A recombinant ribosome-inactivating protein from the plant *Phytolacca dioica* L. produced from a synthetic gene[1]

Francesca Del Vecchio Blanco[b], Valeria Cafaro[a], Antimo Di Maro[a], Roberta Scognamiglio[a], Gesualdo Siniscalco[c], Augusto Parente[b], Alberto Di Donato[a,*]

[a] *Dipartimento di Chimica Organica e Biologica, Università Federico II di Napoli, Via Mezzocannone, 16, 80134 Naples, Italy*
[b] *Dipartimento di Scienza della Vita, Seconda Università di Napoli, Caserta, Italy*
[c] *Dipartimento di Biologia Vegetale, Università Federico II di Napoli, Via Foria, Naples, Italy*

Received 31 August 1998; received in revised form 15 September 1998

Abstract ***Phytolacca dioica*** **L. leaves produce at least two type-I ribosome-inactivating proteins. Each polypeptide chain is subjected to different post-translational modifications giving rise to PD-L1 and PD-L2, and PD-L3 and PD-L4, each polypeptide pair having the same primary structure. With the aim of exploiting the cytotoxic properties of these proteins as potential biological phytodrugs, a gene encoding PD-L4 was designed based on criteria expected to maximize the translation efficiency in tomato. The gene was constructed from 18 oligonucleotides and preliminarily expressed in** ***Escherichia coli*****, using the T7 promoter system. The protein produced was insoluble and accumulated in inclusion bodies to about 300 mg/l of culture. Ribosome-inactivating activity was generated by controlled oxidation of the reduced and denatured protein. The recombinant protein was indistinguishable from natural PD-L4 as isolated from leaves of** ***Phytolacca dioica*****, in both catalytic activity and primary structure.**

© 1998 Federation of European Biochemical Societies.

Key words: Ribosome-inactivating protein; Protein engineering; Toxin

1. Introduction

Ribosome-inactivating proteins (RIPs) [1] are *N*-β-glycosidases which remove a specific adenine residue of the RNA of the large ribosomal subunit (A^{4324} in the 28S rat rRNA) [2–4]. Two classes of RIPs have been described: type-I RIPs, containing a single chain, enzymatically active, and type-II RIPs, made of two chains, one catalytically active, and similar to type-I RIPs, the other containing a cell binding domain [1]. Thus, whereas type-II RIPs are generally extremely toxic to many cell types, type-I RIPs are less toxic, in the absence of a system to enter the cell. Obviously, also type-I RIPs are highly cytotoxic once inside the cell.

Thus, the isolation of RIPs, especially type-I RIPs, is a potential tool (i) for the development of immunotoxins for tumor therapy [5], (ii) for their use as natural phytodrugs [6,7], and (iii) for the production of transgenic plants endowed with specific parasite resistance [8].

Leaves of *Phytolacca dioica* produce several RIPs. Recently (Di Maro et al., submitted), four new type-I RIPs have been isolated and characterized, and named PD-L1, PD-L2, PD-L3, and PD-L4, in decreasing molecular weight.

PD-L1–3 were found to be extensively glycosylated, whereas PD-L4 did not apparently contain any sugar moiety. Moreover, the total amino acid sequence of PD-L4 has been determined [9] and found to be homologous to PAP, PAPII, and PAP-S [10], and to PAP-C [11].

In the light of a potential use of *P. dioica* RIPs to construct single-chain immunotoxins [12,13] and/or transgenic plants [8], we decided to construct a synthetic gene encoding *P. dioica* PD-L4, to facilitate further studies on its mechanism of action, and to express the recombinant protein in *Escherichia coli.*

The choice of PD-L4, instead of the other RIPs purified from leaves of *P. dioica*, was primarily based on the apparent absence of glycosylation, although it has already been shown that the presence of carbohydrates does not interfere with the catalytic activity and the stability of RIPs [1]. However, this should facilitate its heterologous expression both in prokaryotic and eukaryotic systems.

The choice of synthesizing the gene, instead of isolating a cDNA from a *P. dioica* leaf library, is motivated by the advantages given by a designed gene, which contains as many unique restriction sites as possible, compared to its natural counterpart, and uses a specific codon usage. Thus, based on the characteristics above, a synthetic gene is suitable for modular mutagenesis experiments, and is optimized for heterologous expression into selected host cells.

2. Materials and methods

2.1. Proteins

Native PD-L4 was purified as described (Di Maro et al., submitted) from leaves of *P. dioica* L. harvested at the Orto Botanico of the University of Naples Federico II.

2.2. Materials and general procedures

Plasmid pUC118 and *E. coli* strain JM101 were purchased from Boehringer; expression vector pET22b(+) and *E. coli* strain BL21(DE3) from AMS Biotechnology; labelled oligonucleotides from Amersham; reagents for PCR from Perkin-Elmer Cetus. The Gene-Clean kit for elution of DNA fragments from agarose gel was obtained from Bio 101. Enzymes, including restrictases, and other reagents for DNA manipulation were from Promega Biotech. The oligonucleotides for the synthesis of the cDNA coding PD-L4 were synthesized at the Stazione Zoologica 'A. Dohrn' (Naples, Italy). Bacterial cultures, plasmid purifications and transformations were

*Corresponding author. Fax: (39) (81) 552 1217.
E-mail: didonato@unina.it

[1] This paper is dedicated to the memory of Prof. Giacomino Randazzo.

Abbreviations: RIP, ribosome-inactivating protein; PD-L, *Phytolacca dioica* leaf RIP; PD-S, *Phytolacca dioica* seed RIP; EDTA, ethylenediaminetetraacetic acid; PAGE, polyacrylamide gel electrophoresis; SDS, sodium dodecylsulfate; Tris, tris(hydroxymethyl)aminomethane; DTT, dithiothreitol; HPLC, high pressure liquid chromatography; IPTG, isopropyl-β-D-thiogalactopyranoside; dNTP, deoxynucleotide triphosphate; ATP, adenosine triphosphate

0014-5793/98/$19.00 © 1998 Federation of European Biochemical Societies. All rights reserved.
PII: S0014-5793(98)01240-X

performed according to Sambrook [14]. Single- and double-stranded DNA were sequenced with the dideoxy method of Sanger [15] carried out with a Sequenase Sequencing Kit and labelled nucleotides from Amersham.

2.3. Design of the gene encoding PD-L4

PD-L4 contains 261 amino acid residues [9]. A cDNA encoding the protein was designed from its amino acid sequence; the procedure of Libertini and Di Donato [16] was followed with the aim of: (i) including the maximum number of unique restriction sites, useful for future modular mutagenesis, and (ii) privileging vegetable species preferred codons [17]. The prokaryotic cloning vector (see below), required that the gene had to be inserted between *Eco*RI and *Xba*I sites. The cDNA sequence was thus modified, to allow its polar insertion into the cloning vector, by the addition of extra sequences which could be cleaved by *Eco*RI and *Nde*I (the latter needed for the polar insertion into the prokaryotic expression vector) at the 5′ end, and *Xba*I (at the 3′ end). The above mentioned procedure resulted in the design of an 816 bp cDNA (Fig. 1). The sequence was then divided into 18 oligonucleotides, ranging in length from 55 to 63 bases, for a total of 1075 bases. This number is about 34% lower than the number of bases required by other methodologies of total synthesis based on ligation of complementary oligonucleotides. The overlaps between the oligonucleotide pairs were 14–18 bases, with T_m values of 44–46°C.

2.4. Cell-free protein synthesis inhibition assay

Inhibition of protein synthesis was determined using a cell-free rabbit reticulocyte lysate translation system (Promega, WI). RIPs were mixed in a volume of 15 μl with rabbit reticulocyte lysate (10.5 μl), a mixture of all amino acids (minus methionine) at 20 μM final concentration, 1 μl of 10 mCi/ml [^{35}S]methionine, and luciferase RNA as substrate (0.6 μg). The reaction mixture was incubated at 30°C for 1 h. 2 μl was withdrawn and the reaction stopped by addition of 98 μl of 1 M NaOH containing 2% H_2O_2. The translation product was precipitated with 900 μl of ice-cold 25% trichloroacetic acid and 2% casamino acids as carrier for 30 min on ice. The radiolabeled proteins were collected on glass fiber filters, washed with 5% trichloroacetic acid and then with acetone, dried, and counted using a Beckman LS 1701 scintillation counter. Inhibition of protein synthesis, expressed as the percentage of radiolabeled incorporation into translation products of samples containing RIPs compared to untreated samples, was measured at least in duplicate. The procedure as described allowed the detection of at least 5% inhibition.

IC_{50} values were calculated from semi-logarithmic plots constructed using a non-linear fitting procedure, using the program Regression (Blackwell, Oxford).

2.5. In vitro assay for N-glycosidase activity (aniline assay)

Isolated *Saccharomyces cerevisiae* ribosomes [18] (34 μg) were incubated with the RIP (13 μg) at 30°C for 30 min in a total volume of 22 μl of 1.5 mM Tris-HCl pH 7.5, 1.5 mM Na phosphate pH 7.5, 22.5 mM NaCl, 1.5 mM $MgCl_2$, 1.5 mM EDTA, 75 μM DTT, 6% glycerol. Control reactions lacking RIPs were similarly incubated. ATP, at a final concentration of 1 mM, was added to the assay mixture. After incubation, the samples were diluted with 10 μl of 10% SDS and 170 μl of sterile water; RNA was extracted [19] and lyophilized. Aniline treatment of RNA and electrophoresis in agarose/formamide gels were carried out as described previously [20]. Gels were loaded with identical RNA samples treated and untreated with aniline, stained in ethidium bromide and photographed on a UV transilluminator. *N*-β-Glycosidase activity was evidenced by the presence of a ca. 370 nucleotide fragment released from 26S rRNA [21].

2.6. Other methods

Protein sequence determinations were performed on an Applied Biosystems sequencer model 473A, connected on-line with the HPLC apparatus for identification of phenylthiohydantoin derivatives.

The electrospray mass spectrometric analyses were performed on a VG Bio Q mass spectrometer (VG Analytical).

3. Results and discussion

3.1. Synthesis and cloning of the gene for PD-L4

The procedure was essentially that described by Di Donato et al. [22]. The two primer oligonucleotides **A** and **B** (Fig. 1) were used, at a concentration of 0.1 mM, for the synthesis of a central template. A total of 1/20 of this PCR mixture was used as a template for the subsequent PCR experiments. After each PCR run, 1/20 aliquot of the reaction mixture was withdrawn and added to fresh PCR buffer and dNTPs. New primers (i.e. **C** and **D** in the second PCR round) were then added at 1 mM final concentration, along with 1 unit of *Taq* polymerase, and the new elongation continued for additional 20 cycles. The same procedure was repeated until the most extreme primers, i.e. **S** and **T**, were used.

An aliquot of the purified cDNA obtained by the PCR procedure was digested with *Eco*RI and *Xba*I, and ligated to the cloning vector pUC118, previously digested with the same enzymes. The ligation mixture was used to transform competent JM101 cells, plated onto LB/ampicillin plates. Sequencing of the plasmid DNA of recombinant clones revealed that the inserted sequence was correct and corresponded to the gene as designed.

The sequence coding for PD-L4 was excised from the cloning vector using *Nde*I and *Hin*dIII, and ligated to the expression vector pET22b(+), previously cut with the same enzymes. The ligation mixture was used to transform competent JM101 cells, plated onto LB/ampicillin plates. Several clones were subsequently grown in liquid media (supplemented with 50 μg/ml ampicillin), and their plasmid DNA was sequenced. The inserted sequence was correct, and located, as expected, between *Nde*I and *Hin*dIII sites of the vector. The recombinant plasmid was named pPD4.

3.2. Expression of rPD-L4 from E. coli

Plasmid pPD4 (20 ng) was used to transform competent BL21(DE3) cells, plated onto LB/ampicillin plates. One recombinant clone was grown in 5 ml of TB medium, supplemented with ampicillin, at 37°C up to $OD_{600nm} = 0.7$. This culture was used to inoculate 100 ml of TB supplemented with 50 μg/ml ampicillin, and grown at 37°C until optical reading was 4.0 at 600 nm. Then IPTG was added to 0.4 mM final concentration, and growth continued overnight, at 37°C. An SDS-PAGE analysis of induced and non-induced cells extracted in electrophoresis loading buffer showed that a protein with a molecular mass of about 29 kDa, the expected molecular size of recombinant PD-L4, was produced only in the induced cells (Fig. 2).

3.3. Purification and characterization of recombinant PD-L4

Cells from 100 ml of induced culture were harvested by centrifugation for 10 min at $5520 \times g$, at 4°C, suspended in 10 ml of 50 mM Tris-HCl pH 8.0 containing 1 mM Na_2EDTA pH 8.0 (TE) and 20% sucrose, and kept for 15 min on ice. An equal volume of ice-cold water was added, the sample was returned to ice for 15 min, and then centrifuged at $17\,400 \times g$ for 10 min. The pellet was suspended in 10 ml of TE containing 0.25 mg/ml lysozyme, and kept on ice for 30 min. Cells were disrupted by sonication (10×1 min cycle, on ice). The insoluble fraction was separated from the soluble fraction by centrifugation at $17\,400 \times g$ for 20 min, at 4°C. SDS-PAGE analysis of the soluble and insoluble fractions (data not shown) revealed that the protein of interest was present only in the insoluble fraction, hence presumably contained in cell inclusion bodies. Yields were about 0.3 mg of

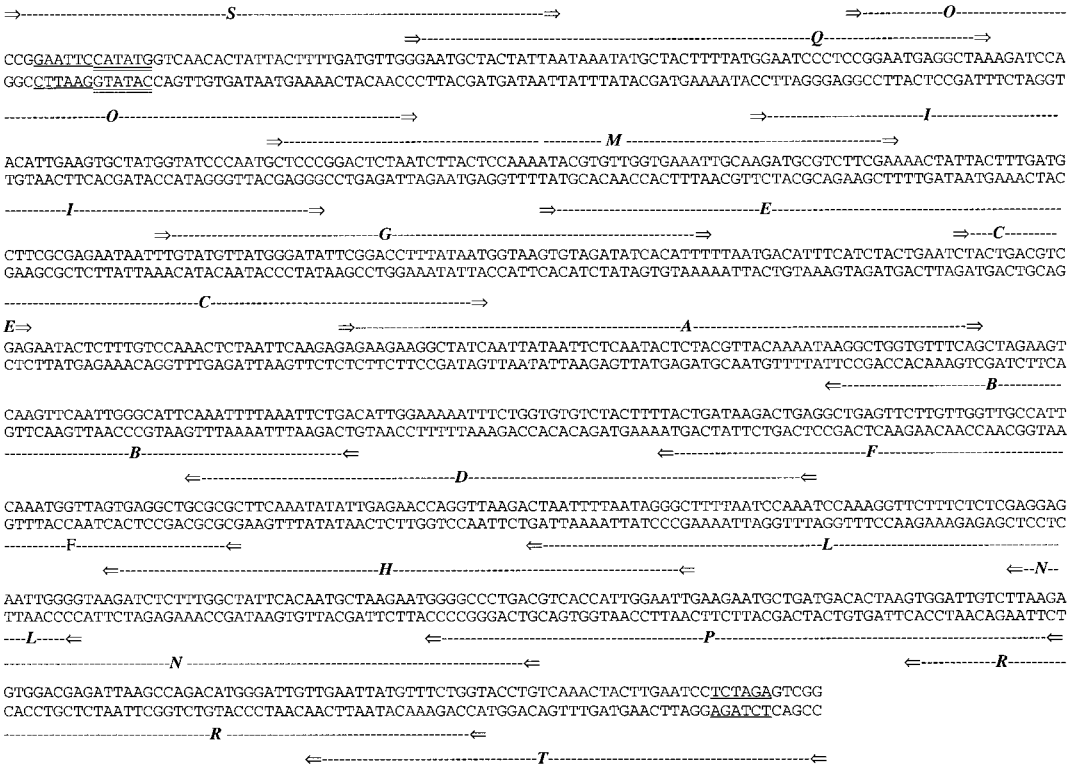

Fig. 1. Sequence of the synthetic gene encoding rPD-L4. The underlined nucleotides indicate the sequences recognized by *Eco*RI and *Nde*I (double underlined), at 5′, and *Hin*dIII, at 3′, respectively. Arrows (⇒) and italic letters indicate the sequence of the oligonucleotides used in the synthesis of the gene.

protein/ml of bacterial culture, on the basis of a densitometric scanning of the electrophoretic profile.

The insoluble fraction after sonication was suspended in 20 ml of TE, containing 2% Triton X-100 and 2 M urea, and sonicated (5×1 min cycle, on ice). Pellet was recovered by centrifugation at 17 400×*g* for 20 min, at 4°C. This procedure was repeated (usually twice) until a clear supernatant was obtained, and repeated two more times after suspending the pellet in TE, without any Triton X-100 and urea.

The insoluble material was dissolved in 2 ml of 0.1 M Tris-HCl, pH 8.0, containing 6 M guanidine-Cl, 5 mM Na_2EDTA, and incubated overnight at 4°C. Total reduction of the proteins present in the mixture was achieved by adding solid reduced glutathione at 0.1 M final concentration, and bringing the pH to 8.4 by small additions of 5 M NaOH. The solution was kept under an N_2 barrier for 2.5 h at room temperature.

Renaturation of the proteins was obtained by dilution of the reduction mixture in the required amount of 10 mM Na phosphate, containing 0.4 M L-arginine, 0.6 mM oxidized glutathione, and 3 mM reduced glutathione, brought to pH 8.4, to a final recombinant protein concentration of 0.07 mg/ml. After protein addition the solution was purged with N_2, and incubated for 24 h at room temperature.

Precipitates formed during incubation were removed by centrifugation at 22 100×*g* for 20 min, at 4°C. On the basis of a densitometric scanning of the electrophoretic profile of an SDS-PAGE analysis of the soluble and insoluble fractions (data not shown) it can be estimated that about 50% of the protein of interest is present in almost pure form in the soluble fraction, whereas the remaining 50% can be found in the precipitate, presumably consisting of misfolded forms of rPD-L4.

The supernatant was concentrated to 10 ml by ultrafiltration using YM-3 membranes, and dialyzed against 5 mM sodium phosphate pH 7.2, at 4°C. This protein preparation was loaded in three consecutive runs on a 1 ml Source 15S (Pharmacia) cation exchange column, equilibrated in 5 mM sodium phosphate pH 7.2. The column was washed, at 1 ml/min, with the same buffer, and eluted by increasing the NaCl concentration (Fig. 3). The eluted protein peaks were collected

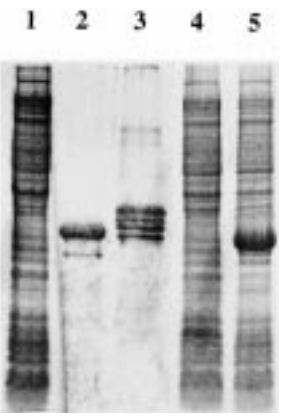

Fig. 2. SDS-PAGE electrophoresis in 15% polyacrylamide gels of induced and non-induced cells carried out under reducing conditions. Cell lysates from non-induced (lane 4) and induced (lane 5) *E. coli* clones harboring plasmid pPD4. In lane 1 non-transformed *E. coli* BL21(DE3) cells, in lane 2 natural PD-S2 (from seeds of *P. dioica*), and in lane 3 a mixture of PD-L1, 2, 3, and 4, are run as a control.

as indicated in Fig. 3, dialyzed and lyophilized. Peak A (the flow-through peak) was found to contain no protein material, as judged by its UV spectrum, and discarded. Peak B (about 10% of peak C) and peak C were analyzed by SDS-PAGE, as shown in the inset of Fig. 3. The electrophoretic analysis revealed that both peaks contained a homogeneous protein, with the same molecular size as PD-L4, under non-reducing conditions. The protein material in peak B was subjected to electrospray mass spectrometry. The results of the analysis showed that the sample contained a heterogeneous mixture of protein species whose molecular mass corresponded to that of natural PD-L4, and to that of natural PD-L4 with an extra methionine, linked to one or two glutathione moieties. These species presumably represented refolding intermediates of the reoxidation procedure, and thus they were not further analyzed.

Sequence analyses of the first 13 residues of the recombinant protein in peak C indicated (data not shown) that two sequences were present, both containing the N-terminal sequence of natural PD-L4. In the first sequence (80%) the N-terminal residue of the protein was methionine, followed by the first 12 residues of PD-L4, whereas in the second sequence (20%), the N-terminal residue was valine, followed by the first 11 residue of natural PD-L4. Thus, the material in peak C is recombinant PD-L4, but it is heterogeneous, containing mainly an unprocessed sequence, with an extra methionine at the N-terminus.

Mature recombinant PD-L4 was obtained by removing the Met(−1) residue using an aminopeptidase from *Aeromonas*

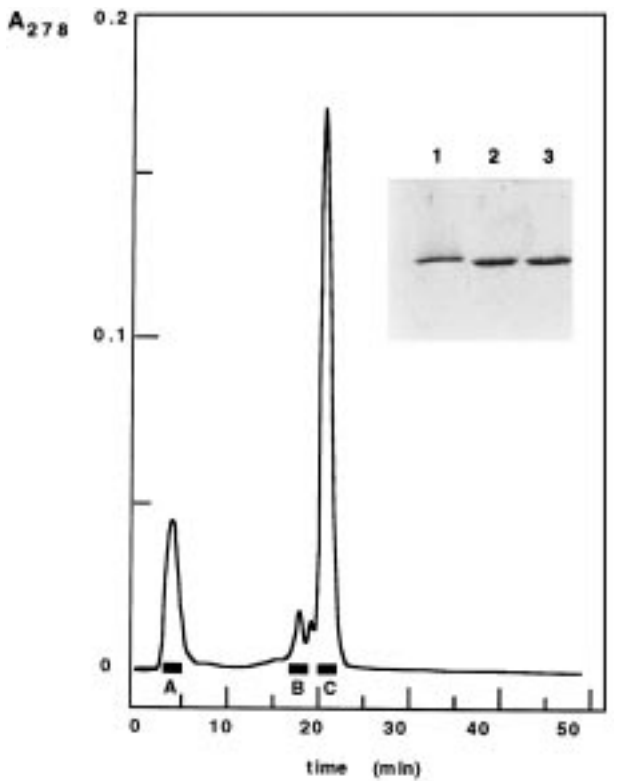

Fig. 3. Ion exchange chromatographic pattern of the reoxidation mixture. The proteins were eluted with a salt gradient from a Source 15S column under the conditions described in the text. In the inset an SDS-PAGE electrophoresis in 15% polyacrylamide gels of the material in peak B (lane 2), and peak C (lane 3) is shown. In lane 1 carbonic anhydrase (MW 29 kDa) was run as a standard.

Table 1
Catalytic activity of recombinant and natural PD-L4

Protein	Aniline assay	IC_{50}[a]
PD-L4	+[b]	151 ± 8 pM
rPD-L4	+	142 ± 7 pM

[a]The IC_{50} value represents the amount of protein required to inhibit 50% of protein synthesis in the rabbit reticulocyte assay. Data were determined by [^{35}S]methionine incorporation into translation products. The IC_{50} values were calculated from two or three experiments. Each data point was measured at least in duplicate.
[b]+ indicates the presence of the ca. 370 nucleotide fragment released from 26S rRNA [21] in the gels run after RIP and aniline treatment, followed by ethidium bromide staining.

proteolytica. Protein material in peak C was dissolved at 13.7 μM final concentration in 5 mM sodium phosphate pH 7.2, containing 10 μM zinc sulfate, and incubated with *Aeromonas* aminopeptidase (13.7 nM) for 72 h at 37°C. The enzyme was then inactivated by addition of EDTA (20 mM), and dialysed against 5 mM sodium phosphate pH 7.2. Sequence analyses after aminopeptidase treatment indicated that more than 96% of N-terminal methionine had been removed.

A sample of mature rPD-L4 was subjected to electrospray mass spectrometry. The average molecular mass of the protein was found to be 29 140.48 ± 1.56 Da, i.e. the mass of natural PD-L4 as calculated from its amino acid composition.

The recombinant protein was also characterized for its enzymic properties, by determining (i) its ability to remove a specific adenine residue of the RNA of the large subunit of *S. cerevisiae* ribosomes in a classical aniline assay, and (ii) the value of its IC_{50} in a cell-free protein synthesis inhibition assay. From the data in Table 1, rPD-L4 was found to be virtually indistinguishable from the natural protein in its activity on intact ribosomes and in its ability to inhibit protein synthesis.

3.4. Concluding remarks

The procedure described for the production of recombinant rPD-L4 is easy and effective, with a yield of about 100 mg of pure recombinant protein per liter of bacterial culture. The recombinant protein is undistinguishable in its catalytic activity from the protein isolated from *P. dioica* leaves. Moreover, the present report has confirmed that the T7 expression system [23] is an effective system for expressing toxic proteins, especially because, as in the present case, recombinant proteins are sequestered into inclusion bodies.

Acknowledgements: This work was carried out under the contract of A.D.D. with Tecnogen S.C.p.A., financed by 'Programma Operativo Plurifondo' 1994–1999 of Regione Campania, Italy, within the framework of the project 'Transgenic tomato plants resistant to CMV and to pathogen fungi'.

References

[1] Barbieri, L., Battelli, M.G. and Stirpe, F. (1993) Biochim. Biophys. Acta 1154, 237–282.
[2] Stirpe, F., Gasperi-Campani, A., Barbieri, L., Falasca, A., Abbondanza, A. and Stevens, W.A. (1983) Biochem. J. 216, 617–625.
[3] Endo, Y., Tsurugi, K. and Lambert, J.M. (1988) Biochem. Biophys. Res. Commun. 150, 1032–1036.
[4] Stirpe, F., Bailey, S., Miller, S.P. and Bodley, J.W. (1988) Nucleic Acids Res. 16, 1349–1357.

[5] Brinkmann, U., Pai, L.H., FitzGerald, D.J., Willingham, M. and Pastan, I. (1991) Proc. Natl. Acad. Sci. USA 88, 8616–8620.
[6] Stirpe, F. (1982) Biochem. J. 202, 279–280.
[7] Stirpe, F., Barbieri, L., Battelli, M.G., Soria, M. and Lappi, D.A. (1992) Biotechnology (NY) 10, 405–412.
[8] Logemann, J., Jach, G., Tommerup, H., Mundy, J. and Schell, J. (1992) Bio/Technology 10, 305–308.
[9] Del Vecchio Blanco, F. (1996) Doctoral Thesis.
[10] Bjorn, M.G., Larrick, J., Piatak, M. and Wilson, K.J. (1984) Biochim. Biophys. Acta 790, 154–163.
[11] Barbieri, L. et al. (1989) Biochem. J. 257, 801–807.
[12] Friedman, P.N., Chace, D.F., Trail, P.A. and Siegall, C.B. (1993) J. Immunol. 150, 3054–3061.
[13] Brinkmann, U. and Pastan, I. (1994) Biochim. Biophys. Acta 1198, 27–45.
[14] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[15] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 76, 5653–5667.
[16] Libertini, G. and Di Donato, A. (1992) Protein Eng. 5, 821–825.
[17] Wada, K., Wada, Y., Doi, H., Ishibashi, F., Gojobori, T. and Ikemura, T. (1991) Nucleid Acids Res. 9, 1981–1986.
[18] Massiah, A.J. and Hartley, M.R. (1995) Plant 197, 633–640.
[19] Hartley, M.R., Wheeler, A. and Ellis, R.J. (1975) J. Mol. Biol. 91, 67–77.
[20] May, M.J., Hartley, M.R., Roberts, L.M., Krieg, P.A., Osborn, R.W. and Lord, J.M. (1989) EMBO J. 8, 301–308.
[21] Endo, Y. and Tsurugi, K. (1987) J. Biol. Chem. 262, 8128–8130.
[22] Di Donato, A., de Nigris, M., Russo, N., Di Biase, S. and D'Alessio, G. (1993) Anal. Biochem. 212, 291–293.
[23] Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) Methods Enzymol. 185, 60–89.