Plant Physiol. (1993) 101: 1415–1416

Plant Gene Register

A Fifth 2S Albumin Isoform Is Present in *Arabidopsis thaliana*

Hilde van der Klei, Jose Van Damme, Peter Casteels[1], and Enno Krebbers*
Plant Genetic Systems, Jozef Plateaustraat 22, B-9000 Gent, Belgium

The 2S albumins, or napin-like proteins, constitute a significant portion of total seed protein in a variety of dicotyledonous species (Youle and Huang, 1981). In mature seed they are found as heterodimers, consisting of two subunits linked by disulfide bridges. They are synthesized as precursors that undergo posttranslational proteolytic processing (Crouch et al., 1983). In most species, they are encoded by multigene families (Josefsson et al., 1987; Scofield and Crouch, 1987; Raynal et al., 1991). Krebbers et al. (1988) reported the isolation and sequencing of an *Arabidopsis* 2S albumin isoform and used the sequence to isolate four genes that encode different isoforms. Hybridization studies suggested that this represented the entire gene family. In the course of experiments to separate individual 2S albumin isoforms, we have now isolated a previously unknown isoform that is not encoded by any of the known genes.

MATERIALS AND METHODS

Seeds of *Arabidopsis thaliana* were ground to a fine powder in a mortar. Lipids were removed by three extractions with hexane (10%, w/v). The residue was lyophilized, suspended in 100 mM citric acid/NaOH buffer, pH 3.0, containing 0.2 M NaCl, and kept at room temperature for 15 min under continuous shaking. To extract all the protein present in the seeds, it was necessary to repeat the extraction three times. The pellet was removed by centrifugation in an Eppendorf centrifuge at full speed. The supernatant of the three extractions was pooled, filtered through a Millipore filter with a pore size of 22 μm, and loaded on a fast protein liquid chromatography gel filtration (Sephadex G-25 medium, Pharmacia) column equilibrated with 50 mM phosphate buffer, pH 7.0. The void volume, which contained the 2S albumin fraction, was loaded on a fast protein liquid chromatography ion-exchange (S-Sepharose High Performance, Pharmacia) column in 50 mM phosphate buffer, pH 7.0. 2S albumins were eluted with a linear gradient: 0 to 50% B in 50 mL, where solvent B is 2 M NaCl in 50 mM phosphate, pH 7.0 (Fig. 1). The fractions eluting from the ion-exchange column were collected separately.

To separate and purify individual isoforms, two HPLC steps were used. In the first step, fractions eluting from the ion-exchange column were each applied in 0.1% (v/v) TFA (solvent A) on a Vydac C4 reversed-phase HPLC column. The column was developed with a two-step linear gradient: 0 to 50% B for 25 min and 50 to 100% B for 12 min, where solvent B is 70% (v/v) acetonitrile in solvent A. The fractions eluting from the C4 column were dried in a Speedvac (Savant). In the second step, the dried fractions were loaded on a phenyl HPLC column (Vydac) in solvent A with a 20 to 60% linear gradient of solvent B over 40 min. The peaks eluting from the phenyl column were prepared for sequencing, which was done on an Applied Biosystems model 470A.

RESULTS AND DISCUSSION

To purify modified 2S albumins (Vandekerckhove et al., 1989), the feasibility of separating individual 2S albumin isoforms was examined. The protein extracted from *Arabidopsis* seeds was applied on an ion-exchange column. Fractions eluting between 0.6 M and 1.0 M NaCl all contained 2S

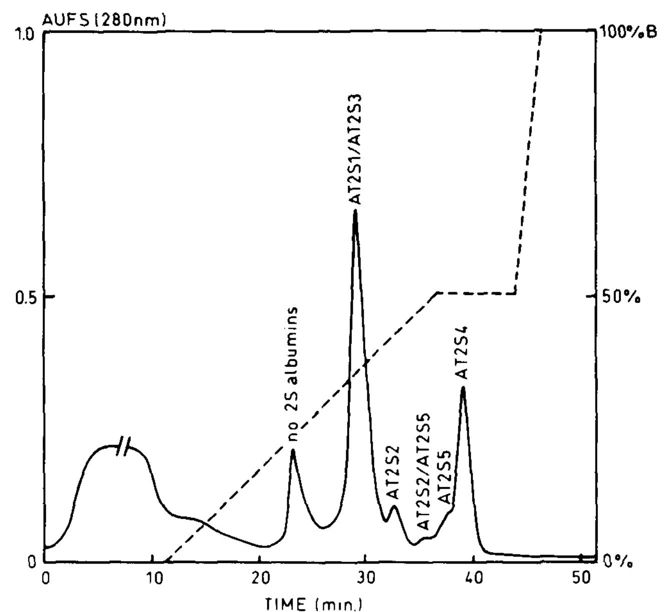

Figure 1. Fast protein liquid chromatography separation of 2S albumin isoforms. Fractionation of the void volume of a Sephadex G-25 column on an ion-exchange column is shown. The dashed line indicates the gradient of solvent B (see "Materials and Methods"). AUFS indicates absorbance units full scale; %B refers to the percentage of solvent B. The 2S albumins found to be present after further purification and sequencing are indicated above the relevant peaks.

[1] Present address: Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10021.

* Corresponding author; fax 32–91–240694.

```
AT2S5 P M G P Q Q - Q S S L K M C C N E L R Q V D K M C V C P T L K Q
AT2S1 P Q G Q Q Q E Q Q L F Q Q C C N E L R Q E E P D C V C P T L K Q
AT2S2 P Q G P Q Q G H Q I L Q Q C C S E L R Q E E P V C V C P T L R Q
AT2S3 F E G P Q Q G Y Q L L Q Q C C N E L R Q E E P V C V C P T L K Q
AT2S4 P - - - - Q R R Q L L Q K C C S E L R Q E E P V C V C P T L R Q
```

Figure 2. The sequence of the first 31 residues of the large subunit of the newly purified isoform (AT2S5). The amino acids shown in bold differ from at least three of the four previously identified isoforms.

albumins (Fig. 1) and these were further purified as outlined in "Materials and Methods." N-terminal sequence analyses of the different fractions revealed that in addition to the four already identified isoforms (Krebbers et al., 1988), a novel isoform had been isolated (Fig. 2). The 31 residues sequenced show significant homology with the other *Arabidopsis* 2S albumin isoforms, including the conservation of the Cys residues at positions 14 and 15. On the basis of its copurification through several steps and the sequence homology, we concluded that a novel isoform, designated AT2S5, had been isolated. It is clear that AT2S5 is the most divergent of the five isoforms. Depending on the isoform, it differs from the other four at 10 or 11 positions, including a deletion of the 31 residues sequenced. This difference should be reflected at the DNA level, and, assuming that a similar level of divergence is found over the rest of the gene, would explain why it was not detected in the hybridization experiments carried out by Krebbers et al. (1988). It is as yet unknown if this gene is linked to the other four, which form a tandem array.

ACKNOWLEDGMENTS

We thank Adri Van Vliet and Joel Vandekerckhove for advice. The authors of our original publication (Krebbers et al., 1988) regret any inconvenience that our initial erroneous conclusion may have caused.

Received December 22, 1992; accepted January 1, 1993.
Copyright Clearance Center: 0032-0889/93/101/1415/02.

LITERATURE CITED

Crouch ML, Tenbarge KM, Simon AE, Ferl R (1983) cDNA clones for *Brassica napus* seed proteins: evidence from nucleotide sequence analysis that both subunits of napin are cleaved from a precursor polypeptide. J Mol Appl Gen **2:** 273–283

Josefson LG, Lenman M, Ericson ML, Rask L (1987) Structure of a gene encoding the 17S storage protein, napin, from *Brassica napus*. J Biol Chem **262:** 12196–12201

Krebbers E, Herdies L, De Clercq A, Seurinck J, Leemans J, Van Damme J, Segura M, Gheysen G, Van Montagu M, Vandekerckhove J (1988) Determination of the processing sites of an *Arabidopsis* 2S albumin and characterization of the complete gene family. Plant Physiol **87:** 859–866

Raynal M, Depigny D, Grellet F, Delseny M (1991) Characterisation and evolution of napin-encoding genes in radish and related crucifers. Gene **99:** 77–86

Scofield SR, Crouch ML (1987) Nucleotide sequence of a member of the napin storage protein family from *Brassica napus*. J Biol Chem **262:** 12202–12208

Vandekerckhove J, Van Damme J, Van Lijsebettens M, Botterman J, De Block M, Vandewiele M, De Clercq A, Leemans J, Van Montagu M, Krebbers E (1989) Enkephalins produced in transgenic plants using modified 2S seed storage proteins. Biotechnology **7:** 929–932

Youle R, Huang AHC (1981) Occurrence of low molecular weight and high cysteine containing albumin storage protein in oilseeds of diverse species. Am J Bot **68:** 44–48