Volume 291, number 2, 285–288 FEBS 10274 October 1991

ADONIS 001457939100984D

A Saporin-6 cDNA containing a precursor sequence coding for a carboxyl-terminal extension

Luca Benatti[2], Gianpaolo Nitti[2], Michela Solinas[1], Barbara Valsasina[2], Alessandro Vitale[3], Aldo Ceriotti[3] and Marco R. Soria[1]

[1]*Laboratories for Biotechnology Transfer, Department of Biotechnology, San Raffaele Research Institute, Milano, Italy,* [2]*Department of Biotechnology, Farmitalia Carlo Erba, Milano, Italy* and [3]*Istituto Biosintesi Vegetali, Consiglio Nazionale delle Ricerche, Milano, Italy*

Received 29 July 1991

Saporin-6 is a single-chain ribosome inactivating protein (RIP) from the seeds and the leaves of *Saponaria officinalis* (Caryophyllaceae) Here we have identified the COOH-terminal end of mature Saporin-6 and, by cDNA sequencing, the predicted carboxyl-terminal sequence of a leaf Saporin-6 primary translation product Our data indicate that the characterized cDNA codes for a precursor containg a 22 amino acid carboxyl-terminal extension not present in mature Saporin-6, that shows similarity to carboxyl-terminal propeptides of vacuolar proteins, suggesting that it may be involved in protein trafficking

Carboxyl-terminal propeptide, Plant toxin, Polymerase chain reaction, Ribosome inactivating protein, *Saponaria officinalis*

1 INTRODUCTION

Single chain RIPs occur throughout the plant kingdom and are distinct from RIPs consisting of two nonidentical subunits (A and B chains) joined by a disulfide bond [1] Several Type 1 RIPs were isolated from the seeds and the leaves of *Saponaria officinalis*, a Caryophyllacea [2] Among these, Saporin-6 (also called SO-6) [3] has yielded promising results as a candidate partner for the synthesis of immunotoxins and of ligand–toxin conjugates [4] Recently, efforts by our group aimed at characterizing the genetic organization of Saporin-6 yielded the coding sequence and deduced amino acid sequence of clones from a leaf cDNA library [5]. However, the cDNA clones coding for Saporin-6 isolated in our laboratory did not contain a translation termination codon Since treatment of purified Saporin-6 with CNBr, pepsin or clostripain yielded putative carboxyl-terminal fragments all terminating at ...Lys-Pro-Lys, we postulated that mature Saporin-6 might derive from a longer precursor cDNA extending beyond the 3′ end of our cDNA clones. The gene product of the longer precursor would then be processed post-translationally by removal of a COOH-terminal extension, as was recently described in other plant proteins [6,7] including another Type 1 RIP, trichosanthin [8] Here we have positively identified . Pro-Lys as the COOH-terminal end of mature seed Saporin-6 We then identified the DNA sequence at the 3′ end of leaf cDNA by amplifying cDNA from the leaves of *Saponaria officinalis* using polymerase chain reaction techniques We thus determined that the precursor form of leaf Saporin-6 has a 22 amino acid carboxyl-terminal propeptide extension showing similarity to putative vacuolar targeting signals

Correspondence address M R Soria, Laboratories for Biotechnology Transfer, Department of Biotechnology San Raffaele Research Institute, Via Olgettina 60, 20132 Milano, Italy Fax (39) (2) 217 02640

2 MATERIALS AND METHODS

Saporin-6 was purified from the seeds of *Saponaria officinalis* as previously described [3] Protein concentration and amino acid composition were determined by phenylthiocarbamyl amino acid (PTC-AA) analysis on an Applied Biosystems (Foster City, CA) automatic system (Models 420/130/920) Vapor-phase hydrolysis with hydrochloric acid was performed at 150°C for 90 min using a Waters (Billerica, MA) Pico-Tag Workstation For hydrazinolysis, samples of Saporin-6 (1–2 nmol) were treated with anhydrous hydrazine at 90°C for 16 h in parallel with other proteins (cytochrome *c*, lysozyme, *β*-lactoglobulin as controls) using the Waters Pico-Tag Workstation Prior to amino acid analysis, hydrazides were treated with benzaldehyde and their adducts removed by ether extraction, then the free amino acid obtained from the COOH-terminus was identified as PTC-AA [9] For carboxypeptidase P (CPP) digestions, 3 nmol of Saporin-6 were denatured at 100°C for 10 min and incubated with CPP at 37°C for 2 h in the presence of 4 M urea in 10 mM sodium acetate, pH 3 8, 0 05% Brij-35 Aliquots withdrawn from the digestion mixture at several incubation times were diluted 1 2 in 10% TFA, vacuum dried, redissolved in 250 ppm Na_4-LDTA and then subjected to PTC-AA analysis

Oligonucleotides were synthesized in a 380B automatic DNA synthesizer (Applied Biosystems). The primers, designed with a *Xho*I-*Sal*I-*Cla*I linker, were oligo-'sense' (5′-GAC-TCG-AGT-CGA-CAT-CGA-TCA-GAT-GAC-GGC-TGA-GGC-AGC-3′), which begins at +513 of the leaf Saporin-6 coding sequence [5] and oligo-'antisense' (5′-GAC-TCG-AGT-CGA-CAT-CGA-TTT-TTT-TTT-TTT-TTT-TT-3′), complementary to the poly(A) tail Reverse transcription and

PCR amplification of the 3′ end of Saporin-6 cDNA from summer leaves of *Saponaria officinalis* [5] was performed as described [10]

3. RESULTS

3.1. *Characterization of the COOH-terminal end of Saporin-6*

Saporin-6 has been found to be remarkably resistant to several proteases commonly employed for structural characterization. Positive identification of the carboxyl-terminal residues was thus hindered by the resistance of Saporin-6 to carboxypeptidase digestion. Treatment with carboxypeptidases Y, A and B, even under strongly denaturing conditions using high enzyme to substrate ratios, did not yield any meaningful result [2,11].

Therefore, appropriate conditions for dealing with such an unusual situation were developed. Initially, the carboxyl-terminal residue was identified as Lys by hydrazinolysis (Fig. 1A). Then, timed digestion experiments with carboxypeptidase P (CPP) were undertaken. This carboxypeptidase exhibits a broader specificity and ability to release almost all COOH-terminal amino acids, is devoid of endopeptidase activity and is active in acidic buffers [12]. Time-course hydrolysis with CPP in strongly denaturing conditions identified -Pro-Lys as the COOH-terminal sequence of mature Saporin-6 (Fig. 1B). boiling the protein for 10 min was required to obtain release of the COOH-terminal residues. Omitting the denaturation step, very slow release was obtained under otherwise identical conditions. By comparison, the same CPP treatment without thermal denaturation led to very effective release of the COOH-terminal amino acids of commercially available proteins (e.g. bradykinin, ubiquitin, lysozyme) used as controls (data not shown).

3.2. *Characterization of a PCR-amplified clone coding for the 3′ end of leaf Saporin-6*

The cDNA clones coding for Saporin-6 isolated in our laboratory [5] were interrupted by the presence of an *Eco*RI site on the coding sequence and thus did not contain a translation termination codon. cDNA amplification experiments were thus performed on the 3′ end of mRNA from leaf tissue of *Saponaria officinalis* using the polymerase chain reaction (PCR) technique. A 5′ end of oligonucleotide primer was synthesized corresponding to the middle of the coding sequence of our cDNA clones, and an oligo(dT)-containing primer was used to amplify the 3′ end, as described in Materials and Methods. After amplification, the amplified DNA band was eluted from the gel and subcloned in *Sal*I-restricted M13mp18. The single-stranded DNA insert from white plaques was sequenced using the universal primer of M13. The DNA sequence and deduced amino acid sequence of the portion of this clone that follows the carboxyl-terminal Lys of mature Saporin-6 is depicted

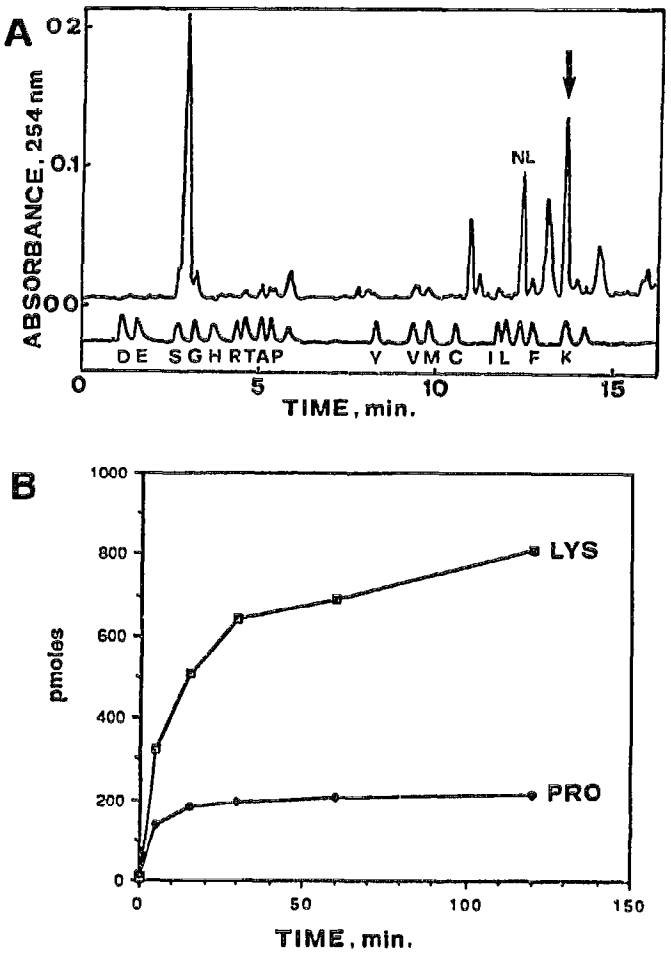

Fig. 1. COOH-terminal analysis of Saporin-6. (A) HPLC analysis of a hydrazinolysate of 1.5 nmol of Saporin-6 after derivatization with phenylisothiocyanate (PITC), compared to standard PTC-AA (lower trace). Arrow indicates the identified COOH-terminal amino acid (Lys), several unidentified peaks present in the chromatogram were detected also in a blank hydrazinolysate. Norleucine (NL) was used as internal standard. (B) Kinetic plot of the amino acids released by carboxypeptidase P digestion of 3.0 nmol of Saporin-6. Automated amino acid analysis of 1/10 aliquots withdrawn from the digestion mixture at different incubation times (as described in Materials and Methods) was performed after derivatization with PITC.

in Fig. 2a. The 5′ end of the clone corresponded to the sequence of our leaf cDNA clones [5] followed by codons for a 22-amino acid extension after the carboxyl-terminal codons of mature Saporin-6, and a translation termination codon (TGA). The latter was followed by a 129 nucleotide long 3′ untranslated sequence containing a hexanucleotide consensus sequence (AATAAA) for polyadenylation (not shown). Thus, leaf Saporin-6 appears to be synthesized as a prepro-protein composed of a 23 amino acid signal sequence, a 253 amino acid mature protein, and a 22-amino acid COOH-terminal extension (Fig. 2b). The deduced sequence of the COOH-terminal extension contains the tripeptide Asn-Ser-Thr, which is a putative site for *N*-glycosylation. No potential glycosylation sites (Asn-Xaa-Ser/Thr) are present in mature Saporin-6

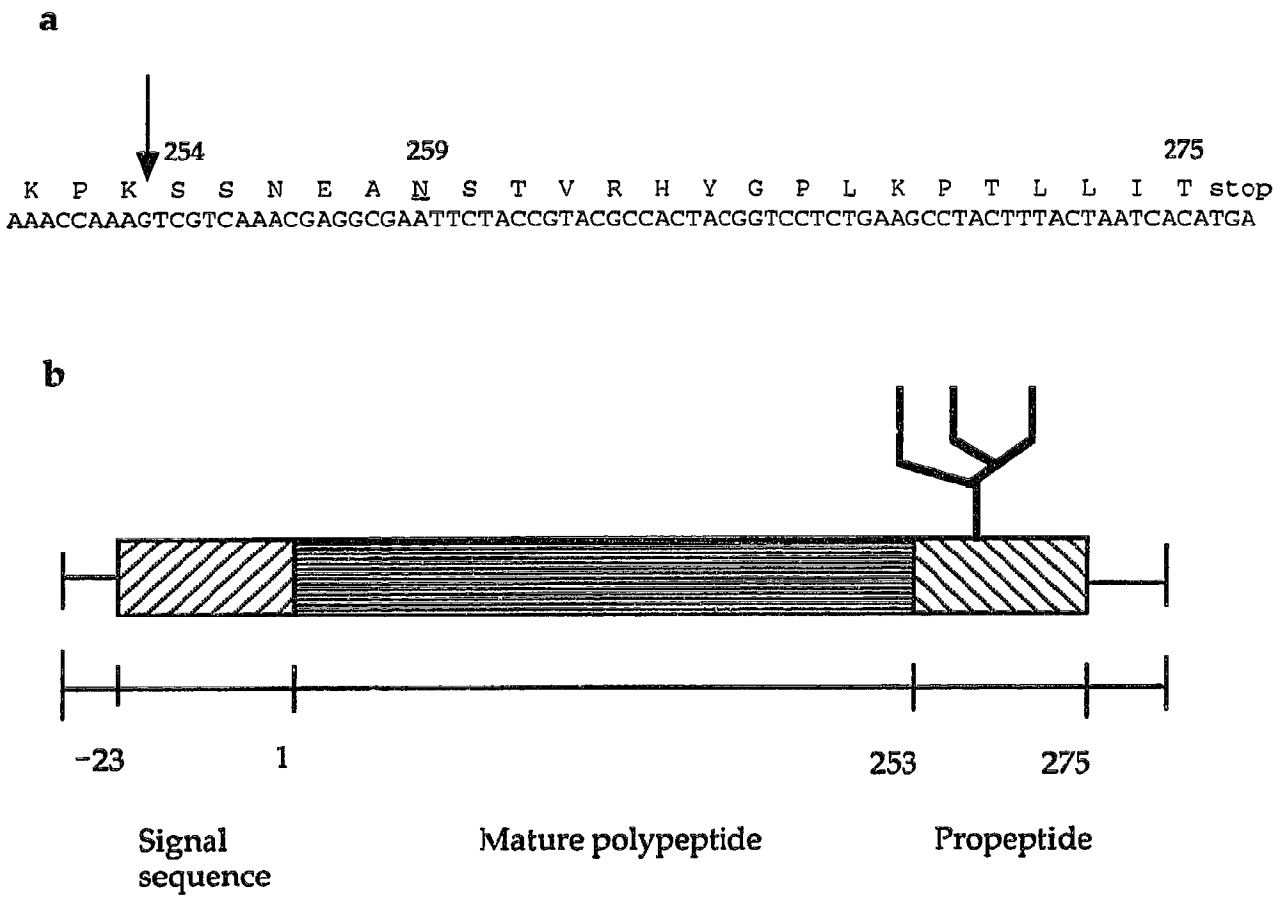

Fig 2 The 3′ end of Saporin-6 cDNA encodes a carboxyl-terminal extension with an N-linked glycosylation site (a) The 22-amino acid COOH-terminal extension of Saporin and the corresponding nucleotide sequence The arrow indicates the end of mature Saporin-6 The potentially glycosylated asparagine residue (Asn-259) is underlined (b) Structure of saporin cDNA The putative N-linked glycosylation site (Asn-259–Ser–Thr-261) is depicted by attachment of a glycan tree to the COOH-terminal extension (c) Comparison of the COOH-terminal propeptides of Saporin, tobacco glucanase and barley lectin Conserved positions around the *N*-linked glycosylation signal are underlined

4 DISCUSSION

In this paper, we have determined the carboxyl-terminal end of mature Saporin-6 and, by sequencing of a leaf cDNA clone, the complete carboxyl-terminal sequence of the Saporin-6 primary translation product The sequence deduced from the cDNA clone extends for 22 amino acids past the end of the mature protein This extension contains a potential glycosylation site. The actual glycosylation of this site still needs to be demonstrated· not all potential glycosylation sites are glycosylated in vivo, because often the conformation of the growing polypeptide hinders access to oligosaccharide transferase [13].

A class of vacuolar proteins, including the basic forms of tobacco β-1,3-glucanase, a plant defense-related enzyme [6], and barley lectin [7], are initially synthesized as glycosylated precursors and subsequently processed to mature, nonglycosylated proteins by the post-translational cleavage of a COOH-terminal glyco-

peptide. The propeptide of barley lectin has been demonstrated to be necessary for targeting of the protein to the vacuole [14]. There is no evidence that the COOH-terminal propeptide of β-glucanase or of other vacuolar proteins contain information for vacuolar targeting However, cDNA and protein sequencing revealed that a group of tobacco β-glucanases that are secreted extracellularly, instead of being targeted to the vacuole, do not possess a COOH-terminal prosequence Instead, they end exactly at the position where the splicing point of the vacuolar counterparts is located [15,16] This led the authors to propose that also the prosequence of vacuolar β-glucanase may contain a vacuolar targeting signal In Fig. 2c we have compared the carboxyl-terminal extension of Saporin-6 to the propeptides of barley lectin and vacuolar β-glucanase. Evident primary sequence similarity between Saporin-6 and barley lectin extends only to an alanine just ahead of the common glycosylation site and is probably not significant However, the similarity between the propeptides of saporin and glucanase extends over a region of 8 amino acids in the area around the glycosylation site Five positions are conserved, while three substitutions take place between uncharged amino acids. Although more data need to be acquired, this similarity may be significant in terms of function, and we believe that the possible involvement of the carboxyl-terminal extension of Saporin-6 as a subcellular location determinant should be taken into consideration.

Certain Type 1 RIPs have been found localized extracellularly in the cell wall matrix of the leaf mesophyll cells [17,18]. The location of other Type 1 RIPs is still unknown. Among the latter are Saporin-6, Dianthin, from *Dianthus caryophyllus*, which belongs to the Caryophyllaceae family like *Saponaria officinalis*, and Trichosanthin, from *Trichosanthes kirilowii*. It has been recently determined through cDNA sequencing that also a Dianthin precursor carries a potentially glycosylated carboxyl-terminal propeptide (G Legname and M Lord, personal communication)

Our data suggest, although do not prove, that the 22-amino acid extension is a propeptide which is removed post-translationally to yield mature Saporin-6 whose carboxyl-terminal end we have determined. That proteolytic processing events involve the carboxyl-terminus of seed Saporin-6 is consistent with results of pulse-chase experiments performed on developing seeds of *Saponaria officinalis*. Saporin-6 was immunoprecipitated from pulse-labelled seeds, yielding an SDS-PAGE pattern consiting of at least four polypeptides in the 30-35 kDa range [11] However, seeds that had been labelled and then chased in cold medium to allow processing of the newly synthesized proteins did not yield the slower-migrating bands. A corresponding increase in intensity of the faster migrating polypeptides was instead observed, confirming the existence of a precursor-product relationship Since mature Saporin-6 is not glycosylated, it is likely that this change reflected the proteolytical processing of the polypeptide chain with removal of a glycosylated propeptide and a reduction in mol wt of about 4000 Da However, we have used leaf mRNA to produce the cDNA clone characterized here, whereas the characterized Saporin-6 protein is obtained from seeds Efforts to obtain mRNA from seeds of *Saponaria officinalis* are currently in progress Thus, it could be that the primary translation products of the leaf and of the seed proteins differ in possessing or not the same carboxyl-terminal extension as that described here If this were the case, it could reflect a different subcellular location of different Saporin-6-polypeptides

Acknowledgements We thank M J Chrispeels for critical reading of the manuscript and for stimulating discussions This work was supported in part by the Progetto Finalizzato Biotecnologie e Biostrumentazione of the Consiglio Nazionale delle Ricerche and by the Ministero della Sanita – Istituto Superiore di Sanita-IV Progetto AIDS 1991

REFERENCES

[1] Stirpe F and Barbieri, L (1986) FEBS Lett 195, 1–8
[2] Stirpe, F , Gasperi-Campani, A , Barbieri, L , Falasca, A , Abbondanza, A and Stevens, W A (1983) Biochem J 216, 617–625
[3] Lapi, D A , Esch, F , Barbieri, L , Stirpe, F and Soria, M (1985) Biochem Biophys Res Commun 129, 934–942
[4] Soria, M (1989) Pharmacol Res 21, 35–46
[5] Benatti, L , Saccardo, B , Dani, M , Nitti, G , Sassano, M , Lorenzetti R , Lappi, D A and Soria, M (1989) Eur J Biochem 183, 465–470
[6] Shinshi, H , Wenzler, H , Neuhaus, J M , Felix, G , Hofsteenge, J and Meins Jr , F (1985) Proc Natl Acad Sci USA 85, 5541–5545
[7] Wilkins, T A , Bednarek, S Y and Raikhel, N V (1990) Plant Cell 2, 301–313
[8] Chow, T P , Feldman, R A , Lovett, M and Piatak, M (1990) J Biol Chem 265, 8670–8674
[9] Strydom, D J (1991) Anal Biochem 174, 679–686
[10] Wang, A M , Doyle, M V and Mark, D F (1989) Proc Natl Acad Sci USA 86, 9717–9721
[11] Soria, M , Benatti, L , Lorenzetti, R , Ceriotti, A , Solinas, M , Nitti, G and Lappi, D A (1991) in Studies on Ribosome Inactivating Proteins from *Saponaria officinalis*, Genetically Engineered Toxins (Frankel, A ed) Marcel Dekker, New York, in press
[12] Lu, S H , Klein, M L and Lai, P H (1988) J Chromatogr 447, 351–364
[13] Kornfeld, R and Kornfeld, S (1985) Annu Rev Biochem 54, 631–664
[14] Bednarek, S Y , Wilkins, T A , Dombrowski, J E and Raikhel, N V (1990) Plant Cell 2, 1145–1155
[15] Van der Bulcke, M , Bauw, G , Castresana, C , Van Montagu, M and Vandekerckhove, J (1989) Proc Natl Acad Sci USA 86, 2673–2677
[16] Payne, G , Ward, E , Gaffney, T , Ahl Goy, P , Moyer, M , Harper, A , Meins Jr , F and Ryals, J (1990) Plant Mol Biol 15, 797–808
[17] Ready, M P , Brown, D T and Robertus, J D (1986) Proc Natl Acad Sci USA 83, 5053–5056
[18] Frotschl, R , Schonfelder, M , Mundry, K W and Adam, G (1990) 8th Int. Cong Virol , Berlin (Abstr)