Cloning and mRNA transcription analysis of five developmentally regulated cDNA families in sunflower immature embryos

Asmâa Kabbaj and André Bervillé[1]

INRA-Montpellier, Station d'Amélioration des Plantes, 2, Place P Viala, 34060 Montpellier Cedex 1, France. [1] To whom correspondence should be sent

Fax 33 67 04 54 15; E-mail berville@msdos.ensam.inra.fr

Introduction

In sunflower, two major classes of storage proteins have been studied, the globulins 11 S (helianthinins) and the albumins 2S. They both represent 85 % of total proteins stored in the seeds []. The helianthinins are soluble in 1M NaCl and particularly rich in Arg, Phe, Gln and Asp amino-acids. They consist of hexameric holoproteins of 300 kDa. Each intermediary subunit (αβ) includes a large (α: 42-39 kDa), acidic polypeptide and a smaller (β: 27-23 kDa), basic polypeptide linked by disulfide bond []. The cDNA cloning and mRNA expression analysis of helianthinin genes have been reported by[]. Transcripts accumulate to maximum levels approximately between 12 and 15 days after flowering (DAF) and disappear in mature seeds. The albumins are soluble in water, basic and consists of a unique polypeptide chain of 10 to 18 kDa []. In all studied species except sunflower, the albumin precursors are cleaved at a conserved site into two subunits (9 and 3 kDa) linked by a disulfide bridge [][][]. The molecular weight and amino-acid composition of albumins in sunflower permits to distinguish at least two fractions. The major fraction includes albumins of « 18-14 kDa », rich in Cys 5-8 % and Lys 7-10 % []. The second fraction consists of albumins of 10 kDa, rich in Cys 8 % and Met 16 % and containing less Lys (4 %) than the major fraction. This Met-rich fraction represents 37 % of total albumins [] and similar albumins have been identified in only Brazil nut . The cloning of two albumin cDNAs, HaG5 and SFA8 clones respectively, from both fractions has been reported [][] Allen *et al.*, 1987. Transcript for the HaG5 clone was detected in the seeds at 5 DAF, reached maximum levels at 12 DAF and then decreased as the seed matures [].

The relative albumin and helianthinin seed content changes between varieties and in response to environmental conditions. Low oil varieties have higher contents of helianthinins than high oil varieties. Plant water supply also effects the quality of seed protein complex. With the increase of soil moistening the helianthinin seed content drops and the relative quantity of albumin increases [].

As a first step towards the investigation of the regulatory mechanisms involved in the developmental expression of genes in maturing seeds, we report here the cloning and mRNA transcription analyses of five different seed specific cDNAs.

Materials and methods

Plant materials

The CANP3 sunflower (*Helianthus annuus* L.) line derived from the Russian population « Armavir 9345 », further designated as normal variety and the high oleic variety provided by Cargill (France), further designated by HOC were used in the experiments. Plants were grown in a greenhouse until the beginning of flowering, they were then transfered to controlled environment cabinets with 12 hr photoperiod at 20°C day / 18°C night or 26°C day / 22°C night. Developing seeds were picked from hand-pollinated capitulum at 6, 12, 16, 20 or 28 DAP. The embryos, 2 g fresh weight at 6

DAP, and 3 to 5 g fresh weight at the other stages, were dissected from achenes and used immediately for RNA extraction.

Seeds were germinated aseptically between two pieces of filter paper imbibed with water in the dark at 20°C during 36 hours. Cotyledon, hypocotyl and root differentiating tissues were obtained from 10 day-old seedlings. All the parts of the seedlings were separated and frozen in liquid nitrogen.

Probes

We used -1) The 1,05kb fragment of the 18S rRNA gene from sunflower []. -2) The cDNA pHa2 corresponding to the Helianthinin gene [].

Total RNA isolation and Northen-blot hybridization

Total RNA was isolated from developing sunflower seeds, seedlings and germinating seeds [].

Construction and differential screening of the cDNA libraries

Poly(A)+ mRNA was purified from total RNA of developing CANP3 seeds at 12 and 20 DAP using Hybond^{Tm}-mAP-messenger affinity paper according to the supplier's instructions (Amersham). Double-stranded cDNA was synthesized from 5μg of Poly(A)+ mRNA using a cDNA synthesis kit (Riboclone cDNA synthesis system, Promega). 3' end labelling of the double stranded-cDNA using Klenow polymerase showed a size ranging from 0.2 to 3 kb. *Eco*RI adaptors were added to the cDNA for ligation into the λ gt10/*Eco*RI arms. Vectors were packaged *in vitro* using commercial extracts (Promega). Recombinant bacteriophages were selected using *E.coli* K12 *C600 Hfl* strain.

For differential screening of the libraries, Mo-MLV reverse transcriptase (Bethesda Research Laboratories), oligodT primers and 70 μCi of [α-^{32}P] dCTP were used to synthesize radiolabelled first strand cDNA using 3μg poly(A)+mRNA. The libraries were sequentially screened with radiolabelled first strand cDNA from 12 and 20 DAP embryos, dry seeds, and leaves. Hybridization was carried as described above except that oligodA (12) "0.5mg/ml" was added to the prehybridization solution. Phage plaques that showed specific hybridization with 12 or 20 DAP embryos probes were isolated.

RESULTS

Isolation and analysis of cDNA expressed during the mid-maturation stage of the seed

We constructed cDNA libraries in λgt10 using poly(A)+mRNA from 12 and 20 day-old embryos. About 10 000 Pfu from each library were screened by differential hybridization using as negative probes poly(A)$^+$mRNA from dry seeds and leaves. We selected 20 clones that hybridized specifically to 12 or 20 day-old embryos probes. By cross-hybridation analysis, inserts were grouped into 5 cDNA families, each family consists of 1 to 5 cDNA carrying related sequences. The longest cDNA representative of each cDNA family, designated A8.1, A7.2, A2, C8.1 and C3 were used for sequence comparison homology to possibly identify their respective encoded proteins. --In order to identify cDNA encoding for helianthinin precursors, we performed hybridization of a southern blot carrying the various isolated cDNAs with the helianthinin pHa2 cDNA clone []. The cDNA A8.1 and two other related cDNAs were revealed. --The complete nucleotide sequence of the A7.2 cDNA was determined. The cDNA is 0,6 kb in length and shows an open reading frame of 424 nucleotides. The encoded polypeptide exhibits 141 amino acids corresponding to a molecular weight of 15,993 kDa. The 3' untranslated region contains

two putative polyadenylation signals. The deduced amino acid sequence shows 91.5 % homology with the SFA8 albumin Met-rich amino acid sequence already isolated from sunflower seeds []. The SFA8 and A7.2 albumin precursors are similar in length and show 12 aminoacid substitutions, most of them are conservatives. --The complete nucleotide sequence of the C3 cDNA was determined. The insert has a lenght of 353 nucleotides which shows significant nucleotide sequence homology with cDNA sequences encoding probable proteinase inhibitors from many species [][].The deduced amino acid sequence shows 57 %, 51,5 %, et 47 % homology sequence with probable proteinase inhibitors from *Glycine max*, *Arabidopsis thaliana* and *Solanum tuberosum*. The sequence homology is the lowest at the peptide 5' region which is strongly hydrophobic and could corresponds to a signal peptide for intramembrane transport. The cleavage site of the putative signal peptide could be predicted according to Von Heijne (1986) [] between the residues Gly-16 and Arg-17 leading to a 47 aminoacid mature protein. All the Cys residues of the mature protein are found in similar positions in the three other peptides compared except for *Arabidopsis thaliana* peptide which is lacking the last Cys residue. This suggests the establishement of similar types of disulphide bridges in all the polypeptides. --The two cDNAs C8.1 and A2 partial sequences did not show significant nucleotide sequence homology with any known genes.

Tissue-specific and temporal mRNA expression analysis

The A8.1, A7, A2, C8.1 and C3 cDNA were labelled *in vitro* and separately hybridized to northern-blots containing total RNA from adult leaves, capitulum without seeds, the cotyledone leaves, hypocotyl and root expanding tissues from 10-day-old plants, mature seeds, germinated seeds for 36 h, and immature embryos at 6, 12, 16, 20 and 28 DAP. The autoradiograms show that each cDNA reacted with a single mRNA. The mRNAs are 1.8, 1.4, 1.1, 0.6 and 0.4 kb long for the cDNA A8.1, C8.1, A2, A7 and C3, respectively. They are accumulated in the seeds between 12 and 20 DAP. The transcripts corresponding to the C8.1, A2 and C3 cDNAs disappeared at 28 DAP whereas those corresponding to the A8.1 and A7 cDNAs were still present but highly reduced in intensity. Furthermore all these transcripts were absent at 6 DAP, in mature seeds and in germinating seeds as well as in other tissues and organs, except for A7 homologous transcript also expressed in the capitulum at the maturity stage.

Variation in intensity of mRNA accumulation along seed growth

For quantitation and comparison of hybridized mRNA signals, all hybridizations were internally standardized by expressing the signal intensity for each cDNA obtained at any seed developmental stage in proportion to the signal obtained with the 18S rRNA gene (Table I). We could in that way display the accumulation kinetics of each mRNA during seed maturation

Comparison of the mRNA accumulation levels between CANP3 and HOC lines

Northern blot signals of each the 5 cDNA used separately as probes were compared between CANP3 and HOC at 12, 16 and 20 DAP using standardized signals (Fig.5, Table II). The A7 and C8.1 homologous mRNA kinetics and accumulation levels were similar in the two lines. The A8.1 and A2 homologous mRNA accumulations were significantly reduced in HOC at 20 DAP. The levels of mRNA accumulation were estimated in HOC at only 63% and 50% of those observed for CANP3, respectively for A8.1 and A2 cDNAs. The C3 homologous mRNA accumulation was in the contrary reduced in CANP3 in comparison with HOC, by 47 % and 36 % respectively at 16 and 20 DAP.

DISCUSSION

The cDNA libraries were constructed from mRNA isolated at 12 or 20 DAP. At these sunflower seed developmental stages, there are a massive synthesis and accumulation of oil and storage proteins[][]. Two of the studied cDNAs, A8.1 and A7.2 encode helianthinin and albumin Met-rich precursors respectively. The A7.2 amino acid deduced sequence shows 91,5% homology with the previously described SFA8 albumin precursor[]. The divergence between SFA8 and A7.2 amino acid deduced sequences indicates that both are part of a gene family with a minimum of two members. This is consistent with other observations on seed storage albumins that seem to be encoded by small multigenic families []. A third cDNA, C3 encodes an undescribed probable proteinase inhibitor. As yet, the physiological fonction of proteinase inhibitors is not clear. They appear to be involved in plant defence mechanisms against parasites[]. Furthermore, they accumulate specifically in storage organs «tubers and seeds» of many species and are supposed to have some additional function as sulphur-rich storage proteins [].

Each cDNA, representing the five families, hybridized one mRNA onto Northern-blots. All these mRNAs were specifically expressed in the immature seeds between 12 and 20 DAP whereas they were absent in all the vegetative tissues. Consequently, the corresponding genes are highly regulated at the transcriptional and/or RNA turnover levels. They exhibit significant transcript accumulation during the mid-maturation stage of seed development and are repressed during most of the plant life. We also noted that the cDNAs exhibit similar expression time frame since they were all not expressed in immature seeds at 6 DAP.. All the five mRNAs appeared between 6 and 12 DAP and are present until 20-28 DAP. Those corresponding to the helianthinin and the Met-rich protein are still present at 28 DAP, whereas all were absent in the mature quiescent seeds.

The persistance of the seed storage protein mRNAs in the late seed maturation stage and in some cases in mature seeds has been reported for many species[][][]. However, it has been suggested that the production of storage protein mRNAs is restricted to the phase of protein accumulation and that the subsequent mRNA levels are due to a relative long-life of these mRNAs [].

The mRNA accumulation kinetics analysis over the seed development revealed that each mRNA accumulation pattern was specific. We suggest that the modulations are the reflect of the order of the encoded protein function or deposit. The unidentified proteins encoded by the cDNAs C8.1 and A2 may have a function related to either minor storage protein or to other storage compounds accumulations. In the case of the helianthinin and the albumin Met-rich mRNAs, we observed that the accumulation kinetics were not synchronous. We also noted that the A7 albumin Met-rich mRNA accumulation occured later than the HaG5 homologous albumin mRNA wich was detected in immature seeds as early as 5 DAP and accumulated at the maximum level at 12 DAP []. However two different albumin fractions have been identified in sunflower: (1) the major 18-14 kDa fraction is similar to the albumin of other plant species, includes the HaG5 deduced polypeptide and accumulates earlier than the helianthinines [][]. (2) the 10 kDa Met-rich fraction has been found only in sunflower and Brazil nut, [][] includes the A7 deduced polypeptide and have been shown to be accumulated between 18 and 20 DAF, later than the 18-14 kDa albumin fraction and the helianthinins[]. These concordance between the mRNA accumulation kinetics and the the deposit of their corresponding protein precursors is in agreement with a primary control of storage protein gene expression at the level of mRNA transcription.

Relationships between oil content and storage protein mRNA accumulation

We observed that The A8.1 helianthinin mRNA accumulation levels were significantly lower in HOC in comparison, suggesting that the HOC seed contains less helianthinin than the CANP3 seed. The studied lines exhibit variable oil content, 45% and 36%, respectively for HOC and CANP3 (unpublished data). It has been reported by Borodulina and Suprinova (1976) [] that the increase of sunflower seed oil content induces an increase of albumin and a relative decrease of helianthinin seed contents. Furthermore a positive correlation were established between the oil and the lysine seed contents. We observed in the high oil variety HOC a lower level of the helianthinin mRNA in comparison with the CANP3 low oil variety. However, for the albumin Met-rich mRNA, the accumulation levels were similar in both lines. We suggest that the high mol weight albumin fraction wich is also the most lys rich is the only one that increases with seed oil content increase. The albumin low mol weight Met-rich could be stable whatever the oil content.

References:

1. Allen RD, Nessler CL, Thomas TL: Developmental expression of sunflower 11S storage protein genes. Plant Mol Biol 5: 165-173 (1985).
2. Allen RD, Cohen EA, Vonder.Haar RA, Orth KA, Ma DP, Nessler CL, Thomas TL: Expression of embryo specific genes in sunflower. In: Davidson EH, Firtel R (eds) UCLA Symposium on molecular approaches to developmental biology. Alan R. Liss, New York, p415 (1987a).
3. Allen RD, Cohen EA, Vonder Haar RA, Adams CA, Ma DP, Nessler CL, Thomas TL: Sequence and expression of a gene encoding an albumin storage protein in sunflower. Mol Gen Genet 210: 211-218 (1987b).
4. Altenbach BS, Pearson KW, Leung FW, Sun SM: Cloning and sequence analysis of a cDNA encoding a Brazil nut protein exceptionally rich in methionine. Plant Mol Biol 8: 239-250 (1987).
5. Boroludina AA, Suprunova LV: Genetic and ecological factors effecting a qualitative change of the sunflower seed protein complexe content. In Proc. 7th Int. Sunflower Conf. p428-432 (1976).
6. Choumane W, Heizmann P:. Structure and variability of nuclear ribosomal genes in the genus Helianthus. Theor Appl Genet 76: 481-489 (1988).
7. Choi *et al.*, 1993
8. Crouch *et al.*, 1983
9. Dalgalarrondo *et al.* 1984 1
10. Kortt AA, Caldwell JB: Low molecular weight albumins from sunflower seed: Identification of a methionine-rich albumin. Phytochemistry 29: 2805-2810 (1990).
11. Kortt AA, Caldwell JB, Lilley GG, Higgins TJV:. Amino acid and complementary DNA sequences of a methionine-rich 2S protein of sunflower (Helianthus annuus L.). Eur J Biochem 195: 329-334 (1991).
12. Laroche-Raynal M, Delseny M: Identification and characterization of the mRNA for major storage proteins from radish. Eur J. Biochem. 157: 321-327 (1986).
13. Mori T, Wakabayashi Y, Takagi S: Occurence of mRNA for storage protein in dry soybean seeds. J. Biochem. 84: 1103-1111 (1978).
14. Odani S, Koide T, Ono T, Ohnishi K: Structural relationship between barley (*Hordeum vulgare*) trypsin inhibitor and castor bean (*Ricinus communis*) storage protein. Biochem J 213: 543-545 (1983).
15. Pernollet J-C: Biosynthesis and accumulation of storage proteins in seeds. Physiol Veg 23: 45-59 (1985).
16. Roberston *et al.* 1978 4
17. Ryan, 1973 4
18. Sharief, F.S.& Li SS-L, 1982. Aminoacid sequence of small and large subunits of seed protein from *Ricinus communis*. J. Biol. Chem. 257: 14753-14759.
19. Sosulki
20. Stiekema WJ, Heidekamp F, Dirkse WG, Van Beckum J, De Haan P, Ten Bosch C, Louwerse JD: Molecular cloning and analysis of four potato tuber mRNAs. Plant Mol Biol 11: 255-269 (1988).
21. Teyssendier de la Serve B, Jouanneau JP: Preferential incorporation of exogenous cytokinin, N6-benzyladenine into 18S and 25S ribosomal DNA of tabacco cells in suspension culture. Biochimie 61: 913-922 (1979).
22. This P, Goffner D, Raynal M, Chartier Y, Delseny M:. Characterization of major storage proteins of sunflower and their accumulation. Plant Physiol Biochem 26 (2): 125-132 (1988).
23. This
24. Von Heijne G: A new method for predicting signal sequence cleavage sites. Nuc Acid Res 14: 4683-4690 (1986).
25. Vonder Haar RA, Allen RD, Cohen EA, Nessler CL,. Thomas T.L: Organisation of the sunflower 11S protein gene family. Gene 74: 433-443 (1988).
26. Youle RJ, Huang AHC: Albumin storage proteins in the protein bodies of castor bean. Plant Physiol 61: 13-16 (1987).
27. Youle RJ, Huang AHC: Occurence of low molecular weight and high cystein containing albumin storage proteins in oil seeds of diverse species. Am J Bot 68: 44-48 (1981).

Table I: Influence of developmental stage in the mRNA accumulation
Hybridization signal intensities in CANP3 line obtained with the 18S rRNA gene and the A8.1, A7, C3, C8.1 or A2 cDNA probes at different developmental seed stages. Plants were maintained at 26°C day/18°C night. Intensities are in arbitrary units.

	Seeds at different developmentally stages					
Probes	6 DAP	12 DAP	16 DAP	20 DAP	28 DAP	dry seeds
18S	1.25	1.82	1.79	1.57	1.30	1.59
A7	0	1.69	2.38	2.31	0.11	0
A7/18S	0	0.93	1.33	1.47	0.08	0
18S	1.25	1.99	2.08	1.99	1.21	2.19
A8.1	0	2.11	2.54	2.31	0.28	0
A8.1/18S	0	1.06	1.22	1.16	0.23	0
18S	1.06	2.41	2.03	1.58	1.50	1.60
C3	0	2.11	1.12	0.87	0	0
C3/18S	0	0.87	0.55	0.55	0	0
18S	1.06	2.41	2.03	1.58	1.50	1.60
C8.1	0	0.89	1.58	1.18	0	0
C8.1/18S	0	0.37	0.78	0.75	0	0
18S	0.69	1.52	1.06	1.14	0.60	0.82
A2	0	0.72	1.16	1.23	0	0
A2/18S	0	0.47	1.09	1.08	0	0

Table II: Comparison of the mRNA accumulation levels between a normal line « CANP3 » and a high oil and oleic line « HOC ». Hybridization signal intensities obtained with the 18S rRNA gene and the A8.1, A7, C3, C8.1 or A2 cDNA probes. Plants were maintained at 26°C day/18°C night. Intensities are in arbitrary units.

	CANP3			HOC		
Probes	12 DAP	16 DAP	20 DAP	12 DAP	16 DAP	20 DAP
18S	1.82	1.79	1.57	1.69	1.59	1.52
A7	1.69	2.38	2.31	1.54	2.18	2.17
A7/18S	0.93	1.33	1.47	0.91	1.37	1.43
18S	1.99	2.08	1.99	2.19	2.05	1.91
A8.1	2.11	2.54	2.31	2.15	2.09	1.40
A8.1/18S	1.06	1.22	1.16	0.98	1.02	0.73
18S	2.41	2.03	1.58	1.66	1.71	1.17
C3	2.11	1.12	0.87	1.41	1.80	1.01
C3/18S	0.87	0.55	0.55	0.85	1.05	0.86
18S	2.41	2.03	1.58	1.66	1.71	1.17
C8.1	0.89	1.58	1.18	0.46	1.44	0.80
C8.1/18S	0.37	0.78	0.75	0.28	0.84	0.68
18S	1.52	1.06	1.14	1.35	1.45	1.56
A2	0.72	1.16	1.23	0.47	1.60	0.84
A2/18S	0.47	1.09	1.08	0.35	1.10	0.54