

Systematic Studies of 12S Seed Storage Protein Accumulation and Degradation Patterns during *Arabidopsis* Seed Maturation and Early Seedling Germination Stages

Qing Li, Bai-Chen Wang, Yu Xu and Yu-Xian Zhu*

The National Laboratory of Protein Engineering and Plant Genetic Engineering, College of Life Sciences, Peking University, Beijing 100871, China

Received 17 October 2006, Accepted 12 December 2006

Seed storage proteins (SSPs) are important for seed germination and early seedling growth. We studied the accumulation and degradation profiles of four major *Arabidopsis* 12S SSPs using a 2-DE scheme combined with mass spectrometric methods. On the 2-DE map of 23 dpa (days post anthesis) siliques, 48 protein spots were identified as putative full-length or partial α , β subunits. Only 9 of them were found in 12 dpa siliques with none in younger than 8 dpa siliques, indicating that the accumulation of 12S SSPs started after the completion of cell elongation processes both in siliques and in developing seeds. The length and strength of transcription activity for each gene determined the final contents of respective SSP. At the beginning of imbibition, 68 SSP spots were identified while only 2 spots were found at the end of the 4 d germination period, with α subunits degraded more rapidly than the β subunits. The CRC α subunit was found to degrade from its C-terminus with conserved sequence motifs. Our data provide an important basis for understanding the nutritional value of developing plant seeds and may serve as a useful platform for other species.

Keywords: 2-DE, Accumulation seed storage proteins, *Arabidopsis*, Degradation, MALDI-TOF-MS, QRT-PCR

Introduction

SSPs are a set of proteins that are synthesized in the developing seed at the stage when cell division is complete (Goldberg *et al.*, 1994) and they play a crucial role in the life cycle of higher plants since they are the major sources of

nitrogen and carbon during subsequent seed germination and early seedling growth. SSPs are primarily synthesized in the late stages of seed development on rough endoplasmic reticulum (ER) as precursors, and are further converted into mature forms and accumulate in protein storage vacuoles (PSVs) of terminally differentiated cells in developing seeds (Herman and Larkins, 1999). Legumin-type globulin and napin-type albumin proteins are the predominant SSPs found in *Arabidopsis* seeds (Heath *et al.*, 1986; Krebbers *et al.*, 1988; Pang *et al.*, 1988; van der Klei *et al.*, 1993). The 12S SSPs, also referred to as cruciferins, usually present in a plant cell as hexameric complexes consisting of six α and six β subunits linked via disulfide bonds.

Four genes encoding 12S SSPs, including *CRA1* (At5g44120, also called cruciferin 4), *CRB* (At1g03880), *CRC* (At4g28520, also called *cruciferin 1*) and cruciferin 2 (*CRU2*, At1g03890) are located on chromosome 5, 1, 4 and 1, respectively in the *Arabidopsis* genome of Columbia ecotype (Pang *et al.*, 1988; Fujiwaraa *et al.*, 2002; Gruis *et al.*, 2002, 2004). During seed maturation, each SSP is proteolytically processed at a conserved Asn-Gly peptide bond by an asparaginyl endopeptidase to produce two mature polypeptides referred to as α and β subunits that are joined together by a disulfide bond (Muntz, 1998; Gruis *et al.*, 2004; Shimada, *et al.*, 2003). The dicysteine residues in CCxQL were reported to play a critical role in protein sorting (Kawagoe *et al.*, 2005). Since all 12S globulin genes are strictly transcribed only in developing seeds at the mid- to late-stages of embryogenesis, they are ideal candidates for exploiting tissue- and temporal-specific regulation of gene expression (Fujiwaraa *et al.*, 2002).

Proteomic analysis offers good opportunity to study large numbers of proteins involved in specific biological responses (Gallardo *et al.*, 2001; Finnie *et al.*, 2002; Canovas *et al.*, 2004; Gruis *et al.*, 2004; Fu *et al.*, 2005). Analysis of developing barley seeds revealed that grain filling and maturation was accompanied by a significant accumulation of low-Mr α -amylase/trypsin inhibitors, serine protease inhibitors and enzymes

*To whom correspondence should be addressed.
Tel: 86-10-6275-4096; Fax: 86-10-6275-4427
E-mail: zhuyx@water.pku.edu.cn

involved in protection against oxidative stress (Finnie *et al.*, 2002). Twenty-four of the 47 identified protein spots from embryos and endosperms of germinating tomato seeds were identified as fragments of different SSPs (Sheoran *et al.*, 2005), whereas 26 of the 74 identified protein spots from germinating *Arabidopsis* seeds were 12S SSP fragments (Gallardo *et al.*, 2001), indicating that SSPs are a major component of the germinating proteome. Here we report a systematic analysis of 12S SSP synthesis and degradation throughout *Arabidopsis* seed maturation and germination processes.

Materials and Methods

Plant growth and sample preparation. Dry mature *Arabidopsis* seeds (Columbia ecotype) were soaked in dark at 4°C for 3 days before being transferred to fully automated growth chambers as reported previously (Wang *et al.*, 2003) and were maintained at 20–22°C with a 16/8 h (light/dark) cycle. Protein samples of various growth stages were prepared from dry seeds, 0 d (harvested upon completion of dark stratification) and 1, 2, 3 or 4 d after beginning of the germination experiments. The siliques were labeled at the time of anthesis and were harvested at 5, 8, 12, 15 and 23 dpa, respectively.

Protein extraction. Two hundred-fifty mg dry seeds or equivalent amount of 0 d seeds and young seedlings harvested at 1 or 2 d after initiation of germination experiment, or 1 g of 3 and 4-d seedlings, or 1 g siliques at different stages were ground in liquid nitrogen in a mortar and pestle. The fine powder was precipitated at –20°C with 10% (w/v) trichloroacetic acid (TCA) in cold acetone containing 0.07% (w/v) 2-mercaptoethanol for at least 2 h. The mixture was further purified following the steps reported previously (Fu *et al.*, 2005; Wang *et al.*, 2006). About 10–20 mg total proteins were obtained from each sample when quantified using the 2-D Quant kit (Amersham Pharmacia Biotech).

Two-dimensional electrophoresis and quantitative analyses. 2-DE was performed as described previously (Fu *et al.*, 2005; Wang *et al.*, 2006) with 200 µg of total proteins loaded onto each gel. Isoelectric focusing was done with the IPGphor system (Amersham Pharmacia Biotech). Immobiline 3–10 and 4–7 linear DryStrips (Amersham Pharmacia Biotech) were run at 30 V for 8 h, 50 V 4 h, 100 V 1 h, 300 V 1 h, 500 V 1 h, 1000 V 1 h and 8000 V 12 h using rehydration buffer (8 M urea, 2% CHAPS, 20 mM dithiothreitol) and 0.5% IPG buffer (Amersham Pharmacia Biotech). SDS-PAGE was performed using 12.5% polyacrylamide gels without a stacking gel in the ETTAN DALT SIX ELECT UNIT 230 (Amersham Pharmacia Biotech). Gels were stained, after SDS-PAGE, with 0.1% (w/v) Coomassie Blue R-350 (CBR350) in 10% acetic acid and destained with 10% acetic acid or were silver stained. Triplicate protein samples were prepared from each designated growth stages, and silver- or CBB-stained gels were run in parallel for each sample for quantitative purposes and for mass spectrometric analysis, respectively. Gel images were processed using 2-D Imagemaster Elit (version 4.01, Amersham Pharmacia Biotech) according to the

manufacturer's instructions. The 2-D gels were aligned, matched, and volumes of each spot were quantitatively determined after normalization of total spot volumes and background subtraction as previously reported (Fu *et al.*, 2005). Details of in-gel digestion, peptide mass fingerprinting and N-terminal sequencing were described in Wang *et al.* (2006).

RNA purification and QRT-PCR analysis. Total RNA from different samples were isolated as previously described (Wang *et al.*, 2003). Five µg total RNA were reverse transcribed in a total volume of 20 µl using SuperScript™ First-Strand Synthesis System for RT-PCR (Invitrogen). Gene-specific primers for different 12S globulins were designed according to sequences available in the *Arabidopsis* genome database (<http://www.arabidopsis.org>). As a loading control, we amplified a 238 bp ubiquitin (*UBQ10*) fragment in parallel experiments. QRT-PCR was carried out using the SYBR green PCR kit (Applied Biosystems) in a DNA Engine Opticon-Continuous Fluorescence Detection System (MJ Research) as reported in (Shi *et al.*, 2006). Samples were analyzed in triplicates and were quantified by the comparative cycle threshold method. All gene-specific PCR primers were as follows: 5'-CAGCCACTTGTCATCATCG-3' and 5'-TAGGAAGCACCATCGCATTA-3' for CRC; 5'-CACGGTAATGGCTTAGAGGA-3' and 5'-ACTGGAATCGGTTGCTTGTT-3' for CRA1; 5'-GCCATCACTCGGATACATTA-3' and 5'-GGTCTCAATCGTGCTAAACT-3' for CRB; 5'-TTTCTTGCCCGCTTTCTTTTA-3' and 5'-TCTTCTTGCGTTCTGCTCCC-3' for CRU2.

Results

Phenotypic analyses of *Arabidopsis* seed/silique expansion and maturation. Phenotypes of developing siliques and seeds taken at 5, 8, 12 and 23 dpa were shown in Fig. 1. Siliques completed most of their elongation growth in the first 5–8 days. Siliques older than 8 dpa did not elongate significantly,

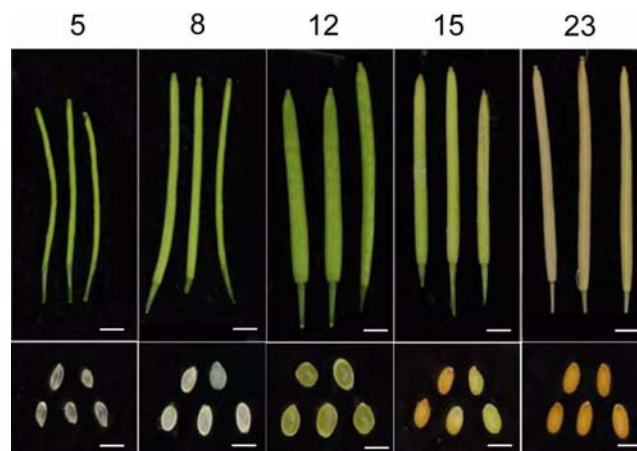


Fig. 1. Phenotypic analysis of developing *Arabidopsis* siliques (Upper panels) and seeds (Lower panels). Photos were taken at the indicated time (dpa). The bars represent 2 mm in upper panels and 0.2 mm in lower panels.

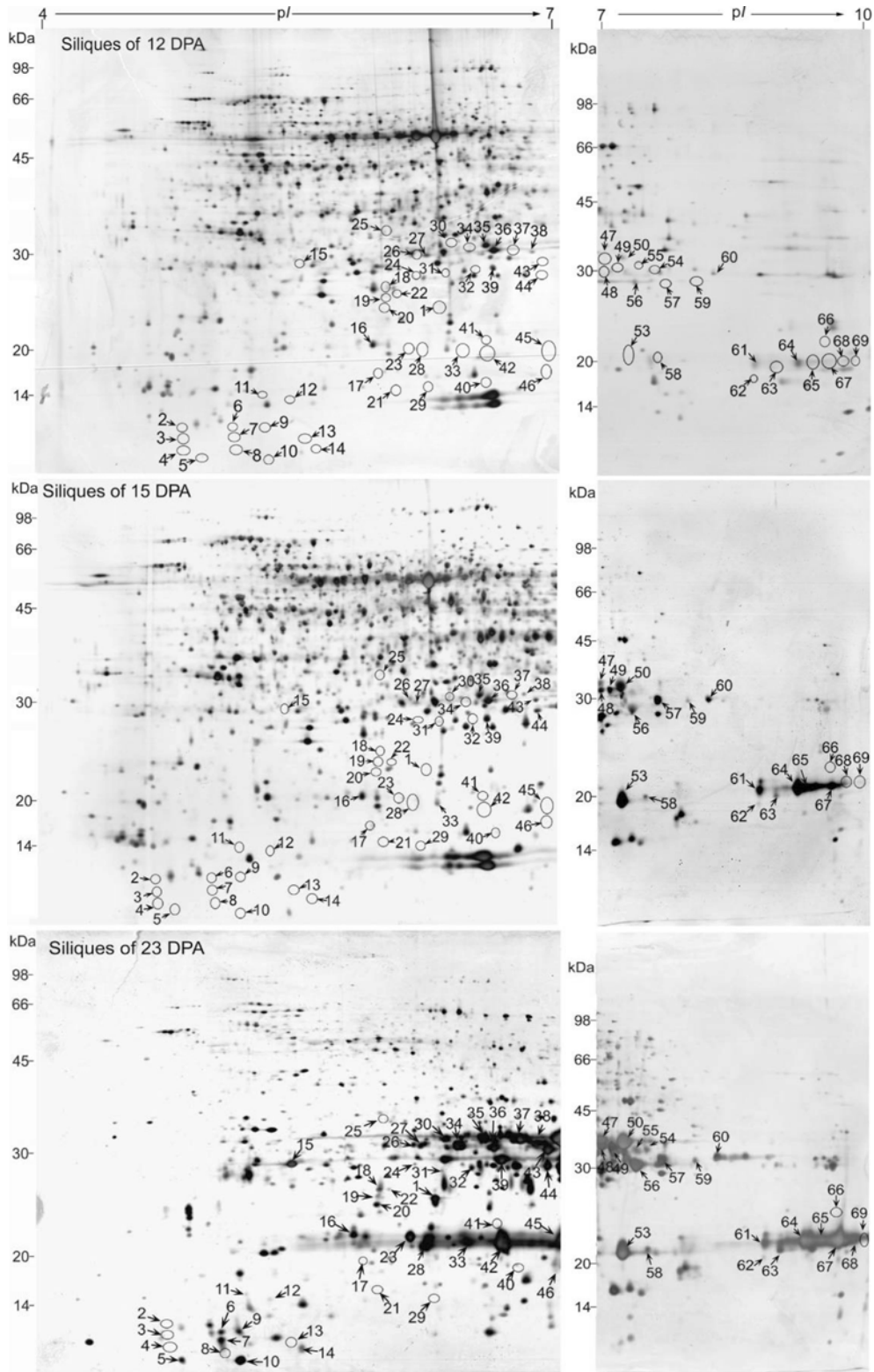


Fig. 2. Silver-stained 2-DE maps of *Arabidopsis* siliques collected at indicated times.

they became thicker in later stages and started yellowing after 15 dpa. By contrast, seeds taken out of 5-8 dpa siliques were small, transparent and full of water. They reached maximum

size and turned green at 12 dpa, started yellowing around 15 dpa and came to full maturity around 23 dpa (Fig. 1).

Table 1. Continued

Spot No.	pI/Mw		Score/ Coverage	Identified peptides	NMP ^a	Siliques (dpa)					Germinating seeds					
	Exp.	Cal				12	15	23	dry seed	0 d	1 d	2 d	3 d	4 d		
27	5.54/31.22	6.27/29.10	74/23	α	8		0.23±0.04	0.18±0.02	0.11±0.01							
34	5.92/31.12	6.27/29.10	76/45	α	12		0.51±0.09	0.79±0.10	1.06±0.21			0.62±0.15	0.35±0.11			
35	5.99/31.26	6.27/29.10	66/22	α	11	0.26±0.03	0.29±0.06	0.34±0.09	0.11±0.03							
38	6.09/31.80	6.27/29.10	80/24	α	10	0.16±0.01	0.12±0.21	0.39±0.09	0.43±0.08			0.26±0.05				
39	6.07/30.98	6.27/29.10	90/37	α	12		0.81±0.11	0.85±0.07	0.41±0.11			0.29±0.07	0.12±0.04			
43	6.19/31.25	6.27/29.10	80/25	α	19		0.19±0.03	0.79±0.11	0.80±0.15			0.21±0.06	0.10±0.03	0.05±0.00		
44	6.23/31.19	6.27/29.10	72/21	α	8		0.21±0.01	0.59±0.03	0.15±0.16			0.17±0.05	0.06±0.01	0.03±0.01		
47	7.01/32.51	6.27/29.10	78/33	α	12		0.26±0.03	0.61±0.27	0.87±0.22			0.31±0.09	0.14±0.02			
48	7.01/32.12	6.27/29.10	80/35	α	14		0.30±0.02	0.74±0.16	1.05±0.25			0.28±0.08	0.11±0.03			
49	7.03/32.50	6.27/29.10	92/29	α	12		0.28±0.05	0.85±0.11	1.24±0.24			0.37±0.03	0.10±0.01			
51	7.05/30.57	6.27/29.10	73/30	α	7		0.31±0.11	0.46±0.17	0.32±0.04			0.15±0.02	0.11±0.02			
52	7.64/30.22	6.27/29.10	84/22	α	4		0.22±0.04	0.29±0.17	0.32±0.01			0.21±0.02	0.13±0.02			
57	7.28/30.22	6.27/29.10	60/21	α	8		0.48±0.07	0.48±0.09	1.22±0.34			0.67±0.23	0.21±0.03			
59	8.14/30.08	6.27/29.10	84/32	α	9		0.10±0.01	0.21±0.07	0.23±0.05			0.18±0.01				
60	8.32/30.22	6.27/29.10	84/32	α	10	0.12±0.01	0.26±0.01	0.30±0.04	1.02±0.01			0.73±0.16	0.41±0.14	0.18±0.03		
58	7.23/20.13	9.06/20.86	60/22	β	8		0.16±0.02	0.21±0.03	0.34±0.06			0.21±0.02	0.24±0.02			
63	8.54/21.33	9.06/20.86	65/22	β	9		0.18±0.02	0.29±0.04	1.72±0.16			1.32±0.23	1.26±0.21			
66	9.56/22.40	9.06/20.86	75/4	β	1			0.11±0.03	0.07±0.01			0.26±0.06	0.11±0.01			
69	9.56/21.33	9.06/20.86	85/44	β	6			0.48±0.11	0.31±0.13							
12	5.24/19.11	5.24/19.11	63/32	f of α	10	CRB (Atlg03880)	0.05±0.01	0.06±0.01	0.01±0.04			0.01±0.00				
13	5.24/13.20	5.24/13.20	89/34	f of α	8			0.05±0.01	0.36±0.04			0.36±0.03	0.27±0.06	0.04±0.01		
20	5.56/25.32	6.34/27.24	71/2	α	1		0.25±0.01	0.35±0.05	0.28±0.02			0.20±0.08	0.12±0.05	0.02±0.01		
26	5.47/31.22	6.34/27.24	84/20	α	7		0.21±0.01	0.11±0.01	0.09±0.02							
31	5.88/29.15	6.34/27.24	107/34	α	9		0.05±0.01	0.09±0.02								
55	7.07/30.12	6.34/27.24	69/20	α	9		0.45±0.07	0.40±0.03	0.25±0.15			0.08±0.03				
56	7.07/29.88	6.34/27.24	91/30	α	11		0.17±0.01	1.23±0.11	1.20±0.37			0.32±0.14	0.21±0.01			
11	5.22/18.60	6.36/20.71	95/31	β	12			0.14±0.02	0.02±0.01							
16	5.45/22.14	6.36/20.71	66/26	β	8		0.22±0.01	0.24±0.05	0.41±0.09			0.33±0.07	0.24±0.05			
17	5.48/21.01	6.36/20.71	100/30	β	11			0.03±0.01	0.12±0.03			0.09±0.02				
28	5.52/21.03	6.36/20.71	101/4	β	1		1.83±0.23	0.32±0.03	0.98±0.21			0.82±0.14	0.50±0.08	0.13±0.02		
41	6.08/23.30	6.36/20.71	79/4	β	1			0.11±0.01	0.10±0.01			0.13±0.02	0.10±0.03	0.08±0.00		
42	6.08/22.19	6.36/20.71	101/4	β	1		1.69±0.24	1.18±0.17	0.94±0.19			0.57±0.15	0.33±0.11	0.07±0.03		
53	7.05/20.15	6.36/20.71	68/21	β	9	CRU2(ATlg03890)	1.26±0.11	1.41±0.19	1.36±0.14			1.39±0.17	0.62±0.11	0.28±0.09		
15	5.30/29.37	5.17/27.03	92/25	α	15		0.82±0.09	0.50±0.04	0.46±0.04			0.32±0.07	0.12±0.01			
23	5.47/22.60	5.61/19.86	90/27	β	6		0.66±0.11	0.65±0.01	0.59±0.10			0.53±0.25	0.23±0.06	0.04±0.02		

^aNMP, number of matched peptides

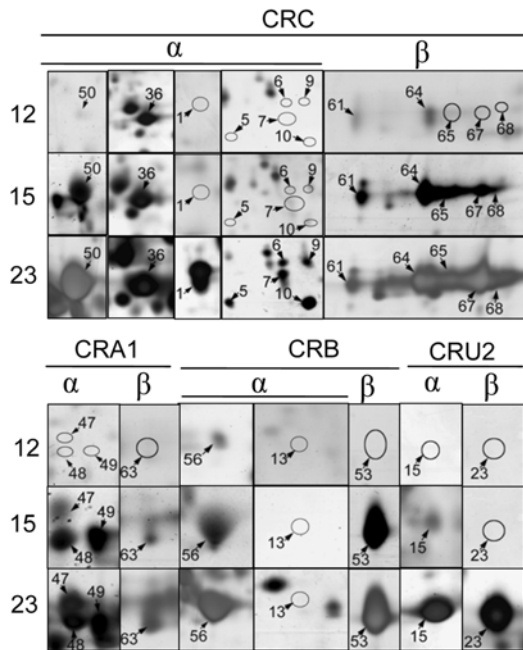


Fig. 3. Accumulation patterns of major *Arabidopsis* 12S SSP fragments during seed maturation. Corresponding gel blocks for representative fragments were extracted from silver stained 2-DEs prepared from 12 to 23 dpa developing siliques as shown in Supplementary Fig. 1. All these 12S SSP fragments were identified by MALDI-TOF-MS or MALDI-TOF-MS/MS methods and were reported in Supplementary Table 1. Triplicate gels were run using independent protein extractions from siliques of the same developmental stages with similar patterns observed.

Proteomic profiles from developing siliques. Silver-nitrate stained gels of proteins extracted from 12 to 23 dpa siliques were shown in Fig. 2. Siliques younger than 8 dpa produced no observable amounts of polypeptides corresponding to any fragments of the 12S SSPs (data not shown). Full-length α subunits for CRC (spot 36) and CRB (spot 56) were first observed on 12 dpa 2-DEs while that of CRA1 (spots 47-49) and CRU2 (spot 15) appeared after 15 dpa (Fig. 3, Table 1). All four polypeptides accumulated to high levels by 23 dpa. β subunits appeared somewhat later on the gel due probably to their smaller sizes. However, they also accumulated to high levels at the end of silique development (Fig. 2). All protein spots identified in 2-DEs of 12 and 15 dpa siliques were putative full-length α or β subunits. Fragments of CRC, CRB and CRA1 α subunits and isoforms of both α and β subunits of these proteins were identified in siliques harvested at 23 dpa (Fig. 2; Supplementary Table 1), indicating that storage protein degradation and modification processes initiated in late stages of seed maturation.

SSP accumulation correlated positively to transcriptional activities of each gene during silique development. Shown in Fig. 4 are the accumulative signal values for all four SSPs. Note that, at full maturity, identified 12S SSP fragments

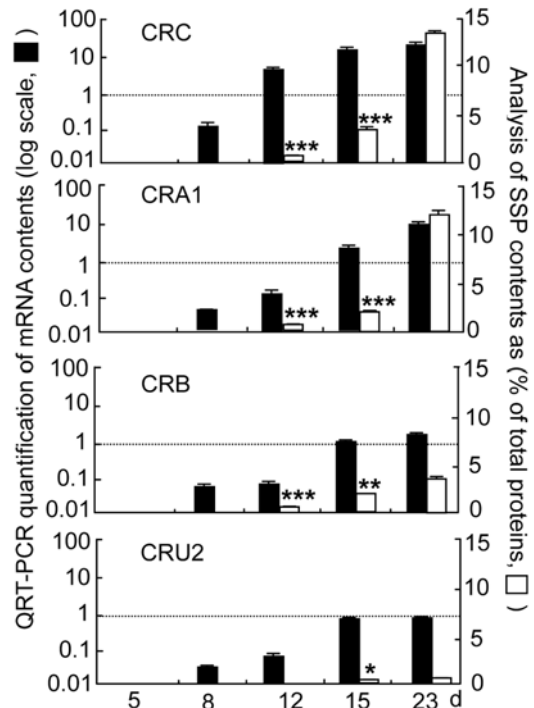


Fig. 4. The transcriptional activities of *Arabidopsis* 12S SSP genes were correlated to the patterns of respective protein accumulation. For comparisons of relative polypeptide contents from 5 to 23 dpa, normalized volumes, expressed as percentages in each 2-DE, for all protein spots corresponding to a particular SSP, were obtained by using the 2D imagemaster software and were put together as one data point in the figure (empty bars). The mean \pm SD for each protein spot was calculated from triplicate gels and the combined SD reported in each column was obtained by evaluating the square root of the summation of squares of all the individual SDs for each data point in a particular gene. For analysis of 12S SSP gene transcript levels over the same developmental period, total RNA isolated from developing siliques was subjected to QRT-PCR by using gene specific primers as listed in “Materials and Methods” section (filled bars). The dashed line indicates the level of *UBQ10* that was included as a loading control and also as a comparison. Data were calculated from three independent RNA extractions. Statistically significant differences were determined using the one-way ANOVA software combined with Tukey’s test throughout this figure. ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$. Significance values were obtained by comparing protein contents obtained from 5, 8, 12 and 15 dpa siliques to that of 23 dpa.

accounted for about 34.5% of total *Arabidopsis* silique proteins with CRC and CRA1 each accumulated to more than 12-15% of the total. We also studied the expression patterns of *Arabidopsis* SSP genes by QRT-PCR (Fig. 4). In RNA samples prepared from 5 dpa siliques, only very low levels of CRC transcripts were detectable. However, in siliques older than 8 dpa, we were able to amplify successfully transcripts for all four SSP genes. The transcription activities for all these four genes increased very rapidly from 8 to 15 dpa and slowed

down significantly in 23 dpa siliques. CRC was the most actively transcribed gene that reached to more than 10 times the amounts of *UBQ10* at 23 dpa, followed by CRA1 and CRB that reached about 5 to 10 times the internal control value (Fig. 4). The close correlation between patterns of protein accumulation and transcript levels reported in Fig. 3 suggest that 12S SSP genes are mainly regulated at the transcription level. We also recorded a significant delay from the time of mRNA transcription to corresponding protein accumulation for all 12S SSPs (Fig. 4). Since we included fruit wall materials in total protein preparations, the observed SSP accumulation may actually be delayed for 1 or 2 days with reference to reality inside the young seeds.

Proteomic and transcription profiles of SSPs in dry and germinating *Arabidopsis* seeds or early seedlings. Previously we reported a series of 2-DEs prepared using total proteins samples extracted from dry seeds to 4-d-old seedlings (Fu *et al.*, 2005). A total of 69 polypeptides were identified as α , β subunits or fragment of different subunits of the four 12S SSP genes using MALDI-TOF-MS and/or MALDI-TOF-MS/MS with reference to the swissprot database (<http://www.expasy.org>). Among the sum, 6 were putative full-length CRC α subunit, 16 fragments of CRC α subunit, 6 CRC putative full-length β subunit; 15 CRA1 α subunit, 6 fragments of CRA1 α subunit and 4 CRA1 β subunit; 5 CRB α subunit, 2 fragments of CRB α subunit and 7 CRB β subunit; 1 each of CRU4 α and β subunits (Table 1).

Comparisons of 2-DE profiles of germinating *Arabidopsis* seeds clearly demonstrated that the full-length α subunits or its major fragments were degraded significantly faster than the β subunits (Fig. 5). Spot 1, a 27 kDa fragment of CRC α subunit, which accumulated to a very high level in dry seeds, disappeared completely during the stratification process while no obvious decrease of all other spots was observed (Fig. 5). By contrast, spot 13, a 13 kDa fragment of CRB α subunit which was not accumulated during seed maturation (Fig. 3), appeared specifically after stratification for two days before being degraded completely by day 4 (Fig. 5). Also, many small fragments, such as spot 2, 3, 4, 6, 7 and 8, were found to increase significantly upon stratification before being degraded completely in later stages. These data suggest that different endo-peptidases may be activated or synthesized during *Arabidopsis* seed germination for the utilization of SSPs. Quantitative analyses showed that mRNAs encoding the four 12S SSP genes disappeared earlier than corresponding polypeptides (Fig. 6).

Prediction of proteolytic cleavage sites. The N-terminal sequences of one putative full-length and five major fragments of CRC α subunit were determined by Edman degradation (Fig. 7). Much to our surprise, all these fragments shared identical N-termini with that of the full-length subunit (data not shown), suggesting that either the full-length α

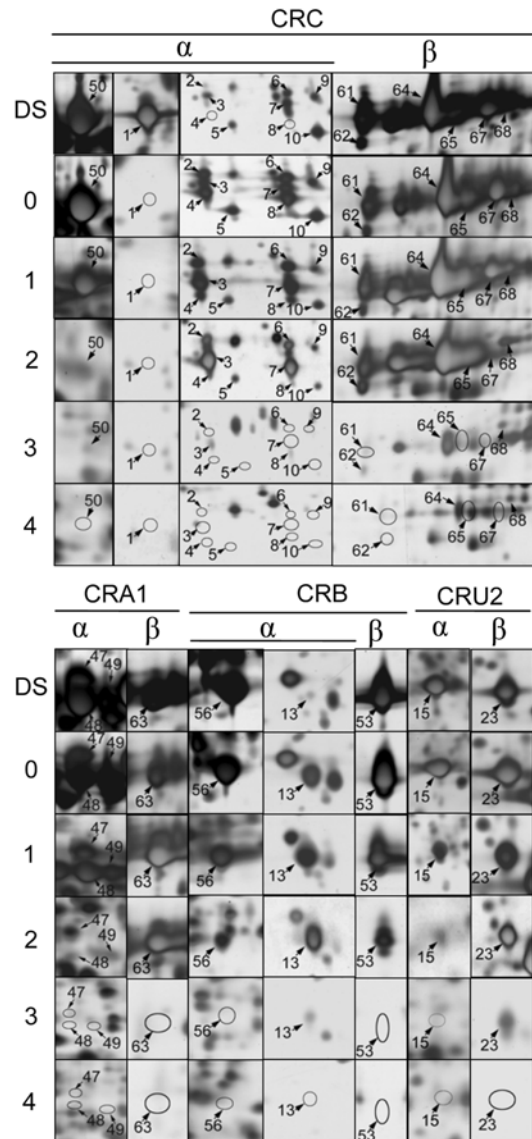


Fig. 5. Degradation patterns of *Arabidopsis* 12S SSP major fragments during seed germination. Corresponding gel blocks for representative fragments were extracted from silver stained 2-DEs prepared from dry or stratified *Arabidopsis* seeds, or early seedlings from d 1 to d 4 as shown in (Fu *et al.*, 2005). Triplicate gels were run using independent protein extractions from the same germination stage plant materials with similar results.

subunit or its major degradation products were first attacked from the C-termini. We next ran the pI/Mw program from the expasy web site (<http://www.expasy.org>) to locate putative cleavage sites. All but one fragment were found to be cleaved between glutamine (Q) and glycine (G) residues. The fragment ended at amino acid residue 288 was cleaved between alanine (A) and glycine (G) residues. We report patterns of CRC α subunit degradation during seedling germination in Fig. 7.

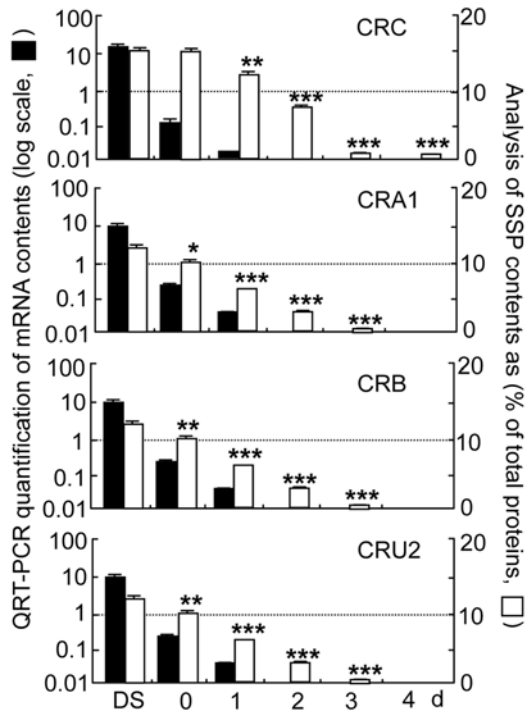


Fig. 6. Transcripts encoding for *Arabidopsis* 12S SSP genes were degraded earlier than respective polypeptides during seed germination. Empty bars represent relative 12S SSP contents over seed germination period. Filled bars represent QRT-PCR analysis of transcript levels for different 12S SSP genes over the same period. See Fig. 3 legend for experimental details. Statistically significant differences were determined using the one-way ANOVA software combined with Tukey's test throughout this figure. ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$. Significance values were obtained by comparing protein contents obtained from seeds or young seedlings 0, 1, 2, 3 or 4 days after germination initiation to that of dry seeds.

Discussion

By using 2-DE coupled with MALDI-TOF-MS, TOF-TOF-MS and QRT-PCR techniques, we identified a large number of *Arabidopsis* 12S SSP fragments, analyzed patterns of gene expression as well as protein accumulation throughout the *Arabidopsis* seed maturation and early germination processes. We showed that, among all four 12S SSP genes in *Arabidopsis* seeds, CRC was expressed at the earliest time (5 dpa), and was followed by CRA1, CRB and CRU2. CRU2 is the last *Arabidopsis* SSP gene that was discovered upon genome database analysis (Gruis *et al.*, 2004), and its expression and accumulation profiles were not known in the literature. All four SSP genes were close to saturation at the transcription level when seeds became mature and started to desiccate at 23 dpa. Messengers of CRC accumulated to the highest level followed by CRA1, CRB and CRU2 (Fig. 4).

Corresponding 12S SSP polypeptides were accumulated significantly later than messenger mRNAs with CRC, CRA1 and CRB subunits first observed in 12 dpa siliques while

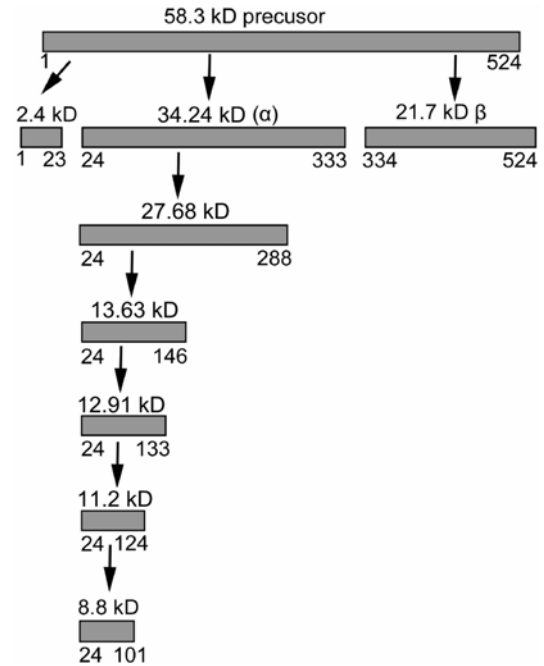


Fig. 7. The CRC α subunit was degraded via its C-termini with conserved hot-spots. The N-terminus of each fragment was determined by Edman sequencing and the putative C-termini were obtained by using the program Mw/pI with consultation to NCBI nr protein library (<http://www.expasy.org>). Theoretical MW of each fragment was labeled on the upper side with their N-terminal and C-terminal amino acid residues marked on the lower side of the bar. Spot 25 was identified as the 34.24 kDa full-length α subunit with spots 1, 6, 3, 8 and 10 representing the 27.68, 13.63, 12.91, 11.20 and 8.8 kDa fragments, respectively.

CRU2 was first detected in 15 dpa siliques (Figs. 3 and 4, Supplementary Table 1). Accordingly, CRC, CRA1 and CRB were found to accumulate much higher than that of CRU2, indicating that production of SSPs in developing *Arabidopsis* seeds was regulated mainly at the transcriptional level. The number of α and β subunit isoforms for CRC, CRA1 and CRB increased significantly as siliques mature, suggesting that modifications of SSPs must have occurred late in seed development (Muntz, 1998).

We also analyzed the degradation patterns of 12S SSPs during seed germination and early seedling development at the proteomic level (Figs. 5 and 6), with a special interest on CRC. Six full-length α and β subunit isoforms each with nearly identical Mw but different pI were identified after 2-DE analysis of dry seeds (Supplementary Table 1), indicating that the two CRC subunits were modified at some point during *Arabidopsis* seed development. During the first two days of imbibition, the CRC α subunit was not degraded continuously. Instead, it was cleaved from one or two hot spots (Q↓G or A↓G) close to its C-terminus. The remaining N-terminal parts of α subunit or all isoforms of full-length β subunit were degraded in a continuous fashion during subsequent seedling growth. Successive C-terminal degradation of 12S

globulins at the glutamic acid residues were recently reported to be repressed under sulfur-deficient conditions, indicating that post-translational modifications of 12S SSPs may bear some physiological importance (Higashi *et al.*, 2006; Manfre *et al.*, 2006). The period of *Arabidopsis* radicle emergence was recently found to correlate with a dramatic new RNA synthesis indicating that a large number of new proteins may be synthesized at this stage (Li *et al.*, 2006). Since fragments of α subunit were found in dry mature seeds (Fig. 3 and Table 1), and endopeptidase activities were reported in protein bodies at late stages of seed development in vetch (Becker *et al.*, 1997; Schlereth *et al.*, 2001), we conclude that partial degradation of CRC α subunit occurred during *Arabidopsis* seed maturation. Our data suggest that CRC degradation may be carried out by different set of proteases at various time of seed germination and *de novo* synthesized enzymes may be recruited for its complete digestion.

Acknowledgments This work was supported by grants from China National Basic Research Program (No. 2004CB117302).

References

- Becker, C., Senyuk, V. I., Shutov, A. D., Nong, V. H., Fischer, J., Horstmann, C. and Muntz, K. (1997) Proteinase A, a storage-globulin-degrading endopeptidase from vetch (*Vicia sativa* L.) seeds, is not involved in early steps of storage protein mobilization. *Eur. J. Biochem.* **228**, 456-462.
- Canovas, F. M., Dumas-Gaudot, E., Recorbet, G., Jorin, J., Mock, H. P. and Rossignol, M. (2004) Plant proteome analysis. *Proteomics* **4**, 285-298.
- Finnie, C., Melchoir, S., Roepstorff, P. and Svensson, B. (2002) Proteome analysis of grain filling and seed maturation in barley. *Plant Physiol.* **129**, 1308-1319.
- Fu, Q., Wang, B. C., Jin, X., Li, H. B., Han, P., Wei, K. H., Zhang, X. M. and Zhu, Y.-X. (2005) Proteomic analysis and extensive protein identification from dry, germinating *Arabidopsis* seeds and young seedlings. *J. Biochem. Mol. Biol.* **38**, 650-660.
- Fujiwara, T., Nambiar, E., Yamagishi, K., Gotoc, D. B. and Naito, S. (2002) Storage proteins, The *Arabidopsis* book. Metabolism, <http://www.bioone.org/bioone>, pp. 1-12.
- Gallardo, K., Job, C., Groot, S. P. C., Puype, M., Demol, H., Vandekerckhove, J. and Job, D. (2001) Proteomic analysis of *Arabidopsis* seed germination and priming. *Plant Physiol.* **126**, 835-848.
- Goldberg, R. B., de Paiva, G. and Yadegari, R. (1994) Plant embryogenesis: zygote to seed. *Science* **266**, 605-614.
- Gruis, D. F., Selinger, D. A., Curran, J. M. and Jung, R. (2002) Redundant proteolytic mechanisms process seed storage proteins in the absence of seed-type members of the vacuolar processing enzyme family of cysteine proteases. *Plant Cell* **14**, 2863-2882.
- Gruis, D. F., Schulze, J. and Jung, R. (2004) Storage protein accumulation in the absence of the vacuolar processing enzyme family of cysteine proteases. *Plant Cell* **16**, 270-290.
- Heath, J. D., Weldon, R., Monnot, C. and Meinke, D. W. (1986) Analysis of storage proteins in normal and aborted seeds from embryo-lethal mutants of *Arabidopsis thaliana*. *Planta* **169**, 304-312.
- Herman, E. M. and Larkins, B. A. (1999) Protein storage bodies and vacuoles. *Plant Cell* **11**, 601-614.
- Higashi, Y., Hirai, M. Y., Fujiwara, T., Naito, S., Noji, M. and Saito, K. (2006) Proteomic and transcriptomic analysis of *Arabidopsis* seeds: molecular evidence for successive processing of seed proteins and its implication in the stress response to sulfur nutrition. *Plant J.* **48**, 557-571.
- Kawagoe, Y., Suzuki, K., Tasaki, M., Yasuda, H., Akagi, K., Katoh, E., Nishizawa, N. K., Ogawa, M. and Takaiwa, F. (2005) The critical role of disulfide bond formation in protein sorting in the endosperm of rice. *Plant Cell* **17**, 1141-1153.
- Krebbes, E., Herdies, L., de Clercq, A., Seurinck, J., Leemans, J., van Damme, J., Segura, M., Gheysen, G., van Montagu, M. and 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.
- Li, Q., Feng, J. X., Han, P. and Zhu, Y.-X. (2006) Parental RNA is significantly degraded during *Arabidopsis* seed germination. *J. Int. Plant Biol.* **48**, 114-120.
- Manfre, A. J., Lanni, L. M., Marcotte, W. R. (2006) The *Arabidopsis* group 1 LATE EMBRYOGENESIS ABUNDANT protein ATEM6 is required for normal seed development. *Plant Physiol.* **140**, 140-149.
- Muntz, K. (1998) Deposition of storage proteins. *Plant Mol. Biol.* **38**, 77-99.
- Pang, P. P., Pruitt, R. E. and Meyerowitz, E. M. (1988) Molecular cloning, genome organization, expression and evolution of 12S seed storage protein genes of *Arabidopsis thaliana*. *Plant Mol. Biol.* **11**, 805-820.
- Schlereth, A., Standhardt, D., Mock, H. P. and Muntz, K. (2001) Stored cysteine proteinases start globulin mobilization in protein bodies of embryonic axes and cotyledons during vetch (*Vicia sativa* L.) seed germination. *Planta* **212**, 718-727.
- Sheoran, I. S., Olson, D. J., Ross, A. R. and Sawhney, V. K. (2005) Proteome analysis of embryo and endosperm from germinating tomato seeds. *Proteomics* **5**, 3752-3764.
- Shi, Y. H., Zhu, S. W., Mao, X. Z., Feng, J. X., Qin, Y. M., Zhang, L., Cheng, J., Wei, L. P., Wang, Z. Y. and Zhu, Y.-X. (2006) Transcriptome profiling, molecular biological, and physiological studies reveal a major role for ethylene in cotton fiber cell elongation. *Plant Cell* **18**, 651-664.
- Shimada, T., Yamada, K., Kataoka, M., Nakaune, S., Koumoto, Y., Kuroyanagi, M., Tabata, S., Kato, T., Shinozaki, K., Seki, M., Kobayashi, M., Kondo, M., Nishimura, M. and Hara-Nishimura, I. (2003) Vacuolar processing enzymes are essential for proper processing of seed storage proteins in *Arabidopsis thaliana*. *J. Biol. Chem.* **278**, 32292-32299.
- Van der Klei, H., van Damme, J., Casteels, P. and Krebbes, E. (1993) A fifth 2S albumin isoform is present in *Arabidopsis thaliana*. *Plant Physiol.* **101**, 1415-1416.
- Wang, B. C., Wang, H. X., Feng, J. X., Meng, D. Z., Qu, L. J. and Zhu, Y.-X. (2006) Post-translational modifications, but not transcriptional regulation, of major chloroplast RNA-binding proteins are related to *Arabidopsis* seedling development. *Proteomics* **6**, 2555-2563.
- Wang, D. Y., Xu, Y. J., Li, Q., Hao, X. M., Cui, K. M., Sun, F. Z. and Zhu, Y.-X. (2003) Transgenic expression of a putative calcium transporter affects the time of *Arabidopsis* flowering. *Plant J.* **33**, 285-292.