Planta (2009) 230:959–971
DOI 10.1007/s00425-009-0999-y

ORIGINAL ARTICLE

Developmental and environmental regulation of soybean *SE60* gene expression during embryogenesis and germination

Yeonhee Choi · Cheol Woong Jeong · Hyonhwa Ohr · Sang-Kee Song · Yang Do Choi · Jong Seob Lee

Received: 23 June 2009 / Accepted: 1 August 2009 / Published online: 19 August 2009

Abstract Soybean SE60 belongs to the γ-thionin family of proteins. We recently demonstrated that SE60 plays a role in defense during soybean development. Here, we show that *SE60* is expressed in a tissue-specific and developmentally regulated manner. The expression of *SE60* is distinct from that of the glycinin (*Gy2*) and extensin (*SbHRGP3*) genes of soybean during embryogenesis and germination. A *SE60::GUS(−809)* transgene, comprising −809 bp of the 5′-flanking region of *SE60* fused to the *GUS* reporter gene, was expressed specifically in developing embryos, but not in the endosperms, from the globular stage of transgenic tobacco and *Arabidopsis* seeds. Furthermore, light affected the *SE60::GUS(−809)* expression pattern in germinating seedlings. Electrophoretic mobility shift assay (EMSA) revealed that soybean nuclear proteins as well as *E. coli*-expressed SB16, a high mobility group protein (HMG), were bound sequence-specifically to the fragment containing AT-rich motifs identified in the *SE60* promoter. Interestingly, the soybean nuclear proteins binding to the two G-boxes and RY repeat were prevalent in seeds of 2–4 mm in size. In contrast, the nuclear proteins binding to the AT-rich motif and *SE60* RNA expression were more prominent in seeds of 4–6 mm in size. Therefore, we propose that factors binding to the G-boxes or RY repeat initiate *SE60* expression during embryogenesis.

Keywords Soybean γ-thionin · Tissue-specific expression · *Cis*-regulatory element · *Trans*-acting factors · Embryogenesis

Abbreviations

bZIP	Basic zipper
bHLH	Basic helix–loop–helix
DAI	Days after imbibition
EMSA	Electrophoretic mobility shift assay
GBF	G-box binding factor
HAI	Hours after imbibition
HMG	High mobility group
NE	Nuclear extracts

Introduction

Seed development in plants results from the multiple divisions of a fertilized egg and central cells, giving rise to the embryo and the endosperm, respectively, and is accompanied by the tissue-specific and developmentally regulated expression of many genes. Soybean embryo-specific (*SE*) genes expressed during the early stage of seed development are considered to be involved in embryogenesis, whereas *SE* genes expressed at the maturation stage encode storage proteins such as glycinin and β-conglycinin (Goldberg

Electronic supplementary material The online version of this article (doi:10.1007/s00425-009-0999-y) contains supplementary material, which is available to authorized users.

Y. Choi (✉) · C. W. Jeong · H. Ohr · J. S. Lee (✉)
School of Biological Sciences,
Seoul National University, Seoul 151-747, Korea
e-mail: yhc@snu.ac.kr

J. S. Lee
e-mail: jongslee@plaza.snu.ac.kr

S.-K. Song
Department of Biology, Yonsei University,
134 Sinchon-dong, Seoul 120-749, Korea

Y. D. Choi
Department of Agricultural Biotechnology and Center
for Agricultural Biomaterials, Seoul National University,
Seoul 151-921, Korea

et al. 1989). Seed development in higher plants requires a large number of genes that are expressed in a developmentally regulated manner. The action of these genes establishes the polarity and morphological pattern of the embryos and prepares the seed for germination (Goldberg et al. 1989). They include seed protein genes that encode a variety of prevalent proteins that accumulate during embryogenesis and are stored in dry seeds. Seed protein genes are expressed temporally during embryogenesis and spatially within the embryonic organ system. Their regulation occurs mainly at the transcriptional level (Verdier and Thompson 2008). However, post-transcriptional processes also play a role in regulating gene expression in seeds (Bayer et al. 2009; Beachy et al. 1985; Walling et al. 1986).

In certain seed protein genes, embryo-specific expression is driven by the 5′-flanking region, and *cis*-regulatory sequences and *trans*-acting factors have been identified by promoter deletion analyses (Ezcurra et al. 1999; Okamuro et al. 1986; Santos-Mendoza et al. 2008; Thomas 1993; Verdier and Thompson 2008). Gene promoters from *napinA* in *Brassica napus* and *legumin B4* in *Vicia faba* have been extensively analyzed (Baumlein et al. 1992; Ezcurra et al. 1999; Stalberg et al. 1996). Sequence comparison of the seed protein genes in combination with binding assays have identified the major factor binding sites, namely, the RY repeat element (CATGCA), the G-box motif (CACGTG) and the B-box (GCCACTTGTC for DisB and CAAACACC for ProxB), which can act synergistically. The RY element and the G-box element are known to interact with B3 protein and basic zipper (bZIP) or basic helix–loop–helix (bHLH) transcription factors, respectively. In addition, the AT-rich motif in the soybean lectin gene was identified as the reactive site for a 60-kDa embryo factor (Jofuku et al. 1987). However, other reports have suggested that the AT-rich motifs act as nonspecific enhancers in most dicot and monocot seed protein genes (Bustos et al. 1991, 1989; Jordano et al. 1989). Moreover, it has been shown that the high mobility group (HMG) and A/T-rich interaction domain (ARID) proteins can bind to the AT-rich domains (Hansen et al. 2008; Pedersen et al. 1991; Ponte et al. 1994; Stros et al. 2007; Wilsker et al. 2002).

In a previous study, we isolated a novel cDNA clone, *SE60*, from soybean seeds, which belongs to the γ-thionin family of proteins (Choi et al. 1993). *SE60* encodes a small cysteine-rich protein, which is synthesized as a preprotein of 75 amino acids. The mature protein containing 47 amino acids is generated following cleavage of a signal peptide. *SE60* is specifically expressed in soybean seeds, and the number of transcripts increases as the seeds mature. Recently, we have discovered that SE60 possesses antimicrobial activity when ectopically expressed in *E. coli* and transgenic tobacco plants (Choi et al. 2008). Therefore, SE60 might play a defensive role during seed development.

Many efforts have been devoted to investigating the expression and regulation of seed protein genes in order to study developmental- and tissue-specific expression in plants. To date, significant advances have been made in understanding the regulation of gene expression during soybean seed development. However, the precise mechanisms controlling their expression are still poorly understood, probably due to the fact that most previous studies have focused on seed protein genes that are expressed at the mid-maturation stage of embryogenesis. In order to gain an understanding of the molecular mechanisms underlying the regulation of the spatiotemporal expression of the *SE60* gene, we analyzed *SE60* promoter activity using several approaches. This work will provide insights into the link between the function and expression of SE60 during soybean development.

Materials and methods

Plant materials

Soybean (*Glycine max* L. cv. HwangKum) was used in this study. For transformation, tobacco (*Nicotiana tabacum* cv. Xanthi) and *Arabidopsis thaliana* Columbia ecotype were used.

Reverse transcription-polymerase chain reaction (RT-PCR) analysis

Total RNA was isolated from soybean tissues using a previously described method (Chirgwin et al. 1979; Glisin et al. 1974). The total RNAs (1.5 μg) isolated from seeds and seedlings at different developmental or germination stages were subjected to cDNA synthesis by AMV reverse transcriptase (Promega) using oligo(dT) as a primer. First-strand cDNA was synthesized and subjected to PCR. The primers used for amplification of the respective cDNAs are as follows:

5′-CCAAGCTTAGAGTGTGTGAGTCACAAAGCCAC-3′ (SE60F)
5′-GGAAGCTTGAGATCTCAACAAATCCTGGTGCA-3′ (SE60R)
5′-CCAACGCCAAAGCAAAAGGAGCAGA-3′ (Gy2F)
5′-CTGCTACCAGCACTAGCTAGAGCCC-3′ (Gy2R)
5′-CTACTAGATAGTGGTTTTGCTTAGC-3′ (SbHRGP3F)
5′-TTGGCTTCACAACCCCTTGCAAACG-3′ (SbHRGP3R).

Construction of *SE60::GUS* chimeric genes

The −809-bp 5′-flanking region of the *SE60* gene was amplified by PCR using soybean genomic DNA as a template and fused to the *β-glucuronidase* (*GUS*) gene in a pBI101 vector (Jefferson et al. 1987). This construct was used for the transformation of *Arabidopsis* using *Agrobacterium*. The chimeric gene was also subcloned into the binary vector pGA482 and used to transform tobacco plants.

The primer set used for amplification of the −809-bp 5′-flanking region of *SE60* is as follows:

GN1 (−809): 5′-GGCCAAGCTTAAGATACCTGAATGAG-3′
GC1: 5′-GGAGATCTCCTCATCTCCATGTTCAC-3′.

Histochemical GUS staining

The GUS histochemical assay using 5-bromo-4-chloro-3-indolylglucuronide (X-Gluc) was carried out using the method of Jefferson et al. (1987).

Preparation of nuclear extracts

Soybean seed (size: 2–4 and 4–6 mm) nuclei were isolated as described by (Luthe and Quatrano 1980a, b) with minor modification. Soybean seeds were frozen in liquid nitrogen and ground to a fine powder with a mortar and pestle. This powder was transferred to a beaker and resuspended with 50 ml of buffer A (0.2 mM EDTA; 0.44 M sucrose; 2.5% Ficoll; 5% Dextran 40; 25 mM Tris, pH 8.0; 10 mM $MgCl_2$; 10 mM β-mercaptoethanol; 2 mM spermidine-HCl; 0.5% Triton X-100; 10 μg/ml leupeptine hemisulfate; 10 μg/ml pepstatin). The slurry was filtered through two layers of 63-μm stainless steel mesh. The filtrates were overlayed on a 2 M sucrose cushion and centrifuged at 2,500 rpm for 5 min. Nuclei, which were found just above the sucrose layer, were pooled and washed with 25 ml of buffer B (buffer A without spermidine-HCl). After a further centrifugation at 2,500 rpm for 5 min, the nuclei were resuspended in 1.5 ml of buffer C (20 mM HEPES, pH 7.5; 0.42 M NaCl; 0.5 mM DTT; 0.2 mM EDTA; 25% glycerol; 10 μg/ml leupeptine hemisulfate, 10 μg/ml pepstatin) and stored at 0°C for 30 min with gentle stirring. After centrifugation for 30 min at 12,000×*g*, the supernatant was dialyzed for 1 h against buffer D (20 mM HEPES, pH 7.5; 0.42 M NaCl; 0.5 mM DTT; 0.2 mM EDTA; 20% glycerol). The dialysate was centrifuged for 10 min at 12,000×*g* and the supernatant was stored at −70°C.

Electrophoretic mobility shift assay (EMSA)

EMSA were performed in a volume of 30 μl containing 10 mM Tris–HCl, pH 7.4, 50 mM NaCl, 1 mM EDTA, 1 mM DTT, 2 μg of double-stranded poly (dI.dC) (Pharmacia, Inc.), 5% glycerol, 30 μg nuclear extracts, and probe DNA (10,000 cpm). The probe DNAs (supporting information Fig. S1) were prepared by PCR, enzyme digestion, and end-labeled with Klenow. For generating the 2G probe, which contains 18 bp from the second G-box sequence of *SE60* plus a few linker bases, two complementary 27 oligomers were synthesized and annealed: GN5: 5′-CACTGAAATCTCACGTGAGATCTGGCC-3′ and GC5: 5′-GGCCAGATCTCACGTGAGATTTCAGTG-3′. The nuclear proteins and DNA probe mixtures were incubated for 20 min at room temperature and separated by 4% polyacrylamide gel electrophoresis in 0.25 × TBE. After electrophoresis, the gel was dried and subjected to autoradiography. *SE60* 5′-flaking sequences from −500 bp to +1 transcription start site are shown in the supporting information Fig. S1.

Results

Expression pattern of the *SE60* gene during seed development and germination

Thionins are known to be synthesized in various tissues of different plants. Purothionins are synthesized in the endosperm of wheat seeds (Jones and Mak 1983). In barley, however, two different thionins are synthesized, one in seed endosperm and the other in the cell wall and vacuole of leaves (Bohlmann et al. 1988; Reimann-Philipp et al. 1989). In petunia, a γ-thionin-like protein is predominantly expressed in pistil tissues (Karunanandaa et al. 1994). The expression of these proteins in various tissues suggests that although all plant thionins are related to each other in terms of overall structure, their evolutionary relationships and mode of expression differ depending on the subgroup of thionins and on the plant species.

In order to gain insights into the developmental regulation of *SE60*, we used RT-PCR to analyze mRNA expression during seed development and germination. We compared *SE60* expression with that of the glycinin (*Gy2*) and extensin (*SbHRGP3*) genes of soybean as controls. The *Gy2* gene is known to be expressed specifically in cotyledons during seed development (Meinke et al. 1981; Nielsen et al. 1989), whereas the *SbHRGP3* gene has been demonstrated to be highly expressed in the mature zone of seedling hypocotyls (Ahn et al. 1998, 1996; Ahn and Lee 2003; Hong et al. 1994). Total RNAs were isolated from developing seeds of 0–2, 2–4, 4–6, and 6–8 mm size; mature green seeds; mature yellow seeds; and dry seeds. A faint band was detected in the 0 to 2-mm seeds and this then gradually accumulated with development (Fig. 1). The highest amount of RNA was detected in 4 to 6-mm seeds, which

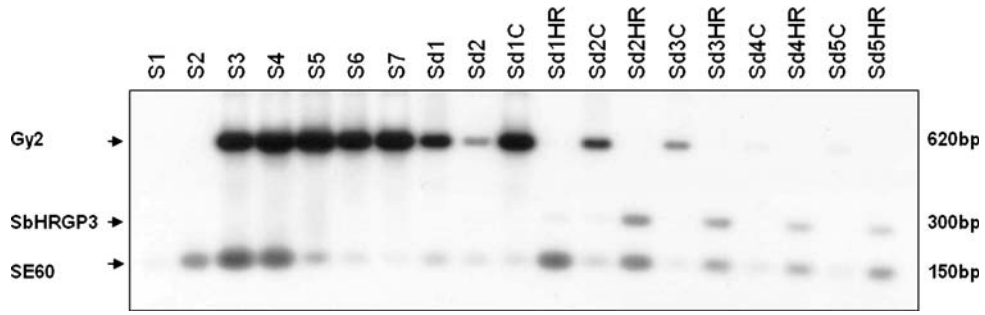

Fig. 1 RT-PCR analysis followed by Southern blot of *SE60*, *Gy2*, and *SbHRGP3* during soybean seed development and germination. Soybean seeds were germinated in the dark condition. *Lanes*: *S1–S4* seeds of 0–2, 2–4, 4–6, and 6–8 mm in size, *S5* mature green seeds, *S6* mature yellow seeds, *S7* dry seeds, *Sd1* and *Sd2* seedlings at 1 and 2 days after imbibition (DAI), *Sd1C–Sd5C* cotyledons of seedlings at 1, 2, 3, 4, and 5 DAI, *Sd1HR–Sd5HR* hypocotyls and root radicles of seedlings at 1, 2, 3, 4, and 5 DAI

corresponds to the cotyledon stage. Thereafter, *SE60* mRNA decreased rapidly as the seed matured. In contrast, *Gy2* mRNA was first detected in seeds of 4–6 mm, increased as maturation proceeded, and then decreased from mature yellow seeds. As expected, *SbHRGP3* mRNA was undetectable in all developing seeds. The RT-PCR analysis reveals that *SE60* is expressed at an earlier stage than the *Gy2* gene during soybean embryogenesis. These results suggest that although both *Gy2* and *SE60* genes are expressed in the developing seeds, their precise regulatory mechanisms are different.

In order to investigate *SE60* expression during germination, total RNAs were isolated from germinating seedlings at 1 and 2 days after imbibition (DAI). In addition, total RNAs from cotyledons and from the hypocotyls plus root radicles of seedlings at 1 to 5 DAI were isolated and subjected to RT-PCR. As shown in Fig. 1, when seeds were imbibed for 1 day, *SE60* mRNA was increased compared with that in dry seeds, but then decreased gradually after 2 DAI. Interestingly, most of the *SE60* mRNA was detected in hypocotyls and root radicles, but not in cotyledons, which differs from the *Gy2* expression pattern. A considerable amount of *Gy2* mRNA was present in dry seeds; however, it decreased more rapidly than *SE60* mRNA as germination commenced. *Gy2* expression was specific to the cotyledons. In contrast, *SbHRGP3* mRNA was barely detectable in seedlings at 1 DAI, but increased in seedlings at 2 DAI, and then decreased gradually as germination proceeded. *SbHRGP3* mRNA was restricted to the seedling axis.

Meinke et al. (1981) also reported that neither glycinin nor conglycinin mRNA was detected until soybean seeds had reached 5 mm in size, even using tenfold more mRNA than we used in the present study. In contrast to the northern blot analyses of Meinke et al. and Goldberg et al. (1989), we detected glycinin mRNA in dry seeds and seedlings at 1 DAI, possibly due to the sensitivity of RT-PCR. On the basis of the radioactivity of the hybridized bands, *Gy2* mRNAs at 1 and 2 DAI were present at approximately 35 and 6% of the levels in dry seeds, respectively. These results clearly indicate that the synthesis of cotyledon-specific *Gy2* mRNA ceased during germination. Taken together, our results suggest that during seed development and germination *SE60* expression is regulated different from that of other well known seed protein genes.

Spatiotemporal expression of the *SE60::GUS(−809)* in transgenic embryos

We previously isolated a 2.3-kb genomic clone containing the *SE60* gene from a screen of a genomic library and determined the nucleotide sequence (Choi et al. 1995). Figure 2a shows the various possible regulatory sequences residing in the 5′-flanking region of the *SE60* gene. These include four AT-rich motifs (Jofuku et al. 1987), two G-boxes (Oeda et al. 1991), an RY repeat element (Dickinson et al. 1988), and a CACA element (Goldberg 1986).

In order to investigate the spatiotemporal expression of *SE60*, −809-base pairs of the 5′-flanking sequence containing a 5′UTR were fused to the *GUS* reporter gene (Fig. 2a). In order to confirm whether this chimeric construct was functional, we performed a transient assay with soybean immature seeds using particle bombardment. Histochemical GUS analysis revealed blue staining in the heart-stage embryos (Fig. 2b), indicating that the GUS construct under the control of the *SE60* promoter was functional in plants.

We next introduced the *SE60::GUS(−809)* transgene into *Arabidopsis* and analyzed the GUS expression during embryogenesis (Fig. 3). GUS was weakly detectable in globular-stage embryos, but then accumulated very rapidly up to the torpedo stage (Fig. 3a–e). Thereafter, GUS activity was maintained, but decreased as the embryos fully matured (Fig. 3f–h). GUS was expressed only in the embryo, not in the endosperm (Fig. 3b–d). Therefore, in *Arabidopsis*, the expression pattern of *GUS* under the

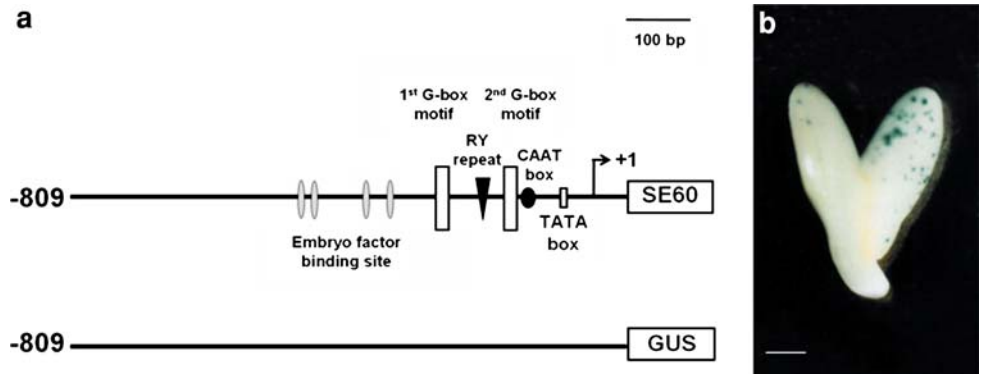

Fig. 2 The putative regulatory elements in the 5′-flanking region of *SE60* and the structure of the *SE60::GUS(−809)* chimeric gene (**a**). Four embryo factor binding sites (AT-rich motifs), two G-box motifs, an RY repeat, and a CACA element were found along with the canonical promoter elements (the CAAT and TATA boxes). **b** Transient GUS assay in the heart-stage immature soybean embryo by particle bombardment with the *SE60::GUS(−809)* transgene. *Scale bar* 500 μm

control of the soybean *SE60* promoter appears to be consistent with the RT-PCR data from soybean plants (Fig. 1).

We further confirmed *SE60* promoter activity by introducing the *SE60::GUS(−809)* transgene into tobacco plants. Histochemical assays of developing transgenic tobacco embryos revealed an expression pattern (Fig. 3i–p) similar to that in *Arabidopsis*. Taken together, these data strongly suggest that the sequence up to −809-bp of the soybean *SE60* 5′-flanking region contains most of the regulatory sequences directing the embryo-specific expression in transgenic *Arabidopsis* and tobacco plants. These data also suggest that *Arabidopsis* and tobacco plants share *trans*-acting factors in common with soybean that can bind to the *SE60* promoter leading to the developmental- and tissue-specific expression.

Expression of the *SE60::GUS(−809)* gene in germinating seedlings

In order to investigate the expression pattern of *SE60::GUS(−809)* in germinating seedlings, seeds from transgenic tobacco plants were imbibed in (MS) salt solution and histochemical assays were carried out at 2, 4, 6, and 8 DAI. We also imbibed seeds under light conditions in order to examine the possibility that light regulates the expression of *SE60::GUS* during germination.

GUS was detectable in the seedling axis at 2 DAI under dark conditions (Fig. 4a). The GUS activity became prevalent mainly in the mature region of hypocotyls at 4 DAI (Fig. 4b), and this expression pattern was maintained under the dark conditions (Fig. 4c, d). Under light conditions, GUS activity was also detected in 2 and 4 DAI seedlings (Fig. 4e, f), exhibiting a similar expression pattern to that of seedling grown under dark conditions. However, in the light-grown seedlings of 6 and 8 DAI, GUS activity disappeared in the hypocotyl region and was maintained mostly in the cotyledons (Fig. 4g, h). We also confirmed the same pattern of GUS expression in light-grown *Arabidopsis* seedlings (supporting information Fig. S2a–c). These results suggest that *SE60* transcripts expressed throughout the embryo during seed formation might be localized either to the cotyledon or hypocotyls with or without light, respectively, during the germination process. This also suggests that light could affect the expression pattern of *SE60::GUS(−809)* gene. The GUS expression pattern in the dark-grown seedlings is consistent with the RT-PCR analysis of soybean seedlings (Fig. 1), in which *SE60* mRNA was detected in the axis of seedlings immediately after the commencement of germination.

Nuclear proteins bind to the *SE60* promoter region

On the basis of our data, we assume that the −809-bp 5′-flanking region of the *SE60* gene contains a few *cis*-regulatory sequences (Table 1), which are presumably responsible for its developmental regulation (Fig. 3). In order to examine the possibility that nuclear proteins bind to these regulatory elements of the *SE60* promoter, we carried out gel mobility shift assays. Two different nuclear extracts, NE1 and NE2, were prepared from immature soybean seeds of 2–4 and 4–6 mm, respectively. Various DNA probes were prepared by PCR amplification or enzyme digestion, and these were radioactively labeled as shown in Fig. 5a. First, we used the Em.1G and Em probes. EMSA with the two nuclear extracts revealed distinctly shifted bands for both the Em.1G (Fig. 5b, lanes 1–3) and Em (Fig. 5b, lanes 5–7) DNA probes. The Em.1G probe exhibited stronger binding affinity with NE1, whereas the Em probe exhibited strong binding with NE2. The Em.1G probe contains four AT-rich sequences and the first G-box element, whereas the Em probe contains only the four embryo factor binding sites.

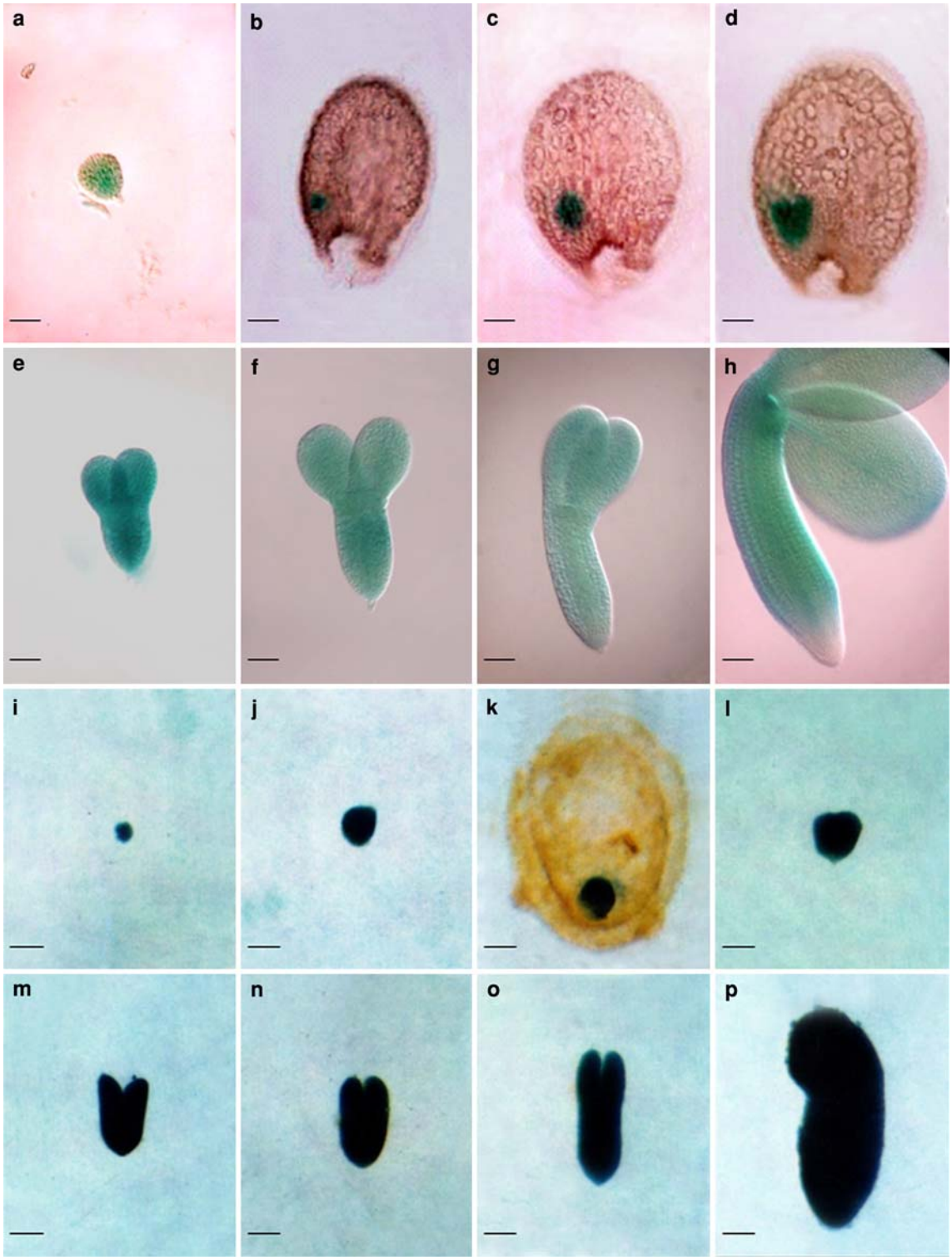

Fig. 3 Histochemical GUS analysis of *SE60::GUS(−809)* expression in transgenic *Arabidopsis* (**a**–**h**) and tobacco embryos (**i**–**p**) at various stages. **a**–**c**, **i**–**k** Embryos at the globular stage. **d**, **l**, **m** Heart-stage embryos. **e**, **f**, **n**, **o** Embryos at the torpedo stage. **g**, **p** Embryo at the walking-stick stage. **h** Mature-stage embryo. *Scale bars* **a**–**h** 50 μm, **i**–**p** 100 μm

In order to examine this different binding affinity, we generated the 1G probe, which contains only the first G-box. EMSA with the 1G probe revealed that the first G-box binding factor (GBF1) is more enriched in NE1 than in NE2 (Fig. 5c). Therefore, this result can explain the stronger band for the Em.1G probe with NE1. Given

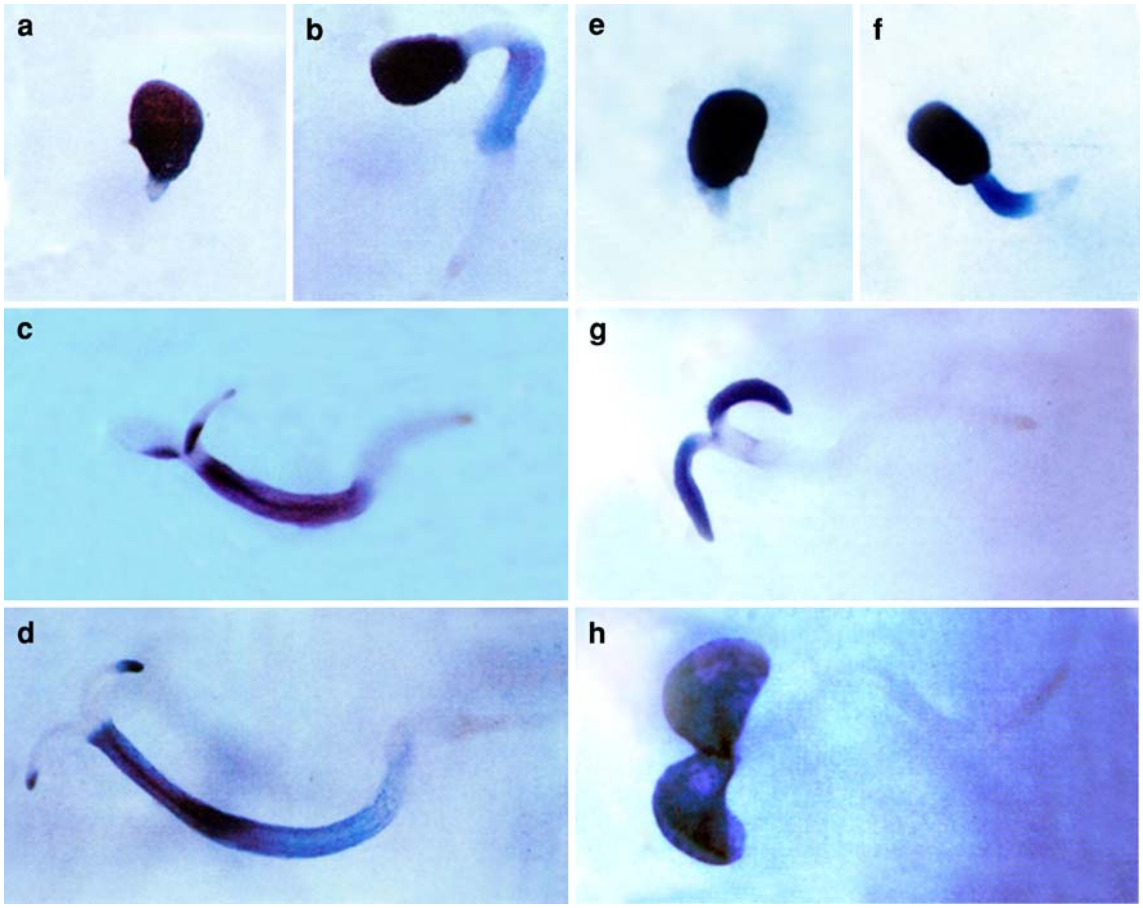

Fig. 4 Histochemical GUS analysis of *SE60::GUS(−809)* gene expression in transgenic tobacco seedlings at various stages of germination. Transgenic seeds were germinated in MS salts without (**a**–**d**) or with (**e**–**h**) light. **a**, **e** Seedlings at 2 days after imbibition (DAI). **b**, **f** Seedlings at 4 DAI. **c**, **g** Seedlings at 6 DAI. **d**, **h** Seedlings at 8 DAI

Table 1 The regulatory sequences identified in the 5′-flanking region of the *SE60* gene

Regulatory element	Sequence	Location	Reference
Embryo factor-binding site (AT-rich motifs)	AATTAAATAA		(Jofuku et al. 1987)
	AATTAAATAA	(−435 to −426)	
	AATTAAATGA	(−419 to −410)	
	AATTTAAAAT	(−348 to −339)	
	AATTAAATTT	(−309 to −300)	
G-box	GCACGTG		(Oeda et al. 1991)
	GCACGTG	(−216 to −210)	
	TCACGTG	(−114 to −108)	
RY repeat	CATGCAT		(Dickinson et al. 1988)
	CATGCAT	(−139 to −133)	
CAAT box	CCACT	(−97 to −93)	
TATA box	TATATA	(−33 to −28)	

that the Em probe exhibits strong binding affinity with NE2, this implies that the binding factors for AT-rich motifs are prevalent in NE2. We also analyzed other regulatory elements, such as the RY repeat and the second G-box. We examined the binding to the RY.2G probe, which contains an RY repeat and the second G-box binding factor (GBF2). EMSA revealed that two major bands shifted (Fig. 5d).

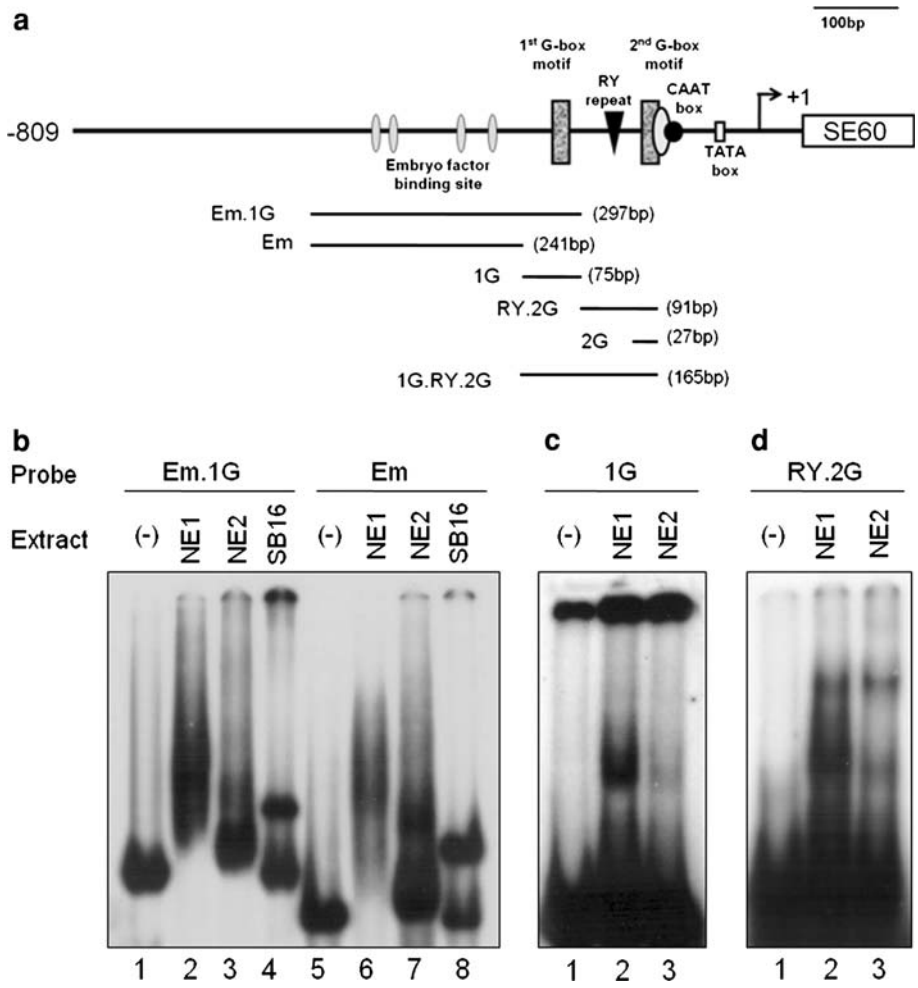

Fig. 5 Various DNA probes and competitors for the gel mobility assays (**a**). **b** EMSA with the Em.1G and Em probes (*lanes 1–4* and *lanes 5–8*, respectively). *Lanes 1* and *5* free DNA probes without any protein binding, *lanes 2* and *6* soybean nuclear extract (NE)1 binding in 2 to 4-mm seeds, *lanes 3* and *7* soybean NE2 binding in 4 to 6-mm seeds, *lanes 4* and *8*: *E. coli*-expressed HMG protein (SB16) binding. **c** EMSA with the 1G probe. *Lane 1* free probe without nuclear extracts, *lane 2* NE1 binding, *lane 3* NE2 binding. **d** Gel mobility assay with the RY.2G probe. *Lane 1* free probe without nuclear extracts, *lane 2* NE1 binding, *lane 3* NE2 binding

Sequence-specific binding of nuclear proteins to *cis*-acting elements

In order to investigate whether binding to the 1G.RY.2G probe is sequence-specific, competition experiments were conducted using various probes as competitors (Fig. 6). The DNA competitors used in these experiments were the 1G.RY.2G, 1G, and RY.2G fragments. As expected, binding to the 1G.RY.2G probe was strongly inhibited by the same cold competitor (Fig. 6a, lanes 3, 4 and 10, 11). The RY.2G competitor was a more effective inhibitor (Fig. 6a, lanes 7, 8 and 14, 15) than the 1G probe (Fig. 6a, lanes 5, 6 and 12, 13), and the extent of inhibition by the RY.2G competitor was slightly less than that by the 1G.RY.2G probe itself. The 1G probe also inhibited the binding of nuclear proteins to the 1G.RY.2G probe at a reduced level compared to the 1G.RY.2G and RY.2G competitors. A remarkable difference in binding was observed between NE1 and NE2, in that one shifted band could be seen only in NE1 (arrowhead), suggesting that certain factors that are enriched in NE1 are almost absent in NE2. Compared with the bands for the RY.2G probe (Fig. 5d), the 1G.RY.2G probe showed at least one more shifted band. The difference in banding patterns between the IG.RY.2G and RY.2G probes indicates the involvement of GBF1, which was shown to be enriched in NE1 (Fig. 5c).

In order to analyze the interaction and relationship between the 1G and RY.2G probes, we used RY.2G as a DNA probe and RY.2G and 1G probes were used as competitors (Fig. 6b). Although the RY.2G competitor effectively inhibited nuclear protein binding to the RY.2G probe, a distinct band decrease was not observed when the 1G probe was used as a competitor. This suggests that the 1G fragment is not an effective competitor and that first G-box binding does not affect second G-box binding.

Two complementary oligonucleotides containing only the second G-box were synthesized and annealed to generate a double-stranded oligomer (2G). When this 2G fragment was used as a competitor, nuclear protein binding to the RY.2G probe was inhibited (Fig. 7a). This indicates that binding to the RY.2G probe is likely to be G-box-specific and confirmed that the second G-box might be an important site for nuclear protein binding in the RY.2G probe.

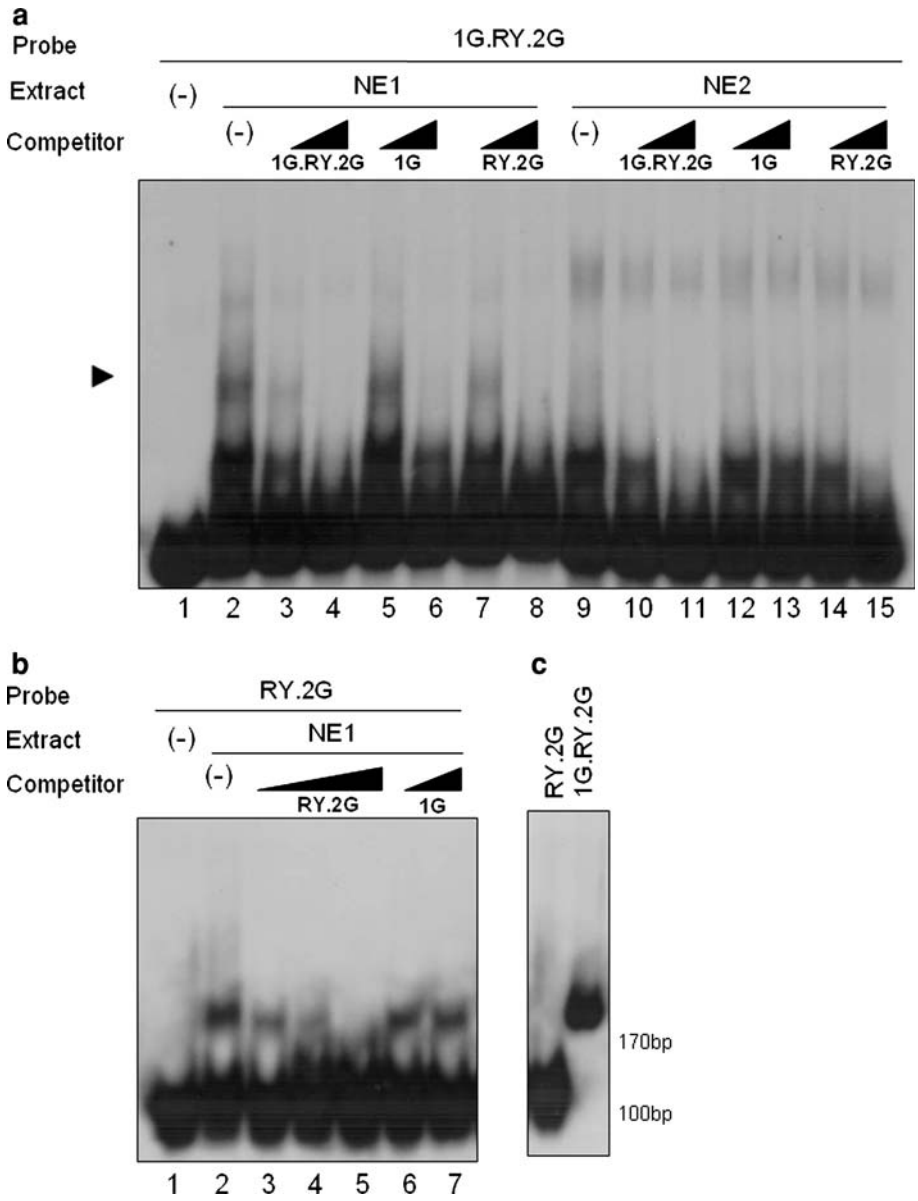

Fig. 6 EMSA with various competitors (**a** and **b**) and SB16 control binding experiment (**c**). The 1G.RY.2G DNA probe (**a**) and RY.2G DNA probe (**b**) were used. Competitors used in **a** and **b** are indicated below the *triangles*, and all probes and competitors are shown in Fig. 5a. **a** *Lane 1* free DNA probe, *lanes 2* and *9* no competitor, *lanes 3*, *5*, *7*, *10*, *12*, and *14* 10 M excess of competitor, lanes *4*, *6*, *8*, *11*, *13*, and *15* 100 M excess of competitors. **b** *Lane 1* free DNA probe, *lane 2* no competitor, *lanes 3* and *6* 10 M excess of competitor, *lane 4* 20 M excess of competitor, *lane 5* and *7* 100 M excess of competitor. **c** *Lane 1* RY.2G probe, *lane 2* 1G.RY.2G probe used

As described above, second G-box binding is not inhibited very effectively by the 1G probe. In order to investigate whether first G-box binding is effectively inhibited by the second G-box, we used the 1G fragment as a DNA probe and the 2G probe as a competitor. Interestingly, binding to the first G-box was successfully inhibited by the 2G competitor (Fig. 7b). Thus, the first G-box cannot function as an effective competitor of the second G-box, but the second G-box can effectively inhibit binding to the first G-box. It is possible that although there is a specific and distinct GBF1 that is unable to bind to the second G-box, GBF2 is a general factor that is able to bind to both G-boxes.

In order to confirm the binding of the nuclear extracts to the synthetic 2G oligomer, EMSA were performed using the 2G DNA probe. Figure 7c shows that nuclear proteins certainly bind to 2G, and that nuclear proteins in NE1 bind to a greater extent than those in NE2, which is similar to the 1G probe. Therefore, both GBF2 and GBF1 appear to be activated at a very early stage of soybean development.

Soybean HMG protein binds to the AT-rich region in the *SE60* promoter

Plant HMG or HMG-like proteins are known to interact with the AT-rich motif of some dicot and monocot genes (Bustos et al. 1989; Grasser 2003; Jordano et al. 1989; Nieto-Sotelo et al. 1994; Pedersen et al. 1991; Stros et al. 2007; Webster et al. 1997). We examined whether a soybean HMG Y-like protein, SB16 (Laux et al. 1991), could indeed bind to the Em.1G and Em probes, each of which contain four AT-rich motifs. We found that the SB16 protein purified from an *E. coli* expression system did indeed bind to these two DNA

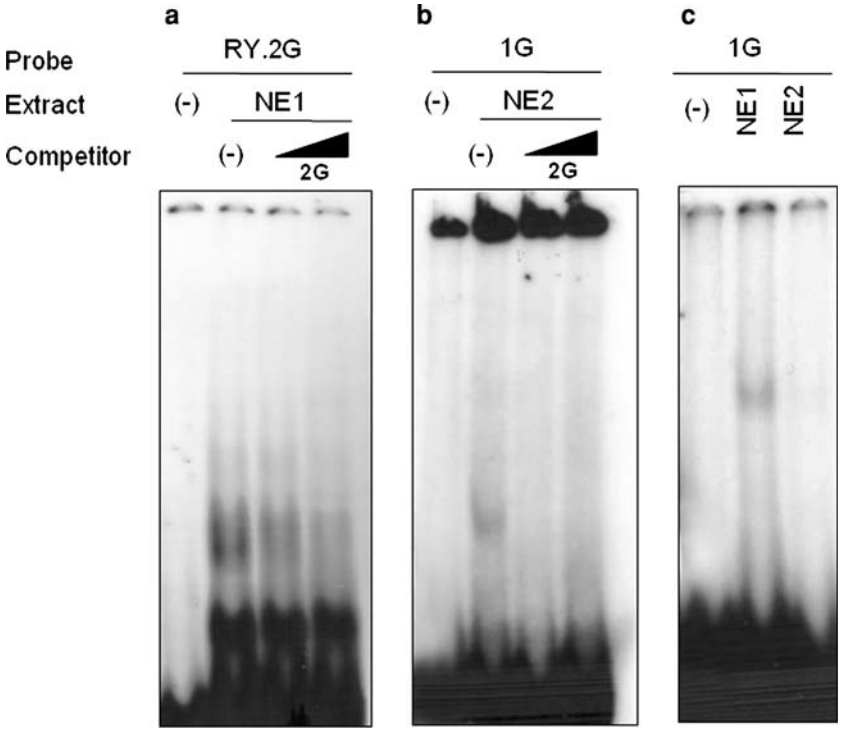

Fig. 7 Competitive EMSA using 2G (**a** and **b**) as a competitor and EMSA with 1G as a DNA probe (**c**). The 2G competitor used in **a** and **b** is indicated below the *triangles*. **a** The RY.2G DNA probe, NE1, and 2G competitor were used. *Lane 1* free DNA probe, *lane 2* no competitor, *lane 3* 50 M excess of competitor, *lane 4* 500 M excess of competitor. **b** The 1G DNA probe, NE2, and 2G competitor were used. *Lane 1* free DNA probe, *lane 2* no competitor, *lane 3* 50 M excess of competitor, *lane 4* 500 M excess of competitor. **c** The 1G DNA probe was used. *Lane 1* no extract, *lane 2* NE1 binding, *lane 3* NE2 binding. In these experiments, 30 μg of NE1 and 15 μg of NE2 were used

probes (Fig. 5b, lanes 4, 8). In order to check the specificity of SB16 binding, we used 1G.RY.2G probe as well as RY.2G probe. As a result, we did not see any binding of SB16 to the both DNA probes (Fig. 6c). Therefore, SB16 binding seems to be sequence-specific.

It is possible that the nuclear proteins that bind to the AT-rich region of the *SE60* gene contain the SB16 protein. However, there is no corresponding EMSA in the NE1 or NE1 (Fig. 5b, lane 2,3,6,7). It is possible that SB16 is not enriched in these NE1 and NE2 extracts or SB16 can form a complex with other proteins and this protein complex could bind to the Em.1G or Em probe. SB16 binding appears to involve only one of the four embryo factor binding sites because only one shifted band was observed in this assay. Multiple shifted bands with NE1 and NE2, however, indicate that other HMG or other AT-rich binding proteins may also bind to the probes (Fig. 5b).

In order to investigate the expression pattern of the *SB16* gene during soybean seed development, RT-PCR was carried out using the same templates that were used in the experiment described in Fig. 1. *SB16* mRNA was first detected in 4–6 mm embryos and accumulated as seed development proceeded (supporting information Fig. S3). This can explain the result, at least in part, that there is no same band appeared with NE1 and NE2 when the Em.1G or Em probe was used compared with SB16 binding to the same probes (Fig. 5b). Although *SB16 mRNA* starts to be expressed from 4 to 6 mm soybean seeds, *SB16* mRNA reaches to the maximal level when the seeds are fully matured. Thus, SB16 is not enriched in the early stage of embryogenesis. This is consistent with the results showing a strong band shift when the Em probe and NE2 were used (Fig. 5b, lane 7). Although *SB16* mRNA is expressed in the 4–6 mm embryos, the nuclear proteins of NE1 exhibit stronger binding to the Em.1G DNA probes than those of NE2. It is probable that GBF1 and GBF2 are more highly expressed in 2–4 mm embryos (Figs. 5c, 7c, respectively) or that there may be other AT-rich binding proteins that are expressed earlier than the SB16 protein.

Discussion

Developmental- and tissue-specific expression of *SE60*

Meinke et al. (1981) classified soybean seed development on the basis of seed size. According to their classification, the seeds of 0–2 mm correspond to the globular and heart stages of the embryo. Seeds of 2–4 mm are in the early cotyledon stage, whereas seeds of 4–6, 6–8, and 8–10 mm are in the cotyledon stage. *Gy2* mRNA appears to be expressed from the 4–6 mm cotyledon stage, reaches a maximum level in mature green seeds, and then declines thereafter (Fig. 1). In the case of *β*-conglycinin, another major storage protein in soybean, embryonic axes and cotyledons accumulate equivalent amounts of the α subunit in >9 mm soybean seeds (Meinke et al. 1981), which is in contrast to glycinin mRNA.

The expression pattern of *SE60* was different to that of both *Gy2* and *SbHRGP3* genes. It appears to be induced and repressed earlier than the *Gy2* gene during embryogenesis. One distinct feature compared with the *Gy2* gene is that *SE60* mRNA is distributed mostly in the seedling axis rather than in the cotyledons. Dietrich et al. (1989) observed germination-related genes in *Brassica napas*. One subset of mRNAs accumulated preferentially in the cotyledons and another mRNA class was seedling axis-abundant. These spatially regulated mRNAs exhibited different mRNA expression patterns; mRNA prevalent in seedling axes accumulated primarily during early embryogenesis, whereas cotyledon-abundant mRNA concentrations increased during late embryogenesis (Dietrich et al. 1989). According to our analysis, the expression pattern of the *SE60* gene appears to be similar to that of axis-abundant genes.

Another example is pSAS10, identified as being synthesized from stored mRNA in the cotyledons of cowpea seeds (Ishibashi and Minamikawa 1990; Ishibashi et al. 1990), which shares significant homology with the SE60 protein. Although *pSAS10* and *SE60* genes can be classified into the same group, the expression of these genes is different. *pSAS10* mRNA is formed at the late stage of seed maturation, is conserved in quiescent seeds, and becomes functional at an early stage of germination. The mRNA levels begin to decline after the onset of imbibition and are very low at 48 h after imbibition (HAI). With respect to mRNA distribution, *pSAS10* mRNA is present in both the cotyledons and embryonic axes of dry cowpea seeds and seeds at 12 HAI, and the level in the axes is approximately half that in the cotyledons (Ishibashi and Minamikawa 1990). Therefore, the expression patterns are gene-specific, even though several genes belonging to the same group may have a similar function.

Light regulation of *SE60::GUS(−809)* in transgenic seedlings

Germination commences with seed imbibition and ends with protrusion of the radicle through the testa. Post-germination represents the re-initiation of sporophytic growth (Bewley and Marcus 1990). Our observations suggest that light affects not only the region of expression, but also the level of *SE60::GUS* gene expression (Fig. 4). When GUS staining was performed on seedlings of over 10 DAI, GUS activity was still maintained in dark-grown seedlings. However, GUS activity declined significantly in light-grown seedling (data not shown). Reimann-Philipp (Reimann-Philipp et al. 1989) reported that, in barley, the intracellular concentration of thionin mRNA was decreased rapidly upon illumination of etiolated barley seedlings and that barley thionin synthesis was downregulated by light. Consistent with these observations, it appears that GUS activity driven by the *SE60* promoter decreases more rapidly in light-grown than in dark-grown seedlings. Since the mechanism whereby light affects *SE60* expression has yet to be determined, further extended experiments will be needed in the near future in order to address this issue. There are many bacteria, fungi, and other pathogens in the soil; therefore, the SE60 protein could act as a defensive chemical after reactivation in the rapidly growing cells of seedlings. This supposition is supported by our previous data showing that the SE60 protein is toxic to *E. coli* and confers resistance against bacterial pathogens in transgenic tobacco plants (Choi et al. 2008).

Nuclear proteins binding to the two G-boxes, RY repeat, and AT-rich motif

The G-box element was first identified as an 11-bp sequence containing a core CACGTG located upstream of many genes encoding the small subunit of ribulose biphosphate carboxylase (*rbcS*) (Giuliano et al. 1988). Several cDNA clones encoding proteins that specifically interact with the G-box element have been isolated (de Pater et al. 1993; Kawagoe et al. 1994; Oeda et al. 1991; Schindler et al. 1992; Tabata et al. 1991). Interestingly, all these proteins belong to the bZIP or bHLH class of transcription factors. Generally, the G-box and its related sequences with their binding proteins function in absisic acid regulation (Chandrasekharan et al. 2003). The RY element is widely distributed in dicot and monocot genes and deletions of the RY repeat in the *Vicia faba* (broad bean) *LeB4* gene and the soybean glycinin gene result in a decrease in reporter gene expression in transgenic tobacco seeds (Baumlein et al. 1992; Lelievre et al. 1992).

EMSA with the DNA probes containing various putative regulatory sequences of *SE60* revealed that nuclear proteins bind to the AT-rich region, the two G-boxes, and the RY motif in a sequence-specific manner (Fig. 6). Furthermore, a soybean HMG protein (SB16) synthesized from cDNA in *E. coli* binds to the AT-rich region. Other DNA probes such as 1G.RY.2G or RY.2G which were used in this study did not reveal any shifted bands with SB16 protein (Fig. 6c). Therefore, binding of the SB16 protein to the AT-rich region is sequence-specific and the HMG actually binds to the AT-rich region. However, whether this HMG is involved in the quantitative expression or seed-specific expression is unknown; consequently, this needs to be addressed in the near future. The HMG mRNA was first detected at the 4 to 6-mm stage and accumulated as seed development progressed. The HMG mRNA was maintained at high levels throughout the course of seed development and the mRNA levels decreased gradually after imbibition. The HMG mRNA was expressed to a greater

extent in the seedling axis than in the cotyledons. This pattern was particularly prominent from the 4-day-old seedlings. The expression pattern of HMG protein is correlated with that of the SE60 protein with respect to expression site during seed development and germination.

We also reported on the difference in binding activity between the two G-boxes and between nuclear extracts prepared from two developmental stages. With the exception of SB16, we confirmed that most of the nuclear proteins binding to the *SE60* promoter are more prevalent in NE1. It is possible that there may be other AT-rich binding proteins that are expressed earlier than the SB16 protein. For example, the two GBFs that are expressed to a greater extent at an earlier stage of seed development might initiate the expression of the *SE60* gene, and thereafter the HMG protein, similar to SB16, could play a role in maintaining the embryo-specific expression of *SE60* at a high level. We are currently generating several constructs, which contain the putative elements found in the *SE60* promoter under the control of a truncated minimal Cauliflower Mosaic Virus (CaMV) 35S promoter, together with the promoter-deleted constructs in order to determine their roles in detail. Our results will provide a useful system for studying embryo development in soybean.

Acknowledgments This work was supported in part by the Crop Functional Genomics Center (CG1121) funded by the Korea Ministry of Science and Technology to J. S. Lee; by a Korea Science and Engineering Foundation (KOSEF) grant funded by the Korea government (MOST) (R01-2007-000-10706-0) to Y. Choi; and by the Brain Korea 21 Project to H. Ohr.

References

Ahn JH, Lee JS (2003) Sugar acts as a regulatory signal on the wound-inducible expression of SbHRGP3::GUS in transgenic plants. Plant Cell Rep 22:286–293

Ahn JH, Choi Y, Kwon YM, Kim SG, Choi YD, Lee JS (1996) A novel extensin gene encoding a hydroxyproline-rich glycoprotein requires sucrose for its wound-inducible expression in transgenic plants. Plant Cell 8:1477–1490

Ahn JH, Choi Y, Kim SG, Kwon YM, Choi YD, Lee JS (1998) Expression of a soybean hydroxyproline-rich glycoprotein gene is correlated with maturation of roots. Plant Physiol 116:671–679

Baumlein H, Nagy I, Villarroel R, Inze D, Wobus U (1992) *Cis*-analysis of a seed protein gene promoter: the conservative RY repeat CATGCATG within the legumin box is essential for tissue-specific expression of a legumin gene. Plant J 2:233–239

Bayer M, Nawy T, Giglione C, Galli M, Meinnel T, Lukowitz W (2009) Paternal control of embryonic patterning in *Arabidopsis thaliana*. Science 323:1485–1488

Beachy RN, Chen ZL, Horsch RB, Rogers SG, Hoffmann NJ, Fraley RT (1985) Accumulation and assembly of soybean beta-conglycinin in seeds of transformed petunia plants. EMBO J 4:3047–3053

Bewley JD, Marcus A (1990) Gene expression in seed development and germination. Prog Nucleic Acid Res Mol Biol 38:165–193

Bohlmann H, Clausen S, Behnke S, Giese H, Hiller C, Reimann-Philipp U, Schrader G, Barkholt V, Apel K (1988) Leaf-specific thionins of barley-a novel class of cell wall proteins toxic to plant-pathogenic fungi and possibly involved in the defence mechanism of plants. EMBO J 7:1559–1565

Bustos MM, Guiltinan MJ, Jordano J, Begum D, Kalkan FA, Hall TC (1989) Regulation of beta-glucuronidase expression in transgenic tobacco plants by an A/T-rich, cis-acting sequence found upstream of a French bean beta-phaseolin gene. Plant Cell 1:839–853

Bustos MM, Begum D, Kalkan FA, Battraw MJ, Hall TC (1991) Positive and negative *cis*-acting DNA domains are required for spatial and temporal regulation of gene expression by a seed storage protein promoter. EMBO J 10:1469–1479

Chandrasekharan MB, Bishop KJ, Hall TC (2003) Module-specific regulation of the beta-phaseolin promoter during embryogenesis. Plant J 33:853–866

Chirgwin JM, Przybyla AE, MacDonald RJ, Rutter WJ (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18:5294–5299

Choi Y, Choi YD, Lee JS (1993) Nucleotide sequence of a cDNA encoding a low molecular weight sulfur-rich protein in soybean seeds. Plant Physiol 101:699–700

Choi Y, Ahn JH, Choi YD, Lee JS (1995) Tissue-specific and developmental regulation of a gene encoding a low molecular weight sulfur-rich protein in soybean seeds. Mol Gen Genet 246:266–268

Choi Y, Choi YD, Lee JS (2008) Antimicrobial activity of gamma-thionin-like soybean SE60 in *E. coli* and tobacco plants. Biochem Biophys Res Commun 375:230–234

de Pater S, Pham K, Chua NH, Memelink J, Kijne J (1993) A 22-bp fragment of the pea lectin promoter containing essential TGAC-like motifs confers seed-specific gene expression. Plant Cell 5:877–886

Dickinson CD, Evans RP, Nielsen NC (1988) RY repeats are conserved in the 5′-flanking regions of legume seed-protein genes. Nucleic Acids Res 16:371

Dietrich RA, Maslyar DJ, Heupel RC, Harada JJ (1989) Spatial patterns of gene expression in *Brassica napus* seedlings: identification of a cortex-specific gene and localization of mRNAs encoding isocitrate lyase and a polypeptide homologous to proteinases. Plant Cell 1:73–80

Ezcurra I, Ellerstrom M, Wycliffe P, Stalberg K, Rask L (1999) Interaction between composite elements in the napA promoter: both the B-box ABA-responsive complex and the RY/G complex are necessary for seed-specific expression. Plant Mol Biol 40:699–709

Giuliano G, Pichersky E, Malik VS, Timko MP, Scolnik PA, Cashmore AR (1988) An evolutionarily conserved protein binding sequence upstream of a plant light-regulated gene. Proc Natl Acad Sci USA 85:7089–7093

Glisin V, Crkvenjakov R, Byus C (1974) Ribonucleic acid isolated by cesium chloride centrifugation. Biochemistry 13:2633–2637

Goldberg RB (1986) Regulation of plant gene expression. Philos Trans R Soc Lond Ser B Biol Sci 314:343–353

Goldberg RB, Barker SJ, Perez-Grau L (1989) Regulation of gene expression during plant embryogenesis. Cell 56:149–160

Grasser KD (2003) Chromatin-associated HMGA and HMGB proteins: versatile co-regulators of DNA-dependent processes. Plant Mol Biol 53:281–295

Hansen FT, Madsen CK, Nordland AM, Grasser M, Merkle T, Grasser KD (2008) A novel family of plant DNA-binding proteins containing both HMG-box and AT-rich interaction domains. Biochemistry 47:13207–13214

Hong JC, Cheong YH, Nagao RT, Bahk JD, Cho MJ, Key JL (1994) Isolation and characterization of three soybean extensin cDNAs. Plant Physiol 104:793–796

Ishibashi N, Minamikawa T (1990) Molecular cloning and characterization of stored mRNA in cotyledons of *Vigna unguiculata*. Plant Cell Physiol 31:39–44

Ishibashi N, Yamauchi D, Minamikawa T (1990) Stored mRNA in cotyledons of *Vigna unguiculata* seeds: nucleotide sequence of cloned cDNA for a stored mRNA and induction of its synthesis by precocious germination. Plant Mol Biol 15:59–64

Jefferson RA, Bevan M, Kavanagh T (1987) The use of the *Escherichia coli* beta-glucuronidase as a gene fusion marker for studies of gene expression in higher plants. Biochem Soc Trans 15:17–18

Jofuku KD, Okamuro JK, Goldberg RB (1987) Interaction of an embryo DNA binding protein with a soybean lectin gene upstream region. Nature 328:734–737

Jones BL, Mak AS (1983) The amino-acid sequence of a purothionin from *Triticum monococcum*, a diploid wheat. Eur J Biochem 133:621–625

Jordano J, Almoguera C, Thomas TL (1989) A sunflower helianthinin gene upstream sequence ensemble contains an enhancer and sites of nuclear protein interaction. Plant Cell 1:855–866

Karunanandaa B, Singh A, Kao TH (1994) Characterization of a predominantly pistil-expressed gene encoding a gamma-thionin-like protein of *Petunia inflata*. Plant Mol Biol 26:459–464

Kawagoe Y, Campbell BR, Murai N (1994) Synergism between CACGTG (G-box) and CACCTG *cis*-elements is required for activation of the bean seed storage protein beta-phaseolin gene. Plant J 5:885–890

Laux T, Seurinck J, Goldberg RB (1991) A soybean embryo cDNA encodes a DNA binding protein with histone and HMG-protein-like domains. Nucleic Acids Res 19:4768

Lelievre JM, Oliveira LO, Nielsen NC (1992) 5′-CATGCAT-3′ Elements modulate the expression of glycinin genes. Plant Physiol 98:387–391

Luthe DS, Quatrano RS (1980a) Transcription in isolated wheat nuclei: I. Isolation of nuclei and elimination of endogenous ribonuclease activity. Plant Physiol 65:305–308

Luthe DS, Quatrano RS (1980b) Transcription in isolated wheat nuclei: II. Characterization of RNA synthesized in vitro. Plant Physiol 65:309–313

Meinke DW, Chen J, Beachy RN (1981) Expression of storage-protein genes during soybean seed development. Planta 153:130–139

Nielsen NC, Dickinson CD, Cho TJ, Thanh VH, Scallon BJ, Fischer RL, Sims TL, Drews GN, Goldberg RB (1989) Characterization of the glycinin gene family in soybean. Plant Cell 1:313–328

Nieto-Sotelo J, Ichida A, Quail PH (1994) PF1: an A–T hook-containing DNA binding protein from rice that interacts with a functionally defined d(AT)-rich element in the oat phytochrome A3 gene promoter. Plant Cell 6:287–301

Oeda K, Salinas J, Chua NH (1991) A tobacco bZip transcription activator (TAF-1) binds to a G-box-like motif conserved in plant genes. EMBO J 10:1793–1802

Okamuro JK, Jofuku KD, Goldberg RB (1986) Soybean seed lectin gene and flanking nonseed protein genes are developmentally regulated in transformed tobacco plants. Proc Natl Acad Sci USA 83:8240–8244

Pedersen TJ, Arwood LJ, Spiker S, Guiltinan MJ, Thompson WF (1991) High mobility group chromosomal proteins bind to AT-rich tracts flanking plant genes. Plant Mol Biol 16:95–104

Ponte I, Guillen P, Debon RM, Reina M, Aragay A, Espel E, Di Fonzo N, Palau J (1994) Narrow A/T-rich zones present at the distal 5′-flanking sequences of the zein genes Zc1 and Zc2 bind a unique 30 kDa HMG-like protein. Plant Mol Biol 26:1893–1906

Reimann-Philipp U, Schrader G, Martinoia E, Barkholt V, Apel K (1989) Intracellular thionins of barley. A second group of leaf thionins closely related to but distinct from cell wall-bound thionins. J Biol Chem 264:8978–8984

Santos-Mendoza M, Dubreucq B, Baud S, Parcy F, Caboche M, Lepiniec L (2008) Deciphering gene regulatory networks that control seed development and maturation in Arabidopsis. Plant J 54:608–620

Schindler U, Beckmann H, Cashmore AR (1992) TGA1 and G-box binding factors: two distinct classes of Arabidopsis leucine zipper proteins compete for the G-box-like element TGACGTGG. Plant Cell 4:1309–1319

Stalberg K, Ellerstom M, Ezcurra I, Ablov S, Rask L (1996) Disruption of an overlapping E-box/ABRE motif abolished high transcription of the napA storage-protein promoter in transgenic *Brassica napus* seeds. Planta 199:515–519

Stros M, Launholt D, Grasser KD (2007) The HMG-box: a versatile protein domain occurring in a wide variety of DNA-binding proteins. Cell Mol Life Sci 64:2590–2606

Tabata T, Nakayama T, Mikami K, Iwabuchi M (1991) HBP-1a and HBP-1b: leucine zipper-type transcription factors of wheat. EMBO J 10:1459–1467

Thomas TL (1993) Gene expression during plant embryogenesis and germination: an overview. Plant Cell 5:1401–1410

Verdier J, Thompson RD (2008) Transcriptional regulation of storage protein synthesis during dicotyledon seed filling. Plant Cell Physiol 49:1263–1271

Walling L, Drews GN, Goldberg RB (1986) Transcriptional and post-transcriptional regulation of soybean seed protein mRNA levels. Proc Natl Acad Sci USA 83:2123–2127

Webster CI, Packman LC, Pwee KH, Gray JC (1997) High mobility group proteins HMG-1 and HMG-I/Y bind to a positive regulatory region of the pea plastocyanin gene promoter. Plant J 11:703–715

Wilsker D, Patsialou A, Dallas PB, Moran E (2002) ARID proteins: a diverse family of DNA binding proteins implicated in the control of cell growth, differentiation, and development. Cell Growth Differ 13:95–106