQ ID N3 sequences.

All the sequences according to the inventions may
15 be derived from either a genomic library, or a cDNA
library or in vitro synthesis.

Another object of the invention is a
substantially purified PGIP protein having an
inhibiting activity for the fungine PG enzyme and being
20 of natural or recombinant origin, or parts thereof.
Preferably said substantially purified PGIP protein
derives from plants, parts thereof or cultured plant
cells, more preferably of the Phaseolus genus, most
preferably of the Phaseolus vulgaris species.

25 According to a preferred embodiment said protein
comprises the amino acid sequence of SEQ ID N2, or
parts thereof, alternatively said protein lacks or is

substituted of one or more amino acids, preferably said protein comprises the amino acid sequence of SEQ ID N4.

Another object of the invention is a recombinant vector comprising anyone of nucleotide sequences of the invention, preferably of plasmid or phage origin, more preferably said recombinant vector is the pAD-1 plasmid (DSM N.6821); alternatively said vector is either the pPT-1 or the pPT-2 plasmid, or the lambda 3.3 PGIP phage.

In a preferred embodiment said recombinant vector comprises a promoter able to express said coding PGIP nucleotide sequence, preferably said promoter is active either in bacteria, or in yeasts, or in high plants.

Another object of the invention are cells transformed by one of vectors above described, preferably said cells are bacterial cells, more preferably belonging to the Escherichia coli species. Alternatively said cells are eucaryotic cells, preferably yeast cells, alternatively a vegetable cells.

Further object of the invention is a process for isolating and cloning PGIP protein coding sequences, or parts thereof, comprising:

- partial sequencing the amino acid sequence of the PGIP protein;

- in vitro synthesizing oligonucleotides, sequence of which derived from said amino acid sequence;
- probing a genomic or a cDNA library from plant cells with said oligonucleotide, previously labelled;
- 5 - isolating positive clones, extracting the recombinant vector and identifying the recombinant insert from said recombinant vector;
- sequencing said recombinant insert.

The invention shall be illustrated in the following examples by making reference to some preferred embodiments, completed with the following figures, where:

Fig. 1 shows the PGIP elution profile on Mono S column;
Fig. 2 shows an agarose gel electrophoresis of amplified DNA, using the following primers: A: mol weight markers; B: control; C: N-A and Int; D: N-B and Int; E: N-A; F: N-B;

Fig. 3 shows a map of primers and of amplified DNAs;
Fig. 4 shows a restriction map of lambda PGIP-3.3 and of pPT-2;

Fig. 5 shows an hydrophoby plot of the PGIP protein;
Fig. 6 shows a Northern blot by using RNAs from different plant tissues, probed with a radiolabelled pPT-1 fragment;

Fig. 7 shows a comparative analysis of PGIP coding sequences isolated from two different cultivars of Phaseolus vulgaris species.

EXAMPLE 1 - Purification and sequencing of PGIP protein

Plants of Phaseolus vulgaris L., Pinto cultivar, are grown on vermiculite under 18 hour light exposure conditions. The callus and cell suspension cultures from Pinto cultivar are obtained as described in (3).

The PGIP protein is purified from hypocotyls (30 $\mu\text{g/kg}$ of fresh weight) by affinity chromatography on Sepharose coupled to endopolygalacturonase, as described in (6). One unit of PGIP protein is the amount able to reduce of 50% the activity of one unit of endopolygalacturonase, isolated from Aspergillus nidulans.

PGIP is then diluted at 1:4 ratio in 20 mM Hepes pH 7.5 buffer, and loaded on Mono S column (Pharmacia), equilibrated with the same buffer. The protein is eluted with a linear gradient 0-0.3 M NaCl in 20 mM Hepes pH 7.5 buffer. Approximately 90% of the loaded material having a 280 nm absorbance elutes with the excluded volume, showing no PGIP activity at all. The remaining 10% elutes with a single peak at a 0.1 M NaCl and shows PGIP activity (fig.1). Fractions containing PGIP activity are collected and loaded again on Mono S column to concentrate the protein to 35 $\mu\text{g/ml}$.

5 μg of the protein are used to determine the N-terminal sequence, according to standard methodology. The remaining amount is digested by 1 mg/ml of TPCK (Sigma) treated trypsin in 1.0 M ammonium bicarbonate

pH 7.8 buffer, at 37°C for 16 hours. The resulting peptides are separated using a binary gradient (solvent A: 0.1% TFA in water; solvent B: 0.085% TFA in 80% aqueous CH₃CN) at a flow rate of 0.2 ml/min. Single peaks are collected, dried, dissolved in 50 µl of solvent A and chromatographed again on the same column.

Peptides, as well as the undigested PGIP protein, are sequenced by a ABI 470A or 477A sequencer. Aminoacid sequences of the N-terminal fragment and of four tryptic fragments of the PGIP protein are shown in the Table 1 herebelow.

15

20

25

Table 1

Aminoacid sequence of the N-terminal fragment and of four tryptic peptides of PGIP protein

Residue number	N-terminal	P-1	P-2	P-3	P-4
5	1	Glu	Ile	Leu	Ile
	2	Leu	His	Phe	Tyr
	3	X	Leu	Thr	X
	4	Asn	Ala	Ser	Thr
	5	Pro	Lys	Met	Leu
	6	Glu(n)			Pro
	7	Asp			Gln
	8	Lys			Gly
	9	Glu(n)			Leu
10	10	Ala			Thr
	11	Leu			Gln
	12	Leu			Leu
	13	X			Lys
	14	Ile			
	15	Lys			
	16	Lys			
	17	Asp			
	18	Leu			
	19	Gly			
15	20	Asn			
	21	Pro			

X means presence of not defined residues;
 aa. 6 and 9 of the N-terminal sequence may be either Glu or Gln.

EXAMPLE 2 - Nucleotide amplification, sequencing and cloning

Two independent oligonucleotides are synthesized on the basis of the N-terminal sequence (N-A and N-B). A third oligonucleotide (int) is synthesized on the basis of the peptide P-3 sequence, starting from the first nucleotide coding for aa. 6 (Pro) to the second

nucleotide coding for aa. 11 (Gln). The nucleotide sequences are shown in the following Table 2.

Table 2

Primer oligonucleotides for amplification reaction and corresponding amino acid sequence

5

4-10 N-terminal sequence (N-A)

amino acid: Asn-Pro-Glu-Asp-Lys-Glu-Ala
(Gln) (Gln)

nucleotide: 5'AAT CCN GAA GAT AAA CAA GC 3'
C C G C G G G

10

14-21 N-terminal sequence (N-B)

amino acid: Ile-Lys-Lys-Asp-Leu-Gly-Asn-Pro

nucleotide: 5'ATT AAA AAA GAT CTN GGN AAT CC 3'
C G G C T C
A

15

6-11 P-3 peptide sequence (Int)

amino acid: Pro-Gln-Gly-Leu-Thr-Gln

nucleotide: 3'GGN GTT CCN AAN TGN GT 5'
C G

N means one of the four nucleotides.

20

N-A, N-B and Int oligonucleotides are used as primers of a polymerase chain reaction (PCR) using genomic DNA, extracted from Phaseolus vulgaris, Pinto cultivar, as template, by conventional methods (7). Polymerizing reactions are carried out on a Perkin Elmer/Cetus equipment according to manufacturer's instructions. The primer hybridization is carried out at 45°C or 50°C, depending upon using respectively

25

redundant or nonredundant nucleotides. The polymerization enzyme (Amplitaq, Perkin Elmer) is used at 1 U/40 μ l of reaction mix. PCR products are fractionated on a 0.8% agarose gel and dyed with ethidium bromide.

A clear DNA band of appr. 0.79 kb, defined as PGIP-1, is evidentiated (fig.2, line c), when using the N-A oligo. The band is missing in control reactions (fig.2, lines b, e, g). By using N-B and int oligos as primers, the DNA product (PGIP-2) is of appr. 0.76 kb, i.e. 30 bp shorter than PGIP-1 (fig.2, line d).

The agarose gel is transferred to nitrocellulose membrane and probed with PGIP-2, previously eluted from the agarose gel and amplified with 6 PCR cycles in presence of P^{32} ATP. PGIP-1 and PGIP-2 sequences hybridized with the same intensity to the radiolabelled probe. Fig.3 shows a scheme of oligonucleotides and PCR obtained products.

The PGIP-2 fragment, eluted from the agarose gel and further amplified by PCR using N-B and Int oligonucleotides, is purified by removing the low melting point agarose and inserted into the pGEM7zf(+) vector (Promega) by conventional methods. The recombinant vector, containing the whole DNA fragment of PGIP-2, is defined as pPT-1.

The pPT-1 insert is sequenced and an open reading frame of 252 amino acids is deducted. As expected the

amino acid sequence comprises sequences corresponding to the 14-21 N-terminal and to the P-3 peptide as well. Moreover it comprises sequences of the P-1, P-2 and P-4 peptides.

5 EXAMPLE 3 - Isolation of a genomic PGIP clone

A genomic DNA library from the Saxa cultivar of Phaseolus vulgaris, obtained by partial digestion with restriction enzyme MboI, into the EMBL3 lambda vector, has been supplied by Clontech (Ca, USA). The library is
10 probed by PCR and plaque-hybridization combination. 5×10^5 recombinant phages are probed by conventional plaque-hybridization, as described in (7), by using the radiolabelled p-PT1 insert. Positive clones are grouped in three different classes, to be further analyzed with
15 PCR. PCR screening is performed using crude phage DNA as template and two non redundant primers (P21 and P22), synthesized on the basis of the p-PT1 insert sequence, whose sequences are as follows:

P21:AAGGCTCTGTTTCAGTTGCTT; P22:CTATACATTGGCGGCATCAAT.

20 Positive clones have an expected 378 bp amplification fragment. The first class shows strong hybridization signals and is able to amplify the 378 bp fragment. The second class also shows strong hybridization signals, but amplification products from PCR are not
25 evidentiated. The third class shows a weaker hybridization, but gives rise to specific amplification products from PCR, other than the 378 bp fragment.

Among positive clones of the first class, one clone (lambda PGIP-3.3. insert = 17 kb) is selected and isolated for a further analysis. A 3.3 kb SalI-SalI fragment shows to contain the PGIP gene, as demonstrated by Southern blot analysis and PCR experiments. The fragment is subcloned in the SalI site of the Bluescript SK+ plasmid (Stratagene), resulting in a recombinant plasmid named pPT-2. The orientation and the position of the PGIP gene in the 3.3 kb cloned fragment is defined by PCR (fig.4), using different combination of both plasmid (T3 and T7) and PGIP specific primers (fig.4). It is calculated that the 3.3 kb SalI-SalI fragment comprises appr. 1860 nt upstream the first ATG, 1026 nt of coding region and 97 nt of the 3' untranslated region, with appr. 450 nt adjacent sequences. A partial restriction map of the 17 kb genomic insert, the orientation and the PGIP position in the pPT-2 subclone are shown in fig.4.

Both DNA strands of the PGIP gene are sequenced; the sequence strategy is shown in fig.4. The whole nucleotide sequence of the PGIP coding region and of the untranslated 3' region (1116 nt) is reported in SEQ ID N1. The polyadenylation site is 97 bp downstream the TAA stop codon. Three sequences, corresponding to putative eucaryotic polyadenylation signals, are at 42, 22 and 16 nt upstream the polyadenylation site respectively. A comparative analysis between the

genomic sequence and the relevant cDNA sequence shows that no introns are present in the 345 nucleotide 3' end gene region. Further analysis shows that no introns are present for the whole PGIP gene, as confirmed by
5 PCR experiments. The PGIP gene DNA sequence defines a 342 amino acid PGIP protein (see SEQ ID N1 and N2). Both the nucleotide and amino acid sequences do not show any significant homology with any known sequence.

A hydrophoby plot of the PGIP protein is shown in
10 fig.5. The sequence includes four potential glycosilation sites.

EXAMPLE 4 - Northern blot analysis

PGIP specific mRNA levels are analysed in different plant tissues. The pPT1 purified insert is
15 used to probe polyA⁺ RNA, isolated from either hypocondils, or leaves, or flowers or suspension cultured cells of Phaseolus vulgaris, Pinto cultivar. A single 1,2 kb transcript is identified in mRNA from suspension cultured cells. The fragment is also present
20 in other tissues, even if minor amount.

PCR experiments, using single-strand cDNA from suspension cultured cells and P-21 and P-22 as templates, shows the expected amplification of the 378 bp fragment, thus indicating that at least a fraction
25 of the PGIP gene expression in cultured cells may be ascribed to the cloned and characterized gene.

EXAMPLE 5 - Analysis of PGIP genes from two different Phaseolus vulgaris cultivars

A cDNA library is made in the lambda gt11 vector from polyA⁺ RNA, purified from suspension cultured cells of Phaseolus vulgaris, Pinto cultivar. By probing the cDNA library with pPT1, a cDNA recombinant clone (lambda AD-1) containing PGIP coding sequences is isolated. The lambda AD-1 insert (842 bp) is subcloned in the Bluescript SK+ vector, named pAD-1, deposited at DSM, Germany, with the n. 6821, and sequenced. The insert represents a whole PGIP cDNA fragment and includes the polyA tail. The sequence of the pAD-1 insert is shown in SEQ ID N3. When compared to SEQ ID N1, the pAD-1 insert refers to sequence from 345 to 1116 nt of SEQ ID N1. A comparative analysis between the genomic clone from Saxa cultivar and the cDNA clone from Pinto cultivar shows a 96.4% homology of nucleotide sequence and a 97.4% homology of amino acid sequence (fig.7).

This invention has been described with reference to some preferred embodiments; however it is understood that any changement and/or modification could be made by anyone skilled in the art, remaining within the scope of the claims.

REFERENCES

- 1) Cooper, R.M. In: Plant Disease: Infection, Damage and Loss. Wood, R.K.S. & Jellis, G.J. eds, Oxford, Blackwell Scientific Publications, pp. 13-27 (1984).
- 2) Hahn, M.G. Bucheli, P., Cervone F., Doares, S.H.,
5 Darvill, A.G. and Albersheim, P. In: Plant-Microbe Interaction, Vol.3, Nester, E. and Kosuge, T.,eds, New York, MacMillan Press, pp. 131-181 (1989).
- 3) Salvi, G., Giarriizzo, F., De Lorenzo, G. and Cervone, F. J. Plant Physiol., 136, 513-518 (1990).
- 10 4) Cervone, F., De Lorenzo, G. Pressey, R., Darvill, A., and Albersheim, P. Phytochemistry 29, 447-449 (1990).
- 5) Cervone F., Hahn, N.G., De Lorenzo, G., Darvill, A. and Albersheim, P. Plant Physiology, 90, 542-548
15 (1989).
- 6) Degrà, L, Salvi, G., Mariotti, D., De Lorenzo, G. and Cervone, F. J. Plant Physiol., 133, 364-366 (1988).
- 7) Sambrook, D., Fritsch, E.F., Nianiat, T. In: Molecular Cloning: A Laboratory Manual, 2nd ed.,
20 vols.1, 2, 3, CSH Lab Press (1989).

- 17 -

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: Consiglio Nazionale delle Ricerche
(B) STREET: P.le Aldo Moro 7
(C) CITY: Roma
(E) COUNTRY: Italy
(F) POSTAL CODE (ZIP): 00185

(A) NAME: University of Georgia Research Foundation
(B) STREET: Boyd Graduate Studies Research Ctr.
(C) CITY: Athens
(D) STATE: Georgia
(E) COUNTRY: USA
(F) POSTAL CODE (ZIP): 30602

(ii) TITLE OF INVENTION: Nucleotide sequences coding an
endopolygalacturonase inhibitor

(iii) NUMBER OF SEQUENCES: 4

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(vi) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: IT RM91/000915
(B) FILING DATE: 06-DEC-1991

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1116 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Phaseolus vulgaris
(B) STRAIN: Saxa

(vii) IMMEDIATE SOURCE:

(B) CLONE: lambda PGIP-3.3

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 1..1026

(ix) FEATURE:

(A) NAME/KEY: 3'UTR
(B) LOCATION: 1027..1116

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATG	ACT	CAA	TTC	AAT	ATC	CCA	GTA	ACC	ATG	TCT	TCA	AGC	TTA	AGC	ATA	48
Met	Thr	Gln	Phe	Asn	Ile	Pro	Val	Thr	Met	Ser	Ser	Ser	Leu	Ser	Ile	
1				5					10					15		
ATT	TTG	GTC	ATT	CTT	GTA	TCT	TTG	AGA	ACT	GCA	CTC	TCA	GAG	CTA	TGC	96
Ile	Leu	Val	Ile	Leu	Val	Ser	Leu	Arg	Thr	Ala	Leu	Ser	Glu	Leu	Cys	
			20					25					30			
AAC	CCA	CAA	GAT	AAG	CAA	GCC	CTT	CTC	CAA	ATC	AAG	AAA	GAC	CTT	GGC	144
Asn	Pro	Gln	Asp	Lys	Gln	Ala	Leu	Leu	Gln	Ile	Lys	Lys	Asp	Leu	Gly	
		35					40					45				
AAC	CCA	ACC	ACT	CTC	TCT	TCA	TGG	CTT	CCA	ACC	ACC	GAC	TGT	TGT	AAC	192
Asn	Pro	Thr	Thr	Leu	Ser	Ser	Trp	Leu	Pro	Thr	Thr	Asp	Cys	Cys	Asn	
	50					55					60					
AGA	ACC	TGG	CTA	GGT	GTT	TTA	TGC	GAC	ACC	GAC	ACC	CAA	ACA	TAT	CGC	240
Arg	Thr	Trp	Leu	Gly	Val	Leu	Cys	Asp	Thr	Asp	Thr	Gln	Thr	Tyr	Arg	
65					70				75						80	
GTC	AAC	AAC	CTC	GAC	CTC	TCC	GGC	CAT	AAC	CTC	CCA	AAA	CCC	TAC	CCT	288
Val	Asn	Asn	Leu	Asp	Leu	Ser	Gly	His	Asn	Leu	Pro	Lys	Pro	Tyr	Pro	
				85					90					95		
ATC	CCT	TCC	TCC	CTC	GCC	AAC	CTC	CCC	TAC	CTC	AAT	TTT	CTA	TAC	ATT	336
Ile	Pro	Ser	Ser	Leu	Ala	Asn	Leu	Pro	Tyr	Leu	Asn	Phe	Leu	Tyr	Ile	
			100					105					110			
GGC	GGC	ATC	AAT	AAC	CTC	GTC	GGT	CCA	ATC	CCC	CCC	GCC	ATC	GCT	AAA	384
Gly	Gly	Ile	Asn	Asn	Leu	Val	Gly	Pro	Ile	Pro	Pro	Ala	Ile	Ala	Lys	
		115					120					125				
CTC	ACC	CAA	CTC	CAC	TAT	CTC	TAT	ATC	ACT	CAC	ACC	AAT	GTC	TCC	GGC	432
Leu	Thr	Gln	Leu	His	Tyr	Leu	Tyr	Ile	Thr	His	Thr	Asn	Val	Ser	Gly	
	130					135					140					
GCA	ATA	CCC	GAT	TTC	TTG	TCA	CAG	ATC	AAA	ACC	CTC	GTC	ACC	CTC	GAC	480
Ala	Ile	Pro	Asp	Phe	Leu	Ser	Gln	Ile	Lys	Thr	Leu	Val	Thr	Leu	Asp	
145					150					155					160	
TTC	TCC	TAC	AAC	GCC	CTC	TCC	GGC	ACC	CTC	CCT	CCC	TCC	ATC	TCT	TCT	528
Phe	Ser	Tyr	Asn	Ala	Leu	Ser	Gly	Thr	Leu	Pro	Pro	Ser	Ile	Ser	Ser	
			165					170					175			
CTC	CCC	AAC	CTC	GGA	GGA	ATC	ACA	TTC	GAC	GGC	AAC	CGA	ATC	TCC	GGC	576
Leu	Pro	Asn	Leu	Gly	Gly	Ile	Thr	Phe	Asp	Gly	Asn	Arg	Ile	Ser	Gly	
			180				185						190			
GCC	ATC	CCC	GAC	TCC	TAC	GGC	TCG	TTT	TCG	AAG	CTG	TTT	ACG	GCG	ATG	624
Ala	Ile	Pro	Asp	Ser	Tyr	Gly	Ser	Phe	Ser	Lys	Leu	Phe	Thr	Ala	Met	
		195					200					205				
ACC	ATC	TCC	CGC	AAC	CGC	CTC	ACC	GGG	AAG	ATT	CCA	CCG	ACG	TTT	GCG	672
Thr	Ile	Ser	Arg	Asn	Arg	Leu	Thr	Gly	Lys	Ile	Pro	Pro	Thr	Phe	Ala	
	210					215					220					
AAT	CTG	AAC	CTG	GCG	TTC	GTT	GAC	TTG	TCT	CGG	AAC	ATG	CTG	GAG	GGT	720
Asn	Leu	Asn	Leu	Ala	Phe	Val	Asp	Leu	Ser	Arg	Asn	Met	Leu	Glu	Gly	
225					230					235					240	
GAC	GCG	TCG	GTG	TTG	TTC	GGG	TCA	GAT	AAG	AAC	ACG	AAG	AAG	ATA	CAT	768
Asp	Ala	Ser	Val	Leu	Phe	Gly	Ser	Asp	Lys	Asn	Thr	Lys	Lys	Ile	His	
				245					250					255		

- 19 -

CTG GCG AAG AAC TCT CTT GCT TTT GAT TTG GGG AAA GTG GGG TTG TCA Leu Ala Lys Asn Ser Leu Ala Phe Asp Leu Gly Lys Val Gly Leu Ser 260 265 270	816
AAG AAC TTG AAC GGG TTG GAT CTG AGG AAC AAC CGT ATC TAT GGA ACG Lys Asn Leu Asn Gly Leu Asp Leu Arg Asn Asn Arg Ile Tyr Gly Thr 275 280 285	864
CTA CCT CAG GGA CTA ACG CAG CTA AAG TTT CTG CAA AGT TTA AAT GTG Leu Pro Gln Gly Leu Thr Gln Leu Lys Phe Leu Gln Ser Leu Asn Val 290 295 300	912
AGC TTC AAC AAT CTG TGC GGT GAG ATT CCT CAA GGT GGG AAC TTG AAA Ser Phe Asn Asn Leu Cys Gly Glu Ile Pro Gln Gly Gly Asn Leu Lys 305 310 315 320	960
AGG TTT GAC GTT TCT TCT TAT GCC AAC AAC AAG TGC TTG TGT GGT TCT Arg Phe Asp Val Ser Ser Tyr Ala Asn Asn Lys Cys Leu Cys Gly Ser 325 330 335	1008
CCT CTT CCT TCC TGC ACT TAACCATTTT CAGATTCGGT AATTATGGAT Pro Leu Pro Ser Cys Thr 340	1056
GCATCATGTT TGCCTTTCTA TGAACATCAA TAATGATACA AGTGTAATAA TAAAAAATTA	1116

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 342 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Thr Gln Phe Asn Ile Pro Val Thr Met Ser Ser Ser Leu Ser Ile 1 5 10 15
Ile Leu Val Ile Leu Val Ser Leu Arg Thr Ala Leu Ser Glu Leu Cys 20 25 30
Asn Pro Gln Asp Lys Gln Ala Leu Leu Gln Ile Lys Lys Asp Leu Gly 35 40 45
Asn Pro Thr Thr Leu Ser Ser Trp Leu Pro Thr Thr Asp Cys Cys Asn 50 55 60
Arg Thr Trp Leu Gly Val Leu Cys Asp Thr Asp Thr Gln Thr Tyr Arg 65 70 75 80
Val Asn Asn Leu Asp Leu Ser Gly His Asn Leu Pro Lys Pro Tyr Pro 85 90 95
Ile Pro Ser Ser Leu Ala Asn Leu Pro Tyr Leu Asn Phe Leu Tyr Ile 100 105 110
Gly Gly Ile Asn Asn Leu Val Gly Pro Ile Pro Pro Ala Ile Ala Lys 115 120 125
Leu Thr Gln Leu His Tyr Leu Tyr Ile Thr His Thr Asn Val Ser Gly 130 135 140

Ala Ile Pro Asp Phe Leu Ser Gln Ile Lys Thr Leu Val Thr Leu Asp
 145 150 155 160

Phe Ser Tyr Asn Ala Leu Ser Gly Thr Leu Pro Pro Ser Ile Ser Ser
 165 170 175

Leu Pro Asn Leu Gly Gly Ile Thr Phe Asp Gly Asn Arg Ile Ser Gly
 180 185 190

Ala Ile Pro Asp Ser Tyr Gly Ser Phe Ser Lys Leu Phe Thr Ala Met
 195 200 205

Thr Ile Ser Arg Asn Arg Leu Thr Gly Lys Ile Pro Pro Thr Phe Ala
 210 215 220

Asn Leu Asn Leu Ala Phe Val Asp Leu Ser Arg Asn Met Leu Glu Gly
 225 230 235 240

Asp Ala Ser Val Leu Phe Gly Ser Asp Lys Asn Thr Lys Lys Ile His
 245 250 255

Leu Ala Lys Asn Ser Leu Ala Phe Asp Leu Gly Lys Val Gly Leu Ser
 260 265 270

Lys Asn Leu Asn Gly Leu Asp Leu Arg Asn Asn Arg Ile Tyr Gly Thr
 275 280 285

Leu Pro Gln Gly Leu Thr Gln Leu Lys Phe Leu Gln Ser Leu Asn Val
 290 295 300

Ser Phe Asn Asn Leu Cys Gly Glu Ile Pro Gln Gly Gly Asn Leu Lys
 305 310 315 320

Arg Phe Asp Val Ser Ser Tyr Ala Asn Asn Lys Cys Leu Cys Gly Ser
 325 330 335

Pro Leu Pro Ser Cys Thr
 340

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 792 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Phaseolus vulgaris
- (B) STRAIN: Pinto

(vii) IMMEDIATE SOURCE:

- (B) CLONE: pAD-1 (DSM N.6821)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..685

(ix) FEATURE:

- (A) NAME/KEY: 3'UTR
- (B) LOCATION: 686..792

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

C AAT AAC CTC GTC GGT CCA ATC CCC CCC GCC ATC GCT AAA CTC ACC	46
Asn Asn Leu Val Gly Pro Ile Pro Pro Ala Ile Ala Lys Leu Thr	
1 5 10 15	
CAA CTC CAC TAT CTC TAT ATC ACC CAC ACC AAT GTC TCC GGC GCA ATA	94
Gln Leu His Tyr Leu Tyr Ile Thr His Thr Asn Val Ser Gly Ala Ile	
20 25 30	
CCC GAT TTC TTG TCA CAG ATC AAA ACC CTC GTC ACC CTC GAC TTC TCC	142
Pro Asp Phe Leu Ser Gln Ile Lys Thr Leu Val Thr Leu Asp Phe Ser	
35 40 45	
TAC AAC GCC CTC TCC GGC ACC CTA CCT CCC TCC ATC TCT TCT CTC CCC	190
Tyr Asn Ala Leu Ser Gly Thr Leu Pro Pro Ser Ile Ser Ser Leu Pro	
50 55 60	
AAC CTC GTA GGA ATC ACA TTC GAC GGC AAC CGA ATC TCC GGC GCC ATC	238
Asn Leu Val Gly Ile Thr Phe Asp Gly Asn Arg Ile Ser Gly Ala Ile	
65 70 75	
CCC GAC TCC TAC GGC TCA TTT TCG AAG CTG TTC ACG TCG ATG ACC ATC	286
Pro Asp Ser Tyr Gly Ser Phe Ser Lys Leu Phe Thr Ser Met Thr Ile	
80 85 90 95	
TCC CGC AAC CGC CTC ACC GGG AAG ATT CCG CCG ACG TTT GCG AAT CTG	334
Ser Arg Asn Arg Leu Thr Gly Lys Ile Pro Pro Thr Phe Ala Asn Leu	
100 105 110	
AAC CTG GCG TTC GTT GAC TTG TCT CGA AAC ATG CTG CAG GGT GAC GCG	382
Asn Leu Ala Phe Val Asp Leu Ser Arg Asn Met Leu Gln Gly Asp Ala	
115 120 125	
TCG GTG TTG TTC GGA TCA GAT AAG AAC ACG CAG AAG ATA CAT CTG GCG	430
Ser Val Leu Phe Gly Ser Asp Lys Asn Thr Gln Lys Ile His Leu Ala	
130 135 140	
AAG AAC TCT CTT GCT TTT GAT TTG GAG AAA GTG GGG TTG TCA AAG AAC	478
Lys Asn Ser Leu Ala Phe Asp Leu Glu Lys Val Gly Leu Ser Lys Asn	
145 150 155	
TTG AAC GGG TTG GAT CTG AGG AAC AAC CGT ATC TAT GGG ACG CTA CCG	526
Leu Asn Gly Leu Asp Leu Arg Asn Asn Arg Ile Tyr Gly Thr Leu Pro	
160 165 170 175	
CAG GGA CTG ACG CAG CTA AAG TTT CTG CAC AGT TTA AAT GTG AGC TTC	574
Gln Gly Leu Thr Gln Leu Lys Phe Leu His Ser Leu Asn Val Ser Phe	
180 185 190	
AAC AAT CTG TGC GGT GAG ATT CCT CAA GGT GGG AAC TTG CAA ACA TTT	622
Asn Asn Leu Cys Gly Glu Ile Pro Gln Gly Gly Asn Leu Gln Thr Phe	
195 200 205	
CAC GTT TCT GCT TAT GCC AAC AAC AAG TGC TTG TGT GGT TCT CCT CTT	670
His Val Ser Ala Tyr Ala Asn Asn Lys Cys Leu Cys Gly Ser Pro Leu	
210 215 220	
CCT GCC TGC ACT TAATCATTTC CAGATTCGGT AATTATGGAT GCATCATGTT	722
Pro Ala Cys Thr	
225	
TGCCTTTCTA TGAACATCAA TAATGATACA AGTTTAAATA AAAATAAAATT TATGAAATAA	782
AAAAAAAAAA	792

- 22 -

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 227 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

Asn Asn Leu Val Gly Pro Ile Pro Pro Ala Ile Ala Lys Leu Thr Gln
 1           5           10           15
Leu His Tyr Leu Tyr Ile Thr His Thr Asn Val Ser Gly Ala Ile Pro
          20           25           30
Asp Phe Leu Ser Gln Ile Lys Thr Leu Val Thr Leu Asp Phe Ser Tyr
          35           40           45
Asn Ala Leu Ser Gly Thr Leu Pro Pro Ser Ile Ser Ser Leu Pro Asn
          50           55           60
Leu Val Gly Ile Thr Phe Asp Gly Asn Arg Ile Ser Gly Ala Ile Pro
          65           70           75           80
Asp Ser Tyr Gly Ser Phe Ser Lys Leu Phe Thr Ser Met Thr Ile Ser
          85           90           95
Arg Asn Arg Leu Thr Gly Lys Ile Pro Pro Thr Phe Ala Asn Leu Asn
          100          105          110
Leu Ala Phe Val Asp Leu Ser Arg Asn Met Leu Gln Gly Asp Ala Ser
          115          120          125
Val Leu Phe Gly Ser Asp Lys Asn Thr Gln Lys Ile His Leu Ala Lys
          130          135          140
Asn Ser Leu Ala Phe Asp Leu Glu Lys Val Gly Leu Ser Lys Asn Leu
          145          150          155          160
Asn Gly Leu Asp Leu Arg Asn Asn Arg Ile Tyr Gly Thr Leu Pro Gln
          165          170          175
Gly Leu Thr Gln Leu Lys Phe Leu His Ser Leu Asn Val Ser Phe Asn
          180          185          190
Asn Leu Cys Gly Glu Ile Pro Gln Gly Gly Asn Leu Gln Thr Phe His
          195          200          205
Val Ser Ala Tyr Ala Asn Asn Lys Cys Leu Cys Gly Ser Pro Leu Pro
          210          215          220
Ala Cys Thr
          225

```

CLAIMS

1. A nucleotide sequence comprising a sequence coding
a PGIP protein having an inhibiting activity of the
5 fungine endo-alpha-1,4-D-polygalacturonase enzyme (PG),
or parts thereof.

2. A nucleotide sequence according to claim 1, wherein
said coding sequence is of vegetable origin, preferably
isolated either from plants, or parts of plants, or in
10 vitro cultured vegetable cells of the Phaseolus genus,
most preferably of the Phaseolus vulgaris species.

3. A nucleotide sequence according to claim 2, wherein
said coding sequence codes a PGIP protein having the
amino acid sequence of SEQ ID N2, or parts thereof.

15 4. A nucleotide sequence according to claim 3, wherein
said amino acid sequence SEQ ID N2 lacks or is
substituted of one or more amino acids.

5. A nucleotide sequence according to claim 4 wherein
said amino acid sequence SEQ ID N2 lacking or
20 substituted of one or more amino acids is a variant of
a PGIP protein.

6. A nucleotide sequence according to claim 5 wherein
said variant of a PGIP protein comprises the amino acid
sequence SEQ ID N4, or parts thereof.

25 7. A nucleotide sequence according to claim 1
comprising the nucleotide sequence of SEQ ID N1,
consisting of:

- a coding region from nucleotide 1 to nucleotide 1026;
- a 3'end untranslated sequence from nucleotide 1027 to nucleotide 1116.

8. A nucleotide sequence according to claim 7 lacking
5 or substituted of one or more nucleotides.

9. A nucleotide sequence according to claim 8
comprising the nucleotide sequence of SEQ ID N3.

10. A nucleotide sequence hybridizing to at least one
of the sequences according to any of previous claims.

10 11. A nucleotide sequence according to claim 10 which
is complementary to at least to one of the sequences
according to any of claims from 1 to 9.

12. A nucleotide sequence comprising an allele
derivative of either SEQ ID N1 or SEQ ID N3 sequences.

15 13. A nucleotide sequence according to any of previous
claims, wherein said sequence is derived from either a
genomic library, or a cDNA library or in vitro
synthesis.

14. A substantially purified PGIP protein having an
20 inhibiting activity for the fungine PG enzyme and being
of natural or recombinant origin, or parts thereof.

15. A substantially purified PGIP protein according to
claim 14, deriving from plants, parts thereof or
cultured plant cells, preferably of the Phaseolus
25 genus, more preferably of the Phaseolus vulgaris
species.

16. A substantially purified PGIP protein according to any of claims 14 or 15 wherein said protein comprises the amino acid sequence of SEQ ID N2, or parts thereof.

17. A substantially purified PGIP protein according to
5 claims 16, wherein said protein lacks or is substituted of one or more amino acids.

18. A substantially purified PGIP protein according to claims 17, wherein said protein comprises the amino acid sequence of SEQ ID N4.

10 19. A recombinant vector comprising a nucleotide sequence according to any of claims from 1 to 13.

20. A recombinant vector according to claim 19, wherein said vector is of plasmid or phage origin.

21. A recombinant vector according to the claim 20,
15 wherein said recombinant vector is the pAD-1 plasmid (DSM N.6821).

22. A recombinant vector according to claim 20, wherein said vector is either the pPT-1 or the pPT-2 plasmid, or the lambda 3.3 PGIP phage.

20 23. A recombinant vector according to claim 19, wherein said vector comprises an expression promoter of said coding PGIP nucleotide sequence.

24. A recombinant vector according to claim 23, wherein said promoter is active either in bacteria, or in
25 yeasts, or in high plants.

25. Cells transformed by a vector according to any of claims from 19 to 24.

26. Cells according to claim 25, wherein said cells are bacterial cells, preferably belonging to the Escherichia coli species.

5 27. Cells according to claim 25, wherein said cells are eucaryotic cells, preferably yeast cells, alternatively vegetable cells.

28. A process for isolating and cloning PGIP protein coding sequences, or parts thereof, comprising:

- 10 - partial sequencing the amino acid sequence of the PGIP protein;
- in vitro synthesising oligonucleotides, sequence of which derived from said amino acid sequence;
- probing a genomic or a cDNA library from plant cells with said oligonucleotide, previously labelled;
- 15 - isolating positive clones, extracting the recombinant vector and identifying the recombinant insert from said recombinant vector;
- sequencing said recombinant insert.

20

25

1/5

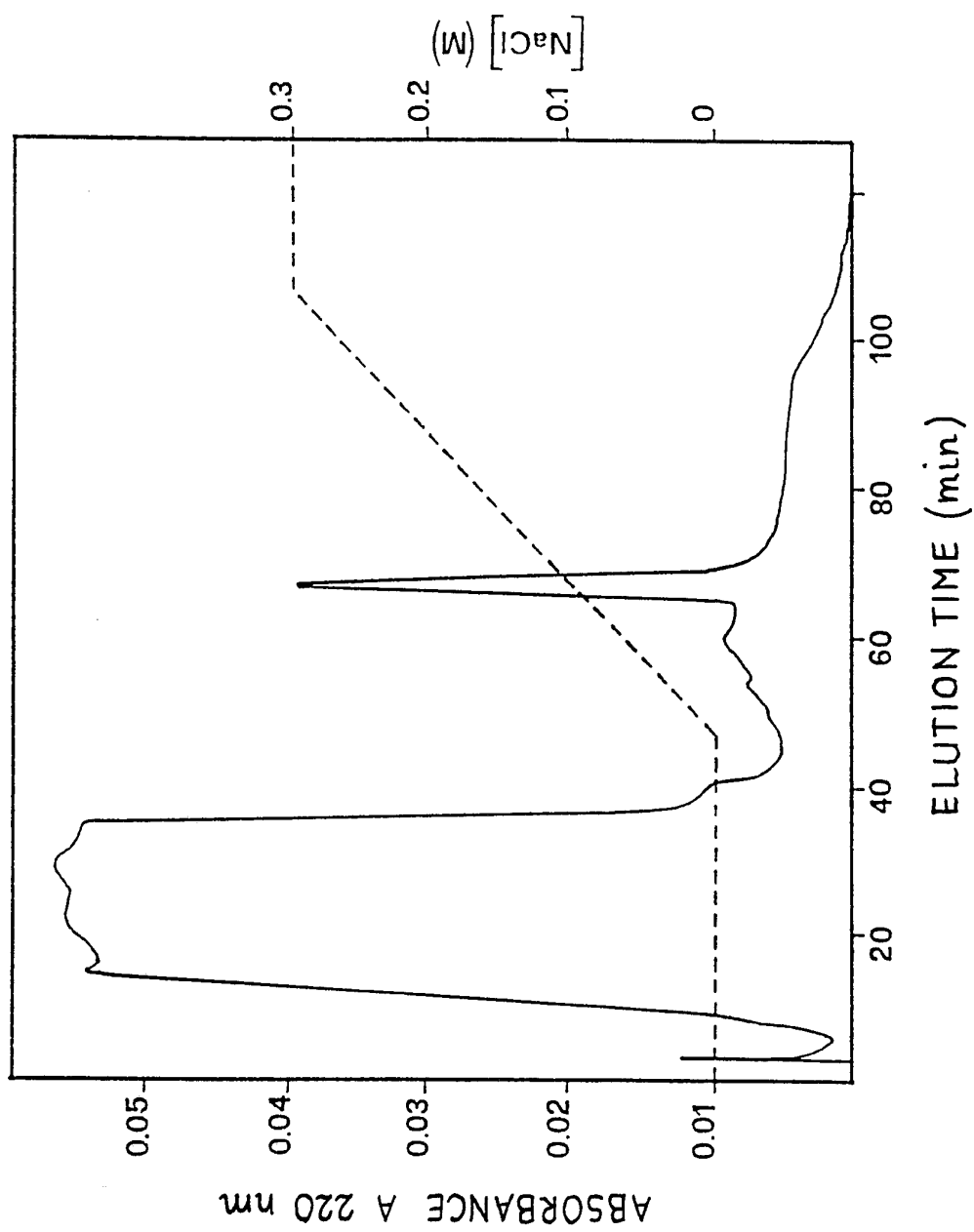


FIG. 1

SUBSTITUTE SHEET

2/5

FIG. 2

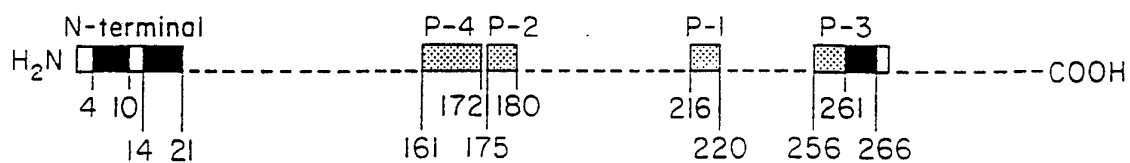
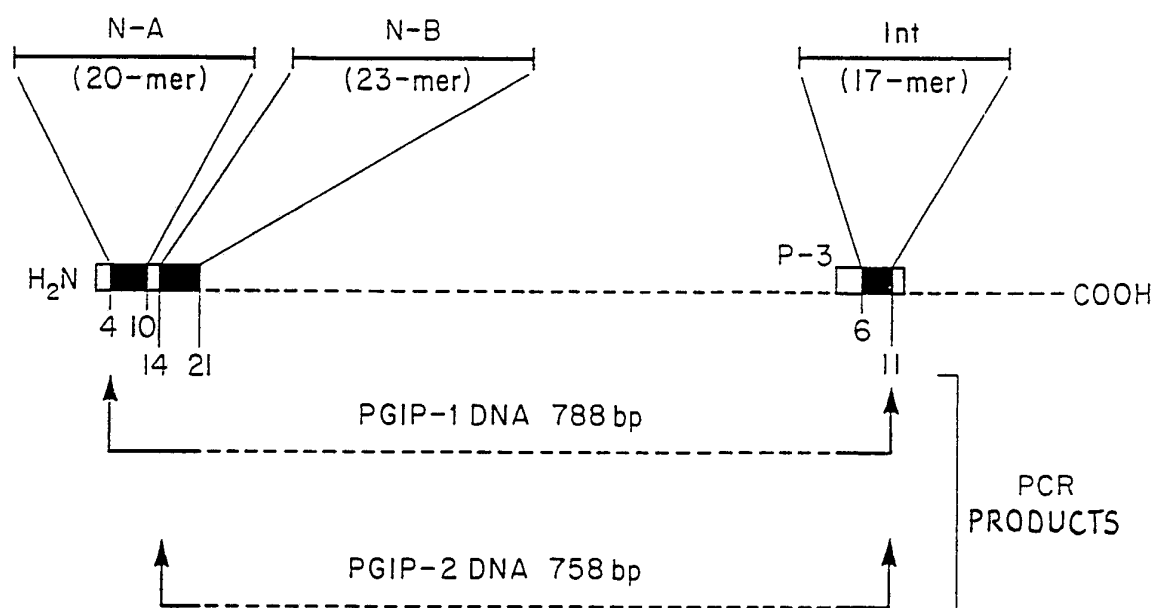
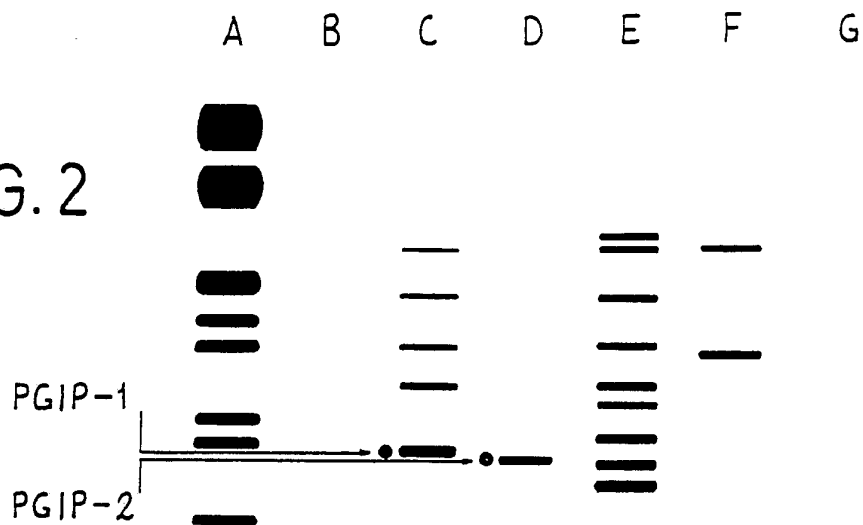


FIG. 3

SUBSTITUTE SHEET

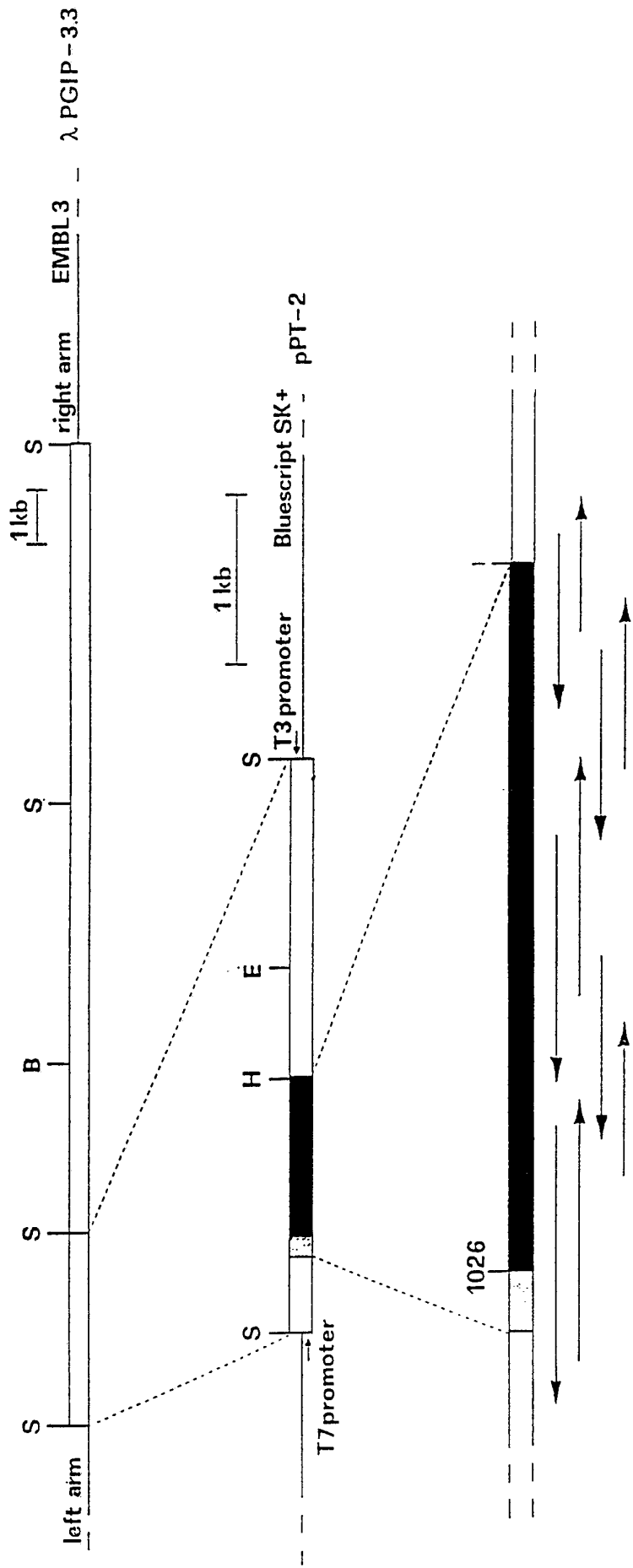


FIG. 4

4/5

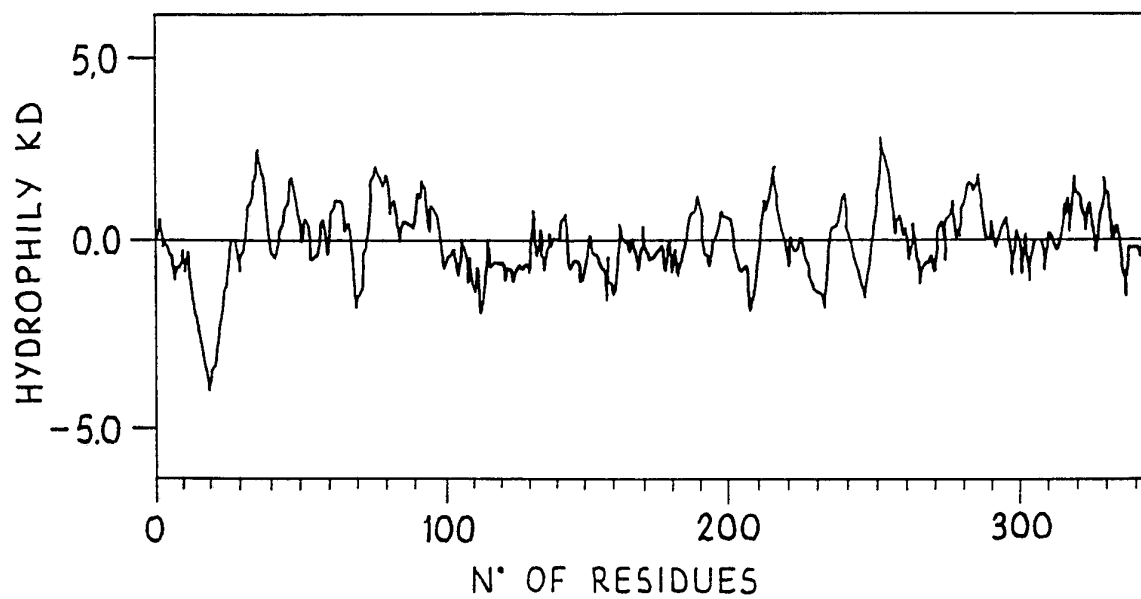
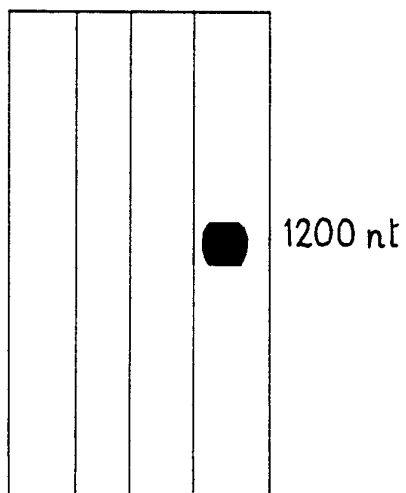


FIG. 5

FIG. 6



5/5

1	C	AAT	AAC	CTC	GTC	GGT	CCA	ATC	CCC	CCC	GCC	ATC	GCT	AAA	CTC	ACC	CAA	CTC	CAC	TAT	CT	60
345	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	404	
61	C	TAT	ATC	ACC	CAC	ACC	AAT	GTC	TCC	GGC	GCA	ATA	CCC	GAT	TTC	TTG	TCA	CAG	ATC	AAA	AC	120
405	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	464	
121	C	CTC	GTC	ACC	CTC	GAC	TTC	TCC	TAC	AAC	GCC	CTC	TCC	GGC	ACC	CTA	CCT	CCC	TCC	ATC	TC	180
465	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	524	
181	T	TCT	CTC	CCC	AAC	CTC	Val GTA	GGG	ATC	ACA	TTC	GAC	GGC	AAC	CGA	ATC	TCC	GGC	GCC	ATC	CC	240
525	-	-	-	-	-	-	-Gly	-	-	-	-	-	-	-	-	-	-	-	-	-	584	
241	C	GAC	TCC	TAC	GGC	TCA	TTT	TGG	AAG	CTG	TTC	ACG	Ser TCG	ATG	ACC	ATC	TCC	CCC	AAC	CGC	CT	300
585	-	-	-	-	-	-	-G	-	-	-	-	-T	-Gly	-	-	-	-	-	-	-	644	
301	C	ACC	GGG	AAG	ATT	CCG	CCG	ACG	TTT	GCG	AAT	CTG	AAC	CTG	GCG	TTC	GTT	GAC	TTG	TCT	CG	360
645	-	-	-	-	-	-	-A	-	-	-	-	-	-	-	-	-	-	-	-	-	704	
361	A	AAC	ATG	CTG	Gln CAG	GGT	GAC	GCG	TGG	GTG	TTG	TTC	GGA	TCA	GAT	AAG	AAC	ACG	Gln CAG	AAG	AT	420
705	G	-	-	-	-Glu	-	-	-	-	-	-	-	-G	-	-	-	-	-	-A--Lys	-	764	
421	A	CAT	CTG	GCG	AAG	AAC	TCT	CTT	GCT	TTT	GAT	TTG	Glu GAG	AAA	GTG	GCG	TTG	TCA	AAG	AAC	TT	480
765	-	-	-	-	-	-	-A	-	-	-	-	-	-Gly	-	-	-	-	-	-	-	824	
481	G	AAC	GGG	TTG	GAT	CTG	AGG	AAC	AAC	CGT	ATC	TAT	GGG	ACG	CTA	CCG	CAG	GGA	CTG	ACG	CA	540
825	-	-	-	-	-	-	-	-	-	-	-	-	-A	-	-	-	-	-A	-	-	884	
541	G	CTA	AAG	TTT	CTG	His CAC	AGT	TTA	AAT	GTG	AGC	TTC	AAC	AAT	CTG	TGC	GGT	GAG	ATT	CCT	CA	600
885	-	-	-	-	-	-A Gln	-	-	-	-	-	-	-	-	-	-	-	-	-	-	944	
601	A	GGT	GGG	AAC	TTG	Gln CAA	AGA	TTT	GAC	GTT	TCT	Ala GCT	TAT	GCC	AAC	AAC	AAG	TGC	TTG	TGT	GG	660
945	-	-	-	-	-	-A Lys	-G	-	-	-	-	-T Ser	-	-	-	-	-	-	-	-	1004	
661	T	TCT	CCT	CTT	CCT	Ala GCC	TGC	ACT	TAA	TCATTTCAGATTCCGGTAATTATGGATGCATCATGTTTGCCTTTCT											731	
1005	-	-	-	-	-	-T Ser	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1075	
732	ATGAACATCAATAATGATACAAAGTTT	AAATAAAATAAATTATGAAATAAAAAAAAAAAAA																			772	
1076	-G-AA-	-T-																			1126	

FIG. 7

INTERNATIONAL SEARCH REPORT

PCT/IT 92/00158

International Application No

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) ⁶		
According to International Patent Classification (IPC) or to both National Classification and IPC		
Int.Cl. 5	C12N15/29; C12Q1/68	C07K13/00; C12N15/70; C12N15/82
II. FIELDS SEARCHED		
Minimum Documentation Searched ⁷		
Classification System	Classification Symbols	
Int.Cl. 5	C12N ; C07K ; C12Q	
Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched ⁸		
III. DOCUMENTS CONSIDERED TO BE RELEVANT⁹		
Category ¹⁰	Citation of Document, ¹¹ with indication, where appropriate, of the relevant passages ¹²	Relevant to Claim No. ¹³
O, X	J. CELL. BIOCHEM. SUPPL., KEYSTONE SYMPOSIUM, THE GENETIC DISSECTION OF PLANT CELL PROCESSES, JAN 10-17, 1991. vol. 15A, 1991, page 54 DE LORENZO, G., ET AL. 'Cloning of the polygalacturonase-inhibiting protein (PGIP) of Phaseolus vulgaris L.' see abstract A208 --- -/--	1-20, 23-28
¹⁰ Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. "&" document member of the same patent family		
IV. CERTIFICATION		
Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report	
25 FEBRUARY 1993	15. 03. 93	
International Searching Authority	Signature of Authorized Officer	
EUROPEAN PATENT OFFICE	MADDOX A.D.	

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)		Relevant to Claim No.
Category °	Citation of Document, with indication, where appropriate, of the relevant passages	
O,X	Biological Abstracts BR43:11387 see abstract & G. BOT. ITAL., 30TH CONGRESS ITALIAN SOCIETY OF PLANT PHYSIOLOGY, STRESA, ITALY, OCTOBER 15-18, 1990. vol. 124, no. 4, 1990, pages 151 - 152 TOUBART, P., ET AL. 'Cloning and expression of a polygalacturonase-inhibiting protein PGIP of Phaseolus-vulgaris L.' ---	1-20, 23-28
X	BIOLOGICAL ABSTRACTS vol. 85 , 1988, Philadelphia, PA, US; abstract no. 51473, CERVONE, F., ET AL. 'Purification and characterization of a polygalacturonase-inhibiting protein from Phaseolus-vulgaris L.' see abstract & PLANT PHYSIOLOGY. vol. 85, no. 3, 1987, ROCKVILLE, MD, USA. pages 631 - 637 ---	14-18
X	BIOLOGICAL ABSTRACTS vol. 90 , 1990, Philadelphia, PA, US; abstract no. 139335, SALVI, G., ET AL. 'A polygalacturonase-inhibiting protein in the flowers of Phaseolus-vulgaris L.' see abstract & J. PLANT PHYSIOL. vol. 136, no. 5, 1990, pages 513 - 518 ---	14-18
X	BIOLOGICAL ABSTRACTS vol. 90 , 1990, Philadelphia, PA, US; abstract no. 68039, DE LORENZO, G., ET AL. 'Host-pathogen interactions: XXXVII. Abilities of the polygalacturonase-inhibiting proteins from four cultivars of Phaseolus vulgaris to inhibit the endopolygalacturonases from three races of Colletotrichum lindemuthianum' see abstract & PHYSIOL. MOL. PLANT PATHOL. vol. 36, no. 5, 1990, pages 421 - 436 ---	14-18
	---	-/--

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)		
Category °	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No.
P,X	THE PLANT JOURNAL vol. 2, no. 3, 1992, pages 367 - 373 TOUBART, P., ET AL. 'Cloning and characterization of the gene encoding the endopolygalacturonase-inhibiting protein (PGIP) of Phaseolus vulgaris L.' see the whole document -----	1-28