cDNA cloning and expression of pokeweed antiviral protein from seeds in *Escherichia coli* and its inhibition of protein synthesis in vitro

Jean-Luc Poyet, Arnd Hoeveler*

Laboratoire de Biochimie et de Biologie Moléculaire, UFR des Sciences et des Techniques, 16 Route de Gray, 25030 Besancon cedex, France

Received 2 January 1997

Abstract Pokeweed antiviral proteins (PAP) represent a family of protein toxins isolated from various organs and at different stages of development of Phytolacca americana (pokeweed). We isolated, sequenced and characterized for the first time a complete cDNA encoding a pokeweed antiviral protein expressed in seeds. The cDNA of PAP-S consists of 1249 nucleotides and encodes a mature 262 amino acid protein. Its predicted amino acid sequence is more similar to PAP (76%) than to PAP II (31%). It is known from literature that PAP-S is more active in inhibiting protein synthesis than other members of the PAP family. Therefore, the cDNA of PAP-S was expressed in Escherichia coli and the biological activity of the recombinant protein was compared with that of PAP purified from spring leaves. In a rabbit translation system, the median inhibitory concentrations (IC50) of recombinant PAP-S and native PAP were determined as 0.07 and 0.29 nM, respectively. Although the PAP-S protein in seeds is glycosylated, PAP-S can be expressed in Escherichia coli in a very active form, indicating that posttranslational modification in pokeweed does not seem to alter its ability to inhibit protein synthesis.

© 1997 Federation of European Biochemical Societies.

Key words: cDNA; Ribosome-inactivating protein; Phytolacca americana; Pokeweed antiviral protein from seeds; Expression in E. coli; Inhibition of in vitro protein synthesis

1. Introduction

Pokeweed antiviral protein is a potent cytotoxin isolated from several organs of the plant pokeweed (*Phytolacca americana*). It attacks both procaryotic and eucaryotic ribosomes, removing a specific adenine base from the rRNA larger subunit and thereby inhibits protein synthesis [1,2]. Thus PAP is a member of a family of RNA N-glysosidases called ribosome-inactivating proteins (RIP) type I which all consist of a single-chain protein [3]. The site attacked by RIPs is a single-stranded loop, which is highly conserved in 28S, 26S, 23S and chloroplast rRNA. In *E. coli* 23S rRNA, this loop has been shown to be involved in interactions with elongation factors EF-G and EF-Tu.

Within the *PAP* gene family, different genes are expressed in various tissues and at different stages of development in *Phytolacca americana*. PAP, PAP II, PAP-S and PAP-R are the forms that appear in spring leaves, summer leaves, seeds and roots, respectively [4–7]. The molecular weight ranges from 29 kDa for PAP to 30 kDa for PAP-S and PAP II. Recently, several members of the PAP family were cloned and the nucleic acid sequences of PAP [8], PAP II [9] and the amino acid sequence of PAP-S [10] were determined.

*Corresponding author. Fax: (32-2) 299-18-60

Although PAP-S was identified as most effective in inhibiting protein synthesis in vitro [6], its nucleic acid sequence is still unknown

PAPs possess antiviral activity on a wide range of plant and human viruses. PAP-S inhibits infection of HEp-2 cells by herpes simplex virus and poliovirus [11] and also the replication of human immunodeficiency virus 1 (HIV-1) in isolated mononuclear blood cells infected in vitro [12]. In the latter study, PAP-S was more effective than other RIPs at inhibiting the expression of reverse transcriptase in infected cells. Dose response experiments indicated that the anti-HIV activity of PAP-S was comparable to AZT. Conjugation of PAP-S with monoclonal antibodies directed against CD4 and the HIV protein gp120 dramatically increase its potency against human T cells infected with two different stains of the HIV virus [13].

Expression of PAP in transgenic plants leads to broad-spectrum resistance to viral infection [14]. Thus, PAPs offer new opportunities for possible application in therapy and as putative protective proteins in transgenic plants against pathogens.

Since it has been shown that isozyme forms of various RIPs [15] 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-S protein.

In this analysis, a cDNA library from pokeweed seeds was constructed and a complete cDNA encoding PAP-S was identified. The protein was expressed in *E. coli* and compared in detail to the activity of PAP produced in leaves of *Phytolacca americana*. The recombinant PAP-S was more active than native PAP in inhibiting eukaryotic in vitro translation.

2. Material and methods

2.1. Construction of the cDNA library

Total RNA was extracted from the seeds of Phytolacca americana as described in reference [16]. Poly(A)+ RNA were purified on an oligo-dT column and used to synthesize cDNA with the 'TimeSaver cDNA Synthesis Kit' from Pharmacia. The cDNAs were inserted between the EcoRI sites of the λZipLox vector (Gibco BRL) and these recombinant DNAs were packaged in vitro using the λ Packaging System (Gibco BRL). The phages were amplified in E. coli Y1090 and subsequently more than 105 independent clones were screened. For the screening, a DNA fragment of PAP-S was amplified by reversed transcription-PCR. A primer (P9) was designed according to the known amino acid sequence of the protein [10], corresponding to the highly conserved region found in all RIPs and involved in the catalytic action of these proteins (5'-CGGGATCCCAT(G,A)-TA(C,T)TT(G,A)AATCTIGCIGC(C,T)T-3') with a BamHI site. Total RNA was reversed-transcribed as described previously [9] and the cDNA was amplified using P9 and a primer (P12) designed to the amino acid sequence of the N-terminus of PAP-S (5'-CCATCGAT-GAT(C,T,A)AA(C,T)ACIAT(C,T,A)ACITT(C,T)GA(C,T)GC-3') with a ClaI site. The amplified fragment (500 bp) was digested and cloned into the pBluescript KS+ vector (Stratagene) digested by ClaI and BamHI. After sequencing, the fragment was used as a probe to

screen the cDNA library. Recombinant pZL1 vectors were rescued from positive bacteriophage clones by in vivo excision according to the procedure provided by the manufacturer. The nucleotide sequence of the largest positive clone was determined in both directions using internal primers. Protein alignments were performed using the CLUSTAL program [17].

2.2. Construction of the PAP-S expression plasmid

Primers P13 (5'-CTTCAGCTGATAAATACGATCACCTTT-GATC-3') and P14 (5'-CGGGATCCGAATCCTTCAAATAGAT-CAC-3') were designed to clone the PAP-S cDNA without the NH₂-terminal extrapeptide into the isopropyl-β-D-thiogalactopyranoside (IPTG)-inductible prokaryotic expression vector pQE60 (Qiagen), giving the vector pPAPS. The complete nucleotide sequence of the PCR reaction product was determined to verify that no mutation had occurred.

2.3. Expression and purification of recombinant PAP-S

The *E. coli* strain JM109 was transformed by pPAPS. The transformants were grown at 37°C in LB medium containing ampicillin (100 µg/ml) to an optical density of 0.5 at 600 nm. After the addition of IPTG to a final concentration of 1 mM, culture was continued for 2 h. The inclusion bodies containing the cytoplasmic recombinant protein were prepared as in [18], except for the last solubilization step done at room temperature for 1 h in a solution containing 10 mM Tris-HCl (pH 8), 8 M urea, 0.2 M NaCl, 1 mM DTT, 10% glycerol and 1 mM EDTA. The solubilized proteins were reade by dialysis against 10 mM Tris-HCl (pH 7.4), 0.1 M NaCl and centrifuged at $15000 \times g$ for 10 min. The supernatant was analyzed by SDS-polyacrylamide gel electrophoresis to estimate the purity of the recombinant protein.

2.4. Inhibition of cell-free protein synthesis

The enzyme activity of the purified recombinant protein was quantitated by a protein synthesis inhibition assay using a rabbit reticulocyte system as described elsewhere [19]. Brome mosaic virus RNA was used as template mRNA. The incorporation of L-[35S]methionine into the TCA-precipitable product was measured by scintillation spectroscopy. Purified PAP from pokeweed leaves was used as positive control.

3. Results and discussion

3.1. Cloning of PAP-S and sequence comparison

Messenger RNA was purified from the seeds of *Phytolacca* americana and the corresponding cDNA library was synthesized and cloned in λ ZipLox. For the screening, the fragment corresponding to the NH₂-terminal part of the protein was used. Several clones were identified and the largest one was sequenced. The complete sequence of PAP-S is shown in Fig. 1.

The PAP-S cDNA clone has an open reading frame of 942 bp and the total length of 1249 bp. The deduced polypeptide has a coding sequence of 314 amino acid residues. A comparison of this sequence with that of PAP-S determined by direct amino acid sequencing [10] suggested that the cDNA encodes a protein containing extrapeptides of 24 and 28 amino acids at the NH₂-terminal and the COOH-terminal, respectively (Fig. 1). According to the sequence, the theoretical molecular weight of mature PAP-S is 29 kDa.

Alignment of the amino acid sequences of PAP-S, PAP, PAP II revealed that PAP-S is more similar to PAP (76% identity) than to PAP II (31% identity) (Fig. 2) indicating a higher degree of relatedness and less evolutionary diversification among PAP and PAP-S. This is in line with recent evidence of our laboratory, that a genomic gene for PAP II encodes two exons separated by one intron, whereas the genomic genes of PAP-S and PAP do not contain an intron (J.L. Poyet and A. Hoeveler, submitted).

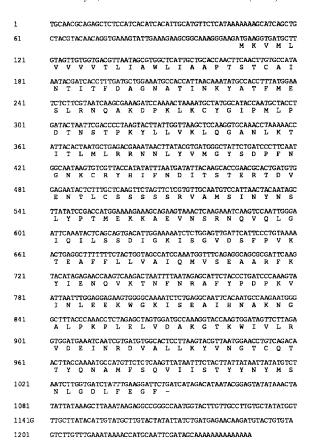


Fig. 1. Nucleotide sequence and the deduced amino acid sequence of the PAP-S cDNA clone. The putative signal peptide (24 amino acids) is encoded by bases 107-178. The putative COOH-terminal extrapeptide (28 amino acids) is encoded by bases 964–1047. Bases 178–963 encode a mature protein of 262 amino acids. The putative polyadenylation signal starts at base 1211 and the poly(A) tail starts at base 1236.

The NH₂-terminal extrapeptide shows features typical among secretory signal peptides [20], leaving the possibility to PAP-S to be targeted to a particular cellular compartment in the seeds. The exact subcellular location of PAP-S in pokeweed seeds has not been determined, although several other seeds RIPs have been localized to protein bodies or vacuoles within endosperm cells [21]. N-glycosylation sites in the COOH-terminal extrapeptide, common to other plant proteins that are targeted to vacuoles of the plant cell, were not observed in PAP-S.

The cDNA deduced sequence of PAP-S identified in our study is very similar to the direct amino acid sequence determination of PAP-S protein isolated from seeds [10]. However, it is not identical (83% identity). The difference could be due to errors in direct protein sequencing as witnessed for ricin [22], to heterogeneity of PAP-S expressed in variants of pokeweed or to different forms of PAP-S in seeds. Interestingly, Barbieri et al. demonstrated the presence of different forms of PAP-S in seeds [6].

In contrast to PAP and PAP II, the amino acid sequence of mature PAP-S contains three potential glycosylation sites at the amino acids residues 10, 44 and 255 (Fig. 2). These sites have been shown to contain a single *N*-acetylglucosamine residue [23]. The biological role of these glycosylations is not known. However, the glycosylations do not seem to be essentiated.

PAP-S	(N-terminal leader)	MKV-ML-VVVVTLIAWLIAAPTSTCA	
PAP	(N-terminal leader)	MKMSML-VVTISIVLILAPTSTDA	
PAP II	(N-terminal leader)	MKMKVLEVVGLAISIWLMLTP-PASS	
		*** ** **	
PAP-S	INTITFDAGNATINKYATFMESL	RNQAKDPKLKCYGIPMLPDTNSTPKYL	50
PAP		RNEAKDPSLKCYGTPMLPNTNTNPKYV	50
PAP II	NIVFDVENATPETYSNFLTSLREAVKDKKLTCHGMIMATTLTEOPKYV		48
	,*,****** .*.*.*. * ***.		-10
PAP-S	TAKLOGANI KTTTI MI DDNNI VA	MGYSDPFNGNKCRYHIFNDITSTERTD	100
PAP		MGYSDPFETNKCRYHIFNDISGTEROD	100
PAP II			92
INI II	LVDLKFGS-GTFTLAIRRGNLYLEGYSDIYNG-KCRYRIFKDSESD 92 **.*.		
PAP-S	VENTI CCCCCDVAM CINVAC	L Y PTMEKKAEVNSRNOVOLGIOILSSD	148
PAP		RYPTLESKAGVKSRSOVOLGIOILDSN	148
PAP II		SYKGMESKGGARTKLGLGKITLKSR	140
PAP II		* * . * * * . * . * . *	140
		^ · · ^ · ^ · · · · ^ · · · ^ · · · ^ · · · ^ ·	
PAP-S	TOVI COMPCEDIV MEXEET	LVAIOMVS E AARFKYIENOVKTNFNRA	194
PAP		LVAIOMVSEAARFKYIENOVKTNFNRA	194
PAP II		-	194
PAP II		LIAVQMVTEASRFKYIENKVKAKFDDA	190
	.^^^	*.*.***.**.*****	
PAP-S	EVENDUITH EDVENTAGE T	HNAKNGALPKPLELVDAKGTKWIV	239
PAP-S PAP I			
PAP II	<u>-</u>	HDAKNGVLPKPLELVDASGAKWIV	239
PAP II	~	AKVGTSGDSTVTLPGDLKDENNKPWTT	240
	. ^.^^	* . * * * *	
D3 D C	I DIDDINDNI ALI WILDIOMOOD		207
PAP-S	 -	TY-QNAMFSQVIISTYYNYMSNLGDLF	287
PAP		TYYQNAMFPQLIMSTYYNYMVNLGDLF	288
PAP II		KSSMFPEIMSYY-YRTSISNLG	282
	*. **** .*	**** **	
PAP-S	EGF 290		
PAP	EGF 291		
PAP II	EFE 285		
	*		

Fig. 2. Alignment of cDNA deduced amino acid sequences of PAP-S, PAP and PAP II. Primary sequences are aligned to show maximum homologies. Identity (*) and similarity () are indicated. Amino acid residues possibly involved in catalytic activity are indicated in bold face and the potential glycosylation sites of PAP-S are underlined.

tial for the inhibitory activity, since PAP-S expressed in *E. coli* demonstrated biological activity similar to the native protein (see below).

The three-dimensional structure of PAP and α -PAP, a protein related to PAP, have recently been elucidated [24,25]. By analogy with PAP, the tyrosine residues at position 72 and 123 of PAP-S are expected to interact with the substrate, whereas glutamic acid at position 176 and arginine at position 179 are involved in the catalytic activity (see Fig. 2). These residues were conserved in all PAP-isoforms studied so far. It can therefore be speculated that the overall structure and mechanism of action of all PAPs is similar. Moreover, a point mutation in the glutamic acid residue of PAP (here at position 176), selected in a yeast system, inactivated the enzymatic function of PAP [26].

3.2. Expression of PAP-S in E. coli

With the help of two primers containing appropriate restriction sites at their 5'-ends, the cDNA of PAP-S without the NH₂-extrapeptide but with the carboxy extension was cloned into the prokaryotic expression vector pQE60. The promoter of pQE60 gives high expression of foreign proteins in *E. coli* after induction by IPTG. The construct was transfected into the *E. coli* strain JM109 and the expression of PAP-S was induced by adding IPTG to the culture medium (Fig. 3). A protein band at approximately 32 kDa was readily visible on SDS-PAGE after induction with IPTG, corresponding to the expected molecular weight of mature PAP-S with the carboxy extension (lane 2, Fig. 3). Moreover, a monoclonal anti-PAP

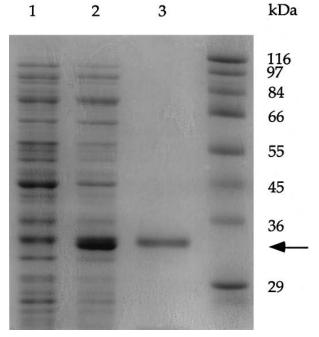


Fig. 3. SDS-Polyacrylamide gel electrophoresis of PAP-S expressed in *E. coli*. Proteins analyzed were from cell containing the expression vector pPAPS with the addition of IPTG (lane 2) or without IPTG (lane 1). Lane 3: PAP-S purified from *E. coli*. The arrow indicates the position of the recombinant protein.

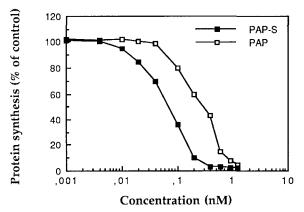


Fig. 4. Inhibition of protein synthesis of recombinant PAP-S (closed boxes) in comparison with purified native PAP (open boxes) from spring leaves in a cell-free translation system from rabbit reticulocytes. The results are expressed as percentage of L-[35S]methionine incorporation relative to that with (100%) and without (0%) template mRNA.

antibody, which was produced in our laboratory against PAP extracted from spring leaves, recognizing an epitope at the C-terminus [27], crossreacted with the recombinant PAP-S (data not shown).

In order to measure and to compare the capacity of PAP-S to inhibit eukaryotic protein translation to that of native PAP, recombinant PAP-S was purified from bacterial cell culture. As can be seen in Fig. 3 (lane 3), a single band was visible after the purification procedure, indicating the homogeneity of the preparation. The final yield of recombinant PAP-S was 5–8 mg/l bacterial cell culture.

3.3. Quantification of protein synthesis inhibition in a cell-free translation system

Tests were carried out to measure the protein synthesis inhibition on a cell-free translation system with various concentrations of recombinant PAP-S or native PAP extracted from the leaves of *Phytolacca americana* (Fig. 4).

In our experiments, PAP-S was more efficient than PAP isolated from spring leaves to inhibit in vitro protein synthesis using a rabbit reticulocyte system. The median inhibitory concentration of L-[³⁵S]methionine incorporation (IC₅₀) of recombinant PAP-S was 0.07 nM, whereas the IC₅₀ value of PAP was determined as 0.29 nM. This is in very good agreement with values given in the literature for PAP-S extracted from seeds (0.04 nM, [6]) and for PAP (0.24 nM, [4]).

It is not known why PAP-S is more efficient at inhibiting protein synthesis than the other members of the PAP family. Recently however, a PAP-S based immunotoxin has been demonstrated to display powerful antitumor activity against CD30+ tumor cells in refractory Hodgkin disease and anaplastic large cell lymphoma [28]. In our studies we clearly show the efficient function of PAP-S as ribosome-inactivating protein in vitro. Furthermore, we were able to produce consistent amounts of purified and highly active material. These results, together with the low aspecific toxicity for intact eukaryotic cells, may make PAP-S more suitable for the construction of hybrid proteins and may facilitate the development of therapeutic applications.

Acknowledgements: We thank P. Millet and his colleagues at the Botanical Garden of the University of Franche-Comté for providing the plant material; Marianne Farrar-Hockley and Marie Chirol for their excellent secretarial assistance.

PAP-S appeared under Accession No. X98079 in the EMBL Data Library.

References

- [1] Y. Endo, K. Mitsui, M. Motizuki, K. Tsurugi, J. Biol. Chem. 262 (1987) 5908–5912.
- [2] Y. Endo, K. Tsurugi, J.M. Lambert, Biochem. Biophys. Res. Commun. 150 (1988) 1032–1036.
- [3] F. Stirpe, L. Barbieri, M.G. Battelli, M. Soris, D.A. Lappi, Biotechnology 10 (1992) 405–412.
- [4] J.D. Irvin, Arch. Biochem. Biophys. 169 (1975) 522-528.
- [5] J.D. Irvin, T. Kelly, J.D. Robertus, Arch. Biochem. Biophys. 200 (1980) 418–425.
- [6] L. Barbieri, G.M. Aron, J.D. Irvin, F. Stirpe, Biochem. J. 203 (1982) 55-59.
- [7] A. Bolognesi, L. Barbieri, A. Abbodanza, A.I. Falasca, D. Caenicelli, M.G. Battelli, F. Stirpe, Biochim. Biophys. Acta 1078 (1990) 293–302.
- [8] Q. Lin, Z.C. Chen, J.F. Antoniw, R.F. White, Plant Mol. Biol. 17 (1991) 609-614.
- [9] J.-L. Poyet, J. Radom, A. Hoeveler, FEBS Lett. 347 (1994) 268– 272.
- [10] S.S. Kung, M. Kimura, G. Funatsu, Agric. Biol. Chem. 54 (1990) 3301–3318.
- [11] L. Foà-Tomasi, G. Campadelli-Fiume, L. Barbieri, F. Stirpe, Arch. Virol. 71 (1982) 323–332.
- [12] M. Olson, S. Ramakrishnan, R. Anand, AIDS Res. Human Retrovir. 7 (1991) 1020–1030.
- [13] Y.W. Kim, M.S.C. Fung, S.C. Sun, C.R.Y. Sun, N.T. Chang, T.W. Chang, J. Immunol. 144 (1990) 1257–1262.
- [14] J.K. Lodge, W.K. Kaniewski, N.E. Tumer, Proc. Natl. Acad. Sci. USA 90 (1993) 7089–7093.
- [15] S. Lee-Huang, P.L. Huang, H.-F. Kung, L. Bao-Qun, P.L. Huang, P. Huang, H.I. Huang, H.C. Chen, Proc. Natl. Acad. Sci. USA 88 (1991) 6570–6574.
- [16] Cashmore, A.R. (1982) Methods in Chloroplast Molecular Biology, Elsevier, Amsterdam, pp. 387–392.
- [17] D.G. Higgins, P.M. Sharp, Gene 73 (1988) 237–244.
- [18] V. Citovsky, D. Knorr, S. Gadi, P. Zambryski, Cell 60 (1990) 637–647.
- [19] J. Radom, J.-L. Schlick, J.-L. Poyet, A. Hoeveler, Plant Sci. 120 (1996) 19–28.
- [20] J. Kataoka, N. Habuka, M. Miyano, Y. Takanami, A. Koiwai, J. Biol. Chem. 266 (1991) 8426–8430.
- [21] R. Carzaniga, L. Sinclair, A.P. Fordham-Skelton, N. Harris, R.R.D. Croy, Planta 194 (1994) 461–470.
- [22] F.I. Lamb, L.M. Roberts, J.M. Lord, Eur. J. Biochem. 148 (1985) 265-270.
- [23] M.R. Islam, S.S. Kung, Y. Kimura, G. Funatsu, Agric. Biol. Chem. 55 (1991) 1375–1381.
- [24] A.F. Monzingo, E.J. Collins, S.R. Ernst, J.D. Irvin, J.D. Robertus, J. Mol. Biol. 233 (1993) 705–715.
- [25] H. Ago, J. Kataoka, H. Tsuge, N. Habuka, E. Inagaki, M. Noma, M. Miyano, Eur. J. Biochem. 225 (1994) 369–374.
- [26] Hur, Y., Hwang, D.J., Zoubenko, O., Coetzer, C., Uckun, F.M. and Tumer, N.E. (1995) Proc. Natl. Acad. Sci. USA 92, 8778– 8452
- [27] B. Desvoyes, P. Dulieu, J.-L. Poyet, P. Adami, Hybridoma 14 (1995) 571–575.
- [28] A. Terenzi, A. Bolognesi, L. Pasqualucci, L. Flenghi, S. Pileri, H. Stein, M. Kadin, B. Bigerna, L. Polito, P.L. Tazzari, M.F. Martelli, F. Stirpe, B. Falini, Br. J. Haematol. 92 (1996) 872–879.