

Isolation and Characterisation of Wheat cDNA Clones Encoding PR4 Proteins

CARLA CARUSO^{a*}, LAURA BERTINI^a, MARINA TUCCI^b, CARLO CAPORALE^a, LUCA LEONARDI^a,
FRANCESCO SACCARDO^b, RAY A. BRESSAN^c, PAOLA VERONESE^c and VINCENZO BUONOCORE^a

^aDipartimento di Agrobiologia e Agrochimica and ^bDipartimento di Produzione Vegetale, Università della Tuscia, I-Viterbo 01100, ITALY and ^cCenter for Plant Environmental Stress Physiology, Department of Horticulture, West Lafayette, Indiana 47907-1165, U.S.A

(Received February 17, 1999)

Two cDNA clones encoding the previously characterised PR4 proteins wheatwin 1 and wheatwin2 from wheat (*Triticum aestivum* cv. S. Pastore) have been identified and named wPR4a and wPR4b, respectively. The clones have been isolated by screening a cDNA library with a specific cDNA probe obtained by RT-PCR. The wPR4a and wPR4b cDNAs contain open reading frames of 441 and 447 bp that encode for wheatwin1 and wheatwin2, respectively.

Keywords: Fungal infection, *Fusarium culmorum*, pathogenesis-related protein, plant defense, PR4 genes, *Triticum aestivum*

EMBL Data Bank Accession No.: AJ 006098 for wPR4a, AJ 006099 for wPR4b

INTRODUCTION

In their natural environment plants are challenged by several pathogenic agents such as viruses, viroids, fungi and bacteria [Hahlbrock and Scheel, 1987]. For their survival plants have developed a complex variety of defence mechanisms that activate a highly co-ordinated resistance response upon recognition of a potential pathogen. Among these mechanisms are the synthesis and accumulation of a set of host-encoded

pathogenesis-related (PR) proteins [Bol et al., 1990; Boller, 1987; Carr and Klessing 1989; Dixon and Lamb 1990; Linthorst, 1991; Stintzi et al. 1993; Van Loon, 1985]. These proteins were first described in tobacco plants infected with tobacco mosaic virus (TMV) [Gianinazzi et al., 1970; Van Loon and Van Kammen, 1970], but since their discovery several new members have been described in other plant species and have been grouped into eleven classes on the basis of their primary structure or DNA-deduced sequence, serological relationships and enzymatic or biological activity [Van Loon et al., 1994]. Several studies on structural-functional characterisation of PR proteins have been reported showing that these proteins display antifungal activity *in vitro* against several plant pathogens; these observations lead to the proposal that the major role for these proteins is defence [Bowles, 1990].

Among the less studied PR proteins are those belonging to the PR4 family. Genes encoding PR4 proteins were studied in potato, tomato, tobacco, *Arabidopsis* and rubber tree [Broekaert et al., 1990; Friedrich et al., 1991; Linthorst et al., 1991; Potter et al., 1993; Stanford et al., 1989],

* Corresponding author: Tel.: +39 0761 357330/357301; Fax: +39 0761 357242; e-mail: caruso@unitus.it

while mature proteins have been characterised only from tobacco [Ponstein *et al.*, 1994] and the monocot barley [Hejgaard *et al.*, 1992; Svensson *et al.*, 1992]. Recently, we isolated two highly homologous PR4 proteins (wheatwin1 and wheatwin2) from wheat seeds showing antifungal activity towards pathogenic fungi of cereals [Caruso *et al.*, 1993; Caruso *et al.*, 1996] and demonstrated that these proteins are specifically induced in wheat seedlings infected with *Fusarium culmorum* [Caruso *et al.*, 1998]. In this paper we report the identification and characterisation of two cDNA sequences encoding the PR4 proteins wheatwin1 and wheatwin2.

MATERIALS AND METHODS

Wheat seeds (*Triticum aestivum* cv. S. Pastore) were germinated for 24 h in the dark at 25 °C, inoculated with *F. culmorum* spores and allowed to continue germination for three more days [Caruso *et al.*, 1998]. Total RNA was extracted from infected wheat seedlings (2 g) as previously described [Prescott and Martin, 1987]. Poly(A)⁺mRNA was prepared from total RNA with the Qiagen Oligotex mRNA mini kit. cDNA synthesis was carried out with M-MLV reverse transcriptase (Promega) following manufacturers' instructions using a degenerate C-terminal primer (5'GTGAAIACIGTGTCCTCAATC 3') corresponding to amino acids 94–100 of the wheatwin1 sequence. After phenol extraction, cDNA fragments encoding PR4 transcripts were PCR-amplified using the same C-terminal primer and the N-terminal primer (5'GCRACGTAYCATTAY TAYCGICCRGCGCARAAY 3') corresponding to the region 8–18 of the protein sequence. Amplification was carried out using a Perkin-Elmer Thermal Cycler 2400 model, Pharmacia Ready to go PCR beads with the following program: 94°C for 1 min, 56°C for 1 min, 72°C for 1 min, 30 total cycles. The PCR fragments were ligated into the pGEM[®]-T vector (Promega, Italy) and sequenced by the dideoxy chain termination method using

universal primers [Sanger *et al.*, 1977]. The 0.3 Kb sequence (RT-5 clone) obtained was found to contain the coding region of wheatwin1 corresponding to position 8 to 100 of the protein sequence. The RT-5 clone labelled with ³²P-dCTP by random priming (prime-it II kit, Stratagene) was used as a probe to screen a λgt11 cDNA library from wheat Chinese Spring endosperm (generous gift from prof. P. Shewry, IACR, Long Ashton, England). Filters were washed 4 times with 5X SSC containing 0.1% SDS at 65°C for 20 min. Positive clones were subcloned in pGEM[®]-4Z vector (Promega, Italy) and sequenced on both strands by the dideoxy chain termination method using universal primers.

RESULTS AND DISCUSSION

A wheat endosperm cDNA library was screened for PR4 proteins using RT-5, a cDNA probe corresponding to a central coding region of wheatwin1. Five positive clones were obtained out of about 1×10^6 plaques screened and sequenced on both strands. One cDNA clone (wPR4a) was found to contain the entire open reading frame for wheatwin1, whereas four cDNA clones differing only for the flanking 5' and 3' regions contained the complete coding sequence for wheatwin2. Among the last cDNAs the one with the shortest non-coding region was chosen for further studies and named wPR4b. The complete nucleotide sequences of wPR4a and wPR4b (deposited in the EMBL Data Bank under the accession number AJ 006098 and AJ 006099, respectively) and the corresponding deduced amino acid sequences are shown in Fig. 1 and Fig. 2, respectively. The wPR4a cDNA (Fig. 1) consists of 751 bp with an open reading frame of 441 nucleotides, starting from the first ATG, and a poly(A)⁺ tail of (A)₂₄ located 195 bp downstream of the stop codon TAG. The deduced amino acid sequence contains a highly hydrophobic N-terminal region with the characteristics of a signal peptide for translocation

GAATTCGCGGCCGCGTCGACCACAACTAGCTAGCAGGTTGTGACTGATCACCGTGGTCG 60

ACGAGACGACGATGGCCGCACGCCCGATGCTGGTGGTGGCGCTGCTGTGCGCGGCGGCGG 120
M A A R P M L V V A L L C A A A A 17

CGGCGGCCACGGCGCAGCAGGCGACCAACGTGCGGGCGACGTACCACTACTACCGGCCGG 180
A A T A Q Q A T N V R A T Y H Y Y R P A 37

CGCAGAACAACCTGGGACCTGGGCGCGCCCGCGGTGAGCGCCTACTGCGCAACCTGGGACG 240
Q N N W D L G A P A V S A Y C A T W D A 57

CCAGCAAGCCGCTGTCGTGGCGGTCCAAGTACGGGTGGACGGCGTTCTGCGGCCCCGCGG 300
S K P L S W R S K Y G W T A F C G P A G 77

GCGCCACGGGCAGGCGTCCTGCGGCAAGTGCCTCCAGGTGACCAACCCGGCGACGGGGG 360
A H G Q A S C G K C L Q V T N P A T G A 97

CGCAGATCACGGCGAGGATCGTGGACCAGTGCGCCAACGGCGGGCTGGACCTGGACTGGG 420
Q I T A R I V D Q C A N G G L D L D W D 117

ACACCGTCTTCACCAAGATCGACACCAACGGGATCGGGTACCAGCAGGGCCACCTTAACG 480
T V F T K I D T N G I G Y Q Q G H L N V 137

TCAACTACCAGTTCGTCGACTGCCGCGACTAGATTACGTTCTCCGTCGATCGATCAAAGG 540
N Y Q F V D C R D * 146

GCTAGCTAGCTATGATAAGAATAAAAGGCTAGCTAAAAGTAAGTTGAGACTGATCCATGC 600

TGTTTAATAAAAGCTAGCACGGGATGCTGCCTTGTATACGTGTTACGTGTCCCAGTACG 660

TACATGGACCCAAATAAAGTGGAAAAGACCAAACGGGTTATCTATTCAAAAAAAAAAAAAA 720

AAAAAAAAAAGGGACGCGTTCCGCGAATTC 751

FIGURE 1 Nucleotide sequences of wPR4a cDNA clone encoding wheatwin1 and the predicted amino acid sequence. The cleavage sites of the N-terminal signal peptides are marked by an arrow, the start and stop codons and the putative polyadenylation signals are underlined. The amino acid sequence corresponding to the RT-5 probe is marked with a dotted line

```

GAATTCGCGGCCGCGTCGACACACAACTAGCTATAGCAGGTTGTCACTGATCATGACGA 60
                                     M T M 3

TGGCCGCACGCCTGATGCTGGTGGCGGCGCTGCTGTGCGCGGCGGCCGCGGCCGACGG 120
A A R L M L V A A L L C A A A A A A T A 23

CGCAGCAGGCGACCAACGTGCGGGCGACGTACCACTACTACCGGCCGCGCAGAACAAC 180
Q Q A T N V R A T Y H Y Y R P A Q N N W 43
↑
GGGACCTGGGCGCGCCCGCGGTGAGCGCCTACTGCGCGACCTGGGACGCCAGCAAGCCGC 240
D L G A P A V S A Y C A T W D A S K P L 63

TGTCGTGGCGGTCCAAGTACGGGTGGACGGCGTTCTGCGGCCCCGCGCGCCACGGGC 300
S W R S K Y G W T A F C G P A G A H G Q 83

AGGCGGCGTGCGGCAAGTGCCTCCGGGTGACGAACCCGGCGACGGGGCGCAGATCACGG 360
A A C G K C L R V T N P A T G A Q I T A 103

CGAGGATCGTCGACCAGTGCGCCAACGGCGGGCTGGACCTGGACTGGGACACCGTCTTCA 420
R I V D Q C A N G G L D L D W D T V F T 123

CCAAGATCGACACCAACGGGATCGGGTACCAGCAGGGCCACCTCAACGTCAACTACCAGT 480
K I D T N G I G Y Q Q G H L N V N Y Q F 143

TCGTGCGACTGCCGCGACTAGATTATGTATCCGTCGATCAAAGGCTAGCTTGCTATGATAA 540
V D C R D * 148

GAATAAGAGGCTAGCTTAGCTAAAAAAAAAAAAAAAAAGAAAAAAAAAGTCGACGCGGCCGCGA 600

```

ATTC 604

FIGURE 2 Nucleotide sequences of wPR4b cDNA clone encoding wheatwin2 and the predicted amino acid sequencend. The cleavage sites of the N-terminal signal peptides are marked by an arrow, the start and stop codons and the putative polyadenylation signals are underlined

through the endoplasmic reticulum and two possible initiation codons at position 72 and 87 of the nucleotide sequence generating two putative signal peptides of 21 and 15 amino acids, respec-

tively. Cleavage of the signal peptide results in a polypeptide of 125 amino acids with a calculated molecular mass of 13670 Da which is in good agreement with that of mature wheatwin1.

The wPR4b (Fig. 2) cDNA consists of 604 bp with an open reading frame of 447 nucleotides, starting from the first ATG and a poly(A)⁺ tail of (A)₁₄G(A)₈ located 61 bp downstream of the stop codon TAG. Similarly to wPR4a, a N-terminal extension typical of a signal peptide is present in the coding region of wPR4b cDNA; in this case, three possible starting codons are present at positions 54, 60 and 75 of the nucleotide sequence giving rise to three putative signal peptides of 23, 21 and 16 amino acids, respectively. Cleavage of this peptide results in a mature protein of 125 amino acid residues with a calculated mass of 13684 Da, corresponding to the size of the mature wheatwin2. The presence of the signal peptide in both cDNAs strongly suggests a secretory pathway for the mature PR4 proteins. Putative poly-adenylation motifs (AATAAA) [30] upstream the poly (A)⁺ tail are present in both cDNA clones. The deduced amino acid sequences of wheatwin1 and wheatwin2, encoded by wPR4a and wPR4b, respectively, differ by only two amino acids out of 125, at positions 62 and 67. In a previous paper we reported three amino acid differences between the two proteins, at positions 45 (Gly→Lys), 62 (Ser→Ala) and 67 (Gln→Arg) for wheatwin1 and wheatwin2, respectively [Caruso et al., 1996]. In order to explain this apparent disagreement, we resubmitted to amino acid sequencing both mature proteins. According to the cDNA deduced sequences, no difference was detected at position 45 where lysine was found for both proteins. In conclusion, the two PR4 proteins differ by two amino acid residues only that probably play a role in determining their different antifungal activity as previously reported [Caruso et al., 1996]. The functional significance of the structural differences among wheatwin1 and wheatwin2 needs further investigation, but the existence of two proteins with different specific antifungal activity might represent a selective advantage towards a wide range of potential pathogens.

The comparison of the nucleotide sequences of wPR4a and wPR4b with sequences stored in several data banks revealed a high degree of

sequence identity (80–90%) with other PR4 genes from tobacco [Friedrich et al., 1991], tomato [Linthorst et al., 1991], maize [EMBL Data Bank a. n. X82184] and barley [EMBL Data Bank a. n. Y10814]. A less striking similarity (60–70%) was found with the potato wound-induced genes *win1* and *win2* [Stanford et al., 1989], the *hev1* gene from rubber tree [Broekaert et al., 1990], the *hel* gene from *Arabidopsis thaliana* [Potter et al., 1993] and the *chp20* cDNA clone from tobacco [Ponstein et al., 1994] that are characterised by the presence of the hevein-like domain, rich in cysteine residues, present also in several lectins [Raikhel and Wilkins, 1987] as well as in the inducible class I basic chitinases [Collinge et al., 1993]. By analogy to chitinase classification based on the presence or absence of the cysteine-rich domain it has been proposed that PR4 proteins with the “hevein-like” domain could be classified as class I whereas those lacking the same domain may be ascribed to the class II [Ponstein et al., 1994]. According to this classification, we propose that the isolated wheat cDNAs belong to PR4 genes of the class II. A general characteristics of PR proteins is the presence of several isoforms localized in different cell compartments; usually, acidic proteins are secreted to the extracellular space while the basic homologues are accumulated in the vacuoles. Moreover, it has been demonstrated that basic PR proteins have a C-terminal extension of few amino acids essential for vacuolar targeting [Bednarek and Raikhel, 1992; Nakamura and Matsuoka, 1993; Neuhaus et al. 1991]. It is interesting to point out that the vacuolar sorting signal is absent in both wheat PR4 genes. Therefore, it is unlikely that either wheatwin1 or wheatwin2 are localised in the vacuoles, confirming that these proteins might be targeted to the extracellular space by the signal peptide. To the best of our knowledge, this is the first report concerning a possible extracellular localisation of basic PR4 proteins. In summary, we have cloned two cDNAs encoding two highly homologous PR4 proteins previously characterised from

wheat seeds [Caruso *et al.*, 1993; Caruso *et al.*, 1996]. The material we have generated is now suitable for the production of recombinant proteins and their use in expression cassettes to obtain transgenic plants. This will allow investigation both on the biological function of PR4 proteins and on the possibility to use them as active deterrent in plants in which the expression of their genes has been manipulated.

Acknowledgements

Dr. Gabriele Chilosi (Dipartimento di Protezione delle Piante, Università della Tuscia, Viterbo, Italy) is greatly acknowledged for the continuous supply of *F. culmorum*-infected wheat seedlings. We thank Prof. P. Shewry, IACR, Long Ashton, England for the use of the Chinese Spring wheat cDNA library and the hospitality given to V.B and L.B. We also greatly acknowledge Drs. Nigel Halford and Sandy Hay, IACR, Long Ashton, England for skilful assistance. This research was supported by a grant from Ministero per le Politiche Agricole (MIPA, former MIRAAF), Piano Nazionale Biotecnologie Vegetali, Progr. N. 152.

References

- Bednarek SK, Raikhel NV: Intracellular trafficking of secretory proteins. *Plant Mol Biol* 20: 133–150 (1992).
- Bol JF, Linthorst JM, Cornelissen BJC: Plant pathogenesis-related proteins induced by virus infection. *Annual Review of Phytopathology* 28: 113–138 (1990).
- Boller T: Hydrolytic enzymes in plant disease resistance. In: Kosuge T, Nester EW (eds.) *Plant-Microbe interactions: Molecular and Genetic Perspectives*, Vol 2, pp 385–413. Macmillan, New York (1987).
- Bowles DJ: Defense-related proteins in higher plants. *Annual Review of Biochemistry* 59: 873–907 (1990).
- Broekaert W, Lee H, Kush A, Chua NH, Raikhel N: Wound-induced accumulation of mRNA containing a hevein sequence in laticifers of rubber tree (*Hevea brasiliensis*). *Proc Natl Acad Sci USA* 87: 7633–7637 (1990).
- Carr JP, Klessing DF: The pathogenesis-related proteins of plants. In: Setlow JK (ed) *Genetic Engineering: Principles and Methods*, pp 65–109. Plenum Press, New York (1989).
- Caruso C, Caporale C, Poerio C, Facchiano A, Buonocore V: The amino acid sequence of a protein from wheat kernel closely related to proteins involved in the mechanisms of plant defence. *J Prot Chem* 12: 379–386 (1993).
- Caruso C, Caporale C, Chilosi G, Vacca F, Bertini L, Magro P, Poerio E, Buonocore V: Structural and antifungal properties of a pathogenesis-related protein from wheat kernel. *J Prot Chem* 15: 35–44 (1996).
- Caruso C, Chilosi G, Caporale C, Leonardi L, Bertini L, Magro P, Buonocore V: Induction of pathogenesis-related proteins in germinating wheat seeds infected with *Fusarium culmorum*. *Plant Sci* 140: 87–97 (1998).
- Collinge DB, Kragh KM, Mikkelsen JD, Nielsen KK, Rasmussen U, Vad K: Plant chitinases. *Plant J* 3: 31–40 (1993).
- Dixon RA, Lamb CJ: Molecular communications in interactions between plants and microbial pathogens. *Annu Rev Plant Physiol Plant Mol Biol* 41: 339–367 (1990).
- Friedrich L, Moyer M, Ward E, Ryals J: Pathogenesis-related protein 4 is structurally homologous to the carboxy-terminal domains of hevein, Win1 and Win2. *Mol Gen Genet* 230: 113–119 (1991).
- Gianinazzi S, Martin C, Vallee JC: Hypersensibilité aux virus, température et protéines solubles chez le *Nicotiana Xanthi* n.c. Apparition de nouvelles macromolécules lors de la répression de la synthèse virale. *Comptes Rendus de l'Académie des Sciences, série D* 270: 2383–2386 (1970).
- Hahlbrock K, Scheel D: Biochemical responses of plants to pathogens. In: Chet I (ed) *Innovative Approaches to Plant Disease Control*, pp 229–254. John Wiley, New York (1987).
- Hejgaard J, Jacobsen S, Bjorn SE, Kragh KM: Antifungal activity of chitin-binding PR-4 type proteins from barley grain and stressed leaf. *FEBS Lett.* 307: 389–392 (1992).
- Linthorst HJM: Pathogenesis-related proteins of plants. *Critical Review of Plant Science* 10: 113–150 (1991).
- Linthorst HJM, Danhash N, Brederode FT, Van Kan JAL, De Witt PJGM, Bol JF: Tobacco and tomato PR proteins homologous to win and pro-hevein lack the "hevein" domain. *Mol Plant-Microbe Interact* 4: 585–592 (1991).
- Nakamura K, Matsuoka K: Protein targeting to the vacuole in the plant cells. *Plant Physiol* 101: 1–5 (1993).
- Neuhaus J-M, Sticher L, Meins F, Boller T: A short C-terminal sequence is necessary and sufficient for the targeting of chitinases to the plant vacuole. *Proc Natl Acad Sci USA* 88: 10362–10366 (1991).
- Ponstein AS, Bres-Vloemans SA, Sela-Burlage MB, Van den Elzen PJM, Melchers LS, Cornelissen BJC: A novel pathogen- and wound-inducible tobacco (*Nicotiana tabacum*) protein with antifungal activity. *Plant Physiol* 104: 109–118 (1994).
- Potter S, Uknes S, Lawton K, Winter AN, Chandler D, DiMaio J, Novitzky R, Ward E, Ryals J: Regulation of a hevein-like gene in *Arabidopsis*. *Mol Plant-Microbe Interact* 6: 680–685 (1993).
- Prescott A, Martin C: A rapid method for the quantitative assessment of levels of specific mRNAs in plants. 4: 219–224 (1987).
- Raikhel NV, Wilkins TA: Isolation and characterization of a cDNA clone encoding wheat germ agglutinin. *Proc Natl Acad Sci USA* 84: 6745–6749 (1987).
- Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain terminating-inhibitors. *Proc Natl Acad Sci USA* 74: 5463–5467 (1977).
- Stanford A, Bevan M, Northcote D: Differential expression within a family of novel wound-induced genes in potato. *Mol Gen Genet* 215: 200–208 (1989).
- Stintzi A, Heitz T, Prasad V, Wiedemann-Merdinoglu S, Kauffmann S, Geoffroy P, Legrand M, Fritig B: Plant 'pathogenesis-related' proteins and their role in defense against pathogens *Biochimie* 75: 687–706 (1993).

- Svensson B, Svendsen I, Hojrup P, Roepstorff P, Ludvigsen S, Poulsen FM: Primary structure of barwin: a barley seed protein closely related to the C-terminal domain of proteins encoded by wound-induced plant genes. *Biochemistry* 31: 8767–8770 (1992).
- Van Loon LC, Van Kammen A: Polyacrylamide disc electrophoresis of the soluble leaf proteins from *Nicotiana tabacum* var. "Samsun" and "Samsun NN". II. Changes in protein constitution after infection with tobacco mosaic virus. *Virology* 40: 199–211 (1970).
- Van Loon LC: Pathogenesis-related proteins. *Plant Mol Biol* 4: 111–116 (1985).
- Van Loon LC, Pierpoint WS, Boller T, Conejero V: Recommendations for naming plant pathogenesis-related proteins. *Plant Mol Biol Rep* 12: 245–264 (1994).
- Wu L, Uede T, Messing J: The formation of mRNA 3'-ends in plants. *Plant J* 8: 323–329 (1995).