FEBS Letters 385 (1996) 154–158 FEBS 16986

# Modified 2S albumins with improved tryptophan content are correctly expressed in transgenic tobacco plants

Lucilia H. Marcellino[a], Goran Neshich[a], Maria F. Grossi de Sá[a], Enno Krebbers[b,**], Eugen S. Gander[a,*]

[a]*EMBRAPA-CENARGEN S.A.I.N., Parque Rural, 70770-930 Brasília DF, Brazil*
[b]*Plant Genetic Systems, J. Plateaustraat 22, B-9000 Gent, Belgium*

Received 29 January 1996; revised version received 2 April 1996

**Abstract Brazil nut 2S albumins lack the essential amino acid tryptophan. In order to improve the protein's nutritional value and create a basis for structural investigations, three separate modified Brazil nut 2S albumin genes were constructed. The first mutant contains five consecutive tryptophan codons, while the other two modified genes encode proteins carrying single tryptophan residues at sites that will allow confirmation of the predicted protein structure through fluorescence quenching techniques. The modified genes, under the regulation of the CaMV 35S promoter, were introduced into *Nicotiana tabacum*. All three modified genes were correctly transcribed and the 2S albumin accumulated in the seeds of transgenic plants.**

*Key words:* 2S storage albumin; *Bertholletia excelsa*; Site-directed mutagenesis; Tryptophan enhancement; Transgenic plant

## 1. Introduction

Brazil nut (*Bertholletia excelsa*) 2S albumins are storage seed proteins rich in the sulfur amino acid methionine [1]. These proteins have been extensively studied since they may be useful for the improvement of the nutritional value of crops. These studies have concentrated on gene characterization [2-4] and expression of Brazil nut 2S albumins in transgenic plants [5-8]. In addition, studies on 2S gene promoter structure [9] and post-translational processing [10] have been performed. However, studies of seed proteins at the structural level, in particular for the 2S albumins, are scarce. Such studies would make possible the evaluation of possible functions of the protein and the definition of regions of the protein which might tolerate changes with minimal biological effect.

A potential approach to determine elements of 3D structure is the use of fluorescence quenching techniques [11]. This method allows the detection of differential quenching of fluorescent light emission from tryptophan residues with distinct solvent accessibility. Since Brazil nut 2S proteins have no tryptophan residues, it would be necessary to introduce them in order to take advantage of this technique. In addition, to allow structural analyses, such a modification would improve the protein's nutritional value since tryptophan is one of the nutritionally essential amino acids.

We have modified a Brazil nut 2S albumin cDNA at three different sites in order to achieve these goals. The modified genes were cloned into a binary plant vector under the control of the CaMV 35S promoter and introduced into *Nicotiana tabacum* via *Agrobacterium tumefaciens*. We show that the wild type as well as the mutated sequences are transcribed in the host plants. The modified proteins were correctly expressed in seeds as judged from their molecular weight and were correctly targeted to protein bodies.

*Corresponding author. Fax: (55) (61) 340 3624.
E-mail: egander@cenargen.embrapa.br

***Present address:* Dupont Agricultural Biotechnology, Experimental Station 402-2253, Wilmington, DE 19880-0402, USA.

## 2. Materials and methods

### 2.1. Choice of the mutagenesis sites

The sites for the introduction of tryptophan residues were chosen based on a predicted theoretical structure obtained through the use of several algorithms: Chou-Fasman, Gor III, Segment 83 and by consulting the HSSP database [12,13] (Fig. 1). In one case a leucine at position 44 facing the interior of the large subunit and in the second case an arginine which, according to the model, faces the outside of the 2S subunit, were replaced by tryptophan residues. A third modification, 5 tryptophan residues in a row between positions 57 and 58, was introduced in the variable region, the least conserved region among 2S albumins from different species.

### 2.2. Site-directed mutagenesis

The starting material was an incomplete cDNA (pBN2S1) lacking part of the sequence encoding the signal peptide [6]. In order to obtain the complete sequence, a 334 bp *Eco*RI/*Bgl*I fragment corresponding to the 5'-end of the Brazil nut 2S albumin gene, containing the complete signal sequences, was isolated from a genomic clone (EMBL GenBank, accession no. X54490). This fragment was ligated to the *Bgl*I/*Eco*RI fragment of the 2S albumin cDNA isolated from pBN2S1. The resulting *Eco*RI fragment was cloned into the *Eco*RI site of pLM 1 (a modified form of pGEM 3Z$^+$ lacking the *Ava*I and *Pst*I sites). This plasmid was called pLM 2, which, after destruction of its 3' *Eco*RI site, became plasmid pLM 6 (Fig. 2), the starting material for all subsequent manipulations.

*2.2.1. Substitution of Leu$^{44}$.* The strategy employed is depicted in Fig. 3A. A 517 bp *Eco*RI/*Pvu*II fragment, spanning the 5'-region down to one triplet before the Leu$^{44}$ codon, was isolated. This fragment was ligated to a 60 bp synthetic double-strand polynucleotide, starting at its 5'-end with the Trp codon TGG instead of CTG (Leu), and extending to the unique *Pst*I site in the gene's coding region. The resulting 577 bp fragment was ligated to the large *Eco*RI/*Pst*I fragment from pLM 6 resulting in plasmid pLM 7. Screening was carried out on the basis of the *Pvu*II site so destroyed.

*2.2.2. Substitution of Arg$^{80}$.* This mutation was made in two steps (Fig. 3B): first a double-stranded synthetic fragment carrying the *Ava*I site at its 5'-end and terminating blunt at the ATG preceding the Arg$^{80}$ AGG codon was synthesized. In the subsequent step a fragment in which the Arg$^{80}$ codon (AGG) was replaced by the Trp codon TGG and extending to the *Bam*HI site was produced via PCR. The ligated fragments were cloned into the large *Bam*HI/*Ava*I fragment of pLM 6. The resulting plasmid, carrying the mutated cDNA (Arg$^{80}$ → Trp$^{80}$), was called pLM 8.

*2.2.3. Insertion of 5 Trp codons between positions 57 and 58 in the variable region of the large subunit.* A 60bp DNA fragment corresponding to the *Pst*I/*Ava*I fragment of pLM 6 containing five Trp

0014-5793/96/$12.00 © 1996 Federation of European Biochemical Societies. All rights reserved.
*PII* S0014-5793(96)00375-4

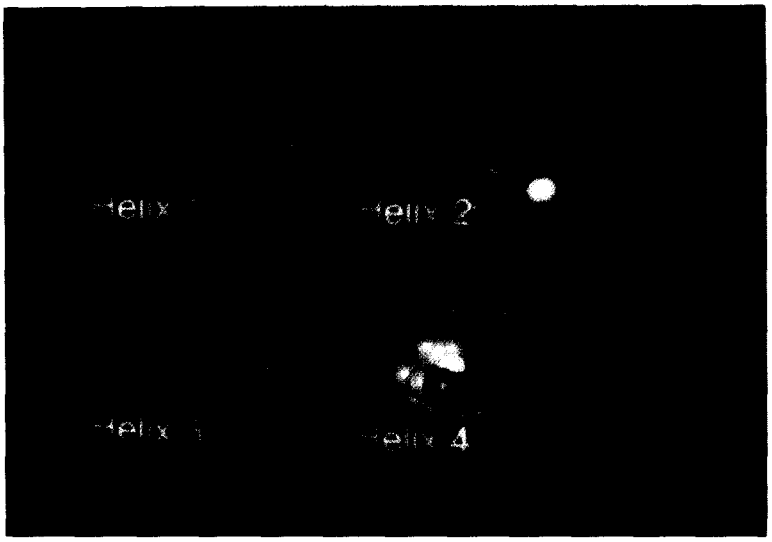


Fig. 1. Theoretical structure of the 2S albumin. The photo shows 4 of the 5 helices presumably present in 2S Brazil nut albumin. Note that $Leu^{44}$ faces the interior of the molecule while $Arg^{80}$ faces outwards – both amino acids are represented in ball model (image produced on SGI ONYX VTX, using the Insight II software).

codons between $Arg^{57}$ and $Met^{58}$ was synthesized (Fig. 3C). This fragment was used to replace the wild type *PstI/AvaI* fragment in pLM 6. Since the mutants were 15 bp longer than the original fragment, the mutated clones, coined pLM 9, were identified by sizing on 6% polyacrylamide gels. All mutations were verified through sequencing.

*2.3. Construction of binary vectors*

The modified gene fragments, as well as a control, were inserted into pMV 25 (Plant Genetic Systems, Gent, Belgium) between the *NcoI* and *XbaI* site. pMV 25 contained the 35S promoter and 3′ OCS cassette flanking those restriction sites. In each case this was done in two steps. First, from each plasmid (pLM 6–9) the 5′ *NcoI/PstI* (278 bp) and the 3′ *PstI/XbaI* fragments were isolated and ligated to each other. The resulting *NcoI/XbaI* fragment was inserted between the *NcoI* and *XbaI* restriction sites of pMV 25. The resulting plasmids were designated pLM 10 (control), pLM 11 ($Leu^{44} \rightarrow Trp^{44}$), pLM 12 ($Arg^{80} \rightarrow Trp^{80}$) and pLM 13 (+five Trp codons). The four resulting 35S-2S albumin-3′ ocs cassettes were isolated with *Eco*RI and *Hind*III and cloned into the *Eco*RI and *Hind*III sites of pGSV 6, an *Agrobacterium* binary vector derived from plasmids described by Deblaere et al. [14]. The resulting plasmids were designated pTLM 14–17, respectively.

*2.4. Transformation of N. tabacum*

pTLM 14–17 were introduced into *N. tabacum* (cv. SR1) as described by Horsh et al. [15]. Kanamycin resistant plants were selected, transferred to the greenhouse and allowed to produce seeds which were used for further analysis.

*2.5. RNA extraction and Northern blot*

Total RNA from leaves was extracted using the method described by Jones et al. [16]. Northern blots were performed according to a standard procedure [17]. The blots were probed using the entire, radiolabelled pLM 10 plasmid.

*2.6. ELISA*

ELISA tests, modified after Voller et al. [18], were performed with 50 μl of total protein extract (1 μg/μl). Proteins were extracted from mature tobacco seeds in PBS (150 mM NaCl; 2.5 mM KCl; 10 mM $Na_2HPO4$; 1.7 mM $KH_2PO_4$; pH 7.4) and 0.5 mM PMSF. The amount of protein in each sample was quantified using the BioRad assay. Brazil nut 2S albumin adsorption was determined using polyclonal mouse antibodies against Brazil nut 2S albumin, alkaline phosphatase coupled to mouse anti-IgG (Promega) and *p*-nitrophenyl phosphate as a chromogenic substrate. The samples were read in a 'Metertech Σ 960' Microplate reader at 405 nm.

*2.7. Immunoblot*

Immunoblot analyses were carried out according to Karey and Sirbasku [19]. Polypeptides were transferred onto a Zeta probe membrane (Bio-Rad) and probed with the Brazil nut 2S albumin specific polyclonal mouse antibodies. The bands were visualized with alkaline phosphatase coupled to mouse anti-IgG (Promega), BCIP (5-bromo-4-chloro-3-indolyl phosphate *p*-toluidine salt) and NBT (*p*-nitrotetrazolium chloride). Quantification of 2S albumins was performed by comparing the band intensities with that of a purified 2S Brazil nut albumin standard of known concentration using MacIntosh 'Image' software.

*2.8. Immunocytochemical localization of 2s proteins*

Seeds from plants transformed with pTLM 16 were analyzed by in situ immunocytochemistry according to Grossi de Sá et al. [20].

## 3. Results and discussion

*3.1. Theoretical 2S albumin structure and choice of mutagenesis sites*

The goal of our work was twofold, to improve the nutritional value of the Brazil nut 2S proteins and to create a basis for future modeling work. Brazil nut 2S storage albumins, while rich in methionine, are deficient in tryptophan which is an essential amino acid in human and animal diets. Furthermore, the fluorescence properties of this amino acid are useful in protein structure determination. To select possible positions

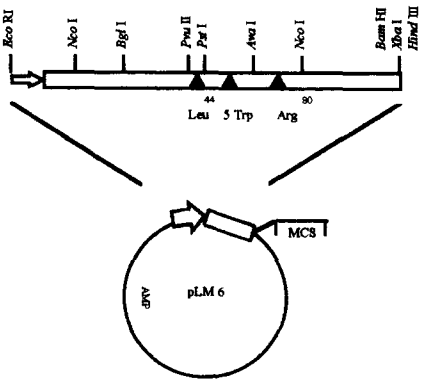


Fig. 2. Plasmid pLM6. ⇒: 2S promoter. The main restriction sites in the 2S coding region are indicated and the sites where mutations have been induced are marked by triangles.

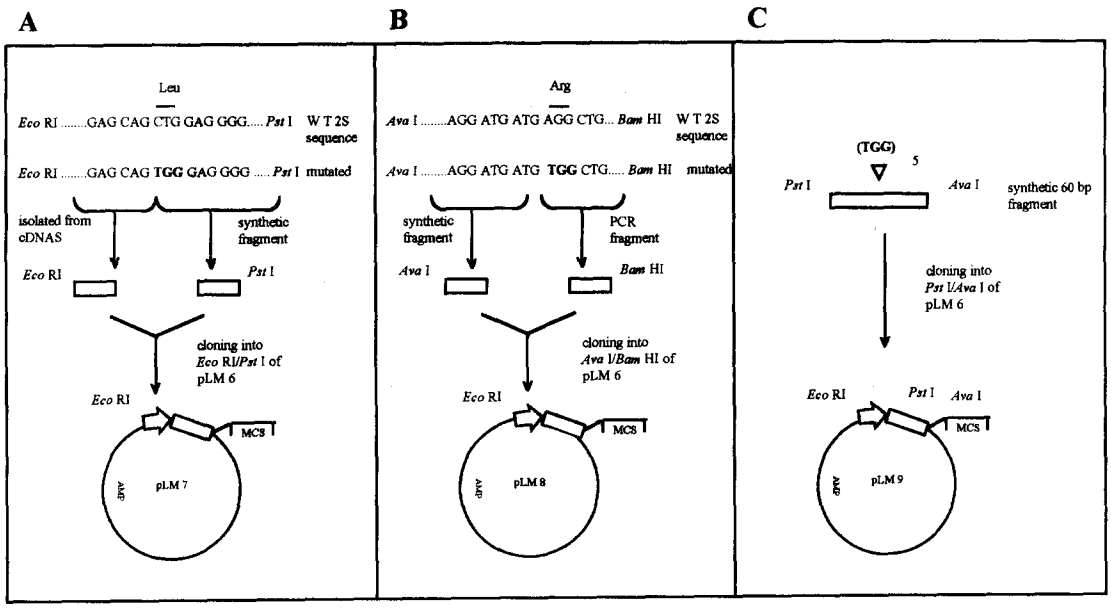


Fig. 3. Strategies of mutagenesis. (A) Substitution of Leu[44] by tryptophan. A synthetic fragment containing the TGG codon was ligated to the *Eco*RI/*Pvu*II fragment and introduced into pLM6. (B) Substitution of Arg[80] by tryptophan. This mutagenesis was performed in two steps: first a 500 bp fragment was amplified by PCR using an oligonucleotide containing the desired TGG codon. The PCR fragment was then ligated to a synthetic oligonucleotide in order to complete the *Ava*I/*Bam*HI fragment and subsequently ligated into pLM6. (C) Insertion of 5 Trp residues. A synthetic *Pst*I/*Ava*I fragment containing 5 Trp codons between Met[57] and Arg[58] was used to replace the original fragment in pLM 6.

in the protein for tryptophan insertion, it was necessary to establish a plausible theoretical tertiary structure of the 2S protein. In our model (Fig. 1) the 2S albumin assumes the conformation of a 'protein globule' formed by 5 α-helices and connecting turns. In this model, Leu[44] faces the molecule's hydrophobic core, while Arg[80] faces the surface. Replacement of these two residues with tryptophan would allow the accuracy of the model to be tested by fluorescence quenching techniques. These mutations were predicted not to influence structural features of the model.

Similar reasoning led to the design of a Brazil nut 2S albumin containing five tryptophan residues in the variable region previously used by Vandekerckhove et al. [21] for the insertion of the neuro-pentapeptide Leu-enkephalin. The success of that and several other insertions in that region [22] in the 2S albumin of *Arabidopsis* suggested that the tryptophan insertion would also be tolerated.

### 3.2. Analysis of transgenic plants

*N. tabacum* was transformed with plasmids pTLM 14 (the unmodified control) and pTLM 15 (Leu[44] → Trp[44]), 16 (Arg[80] → Trp[80]) and 17 (5 Trp insertion) resulting in the transgenic plants, types 14–17. Several previous reports of the introduction and expression of modified seed storage protein genes into transgenic plants have appeared. A modified β-phaseolin gene containing 6 additional methionine codons was introduced in *N. tabacum* [23]. The modified protein was degraded, either in the Golgi apparatus or in the protein bodies [23]. As mentioned above, modified forms of the *Arabidopsis* 2S albumins have been expressed in *A. thaliana*, *Brassica napus* and *N. tabacum* [21,22]. Those experiments were successful, the proteins being synthesized and stably accumulated in seeds. The differing results show that, despite the lack of an enzymatic function, seed protein modifications are not guaranteed to succeed. Therefore, the transgenic plants carrying the modified Brazil nut 2S albumin genes were analyzed

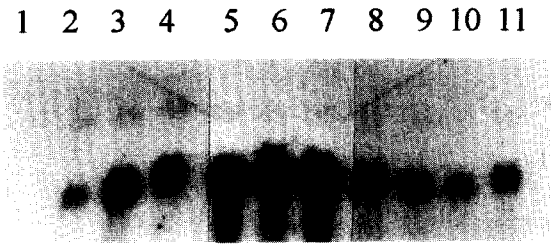


Fig. 4. Northern blot analysis of total leaf RNA from transgenic plants. Lanes: 1, untransformed plant; 2,3, transgenic plants type 14 (normal 2S albumin); 4, transgenic plant type 15 ( mutation Leu[44] → Trp[44]); 5–8, transgenic plants type 16 (mutation Arg[80] → Trp[80]); 9–11, transgenic plants type 17 (modified 2S albumin containing 5 additional tryptophan codons).

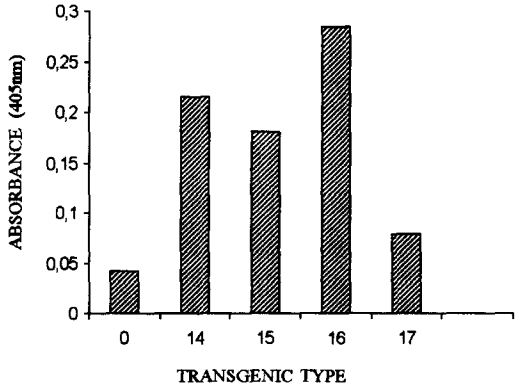


Fig. 5. ELISA test of seed proteins. Proteins were extracted from seeds of transgenic plants type 14, 15, 16 and 17 (R1) and were analyzed using mouse anti-2S polyclonal antibody as described. The column marked '0' shows the background level in untransformed plants.

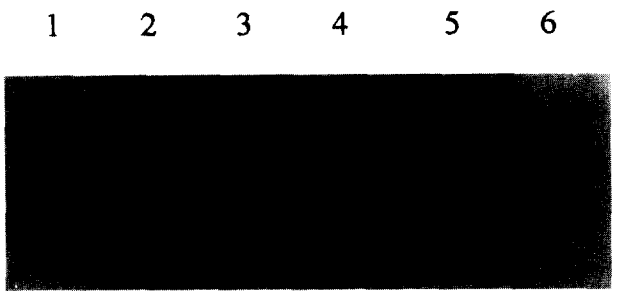


Fig. 6. Immunoblot analysis of R1 seed extracts from transgenic plants. Lanes: 1, 50 ng purified Brazil nut 2S albumin; 2, seed extracts (10 μg total protein) from an untransformed control plant; 3–6, extracts from transgenic types 14, 15, 16 and 17, respectively (10 μg total protein).

with respect to RNA level expression as well as protein accumulation and compartmentalization.

*3.2.1. RNA analysis.* To evaluate in which of the kanamycin resistant plants the modified genes were being transcribed, Northern blots using total RNAs from leaves of the four transgenic lines were performed. In all cases the RNAs hybridizing with the radiolabeled Brazil nut 2S albumin probe show the same electrophoretic mobility as the RNA from the plants transformed with the unmodified 2S albumin (Fig. 4, lane 2). In this assay the 15 nucleotide difference in the case of pTLM 17 transformed plants is undetectable. The variable amount of 2S albumin mRNA detected in the various RNA preparations was presumed to be due to normal position effects. However, plants with pTLM 17 consistently had lower levels of 2S albumin mRNA. The sample size and precision of the technique used were not sufficient to determine if that was due to lower stability of the mRNA encoded by pTLM 17.

*3.2.2. Protein analysis.* R1 seeds originating from plants found to be positive for Brazil nut mRNA expression in the leaves were analyzed for the presence of the corresponding 2S albumins. Since it was not clear whether the modified proteins would react with antibodies against normal Brazil nut 2S albumin, initially ELISA techniques were used with protein extracts from untransformed plants, plants transformed with the wild type sequences and plants transformed with the constructs containing three mutations. As can be seen in Fig. 5, extracts from seeds of progeny of Northern-positive plants, regardless of the type of modification, react with antibodies raised against unmodified 2S albumin. Extracts from seeds from transgenic plants type 16 gave the strongest responses, while those from type 17 showed consistently lower signals. This is consistent with the results of the Northern analysis, where leaves of plants carrying that construct had lower 2S albumin mRNA levels than the other lines.

When total protein extracts were analyzed on SDS-PAGE followed by immunoblots, the extracts from plants type 14–16 showed only one band reacting with the anti-2S-antibody (Fig. 6). This band had an identical electrophoretic mobility to that of the purified Brazil nut 2S albumin. The overall pattern of the immunoblot analysis was in agreement with the ELISA tests; again extracts from seeds of type 14, 15 and 16 gave stronger signals than those from type 17, where the Brazil nut 2S albumin band is barely visible. To exclude the possibility that apparent low expression levels of modified Brazil nut 2S albumin in type 17 plants were unique to the individual plants analyzed, seeds from 22 independent R0 transformants of this type were analyzed using ELISA. All type 17 transformants showed low levels of 2S albumin (Fig. 7). Therefore, consistently low expression in type 17 transformants is most likely due to a lower mRNA stability and not to incorrect processing or storage of the protein, since the protein product of the pTLM 17, though only weakly visible, shows the same electrophoretic behavior as the wild type protein.

Quantification of the immunoblot shown in Fig. 6, using the known amount of purified Brazil nut 2S albumin in lane 1 as a standard, indicates that the transgenic plant lines used in that experiment of types 14, 15 and 16 express the transgenic proteins at 0.30, 0.16 and 0.27% of the total extractable proteins. This expression level is better than that obtained with chimeric Brazil nut 2S genes in *B. napus* [6], but less than the levels with modified *Arabidopsis* 2S albumins expressed in *Arabidopsis* [22], and the levels (1.7–4%) obtained by Altenbach et al. [24] in *B. napus*. Part of the difference may be due to our use of the relatively constitutive 35S promoter instead of a strong seed promoter. The former was used because the primary purpose of the present work was to determine if the different modifications made would result in stable protein accumulation, and because it was thought that isolation of pure modified 2S albumins for structural studies might be more straightforward from leaf tissue.

*3.2.3. In situ immunocytochemical localization.* The results presented so far indicate that all the mutated 2S sequences are expressed and result in proteins with electrophoretic behavior on SDS-PAGE identical to that of wild type Brazil nut 2S albumins. This suggests that the mutated proteins are correctly processed. Most dicot storage proteins are synthesized on the rough endoplasmic reticulum, pass through the lumen of the ER and the Golgi complex and subsequently are stored in vacuolar derived protein bodies, also called protein storage vacuoles (PSV). In order to verify whether the modified proteins were correctly compartmentalized in the PSV, immunocytochemistry was performed using seeds from type 16 plants. Fig. 8B shows transgenic tobacco seed storage parenchymatic cells after immunogold labeling with Brazil nut 2S albumin-specific antibodies. Gold particles are observed over the protein bodies, but not over the lipid bodies or in the untransformed control (Fig. 8A). Therefore, at least in the case of the

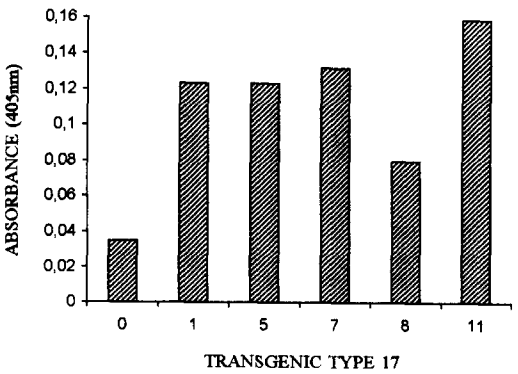


Fig. 7. ELISA test of type 17 transformants. R0 seed extracts (1 μg/μl of protein) were analysed using anti-2S polyclonal antibodies. The column marked '0' shows the background level in untransformed plants.

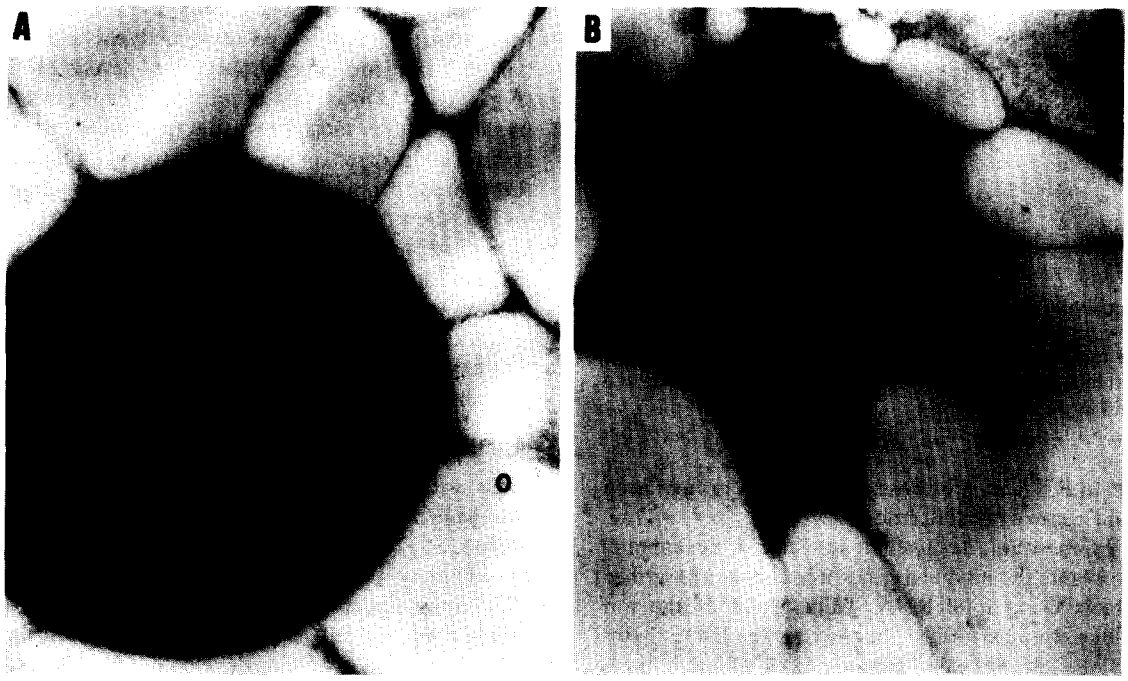


Fig. 8. Immunocytochemical localization of modified 2S albumin in mature seeds from transformed tobacco. (A) Untransformed seed control (×21 500). (B) Endosperm from type 16 seeds (×28 000). P, protein body; O, oil body.

$Arg^{80} \rightarrow Trp^{80}$ modification the modified proteins are correctly targeted to the storage organelles.

Brazil nut methionine-rich 2S albumins contain no residues at all of the essential amino acid tryptophan. This is also the case for other methionine-rich proteins from corn and rice [25]. We show here that modified Brazil nut 2S sequences designed in part to correct this deficiency and, simultaneously to provide the starting material for structural studies of the 2S albumins, are correctly expressed and targeted in transgenic tobacco seeds. Furthermore, the transgenes were stably inherited by R1 progeny.

*Acknowledgements:* We thank Drs. Martha Cecília Laboissiere and Márcio Elias Ferreira for critical reading of the manuscript. We thank Dr. Jan Leemans of Plant Genetic Systems, Gent, Belgium for support during the site-directed mutagenesis work. This work was supported by the Brazilian National Research Council (CNPq) and the Brazilian Enterprise for Agricultural Research (EMBRAPA).

## References

[1] Youle, R. and Huang, A.H.C. (1981) Am. J. Bot. 68, 44–48.
[2] Altenbach, S.B., Pearson, K.W., Leung, F.W. and Sun, S.S.M. (1987) Plant Mol. Biol. 8, 239–250.
[3] De Castro, L.A.B., Lacerda, Z., Aramayo, R.A., Sampaio, M.J.A.M. and Gander, E.S. (1987) Mol. Gen. Genet. 206, 338–343.
[4] Gander, E.S., Holmstroem, K.-O., De Paiva, G.R., De Castro, L.A.B., Carneiro, M. and Grossi de Sá, M.F. (1991) Plant Mol. Biol. 16, 437–448.
[5] Altenbach, S.B., Pearson, K.W., Meeker, G., Staraci, L.C. and Sun, S.S.M. (1989) Plant Mol. Biol. 13, 513–522.
[6] Guerche, P., De Almeida, E.R.P., Schwarzstein, M.A., Gander, E.S., Krebbers, E. and Pelletier, G. (1990) Mol. Gen. Genet. 221, 306–314.
[7] Aragão, J.L., Grossi de Sá, M.F., Almeida, E.R., Gander, E.S. and Rech, E.L. (1992) Plant Mol. Biol. 20, 357–359.
[8] Saalbach, I., Pickardt, T., Machemehl, F., Saalbach, G., Schieder, O. and Muntz, K. (1994) Mol. Gen. Genet. 242, 226–236.
[9] Grossi de Sá, M.F., Weinberg, D.F., Rech, E.L., Barros, L.M.G., Aragão, F.J.L., Holmstroem, K.-O. and Gander, E.S. (1994) Plant Sci. 103, 189–198.
[10] Krebbers, E., Herdies, L., De Clercq, A., Seurinck, J., Leemans J.,Van Damme, J., Segura, M., Gheysen, G., Van Montagu, M. and Vandekerckhove, J.(1988) Plant Physiol. 87, 859–866.
[11] Jang, D.J. and El-Sayed, M.A. (1989) Proc. Natl. Acad. Sci. 86, 5815–5819.
[12] Chou, P.Y. and Fasman, G.D. (1979) Annu. Rev. Biochem. 47, 251–276.
[13] Gibrat, J.F., Garnier, J. and Robson, B. (1987) J. Mol. Biol. 198, 425–443.
[14] Deblaere, R., Reynaerts, A., Hoefte, H., Hernalsteens, J.-P., Leemans, J. and Van Montagu, M. (1987) Methods Enzymol. 153, 277–292.
[15] Horsh, R.B., Fry, J.E., Hoffman, N.L., Eichholtz, D., Rogers, S.G. and Fraley, R.T. (1985) Science 227, 1229–1231.
[16] Jones, J.D.G., Dunsmuir, P. and Bedbrook, J. (1985) EMBO J. 4, 2411–2418.
[17] 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.
[18] Voller, A., Bidwell, D. and Bartlett, A. (1973) Microplate immunoassay for the immunodiagnosis of virus infections. In: Manual of Clinical Immunology (Rose, N. and Friedman, H. eds.) pp. 506–512, American Society for Microbiology, Washington, DC.
[19] Karey, K.P. and Sirbasku, D.A. (1989) Anal. Biochem. 178, 255–259.
[20] Grossi de Sá, M.F., Martins de Sá, C., Harper, F., Coux, O., Akayat, O., Pal, J., Florentin, Y. and Scherrer, K. (1988) J. Cell Sci. 89, 151–165.
[21] Vandekerckhove, J., Van Damme, J., Van Lijsebettens, M., Botterman, J., De Block, M., Vandewiele, M., De Clercq, A., Leemans, J., Van Montagu, M. and Krebbers, E. (1989) Bio/technology 7, 929–932.
[22] De Clercq, A., Vandewiele, M., Van Damme, J., Guerche, P., Van Montagu, M., Vandekerckhove, J. and Krebbers, E. (1990) Plant Physiol. 94, 970–979.
[23] Hoffman, L.M., Donaldson, D.D. and Herman, E.M. (1988) Plant Mol. Biol. 11, 717–729.
[24] Altenbach, S.B., Kuo, C.-C., Staraci, L.C., Pearson, K.W., Wainwright, C., Georgescu, A. and Townsend, J. (1992) Plant Mol. Biol. 18, 235–245.
[25] Altenbach, S.B. and Simpson, R.B. (1990) TIBTECH 8, 156–160.