

Isolation and characterization of a cDNA clone encoding the pokeweed antiviral protein II from *Phytolacca americana* and its expression in *E. coli*

Jean-Luc Poyet, Jean Radom, Arnd Hoeveler*

Laboratoire de Biochimie et Biologie Moléculaire, UFR Sciences et Techniques, 16 Route de Gray, 25030 Besançon Cedex, France

Received 25 April 1994

Abstract

Three distinct ribosome-inactivating proteins (RIPs) were isolated from pokeweed (*Phytolacca americana*). We identified and sequenced for the first time a complete cDNA encoding the pokeweed antiviral protein II (PAP II), which is expressed in the late summer leaves of pokeweed. The cDNA of PAP II consists of 1,187 nucleotides and encodes a mature protein of 285 amino acids. Its predicted amino acid sequence is only 33% similar to PAP and PAP-S. The NH₂ terminal extrapeptide (25 amino acid residues) was similar but not identical to that of PAP's extrapeptide. The cDNA of PAP II was expressed in *E. coli*. The growth of the transformants was strongly inhibited after induction of the gene. Furthermore, PAP II, which was produced in *E. coli*, inhibited protein synthesis in a rabbit reticulocyte translation system. Thus, recombinant PAP II would appear to be as functional as native PAP in inhibiting protein synthesis in both prokaryotes and eukaryotes.

Key words: cDNA; Ribosome-inactivating protein; *Phytolacca americana*; Pokeweed antiviral protein from summer leaves; Expression in *E. coli*; Inhibition of in vitro translation

1. Introduction

The plant pokeweed (*Phytolacca americana*) produces at least three ribosome-inactivating proteins (RIPs) in different tissues and at various stages of its development. PAP, PAP II and PAP-S are the forms of pokeweed antiviral proteins (PAPs) that appear in spring leaves, summer leaves and seeds respectively [1–3]. Whereas PAP is expressed at all stages of development, PAP II is synthesized progressively with the ageing of the plant [4]. PAP II does not react with anti PAP serum, whereas PAP-S shows a partial reaction [1,3].

PAPs inhibit protein synthesis by cleaving the N-glycosidic bond at adenine 4324 of rat liver 28S-rRNA and prevent binding of elongation factors [5,6]. Their antiviral effect is not understood in molecular terms [7,8]. However, it has been suggested that PAP entry is mediated by changes in the cellular membrane induced by the adsorption of viral particles on the cell surface [7]. In this process, virus entry is not required. A change in the cell membrane integrity could lead to the entry of PAP and thus would provide a mechanism of cellular suicide.

Because of this dual inhibitory activity, PAPs (as other RIPs) have become the subject of a wide range of investigations concerning their potential application as novel therapeutic agents and as putative protective proteins used by plants as a defense against viruses [9,10].

Recently, studies have demonstrated the use of PAP as the killing component of immunotoxins for the treatment of acute lymphoblastic leukemia [11–13]. An immunotoxin composed of PAP and B43 anti-CD19 mono-

clonal antibody with antileukemic efficiency has been extensively characterized and approved for a phase I clinical trial [12,13].

PAP, as a single chain RIP, has some advantages when compared with double chain RIP (i.e. ricin). It is smaller in size, extremely thermostable and has very low aspecific toxicity for intact eukaryotic cells, due to the lack of a galactose binding unit.

Recently, the nucleic acid sequence of PAP [14], PAP α [15], a sequence related to PAP and the amino acid sequence of PAP-S [16] were determined. However, the sequence of PAP II is not known. An amino acid sequence comparison between PAP and PAP-S revealed 76% homology. It is known from N-terminus protein sequencing that PAP II is less homologous [17].

Since it has been shown that isozyme forms of various RIPs [18] may have different properties that may make one form more valuable than another for therapeutic applications, it was decided to clone and sequence the cDNA encoding the PAP II protein.

In this analysis RNA was extracted from pokeweed leaves of the late summer months and a cDNA library was constructed. A complete cDNA encoding PAP II was identified, sequenced and expressed in *E. coli*. Finally, the functional efficiency of PAP II to inhibit in vitro protein synthesis was demonstrated.

2. Materials and methods

2.1. Isolation of PAP-II cDNA

Total RNA was extracted from the leaves of *Phytolacca americana* in late summer according to the method of Cashmore [19]. Poly (A)⁺ RNA were purified on oligo-dT column (Pharmacia). Subsequently, a cDNA library was constructed using the 'TimeSaver cDNA Synthesis Kit' from Pharmacia and pUC18 as cloning vector. In order to screen

* Corresponding author. Fax: (33) 81 66 62 67.

this cDNA library a DNA-fragment of PAP II was amplified by reversed transcriptase PCR (RT-PCR). It was postulated that a highly conserved region existed in PAP II found in all RIPs and involved in the catalytic action of these proteins [9]. The corresponding degenerated nucleic acid sequence to amino acid sequence was designed, primer P9 (5'-CGGGATCCCAT(G,A)TA(C,T)TT(G,A)AATCTIGCIGC(C,T)T-3'), with a *Bam*HI site. Five mg of RNA were hybridized with 5 pmol of primer P9, and RNAs were reverse-transcribed in reaction volume of 20 μ l containing 50 mM Tris-HCl, pH 8.3, 40 mM KCl, 10 mM DTT, 6 mM MgCl₂, 700 mM of each dNTP, 20 U of RNasin and 30 U of M-MLV reverse transcriptase (Boehringer). The reaction products were incubated for 1 h at 37°C. A second primer, P8, was designed by reference [17] to the amino acid sequence of the N-terminus of PAP II (5'-CCATCGATGAA(C,T)AT(C,T,A)GTITT(C,T)GA(C,T)GTIGA(G,A)AA-3') with a *Clal* site. Five μ l of reverse transcription reaction mixture were added to 45 μ l of 10 mM Tris-HCl, pH 8.8, 50 mM KCl, 1.5 mM MgCl₂, 0.1% Triton X-100, 30 pmol of primer P8 and P9, 200 mM of each dNTP and 2 units of *Taq* polymerase (Promega). The reaction was performed through 30 cycles of 30 s at 95°C, 1 min at 54°C and 1 min at 72°C. A 500 bp fragment was amplified, gel purified and cloned into the pBluescript KS+ vector (Stratagene) digested by the appropriate enzymes. The nucleotide sequence was determined using the '77 Sequencing Kit' (Pharmacia). After confirmation of its sequence the fragment was used as a probe to screen the cDNA library. The largest positive cDNA clone was sequenced in both directions using internal primers. Protein alignments were performed using the CLUSTAL program [20].

2.2. Cloning of PAP II into a procaryotic expression vector

The primer P10 (5'-CAGCTGCAGAACATAGTGTGACGTTGAG-3') and the primer P11 (5'-CTGGGATCCCTCGAATTCACCAAGGTTAC-3') were designed to clone the PAP II cDNA without the NH₂-extrapeptide into the expression vector pFv90 (a generous gift from Dr. M. Little, Deutsches Krebsforschungszentrum Heidelberg). The plasmid pFv90 was a derivative of the expression vector pSEX [21]. It is inducible by IPTG and contains a procaryotic signal peptide of 26 amino acid residues allowing the secretion of the recombinant protein into the periplasm of the bacteria.

The complete nucleotide sequence of the PCR reaction products were compared with the cDNA clone.

2.3. Expression of PAP II

The *E. coli* strain JM109 was transformed by pFv90 or pUC18 with a PAP II insert. The transformants were designated pFv90/PAP II and pUC18/PAP II. pFv90/PAP II was grown at 37°C in LB medium containing ampicillin (50 μ g/ml) to an optical density of 0.4 at 600 nm. After induction with 1 mM IPTG, the culture was incubated for an additional hour at 37°C. Cells were harvested by centrifugation, resuspended in 30 mM Tris-HCl, pH 8, 20% sucrose, 1 mM EDTA and subsequently sonicated. Proteins were separated by 12.5% SDS-polyacrylamide gel electrophoresis.

2.4. Growth inhibition of bacteria

A preculture of pFv90/PAP II or pUC18/PAP II was used to initiate a culture in LB medium containing 1 mM IPTG. The OD₆₀₀ of the cultures were measured every 20 min.

2.5. Inhibition of in vitro translation

The activity of the recombinant PAP II to inhibit protein synthesis was measured using a rabbit reticulocyte system (Promega). Bromo mosaic virus RNA (Boehringer) was used as template. The reactions were in 12.5 μ l.

3. Results and discussion

3.1. Cloning the coding sequence of PAP II

Late summer leaves of *Phytolacca americana* were harvested, mRNA was extracted and the corresponding cDNA was synthesized. With the help of two degenerated primers, one for the N-terminus and the second for

```

1  CTTTCTCTTCTTTTCTGTTCAATTACAAAAGCTTCAGCTGCGCAACAGGGAATGAAG
61  ATGAAGGCTGTAGAAAGTAGTTGGGTGGCAATATCGATATGGCTGATGCTTACACCACCA
    M K V L E V V G L A I S I W L M L T P P
121  GCTTCTTCAAAACATAGTGTGTTGACGTTGAGAATGCCACACAGAAACCTACTCTAATTTT
    A S S N I V F D V E N A T P E T Y S N F
181  CTGACTAGTTTGGCAGAGCTGTGAAGACAGAAATGACATGCCATGGAATGATAATG
    L T S L R E A V K D K K L T C H G M I M
241  GCCACACCTCAGTGAACAACCCAGTATGTTGGTTGACCTCAAAATCCGATCTGGA
    A T T L T E Q P K Y V L V D L K F G S G
301  ACATTCACATTAGCAATCAGAAGGGGAACTTATATTGGAGGGCTATTCTGACATTAC
    T F T L A I R R G N L Y L E G Y S D I Y
361  AATGAGAAATGCTGTTATCGGATCTTCAAGGATTCAGAATCCGATGCCCAAGAGACCGTT
    N G K C R Y R I F K D S E S D A Q E T V
421  TGCCCCGGGACAAAAGCAAGCTGGCACTCAGAATAATATCCCTATGAAAGAGTTAC
    C P G D K S K P G T Q N N I P Y E K S Y
481  AAAGGGATGGAATCAAAGGTGGGCTAGAACTAAATAGGGTTAGGAAGATAACACTC
    K M E S K G G A R F D I N G Y Q P D P
541  AAGAGTCGAATGGTAAATCTACGGCAAGGATGCAACGGATCAGAAGCAGTATCAAAA
    K S R M G K I Y G K D A T D Q K Q Y Q K
601  AATGAGGCTGAATTTCTTCTATAGCGTTCAAATGGTTACTGAGGCATCAAGGTTCAAA
    N E A E F L L I A V Q M V T E A S R F K
661  TACATTGAGAACAAGGTAAGGCTAAATTTGATGATGCCAATGGTATCAGCCAGATCCT
    Y I E N K V K A K F D I N G Y Q P D P
721  AAAGCTATTTCCCTAGAGAAAATGGGACAGTGTGTTCTAAGGTCATTGCAAAAAGTTGGC
    K A I S L E K N W D S V S K V I A K V G
781  ACCTCCGGTATAGTACTGTCTTACTTTACCTGGAGACCTAAAAGATGAGAATAATAAACCT
    T S G D S T V T L P G D L K D E N N K P
841  TGGACTACGGCCACCATGAACGACCTTAAAGACGACATTATGGCACTCTCAACCCAGTT
    W T T A T M N D L K N D I M A L L T H V
901  ACTTGCAAGGTTAAAGTTCCATGTTCCCTGAAATATATGCTTATTATAGGACTAGT
    T C K V K S S M F P E I M S Y Y R T S
961  ATTAGTAACCTTGGTGAATTCGAGTGATTCAAATCATATCAATGTGAATAAGGAGGCA
    I S N L G E F E -
1021  AGCAACAATAGTATATTGTTGACCGGCTCTATGTATCTTATGTGTTTTATCTTCTC
1081  GTACTACCTTGTCTTACTACTGTACCTAATAAGTACCGGTGAATCATATTGGAATAAAA
1141  TTTGGCTTATAGAATCCAATGTTTAAATTAATTTCTAAAAA

```

Fig. 1. Nucleotide sequence and the deduced amino acid sequence of the PAP II-cDNA clone. The putative signal peptide (25 amino acids) is encoded by bases 55–129. Bases 130–984 encode a mature protein of 285 amino acids. The putative polyadenylation signal starts at base 1,134 and the poly(A) tail starts at base 1,179.

a highly conserved sequence found in all RIPs and located approximately in the middle of the PAP II protein, one fragment was amplified corresponding to the NH₂ terminal part of the protein. This 500 bp fragment was used to screen a cDNA library. Several clones were identified, the largest one was sequenced. The complete sequence of PAP II is shown in Fig. 1.

The PAP II cDNA clone was 1,187 bp in length with an open reading frame of 932 bp. The deduced polypeptide had a putative signal peptide of 25 amino acid residues and a coding sequence of 285 amino acid residues. According to the sequence, the theoretical molecular weight of PAP II was 32 kDa. The first 30 amino acid residues from the deduced mature protein were identical to those obtained by amino-terminal sequencing of purified PAP II [17] which confirmed that the cDNA clone encoded the PAP II polypeptide. As in most plant and animal mRNAs, a consensus polyadenylation signal (AATAAA) was found 38 bases upstream of a putative polyadenylation site.

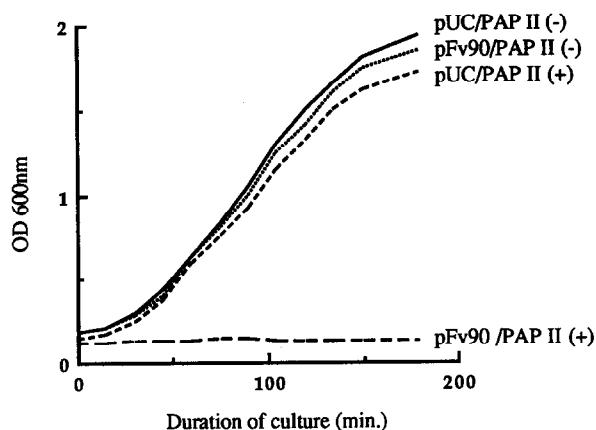


Fig. 4. Growth curves of transformants. (+) and (-) indicate the presence or absence of 1 mM IPTG, respectively.

The sequence also suggested the presence of extrapeptides at the N- and probably at the C-terminus. The signal peptide of PAP II was homologous, but not identical to that of PAP's, documenting an important role of this sequence to address PAPs to the cell wall matrix. Since the theoretical molecular weight of PAP II is 2 kDa higher than the purified form from leaves [3], the cDNA may encode a protein harboring a C-terminus. The presumed carboxy-extension of PAP II contained the motive (MFP) which was also found in PAP and PAP α . Comparisons with sequences found in other RIPs show no obvious similarities. N-glycosylation sites common to other plant proteins that are targeted to vacuoles of the plant cell were not observed in PAP II.

3.3. Expression of PAP II in *E. coli*

The cDNA sequence of PAP II was modified in order to clone it into the procaryotic expression vector pFv90. The promoter of pFv90 gives high expression of foreign proteins in *E. coli* after induction by IPTG. The expression vector pFv90/PAP II was introduced into the *E. coli* strain JM109 and the expression of PAP II was induced by adding IPTG to the culture medium.

A protein band at approximately 35 kDa was readily detectable on SDS-PAGE after induction with IPTG (lane 4, Fig. 3). The recombinant PAP II protein was approximately 3 kDa larger in size than the native mature PAP II polypeptide, since pFv90 contains a signal peptide of 26 amino acid residues which permits the secretion of the recombinant protein into the periplasm of the bacteria.

To demonstrate that the cDNA encoding recombinant PAP II protein is biologically active two different sets of experiments were performed.

Firstly, the influence of PAP II protein expression on *E. coli* growth was tested. Cultures were initiated from overnight precultures and the cell growth was followed by measuring the optical density at 600 nm. PAP II

expression was induced by adding IPTG to the medium. As shown in Fig. 4, in the presence of IPTG, bacterial culture harboring pFv90/PAP II stopped growth while several control cultures continued to grow. From these experiments it can be concluded that the recombinant PAP II was active on *E. coli* ribosomes and inhibited protein synthesis in the same way as native PAP and MAP (Mirabilis antiviral protein, [26]) do. Single chain RIPs can cleave the bond between A-2660 and the ribose of the prokaryotic 23S-rRNA, a position equivalent to the A-4324 of the eucaryotic 28S-rRNA [25,27]. Therefore single chain RIPs are toxic for *E. coli* inhibiting protein synthesis and leading to growth arrest.

Secondly, tests were carried out to determine whether PAP II inhibits in vitro translation of rabbit reticulocyte lysates (Fig. 5). Protein synthesis was obtained with bacterial lysates without plasmid (lane 1), with pFv90/PAP II but without induction (lane 2), and in the presence of water (lane 4). However, bacterial lysates containing PAP II protein completely inhibited protein synthesis (lane 3). These results clearly show the efficiency of recombinant PAP II to inhibit eukaryotic protein synthesis.

In recent years, there has been a growing interest in using single chain RIPs, such as PAP, as the killing component of immunotoxins [9–13]. Frequently one plant produces a family of RIPs. The proteins TAP and TAP29 from *Triochosanthos kirilowi* are isozyme forms which exhibit different properties making one form more

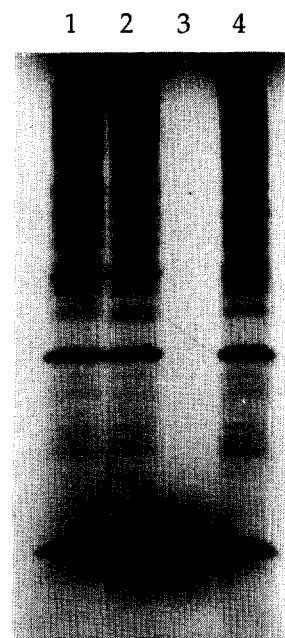


Fig. 5. Inhibitory activity of PAP II, on in vitro protein synthesis in a rabbit reticulocyte system using Brome mosaic virus RNA as template: autoradiography of the in vitro synthesized proteins analysed by SDS-PAGE. Protein syntheses were carried out in the presence of 100 ng of proteins of cell lysate from *E. coli* without an expression vector (lane 1), with pFv90/PAP II after induction by IPTG (lane 3), with pFv90/PAP II without IPTG (lane 2) or in the presence of water (lane 4).

suitable than another for a certain clinical purpose [18]. The potential and ability of PAP II as a therapeutic agent is not known, but it can now be used to construct toxic hybrid proteins. This may have useful clinical applications in the future.

Acknowledgements: We thank P. Millet and his colleagues at the Botanical Garden of the University of Franche-Comté, for providing the plant material; Dr. M. Little and Dr. J. Dübel for the generous gift of the pFv90 expression vector and Y. Gaudemer for supporting this work and reading the manuscript. J.L. Poyet has a fellowship from the Association pour la Recherche sur le Cancer (A.R.C.). This work was supported by the Ministère de la Recherche et de l'Éducation. PAP II will appear under Accession No. X78628 in the EMBL Data Library.

References

- [1] Barbieri, L., Aron, G.M., Irvin, J.D. and Stirpe, F. (1982) *Biochem. J.* 203, 55–59.
- [2] Irvin, J.D. (1975) *Arch. Biochem. Biophys.* 169, 522–528.
- [3] Irvin, J.D., Kelly, T. and Robertus, J.D. (1980) *Arch. Biochem. Biophys.* 200, 418–425.
- [4] Houston, L.L., Ramakrishnan, S. and Hermodson, M.A. (1983) *J. Biol. Chem.* 258, 9601–9604.
- [5] Endo, Y., Tsurugi, K. and Lambert, J.M. (1988) *Biochem. Biophys. Res. Commun.* 150, 1032–1036.
- [6] Endo, Y., Mitsui, K., Motizuki, M. and Tsurugi, K. (1987) *J. Biol. Chem.* 262, 5908–5912.
- [7] Ready, M., Brown, D.T. and Robertus, J.D. (1986) *Proc. Natl. Acad. Sci. USA* 83, 5053–5056.
- [8] Frötschl, R., Schönfelder, M. and Adam, G. (1990) *Arch. Phytopathol. Pflanzenschutz* 26, 319–328.
- [9] Irvin, J.D. and Uckun, F.M. (1992) *Pharmacol. Ther.* 55, 279–302.
- [10] Pastan, I. and FitzGerald, D. (1991) *Science* 254, 1173–1177.
- [11] Myers, D.E., Irvin, J.D., Smith, R.S., Kuebelbeck, V.M., Tuel-Ahlgren, L. and Uckun, F.M. (1991) *J. Immunol. Methods* 136, 221–238.
- [12] Uckun, F.M., Manivel, C., Arthur, D., Chelstrom, L.C., Finnegan, D., Irvin, J.D., Kuebelbeck, V., Myers, D.E. and Gunther, D. (1992) *Blood* 79, 2201–2214.
- [13] Jansen, B., Kersey, J.H., Jaszc, W.B., Gunther, R., Nguyen, D.P., Chelstrom, L.M., Tuel-Ahlgren, L. and Uckun, F.M. (1993) *Leukemia* 7, 290–297.
- [14] Lin, Q., Chen, Z.C., Antoniwi, J.F. and White, R.F. (1991) *Plant Mol. Biol.* 17, 609–614.
- [15] Kataoka, J., Habuka, N., Masuta, C., Miyano, M. and Koiwai, A. (1992) *Plant Mol. Biol.* 20, 879–886.
- [16] Kung, S.S., Kimura, M. and Funatsu, G. (1990) *Agric. Biol. Chem.* 54, 3301–3318.
- [17] Bjorn, M.J., Larrick, J., Piatak, M. and Wilson, K.J. (1984) *Biochem. Biophys. Acta* 790, 154–163.
- [18] Lee-Huang, S., Huang, P.L., Kung, H.-F., Bao-Qun, L., Huang, P.L., Huang, P., Huang, H.I. and Chen, H.C. (1991) *Proc. Natl. Acad. Sci. USA* 88, 6570–6574.
- [19] Cashmore, A.R. (1982) *Methods in Chloroplast Molecular Biology*, Elsevier, Amsterdam, pp. 387–392.
- [20] Higgins, D.G. and Sharp, P.M. (1988) *Gene* 73, 237–244.
- [21] Breitling, F., Dübel, S., Seehaus, T., Klewinghaus, I. and Little, M. (1991) *Gene* 104, 147–153.
- [22] Frankel, A., Welsh, P., Richardson, J. and Robertus, J.D. (1990) *Mol. Cell. Biol.* 10, 6257–6263.
- [23] Kim, Y. and Robertus, J.D. (1992) *Protein Eng.* 5, 775–779.
- [24] Monzingo, A.F., Collins, E.J., Ernst, S.R., Irvin, J.D. and Robertus, J.D. (1993) *J. Mol. Biol.* 233, 705–715.
- [25] Hartley, M.R., Legname, G., Osborn, R., Chen, Z. and Lord, J.M. (1991) *FEBS Lett.* 290, 65–68.
- [26] Kataoka, J., Habuka, N., Furuno, M., Miyano, M., Takanami, Y. and Koiwai, A. (1991) *J. Biol. Chem.* 266, 8426–8430.
- [27] Dore, J.M., Gras, E., Depierre, F. and Wijdenes, J. (1993) *Nucleic Acids Res.* 21, 4200–4205.