Plant Molecular Biology **22**: 1187–1190, 1993.

Update section

Short communication

Cloning and characterization of a cDNA encoding an antimicrobial chitin-binding protein from amaranth, *Amaranthus caudatus*

Miguel F.C. De Bolle[1], Katrien M.M. David[1], Sarah B. Rees[2], Jozef Vanderleyden[1], Bruno P.A. Cammue[1] and Willem F. Broekaert[1]*

[1] *F.A. Janssens Laboratory of Genetics, Catholic University of Leuven, Willem de Croylaan 42, B-3001 Heverlee, Belgium (* author for correspondence);* [2] *Zeneca Agrochemicals, Jealott's Hill Research Station, Bracknell, Berkshire RG12 6EY, UK*

Received 16 April 1993; accepted in revised form 8 June 1993

Key words: Amaranthus caudatus, chitin-binding protein, antifungal protein, antimicrobial protein, cDNA clone

Abstract

A cDNA clone encoding an antimicrobial chitin-binding protein from amaranth (*Amaranthus caudatus* L.) was isolated using a cDNA library constructed from near-mature seed poly(A)$^+$ mRNA. The deduced amino acid sequence of the cDNA clone encodes a predicted polypeptide of 86 amino acids. This polypeptide has three distinct domains: an amino-terminal putative signal peptide (25 amino acids), a domain corresponding to the mature protein (30 amino acids), and a carboxyl-terminal propeptide (31 amino acids) containing a putative N-glycosylation site. The encoded protein differs from all known members of the family of chitin-binding proteins. Transcripts of the expected size (650 bp) are present in developing seeds but not in roots, leaves or stressed leaves.

Seeds of amaranth (*Amaranthus caudatus* L.) contain two chitin-binding proteins, designated Ac-AMP1 and Ac-AMP2, that are respectively 29 and 30 amino acids in length. Ac-AMP2 differs from Ac-AMP1 only by the presence of an additional arginine residue at the carboxyl terminus. The Ac-AMPs show striking homology to the cysteine/glycine-rich domains occurring in many types of chitin-binding proteins, including lectins, chitinases and hevein (for a recent review on chitin-binding proteins, see [11]). Both Ac-AMP1 and Ac-AMP2 are potent inhibitors of filamentous fungi and Gram-positive bacteria and are therefore also classified as antimicrobial proteins [3]. We now report on the isolation and characterization of a cDNA clone encoding Ac-AMP2.

Near-mature seeds were collected from inflorescences of amaranth plants at the stage of spontaneous shedding. Total RNA was extracted from 5 g of seeds by the method of de Vries *et al.* [6] except that 6 ml of a 1:2 phenol/RNA extraction buffer mixture was used per g tissue. Poly(A)$^+$ RNA was purified by oligo (dT)-cellulose affinity chromatography [13], yielding about 7 μg of poly(A)$^+$ RNA. Double stranded cDNAs were

prepared from 1.5 μg of poly(A)$^+$ RNA and ligated to *Eco* RI/*Not* I adaptors using the cDNA Synthesis Kit of Pharmacia. The cDNAs were cloned into the λZAP II phage vector (Stratagene) and packaged *in vitro* with the Gigapack II Gold packaging system (Stratagene) according to the manufacturer's instructions. A DNA probe for screening of the cDNA library was produced by polymerase chain reaction (PCR) as follows. Two degenerate oligonucleotides were synthesized: OWB13 (5′-GTNGGNGARTGKGTNMGNGG) and OWB14 (5′-CCRCARTAYTTNGGNCCYTTNCC). OWB13 corresponds to amino acids 1 to 7 of *Ac*-AMP1 and has a sense orientation. OWB14 corresponds to amino acids 22 to 29 of *Ac*-AMP1 and has an antisense orientation. PCR was performed with the *Taq* polymerase as previously described [2] using OWB13 and OWB14 as amplimers and 25 ng of cDNA as target DNA. The 100 bp PCR amplification product was purified on a 3% agarose (NuSieve, FMC) gel and reamplified by PCR under the same conditions except that the reaction mixtures contained 130 μM dTTP and 70 μM digoxigenin-11dUTP instead of 200 μM dTTP. The digoxigenin-labelled PCR product was purified on a 3% NuSieve agarose gel. About 10000 plaque-forming units of the λZAP II cDNA library were screened with the digoxigenin-labelled PCR product by *in situ* plaque hybridization. Detection of hybridized probe was done using anti-digoxigenin antibodies linked to alkaline phosphatase (Boehringer Mannheim) and its substrate 5-bromo-4-chloro-3-indolyl phosphate (Boehringer Mannheim) according to the manufacturer's instructions. Inserts from purified plaques were excised *in vivo* into pBluescript phagemid form with the aid of the helper phage R408, according to the instructions of Stratagene. Nucleotide sequencing was done by the dideoxy chain-termination method with an ALF automated sequencer (Pharmacia) using fluorescein-labelled M13 forward and reverse primers (Pharmacia). Sequence analysis was performed by the PC-Gene software (Intelligenetics).

The inserts from 10 different positive clones were released by *Eco* RI digestion and their sizes compared by agarose electrophoresis. The clone with the longest insert (AC1) was subjected to nucleotide sequence analysis. AC1 is 590 nucleotides long and contains an open reading frame of 86 amino acids (Fig. 1). The 25 amino-terminal amino acids have a predicted signal peptide structure obeying the (-1, -3)-rule [17]. The deduced amino acid sequence of the region following the putative signal peptide is identical to the 30 amino acid sequence of mature Ac-AMP2 as determined by protein sequencing [3]. In addition, the mature protein domain is extended by a 31 amino acids carboxyl-terminal domain, of which the asparagine at position 74 represents a putative site for N-glycosylation [7]. AC1 has 5′- and 3′-untranslated regions of 45 and 284 nucleotides, respectively. The 3′-end untranslated region is not terminated by a poly(A) tail, indicating that AC1 is not a full-length cDNA clone.

All of 5 other sequenced positive clones have deduced amino acid sequences identical to that of AC1. They differed only from each other by different degrees of truncation at the 5′ or 3′ end. The fact that the 6 sequenced clones all contained an arginine at position 30 of the mature protein domain, suggests that Ac-AMP1 and Ac-AMP2 are both derived from the same precursor preproprotein.

```
1    CAAAAAAAAAAAATAAAGTCAAGAGTATTAATTAGGTGAGAAAAAATGGTGAACATGAAG
                                                 M  V  N  M  K
61   TGTGTTGCATTGATAGTTATAGTTATGATGGCGTTTATGATGGTGGATCCATCAATGGGA
      C  V  A  L  I  V  I  V  M  M  A  F  M  M  V  D  P  S  M  G
121  GTGGGAGAATGTGTGAGAGGACGTTGCCCAAGTGGGATGTGTTGCAGTCAGTTTGGGTAC
     [V  G  E  C  V  R  G  R  C  P  S  G  M  C  C  S  Q  F  G  Y
181  TGTGGTAAAGGCCCAAAGTACTGTGGCCGTGCCAGTACTACTGTGGATCACCAAGCTGAT
      C  G  K  G  P  K  Y  C  G  R] A  S  T  T  V  D  H  Q  A  D
241  GTTGCTGCCACCAAAACTGCCAAGAATCCTACCGATGCTAAACTTGCTGGTGCTGGTAGT
      V  A  A  T  K  T  A  K  N  P  T  D  A  K  L  A  G  A  G  S
                              △
301  CCATGAAAGTAGTAGCTAGCTAGGTTCACGTTGGATTACCAAGCCGTGCCAGTACTACTG
      P  *
361  TGGCCGTGCCAGTACTAATGTTCTCTTATATGTCTGAAATAAGCTCCTATATAAATACTA
421  GTATCTTGATGTAATGGAGTATTTTCATTTTGTTTTTATTTGAGTTATGATCGTGACTTC
481  CTTGTGTTGGTTTAACTTGTATATTGTAATGCATCTTAAATGCTGTCTCAAATAATTTGA
541  TGTATTAAACACTTGTTTTGTTTTTAATACATACTAAGTGCTGTAAATTC
```

Fig. 1. Nucleotide sequence and deduced amino acid sequence of the cDNA insert from clone AC1. The putative signal sequence is underlined and the sequence of mature Ac-AMP2 is boxed. Imperfect direct repeats flanking the mature Ac-AMP2 domain are indicated by dashed underlining. The putative N-glycosylation site is marked by an open triangle and the stop codon is represented by an asterisk.

All chitin-binding proteins so far characterized at the cDNA or gene level have a carboxyl-terminal domain adjacent to the cysteine/glycine-rich domain. The former domain can be either a chitinolytic domain as for class I chitinases [16], a domain homologous to pathogenesis-related PR4 proteins [9] or a short peptide responsible for vacuolar targeting in the case of Gramineae lectins [1]. The carboxyl-terminal portion encoded by Ac-AMP2 transcripts show no sequence homology to previously characterized carboxyl-terminal domains of chitin-binding proteins, nor with any of the entries in the amino acid sequence databank Swiss-Prot (release 20). Hence, the structure of Ac-AMP2 cDNA appears to be unique. The carboxyl-terminal domain of Ac-AMP2 cDNA has some characteristics in common with the carboxyl-terminal propeptide of Gramineae lectins [5] and tobacco glucanase [15]: (1) the absence of the peptide from the mature proteins, (2) the occurrence of a single N-glycosylation site, (3) the presence of a valine residue 8 amino acids ahead of the glycosylation site, and (4) the relative abundance of amino acids with short hydrophobic side-chains (alanine and valine). It is known that the carboxyl-terminal propeptide of Gramineae lectins acts as a sorting determinant for transport to the vacuoles [1], suggesting a similar function for the Ac-AMP2 propeptide. However, the intracellular location of Ac-AMPs remains to be determined, and also it is not shown that the N-glycosylation site of the Ac-AMP2 propeptide is used *in vivo*.

It has previously been suggested that the different types of chitin-binding proteins have arisen from the insertion of a cysteine/glycine rich domain in unrelated genes by a transposition-like mechanism [16]. In support of this hypothesis, Shinshi *et al.* have demonstrated the presence of perfect or imperfect direct repeats, 9 to 10 bp long, flanking the cysteine glycine-rich regions in various basic chitinase genes and in potato *win* genes [16]. Imperfect 10 bp repeats also occur in the Ac-AMP2 cDNA sequence, namely 21 bp upstream of the first nucleotide of the cysteine/glycine-rich domain (100 ATGGTGGATC) and 10 bp downstream of the last nucleotide of the cysteine/glycine-rich domain (200 ACTGTGGATC) (Fig. 1). Such direct repeats are believed to be remnants of sequence duplications created at the target site of insertion by a transposon-like element [12].

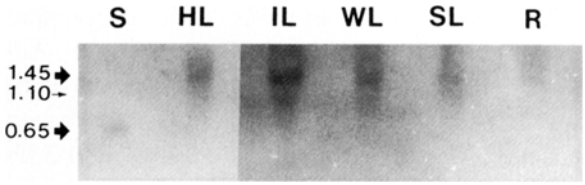

Fig. 2. Northern blot analysis of AC-1 cross-hybridizing transcripts in different healthy and stressed tissues of *A. caudatus*. Near-mature seeds (S) and roots (R) were collected from adult *A. caudatus* plants. Leaves of five-week-old *A. caudatus* plants were infected with a *Botrytis cinerea* spore suspension (5×10^5 spores/ml in 0.5 M glucose) (IL), wounded with a sterile steel brush (WL) or sprayed with 10 mM sodium salicylate (SL) and harvested three days after treatment, together with healthy leaves from untreated control plants (HL). Total RNA was isolated as described [10]. Glyoxylated RNA (20 μg per sample) was fractionated on a 1.4% agarose gel and transferred to nylon membranes as described [13]. A digoxigenin-labelled RNA antisense probe corresponding to AC1 was prepared using T3 RNA polymerase and *Hin*d III-linearized pBluescript II SK containing the AC1 insert. The membranes were
hybridized, washed and developed with Lumigen PPD (Boehringer Mannheim) according to the manufacturer's instructions.

Ac-AMPs were originally isolated from seeds of amaranth [3]. In order to examine whether Ac-AMPs or Ac-AMP-like proteins are expressed in other tissues as well, northern blotting analysis was performed on RNA extracted from seeds, roots, leaves and leaves subjected to various stress treatments. As expected, near-mature seeds of *A. caudatus* contain a 650 bp transcript that hybridizes with the Ac-AMP2 cDNA probe (Fig. 2). However, the 650 bp transcripts were not detected in roots, nor in leaves of healthy *A. caudatus* plants, neither in leaves stressed by infection with the fungus *Botrytis cinerea*, by wounding or by salicylic acid treatment. In contrast, roots and particularly healthy or stressed leaves contain two transcripts of 1.45 kb and 1.1 kb, respectively, that cross-hybridize with the Ac-AMP2 cDNA probe. These transcripts, which are not present in near-mature seeds, may encode

chitinases or other proteins with a cysteine/glycine-rich domain homologous to the Ac-AMPs. It has previously been shown that chitinase transcripts from bean leaves [4] and tobacco leaves [14] have a size of ca. 1.2 kb and 1.3 kb respectively. We conclude from the available data that true Ac-AMPs are most probably expressed only in seeds of amaranth. These peptides do not seem to have homologues that are expressed in vegetative tissues, not even after stress treatments known to induce pathogenesis-related proteins [8]. Hence, if Ac-AMPs play a role in plant defense, this role is likely to be restricted to seed protection.

Acknowledgements

This work was supported in part by the ECLAIR programme (AGRE-0005) of the Commission of the European Community. The authors are indebted to K. Eggermont and I. Goderis for excellent technical assistance. W.F.B. is Research Associate of the Belgian 'Nationaal Fonds voor Wetenschappelijk Onderzoek'.

References

1. Bednarek SY, Raikhel NV: The barley lectin carboxyl-terminal propeptide is a vacuolar protein sorting determinant in plants. Plant Cell 3: 1195–1206 (1992).
2. Broekaert W, Lee H-I, Kush A, Chua N-H, Raikhel N: Wound-induced accumulation of mRNA containing hevein sequence in laticifers of rubber tree (*Hevea brasiliensis*). Proc Natl Acad Sci USA 87: 7633–7637 (1990).
3. Broekaert WF, Mariën W, Terras FRG, De Bolle MFC, Proost P, Van Damme J, Dillen L, Claeys M, Rees SB, Vanderleyden J, Cammue BPA: Antimicrobial peptides from *Amaranthus caudatus* seeds with sequence homology to the cysteine/glycine-rich domain of chitin-binding proteins. Biochemistry 31: 4308–4314 (1992).
4. Broglie KE, Gaynor JJ, Broglie RM: Molecular cloning of the genes encoding an endochitinase from *Phaseolus vulgaris*. Proc Natl Acad Sci USA 83: 6820–6824 (1986).
5. Chrispeels MJ, Raikhel NV: Short peptide domains target proteins to plant vacuoles. Cell 68: 613–616 (1992).
6. de Vries S, Hoge H, Bisseling T: Isolation of total and polysomal RNA from plant tissues. In: Gelvin SB, Schilperoort RA, Verma DPS (eds), Plant Molecular Biology Manual, pp. B6/1–B6/13. Kluwer Academic Publishers, Dordrecht (1988).
7. Faye L, Johnson D, Sturm A, Chrispeels MJ: Structure, biosynthesis, and function of asparagine-linked glycans on plant glycoproteins. Physiol Plant 75: 309–314 (1989).
8. Linthorst HJM: Pathogenesis-related proteins of plants. Crit Rev Plant Sci 10: 123–150 (1991).
9. Linthorst HJM, Danhash N, Brederode FT, van Kan JAL, de Wit PJGM, Bol JF: Tobacco and tomato PR proteins homologous to *win* and pro-hevein lack the 'hevein' domain. Mol Plant-Microbe Interact 4: 586–592 (1991).
10. Logemann J, Schell J, Willmitzer L: Improved method for the isolation of RNA from plant tissues. Anal Biochem 163: 16–20 (1987).
11. Raikhel NV, Broekaert WF: Chitin-binding proteins: their biology and possible function. In: Verma DPS (ed) Control of Plant Gene Expression, pp. 407–423. CRC Press (1993).
12. Saedler H, Nevers P: Transposition in plants: a molecular mode. EMBO J 4: 585– 590 (1985).
13. Sambrook J, Fritsch EF, Maniatis T: Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989).
14. Shinshi H, Mohnen D, Meins F: Regulation of a plant pathogenesis-related enzyme: inhibition of chitinase and chitinase mRNA accumulation in cultured tobacco tissues by auxin and cytokinin. Proc Natl Acad Sci USA 84: 89–93 (1987).
15. Shinshi H, Wenzler H, Neuhaus J-M, Felix G, Hofsteenge J, Meins F Jr: Evidence for N- and C-terminal processing of a plant defense-related enzyme: primary structure of tobacco prepro-β-1,3-glucanase. Proc Natl Acad Sci USA 85: 5541–5545 (1988).
16. Shinshi H, Neuhaus J-M, Ryals J, Meins F Jr: Structure of tobacco endochitinase gene: evidence that different chitinase genes can arise by transposition of sequences encoding a cysteine-rich domain. Plant Mol Biol 14: 357–368 (1990).
17. Von Heijne G: Transcending the impenetrable: how proteins come to terms with membranes. Biochim Biophys Acta 974: 307–333 (1988).