Plant Molecular Biology **19**: 937–949, 1992.

The legumin boxes and the 3′ part of a soybean β-conglycinin promoter are involved in seed gene expression in transgenic tobacco plants

Sylvain Chamberland, Nathalie Daigle and François Bernier[1]*
Département de Biologie, Université Laval, Ste-Foy, Québec, Canada G1K 7P4; [1] *Present address: Intitut de Botanique, Université Louis Pasteur, 28 rue Goethe, 67083 Strasbourg Cedex, France (*author for correspondence)*

Received 11 December 1991; accepted in revised form 10 April 1992

Key words: β-conglycinin, embryogenesis, *Glycine max*, legumin box, regulatory elements, seed storage proteins

Abstract

β-conglycinin is one of the major seed storage proteins in soybean. It is composed of three subunits, namely α, α′ and β. The expression of β-conglycinin is highly regulated, being restricted to the embryo during the mid-maturation phase of embryogeny. Two series of constructs were made with the α′ subunit promoter and the GUS reporter gene to investigate the *cis*-acting elements involved in the regulated expression of this promoter. The activity of each construct was tested in transgenic tobacco plants.

In the first series of constructs, we checked if the 'legumin box', a sequence found in most legume seed storage protein genes as well as in other seed-specific genes, is involved in the regulated expression of the α′ subunit of the β-conglycinin gene in tobacco. To this end, both copies of the α′ subunit promoter legumin boxes were mutagenized *in vitro*. The transcriptional activity of the single mutants and the double mutant were compared with that of the wild-type promoter. Our results show that the legumin boxes act together to increase transcription of the β-conglycinin α′ subunit gene by about a factor of ten. This is the first demonstration of a function for the legumin box in transcriptional regulation.

In the second series of experiments, we wished to determine if the 3′ part of the promoter (the CCAAT and TATAA region) contains important regulatory elements. We found that this small fragment (−82 to + 13 bp) can confer by itself a low level of seed-specific gene expression. Chimaeric promoters constructed from parts of the α′ subunit promoter and of the constitutive CaMV 35S promoter were also analysed. These constructs also revealed the importance of the CCAAT and TATAA region of the α′ subunit promoter in seed-specific gene expression.

Introduction

The seed storage proteins constitute the majority of the proteins of a mature seed and are usually represented by a limited number of different polypeptides. In soybean (*Glycine max*), β-conglycinin is one of the two major seed storage proteins and is composed of three subunits, α, α′ and β. The genes encoding these subunits are expressed only in seeds (embryonic axis and cotyledons) and only

during the mid-maturation stage, making them an attractive model for the study of seed-specific gene expression.

Genes encoding the α' and β subunits of β-conglycinin have been cloned and expressed in transgenic petunia and tobacco plants [7, 25]. In all cases, correctly regulated expression of these genes was observed, indicating that transgenic plants can be used to investigate the regulatory elements involved in gene expression during soybean embryogenesis.

A deletion series of the α' subunit promoter analyzed in transgenic tobacco plants has revealed that the first 257 bp upstream of the transcription start site are sufficient to give a level of regulated expression almost equal to that of the complete promoter [9]. In addition, the −257 to −82 bp fragment was shown to possess a seed-specific enhancer activity when inserted into the constitutive 35S promoter of cauliflower mosaic virus (CaMV) [10].

In vitro studies have revealed the presence of several proteins binding to the promoters of the α' and the β subunit genes of β-conglycinin [1, 19]. In the α' promoter, two of these factors (SEF 3 and SEF 4) have binding sites in the fragment that was shown to give full promoter activity in transgenic tobacco. A third factor (SEF 1) binds to two AT-rich sequences located in a more distal region of the α' subunit promoter (around −660 and −780 bp) whereas the other factor identified so far (SEF 2) binds to almost any promoter fragment with little or no sequence specificity.

The next step in the analysis of the α' subunit promoter is to determine the *in vivo* function of the *cis*-acting elements thought to be involved in seed-specific gene expression. To do so, we have decided to follow two different approaches. In the first one, specific mutations were introduced into one or both of the 'legumin boxes' in order to verify their role in seed-specific gene expression. This sequence has been found in the promoters of most legume seed storage protein genes, as well as other seed-specific genes [2, 6, 11, 14, 15], and this conservation has led to the hypothesis that it might be involved in regulated gene expression in seeds. In the second approach, we wanted to address the possible role of the 3′ part of the promoter (−82 to +13; including the CCAAT and TATAA boxes and one copy of the legumin box) by testing its activity in transgenic tobacco. In addition, parts of the CaMV 35S promoter were used to construct chimaeric promoters to further assess the role of the different parts of the α' subunit promoter in seed-specific gene expression. The 3′ part of this promoter has not been extensively studied since the discovery of a seed-specific enhancer in the region just upstream of it [10].

Our results show that the legumin box plays an important role in β-conglycinin transcription: mutation of both copies of these boxes in the α' promoter results in a significant tenfold reduction in transcription of the GUS reporter gene. We also show that the small 3′ part of the α' subunit promoter contains sufficient information to direct a very low level of seed-specific gene expression.

Materials and methods

Origin of the clones

The 35S promoter fragments were isolated from the pBI121 vector (Clontech). The α' promoter fragments were isolated from the Gmg 17.1 clone [29].

In vitro *mutagenesis*

In vitro mutagenesis was carried out according to the procedure of Lewis and Thompson [20] using the Altered Sites kit from Promega. The α'C6 promoter fragment (−257 to +13) was first subcloned as a *Bam* HI-*Xmn* I fragment into the vector pSELECT-1 (Promega) cut with *Bam* HI and *Sma* I. The mutagenesis reaction was performed with oligonucleotide X to create a *Xho* I site in the 5′ legumin box (mutant X) and with oligonucleotide B to create a *Bgl* II site in the 3′ legumin box (mutant B). A double mutant (mutant XB) was obtained by the simultaneous use of oligonucleotides X and B (see Fig. 1B):

oligo X: 5′-ACCCCAAAAGCCCTCGAGAACAACACGTACTC-3′
Xho I

oligo B: 5′-TGGCCAAATGTTCAGATCTGTTAACAAGACCTA-3′
Bgl II

The mutant and wild-type promoters were then excised from pSELECT-1 by a *Hin*d III and *Eco* RI digestion (the *Eco* RI site was blunted) and inserted in front of the GUS reporter gene in the vector pBI101 (Clontech) cut with *Hin*d III and *Sma* I.

Chimaeric promoter constructs

The different fragments isolated from the β-conglycinin α′ subunit promoter and from the 35S promoter were assembled in pUC19 according to standard molecular biology procedures [28]. These constructs were then excised from pUC19 by a *Hin*d III-*Xba* I digestion and inserted in front of the GUS reporter gene of pBI101 (Clontech) cut with these same enzymes (see Fig. 1C).

Plant transformation

The different constructs in pBI101 were mobilized from the *Escherichia coli* strain DH5α to *Agrobacterium tumefaciens* LBA 4404 by a triparental mating procedure using an *E. coli* strain carrying the plasmid pRK2013 as a helper [21]. For tobacco (*Nicotiana benthamiana*) transformation and regeneration, the standard leaf disc protocol was used [16].

Determination of GUS activity

Determination of GUS activity was performed as described by Jefferson [17]. For the standard assay, 50 seeds were ground in 1.25 ml of extraction buffer and 50 μl of this extract was added to 650 μl of extraction buffer containing MUG at 1.25 mM. The reaction was carried out at 37 °C and 200 μl aliquots (40 μl for plants giving a high activity) were removed at 5, 20 and 35 min and added to 2 ml of stop buffer. The relative fluorescence was determined (excitation 355 nm; emission 455 nm) using 0 to 1.0 μM methylumbelliferone (MU) as a standard. The activity in leaves was determined from a disc (diameter 6 mm) taken from a young leaf and ground in 500 μl of extraction buffer. 50 μl of this extract was used in a 650 μl reaction and 200 μl aliquots were removed at 5, 20 and 35 min. Protein concentrations in these extracts were determined by the method of Bradford (Bio-Rad kit). Determination of GUS activity in the roots was done in the same conditions as for leaves using about 5 μg of protein. For some assays, larger amounts of materials and longer reaction times were used to increase the sensitivity (see Results).

Histochemical localization of GUS

The localization of GUS activity was performed as described [17] with minor modifications. Mature seeds were deposited in dense layers in monocyanoacrylate adhesive (Krazy Glue) and sectioned using a Vibratome. Sections of 100 to 200 μm were fixed in 0.3% formaldehyde, 3 M mannitol, 10 mM MES pH 5.6 for 45 min. The fixation solution was then replaced by a 50 mM sodium phosphate buffer (pH 7.0) containing 50 μg/ml of the substrate 5-bromo-4-chloro-3-indolyl glucuronide (X-gluc). Reactions were carried out for periods ranging from 12 h to 6 days at 37 °C after which the embyos were placed in glycerol and examined under the microscope.

Results

Constructs and in vivo *mutagenesis*

Figure 1A is a schematic representation of the α′ subunit promoter showing the position of the CCAAT, TATAA and legumin boxes as well as the predicted transcription start site [12]. The binding sites for the previously described proteins SEF 1, SEF 3 and SEF 4 [19] are also indicated in Fig. 1A. The −257 to + 13 fragment of the α′ subunit promoter was used as the starting material for the *in vitro* mutagenesis of the legumin boxes. Mutations were introduced into each of the legumin boxes to create construct X (mutation of the 5′ legumin box) and construct B (mutation of the 3′ legumin box) or in both legumin boxes, creating construct XB. Figure 1B shows that, in each case, 4 bp were substituted, creating new restriction enzyme cutting sites and thus facilitating identification of the mutants. This number of changes also insures that the legumin box consensus sequence is completely lost. More specifically, these mutations destroy the alternating purine-pyrimidine sequence typical of this element. The wild-type fragment (construct N) was used as a control in these experiments.

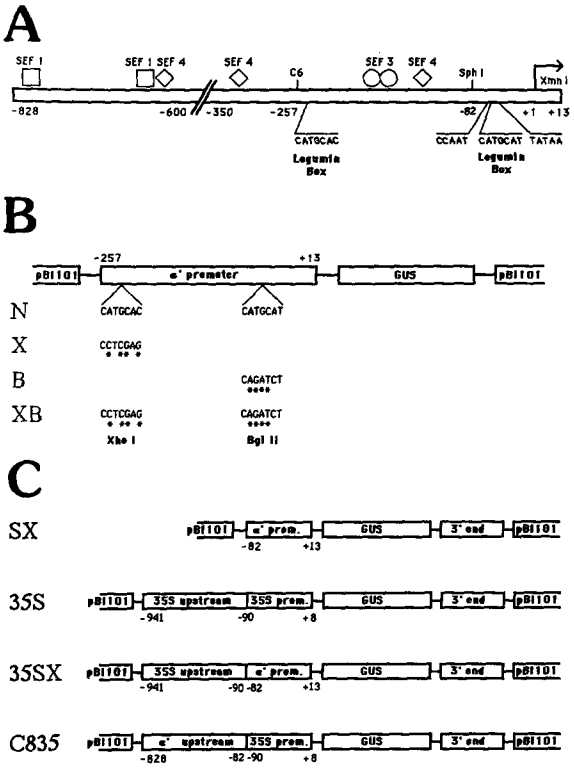

Fig. 1. Maps of the different constructs used int his work. A. α′ promoter fragment showing the position of the CCAAT and TATAA boxes as well as the legumin boxes and the binding sites for the DNA binding proteins previously identified (SEF 1, SEF 3 and SEF 4) [19]. B. Maps of the legumin box mutant constructs in pBI101: the sequences of both wild-type legumin boxes are shown (construct N). For the mutant promoters (constructs X, B and XB), the nucleotides that were changed are indicated by asterisks. C. Maps of the chimaeric promoter constructs in pBI101.

The 3′ part of the α′ subunit promoter (−82 to + 13) was fused to the GUS reporter gene to test its promoter activity (construct SX). It was also fused to the −941 to −90 fragment of the 35S promoter to generate a chimaeric promoter (construct 35SX). The intact 35S promoter (−941 to + 8; construct 35S) was used as a control. Finally, the upstream part of the α′ promoter (−828 to −82) was fused to the 3′ part of the 35S promoter (−90 to + 8; construct C835). All these constructs are represented schematically in Fig. 1C.

The legumin box mutant and chimaeric promoter constructs were then fused to the GUS reporter gene in the vector pBI101 in order to test their transcriptional activity in transgenic plants.

Tobacco transformation

All the constructs described in the previous section were introduced into tobacco (*N. benthamiana*) by *A. tumefaciens*-mediated transformation [16]. Usually, 10 to 20 independent transformants were regenerated for each construct. The SX construct was expected to give very low, if any, GUS activity. For this reason, transformation was confirmed by allowing each of the regenerated plants to self-fertilize and examining the frequency of kanamycin resistance among the progeny (data not shown).

Transcriptional activities of the legumin box mutants

For this series of constructs, GUS activities were first measured in seeds collected at different times during embryogenesis (12, 16, 20 and 24 days after anthesis and dry seeds) for at least five plants transformed with each construct (results not shown). In all cases, GUS activity was first detected at 20 days and was shown to be stronger at 24 days and even slightly stronger in mature seeds. Since the kinetics of GUS accumulation were exactly the same for all four constructs, GUS activities in dry seeds could be considered as an accurate reflection of the relative transcriptional activities of the four promoters during embryogenesis. So, all the subsequent work and comparisons were carried out with dry seeds.

As expected, important plant-to-plant variations were observed for a given construct because of the so-called 'position effect' [26]. Figure 2 shows the distribution of GUS activities in dry seeds for all the plants transformed with the four constructs. The average activities are shown in Table 1 and indicate that mutation of each legumin box reduced transcription about twofold whereas mutations of both boxes produced a more pronounced effect. Since the reduction in the average GUS activity observed for the double mutant is about sevenfold (Table 1), it is likely

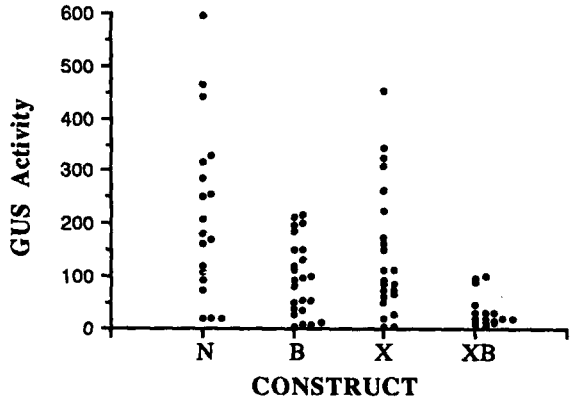

Fig. 2. GUS activities in dry seeds from individual tobacco plants transformed with the legumin box mutant constructs. Each point represents the GUS activity (expressed in pmol of MU formed per seed per minute) for an independently transformed plant.

Table 1. Average and median GUS activities in dry seeds for the legumin box mutant constructs.

Construct	Number of plants	GUS activity (pmol MU formed per seed per minute)		*P**
		Average	Median	
N	19	218	183	-
X	23	142	92	0.056
B	24	98	96	0.005
XB	17	33	19	0

* Probability that the lower median activity of this construct compared to the N construct is only accidental.

that both copies of the legumin box do not act independently from each other.

The differences in the transcriptional activities between the mutant and the wild-type promoters look convincing, especially in the case of the double mutant. However, the important plant-to-plant variations prompted us to perform statistical tests to confirm that these differences were not only accidental. In a case like this, where the different values do not follow a normal distribution (see Fig. 2), the median values are used instead of the averages for statistical comparisons. The 'median test' [24] indicates the probability that a difference between the two sets of data is only accidental. Table 1 shows that the lower transcription activities (about twofold) in the single mutant constructs is very likely to be caused by the mutations. In the case of the double mutant, the almost tenfold decrease in median GUS activity has to be attributed to the mutations themselves rather than to chance. Another statistical test, the Mann-Whitney test [24], uses the ranks to compare sets of non-parametric values. The results of this test perfectly agree with those obtained with the median test (not shown). Therefore, our results indicate that each copy of the legumin box increases transcription of the β-conglycinin α' subunit gene but that two copies of this element are required to obtain full promoter activity in transgenic tobacco plants.

We next wanted to verify if the legumin boxes are acting as negative regulatory elements in or-

gans other than seeds, preventing β-conglycinin expression in these organs. GUS activities were measured in leaves of all the transformed plants as well as in leaves of untransformed tobacco plants. We were unable to detect activity in any case, even by increasing the sensitivity of the assay by a factor of around 250 (by increasing the amount of material used in the assay as well as the reaction times). Since the inactivation of both legumin boxes does not result in unregulated expression of the α′ subunit promoter, it is likely that these DNA sequences do not act as negative regulatory elements.

Histological localization of GUS activity in seeds expressing the legumin box mutants

The *in vitro* determinations of GUS activity have shown that expression of the different legumin box mutant promoters is still restricted to seeds. In order to determine whether mutations in the legumin boxes would change the specificity of expression in the seed itself, we performed histochemical localization of GUS activity [17]. Mature seeds obtained by self-fertilization of plants transformed with each of the four constructs were sectioned and treated with the histological stain for GUS, X-gluc. As shown in Fig. 3B, the wild-type α′ subunit promoter generated a uniform pattern of staining in the embryo, indicating expression in the cotyledons and in the embryonic axis. A very light blue coloration was also observed in the endosperm (data not shown). This corresponds to the pattern already observed for the expression of this same promoter in transgenic tobacco [22]. As a control, seeds coming from untransformed plants were stained in the same conditions. No GUS expression could be detected in any part of these seeds (Fig. 3A).

With the double mutant promoter driving the GUS reporter gene (construct XB), a uniform blue staining of the embryo and a very weak staining of the endosperm were also observed (Fig. 3C). This pattern was undistinguishable from the pattern obtained with the wild-type promoter (Fig. 3B). As expected, the single mutant promoter constructs (constructs X and B) also displayed the same pattern of GUS activity (data not shown). Therefore, mutations of the legumin boxes do not alter the specific localization of the α′ subunit promoter transcription in the seeds of transgenic tobacco plants.

Transcriptional activities of the chimaeric promoter constructs

GUS activities were measured in different organs (young leaves, roots and dry seeds) for all the tobacco plants transformed with the chimaeric promoter constructs and the average and median

Table 2. Average and median GUS activities in different organs for the chimaeric promoter constructs.

Construct	Number of plants	GUS activity in roots (pmol MU formed per 100 μg protein per minute)		GUS activity in leaves (pmol MU formed per 10 μg protein per minute)		GUS activity in dry seeds (pmol MU formed per seed per minute)	
		Median	Average	Median	Average	Median	Average
SX	5	0[a]	0	0[b]	0[b]	0.085[b]	0.14[b]
35S	10	22	35	24	39	1.2	15
35SX	19	50	60	29	49	29	36
C835	13	0	0	0	0	28	47

[a] In this table, 0 means: no significant activity above background (untransformed tobacco) using the standard assay, unless otherwise noted; assays done in duplicate.

[b] Values obtained by a more sensitive assay (see Results); 10 untransformed tobacco plants and 10 plants transformed with the vector pBI101 were used as controls; all assays done in triplicate.

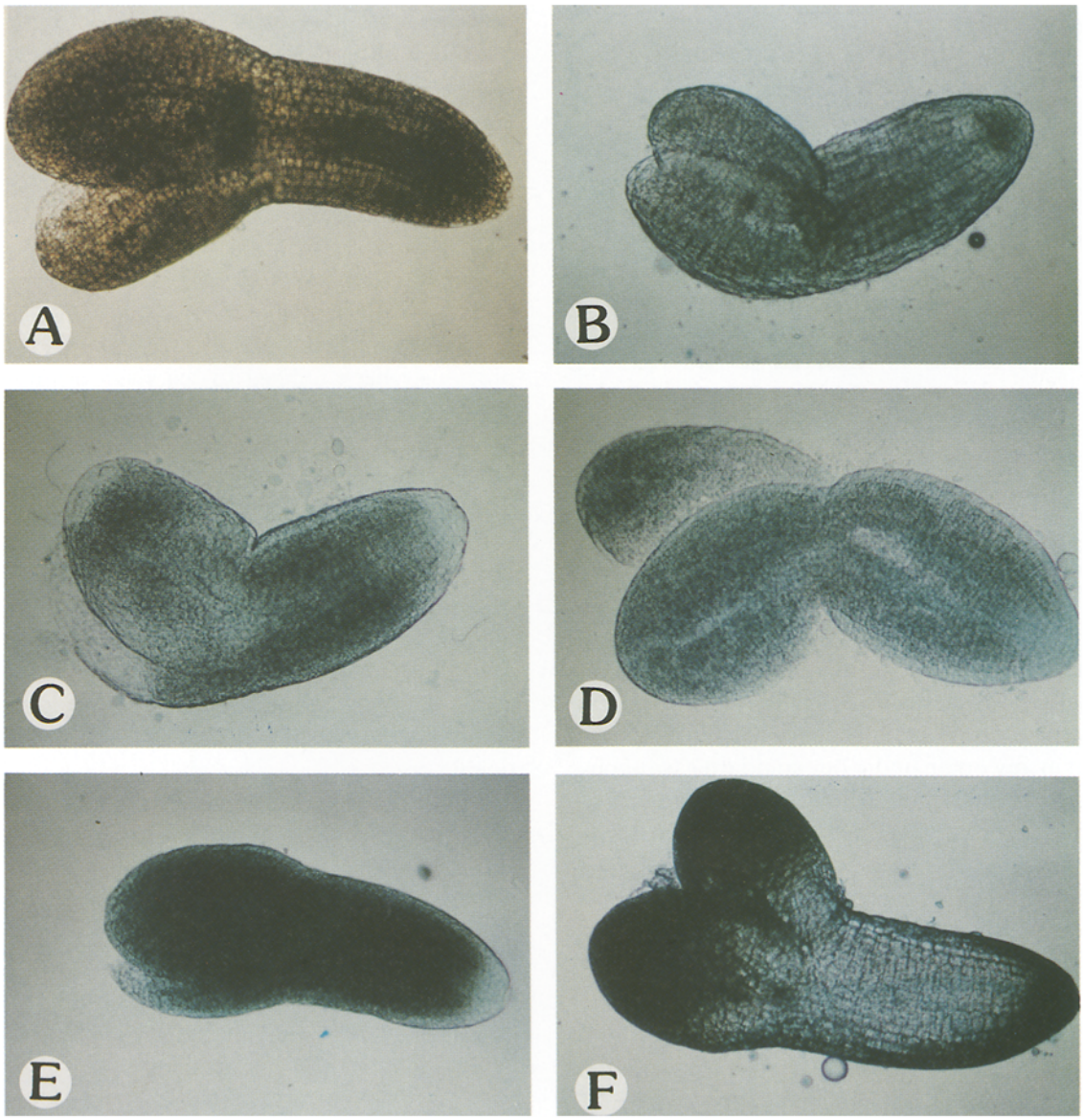

Fig. 3. Histochemical localization of GUS activity in embryos for some of the legumin box mutant constructs and some of the chimaeric promoter constructs. A. Untransformed *N. benthamiana*. B. Construct N. C. Construct XB. D. Construct 35S. E. Construct 35SX. F. Uncharacteristic expression pattern observed in a few seeds transformed with construct SX.

activities are given in Table 2. As with the legumin box mutant constructs, important plant-to-plant variations in the level of GUS expression were observed (data not shown). Because of that, the median values were also used for statistical analyses and are represented graphically in Fig. 4. In this case also, no differences were observed between the kinetics of GUS accumulation for the different constructs and the activity in dry seeds was shown to correctly reflect the level of expression during embryogenesis and was thus used for comparisons.

The 35S construct was used as a control for these experiments and, as expected, expression

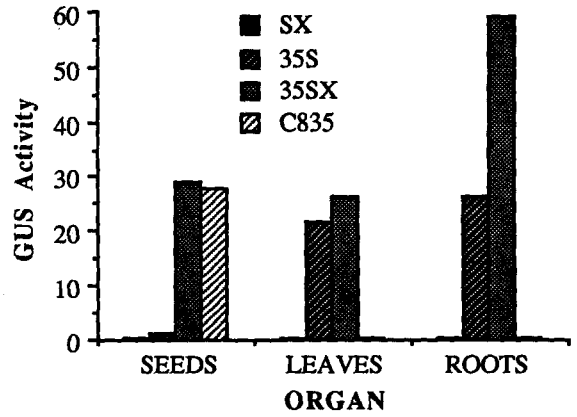

Fig. 4. Histogram showing the median GUS activities in three different organs for the chimaeric promoter constructs. Activities are expressed in pmol MU formed per seed per minute (seeds), pmol MU formed per 10 μg protein per minute (leaves), or pmol MU formed per 100 μg protein per minute (roots).

from the CaMV 35S promoter was detected in all tissues tested (see Table 2). On the other hand, no expression could be detected from the SX construct in any organ under our standard assay conditions. To determine whether this small part of the α' promoter can direct a low level of seed-specific gene expression, more sensitive assays were performed by increasing the amount of material and the reaction times. This represented an overall increase in sensitivity by a factor of about 200. In this manner, reproducible GUS activities were detected in all five tobacco plants transformed with the SX construct (Table 2). To make sure that this activity was significant, measurements were done in the same conditions with several samples of seeds coming from 10 untransformed plants and with seeds from 10 different plants transformed with pBI101 (vector containing the GUS reporter gene without a promoter). In all these control experiments, GUS activities were at least 20 times lower than in the SX plants (results not shown). This clearly indicates that GUS transcription in the SX plants is under the control of the small α' subunit promoter fragment. To determine whether the small 3′ part of the α' subunit promoter retains its specificity, more sensitive GUS assays were also performed in leaves of the SX plants as well as of untransformed tobaccos and plants transformed with pBI101. Increasing the sensitivity by the same factor as for the assays in seeds did not reveal any GUS activity above background in leaves of any of the SX plants.

In an attempt to increase the level of regulated expression from the 3′ part of the α' subunit promoter, we added in front of it the upstream part (−941 to −90) of the constitutive CaMV 35S promoter (construct 35SX). This approach has previously been utilized to reactivate promoters that had been deleted to the point where they could no longer give rise to detectable transcription. In several cases, these chimaeric promoters displayed the regulated expression typical of their corresponding intact promoter: the soybean leghaemoglobin *lbc3* gene [33], the maize *Adh-1* gene [13], the photosynthetic gene *ST-LS1* [32] and the potato proteinase inhibitor II gene [18]. This indicated that these small promoter fragments contained all the necessary regulatory elements but needed strong enhancers to be functional.

On the other hand, a construct comprising the 35S upstream region and the 3′ part of the *Phaseolus vulgaris* seed storage protein phaseolin gene gave high transcriptional activity in mature seeds, leaves and roots [8]. Like the 35S-phaseolin chimaeric promoter construct, the 35SX construct was found to be expressed in all organs tested and, in all three cases, its level of expression was higher than the 35S construct (Table 2). In the case of seeds, the very low median activity of the 35S construct is explained by the fact that several plants gave values barely above background. However, these same plants gave good activities in roots and leaves. The reason for the behaviour of these plants is unknown. Overall, the 35SX promoter was nevertheless much more efficient than the 35S promoter in seeds.

The other chimaeric promoter construct (construct C835) comprising the upstream part of the α' subunit promoter and the 3′ part of the 35S promoter gave the highest activity in seeds and no activity in leaves, which was expected since nei-

ther the α′ subunit promoter nor the 3′ part of the 35S promoter are expressed in this organ in tobacco [4, 9]. No activity above background was detected in roots of any of the plants transformed with the C835 construct, indicating a transcriptional activity at least 50- to 100-fold lower than the 35S construct. This was unexpected since it was shown by Benfey *et al.* [5] that the 3′ part of the 35S promoter can by itself direct a measurable level of transcription in roots. It is interesting to point out that a chimaeric promoter comprising the upstream part of a phaseolin promoter and the −90 to +8 fragment of the 35S promoter did not yield any transcription activity in seeds of transgenic tobacco plants either [8].

Histochemical localization of GUS activity in plants expressing the chimaeric promoter constructs

In addition to determining the total GUS activity in seeds, we wished to determine whether the location of this transcriptional activity was exactly the same in seeds expressing the different chimaeric promoter constructs. To this end, dry seeds from several plants for each construct were cut and stained with X-gluc as described by Jefferson [17]. Fig. 3D and 3E show typical results obtained with the 35S and the 35SX constructs, respectively: for both these constructs, a uniform blue staining pattern was observed in the embryos. A very slight blue coloration was also observed in the endosperm (not shown). The same results were also obtained with the C835 construct (not shown). For the SX construct, prolonged incubations with X-gluc (up to 6 days) allowed us to localize GUS activity even though the activity was much weaker than with the other constructs (see Table 2). Most of the seeds expressing GUS under the control of this small part of the α′ subunit promoter exhibited exactly the same expression patterns as we observed with a longer fragment of this same promoter (construct N, Fig. 3B) or with the 35S promoter (Fig. 3D) and both chimaeric promoters (construct 35SX, Fig. 3E; construct C835, not shown). A few of the seeds coming from plants transformed with the SX construct exhibited a slightly different staining pattern (Fig. 3F). In these seeds, there seemed to be more GUS activity in the radicle and in the tips of the cotyledons than in the rest of the embryo. Since this staining pattern was only observed in very few of the SX seeds (among about 100 seeds which were observed), it probably does not represent a characteristic of this short promoter fragment although this possibility can not be completely ruled out at this time.

Discussion

Role of the legumin box

Our results clearly show that the legumin box is involved in the transcription of the α′ subunit gene of soybean β-conglycinin. The importance of this element in seed storage protein expression has been suggested many times because of its conservation in many legume seed storage protein genes as well as in other seed-specific genes [2, 6, 11, 14, 15]. However, this is, to our knowledge, the first demonstration of a transcriptional function for this conserved element.

From our results, we can conclude that the legumin boxes are *cis*-acting elements which specifically increase the transcription of the β-conglycinin α′ subunit gene in tobacco embryos by a factor of about ten. The absence of legumin boxes in most non-seed genes [11] in addition to the fact that the α′ subunit gene is specifically expressed in seeds make it very unlikely that this element constitutes a general enhancer of transcription. However, we cannot formally conclude from our results that the legumin box is a seed-specific enhancer. Other constructs will be needed to conclusively answer this question.

We can also conclude from our results that the legumin boxes do not act as negative regulatory elements since a promoter lacking both intact copies of this sequence (construct XB) still shows the same exact pattern of regulated expression as the wild-type promoter (construct N). This also means that the legumin boxes are not the sole

determinant of the α′ subunit gene seed-specific expression.

The legumin box was also suggested to be involved in *Phaseolus vulgaris* phytohaemagglutinin L gene expression [27]. Analysis of a promoter deletion series showed that the region between −345 and −125 enhances the level of transcription by a factor of about 25 in transgenic tobacco plants. This region comprises the 63 bp fragment (containing a copy of the legumin box) that was shown to be lacking in the weakly expressed phytohaemagglutinin *lec2-P* allele [34]. However, other DNA sequences are contained in this fragment, making it impossible to conclude that the legumin box itself was responsible for the increase in transcription.

Two other studies have addressed the role of the legumin box in seed storage protein gene expression. In both cases, the legumin box was shown to have no effect on the transcription level of deleted legumin gene promoters. In the first of these studies, a construct was made by adding an oligonucleotide corresponding to the legumin box of the *Pisum sativum legA* gene to a construct containing the first 102 bp of the 5′ flanking sequence of this gene. Both constructs were introduced into tobacco plants [31]. Neither of these two constructs gave a measurable transcriptional activity in seeds whereas it was shown previously that 549 bp of 5′ flanking sequence of this same promoter would direct a good level of seed-specific transcription [30]. From this, the authors concluded that the pea *legA* legumin box is unable to activate transcription of the *legA* gene in tobacco. In the second study, different deletions of the *Vicia faba* legumin gene *LeB4* were made and their activities were compared in transgenic tobacco plants [3]. It was found that a deletion mutant containing only the first 0.2 kb of the *LeB4* promoter gave less than 10% of the activity of the complete promoter. A slightly shorter deletion, destroying the only legumin box present in this promoter, did not reduce further the transcriptional activity.

Two reasons can explain the differences between the results obtained with the *legA* and the *LeB4* genes and our results with the α′ subunit of β-conglycinin. Firstly, all our constructs contained exactly the same promoter fragment, in which some specific point mutations were introduced. Secondly, and more importantly, we used a promoter fragment which was shown to give a transcription level almost equal to the complete promoter [9]. In experiments dealing with promoters giving no expression or levels of expression much lower than the corresponding intact promoters, the only conclusion that can be drawn from a lack of activation is that the element under study is not active by itself, but it cannot be concluded that this element is inactive in the intact promoter.

Interactions between *cis*-acting elements are mediated by *trans*-acting factors binding to them. In soybean, several proteins (SEF 1, SEF 2 and SEF 3) have been shown to bind to specific sequences of the β-conglycinin α′ subunit promoter but none of them to the legumin box [1, 19]. Gel retardation assays performed with protein extracts from pea also failed to detect a factor binding to the *legA* promoter legumin box [23, 31]. However, it is still premature to conclude that there is no specific legumin box binding factor.

Chimaeric promoter constructs

The chimaeric promoter constructs were made in order to determine if the 3′ part of the α′ subunit promoter can direct by itself a low level of regulated transcription and also to check for the presence of negative regulatory elements. Because the 35S promoter is the best characterized 'constitutive' plant promoter [4], it was chosen as a control for these experiments as well as the source of some fragments for the construction of the chimaeric promoters. It is already known that, in the seed, the complete 35S promoter will direct a high level of gene expression in the embryonic axis and in the cotyledons and low expression in the endosperm [5].

In our 35S transformed plants, GUS expression was indeed detected in the embryonic axis and the cotyledons, with only a very weak activity in the endosperm. In several of these plants, there

was in fact a very low activity in the seeds whereas activities in leaves and roots matched those of the other plants transformed with the same construct. For the moment, we cannot explain the behaviour of these plants.

Role of the upstream part of the β-conglycinin promoter

The C835 construct gave a significantly higher activity than the 35S construct in seeds. It can be argued that some of the enhancers present in the 5′ part of the α′ promoter are seed-specific and thus confer a better activity than the enhancers present in the 35S promoter. The lack of C835 expression in leaves can easily be explained by the fact that this chimaeric promoter does not include any element able to give rise to transcription in this organ. On the other hand, because of the presence of the 3′ part of the 35S promoter in this construct (−90 to +8), some expression was expected to be detected in the roots. A previous study has shown that this part of the 35S promoter (called the A domain) gives about 10% of the activity of the complete 35S promoter in roots [5]. In our case, the activity of the C835 construct was not detectable by our standard assay meaning that it was at least 50- to 100-fold lower than the activity of the 35S construct. Possibly, there is a negative regulatory element in the upstream part of the α′ subunit promoter (between −828 and −82) preventing expression from this promoter in organs other than seeds. Alternatively, the difference between our results and those obtained by Benfey *et al.* [5] could be attributed to the fact that the constructs and vectors involved were not exactly the same. Other constructs will have to be analysed to confirm the presence of this regulatory element.

Role of the 3′ part of the β-conglycinin promoter

Although very weak compared with the other constructs, a reproducible GUS activity could be detected in all five plants transformed with the SX construct. This very weak activity cannot be attributed to a background activity in tobacco because the activity in the SX plants was significantly stronger than in untransformed tobacco plants and also than in plants transformed with the pBI101 vector. Also, this activity cannot be attributed to read-through transcription from the NPTII gene just upstream of the GUS gene in PBI101 because there was no detectable GUS activity in leaves of plants transformed with the SX construct or with pBI101, even when using a very sensitive assay. Since NPTII expression is under the control of the NOS promoter, this gene is normally transcribed in leaves and any read-through transcription would also be observed in this organ. The lack of expression of the SX constructs in leaves and roots also indicates that this small promoter fragment still retains its specificity and can still confer correct developmental regulation of transcription.

In a previous study, it was found that the proximal 159 bp of the α′ subunit promoter were required to obtain any detectable activity of a marker gene (chloramphenicol acetyl transferase, in this case); no activity could be detected from the proximal 69 bp of this promoter [9]. In this work, we show a weak but reproducible transcription from the proximal 77 bp of the same promoter. However, we had to use a much more sensitive assay with this construct than with the other ones to detect any activity. In the previous work [9] the same conditions were used for all the assays in all the plants, which explains that no activity could be detected with the short promoter fragment.

The role of the 3′ part of the α′ subunit promoter was further evidenced by the results obtained with the chimaeric promoter comprising the upstream part of the 35S promoter and the 3′ part of the α′ subunit promoter (construct 35SX). This construct gave a significantly stronger activity than the 35S construct in seeds. This can easily be explained by the presence in the 3′ part of the α′ subunit promoter of seed-specific *cis*-acting elements (the legumin box, for example) which work better in this organ than the elements present in the equivalent part of the 35S promoter. On the other hand, 35SX construct expression was not restricted to seeds so it seems that there is no negative regulatory element in the 3′ part of the α′ promoter or, at least, none that is strong enough

to prevent expression from the very strong 35S enhancers.

The expression levels conferred by the SX and the 35SX constructs reveal that the 3′ part of the promoter of the β-conglycinin α′ subunit gene is involved in seed-specific gene expression. This result is especially interesting since most of the studies with this promoter have concentrated so far on its upstream part, and especially on the −257 to −82 region [1, 10]. Our results indicate that seed-specific regulatory elements can now be looked for in a very short DNA fragment (90 bp). Apart from the ubiquitous CCAAT and TATAA boxes, the only putative regulatory element identified in this fragment is the legumin box. It is worth pointing out that this region contains no binding site for any of the previously characterized β-conglycinin promoter DNA-binding proteins (SEF 1, SEF 3 and SEF 4) [19]. So, neither a single SEF factor nor a combination of SEF factors constitutes the sole determinant of seed-specific gene expression. On the other hand, we have also shown that a longer promoter fragment (−257 to +13) lacking both intact copies of the legumin box but containing binding sites for SEF 3 and SEF 4 is still correctly regulated. This means that the legumin box is not the sole determinant of seed specificity either. The analysis of other mutants will be needed to determine whether there is one yet unidentified element mediating seed-specific gene expression or many regulatory elements can independently give rise to regulated gene expression during embryogenesis.

Acknowledgements

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) and from the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche du Québec (Fonds FCAR). S.C. was supported by a studentship from the Fonds FCAR du Québec. We thank P.A. Lessard and R.N. Beachy for providing us with the α′ subunit promoter fragment. We also thank Dr Guy Bellemare and Mr Pierre Bilodeau for critical reading of the manuscript and Mr Alain Goulet for help with histological stainings.

Note added in proof

During the time this paper was in the review process, the results of two other studies describing the function of the legumin box have been published. In one of these, 6 bp from the core CATGCAT motif were deleted in a 2.4 kb *Vicia faba LeB4* legumin promoter fused to the GUS reporter gene (Bäumlein *et al.* The Plant J 2: 233–239, 1992). In the other, 0.4 and2.3 kb *Gy2* glycinin promoters containing a 7 bp deletion in one of the two legumin boxes were also introduced in front of the GUS gene (Lelievre *et al.* Plant Physiol 98: 387–391, 1992). In both cases, the deletion drastically reduced the reporter gene expression in transgenic tobacco plants (by a factor of at least ten). The much more pronounced effect that these two groups observed by the mutagenesis of one legumin box as compared with our results could be explained by the fact that we studied substitution mutations whereas they used deletions, in which the spacing between the different promoter elements is not kept intact. This spacing could be critical for the optimal activity of the promoter.

References

1. Allen RD, Bernier F, Lessard PA, Beachy RN: Nuclear factors interact with a soybean β-conglycinin enhancer. Plant Cell 1: 623–631 (1989).
2. Baszczynski CL, Fallis L: Isolation and nucleotide sequence of a genomic clone encoding a new *Brassica napus* napin gene. Plant Mol Biol 14: 633–635 (1990).
3. Bäumlein H, Boerjan W, Nagy I, Panitz R, Inzé D, Wobus U: Upstream sequences regulating legumin gene expression in heterologous transgenic plants. Mol Gen Genet 225: 121–128 (1991).
4. Benfey PN, Chua N-H: The cauliflower mosaic virus 35S promoter: combinatorial regulation of transcription in plants. Science 250: 959–966 (1990).
5. Benfey PN, Ren L, Chua N-H: The CaMV 35S enhancer contains at least two domains which can confer different developmental and tissue-specific expression patterns. EMBO J 8: 2195–2202 (1989).

6. Bogue MA, Vonder Haar RA, Nuccio ML, Griffing LR, Thomas TL: Developmentally regulated expression of a sunflower 11S seed protein gene in transgenic tobacco. Mol Gen Genet 222: 49–57 (1990).
7. Bray EA, Naito S, Pan N-S, Anderson E, Dubé P, Beachy RN: Expression of the β-subunit of β-conglycinin in seeds of transgenic plants. Planta 172: 364–370 (1987).
8. Bustos MM, Begum D, Kalkan FA, Battraw MJ, Hall TC: 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 (1991).
9. Chen Z-L, Schuler MA, Beachy RN: Functional analysis of regulatory elements in a plant embryo-specific gene. Proc Natl Acad Sci USA 83: 8560–8564 (1986).
10. Chen Z-L, Pan N-S, Beachy RN: A DNA sequence element that confers seed-specific enhancement to a constitutive promoter. EMBO J 7: 297–302 (1988).
11. Dickinson CD, Evans RP, Nielsen NC: RY repeats are conserved in the 5′-flanking regions of legume seed-protein genes. Nucl Acids Res 16: 371 (1988).
12. Doyle JJ, Schuler MA, Godette WD, Zenger V, Beachy RN: The glycosylated seed storage proteins of *Glycine max* and *Phaseolus vulgaris*. J Biol Chem 261: 9228–9238 (1986).
13. Ellis JG, Llewellyn DJ, Dennis ES, Peacock WJ: Maize *Adh-1* promoter sequences control anaerobic regulation: addition of upstream promoter elements from constitutive genes is necessary for expression in tobacco. EMBO J 6: 11–16 (1987).
14. Ericson ML, Murén E, Gustavsson H-O, Josefsson L-G, Rask L: Analysis of the promoter region of napin genes from *Brassica napus* demonstrates binding of nuclear proteins *in vitro* to a conserved sequence motif. Eur J Biochem 197: 741–746 (1991).
15. Gatehouse JA, Evans IM, Croy RRD, Boulter D: Differential expression of genes during legume seed development. Phil Trans R Soc Lond B 314: 367–384 (1986).
16. Horsch RB, Fry JE, Hoffmann NL, Eichholtz D, Rogers SG, Fraley RT: A simple and general method for transferring genes into plants. Science 227: 1229–1230 (1985).
17. Jefferson RA, Kavanagh TA, Bevan MW: GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J 6: 3901–3907 (1987).
18. Keil M, Sánchez-Serrano J, Schell J, Willmitzer L: Localization of elements important for the wound-inducible expression of a chimeric potato proteinase inhibitor II-CAT gene in transgenic tobacco plants. Plant Cell 2: 61–70 (1990).
19. Lessard PA, Allen RD, Bernier F, Crispino JD, Fujiwara T, Beachy RN: Multiple nuclear factors interact with upstream sequences of differentially regulated β-conglycinin genes. Plant Mol Biol 16: 397–413 (1991).
20. Lewis MK, Thompson DV: Efficient site directed *in vitro* mutagenesis using ampicillin selection. Nucl Acids Res 18: 3439–3443 (1990).
21. Matzke AJM, Matzke MA: A set of novel Ti plasmid-derived vectors for the production of transgenic plants. Plant Mol Biol 7: 357–365 (1986).
22. Matzke AJM, Stöger EM, Schernthaner JP, Matzke MA: Deletion analysis of a zein gene promoter in transgenic tobacco plants. Plant Mol Biol 14: 323–332 (1990).
23. Meakin PJ, Gatehouse JA: Interaction of seed nuclear proteins with transcriptionally-enhancing regions of the pea (*Pisum sativum* L.) *legA* gene promoter. Planta 183: 471–477 (1991).
24. Mood MA: Introduction to the Theory of Statistics. McGraw-Hill, New York (1950).
25. Ratio S, Dubé PH, Beachy RN: Differential expression of conglycinin α′ and β subunit genes in transgenic plants. Plant Mol Biol 11: 109–123 (1988).
26. Peach C, Velten J: Transgene expression variability (position effect) of CAT and GUS reporter genes driven by linked divergent T-DNA promoters. Plant Mol Biol 17: 49–60 (1991).
27. Riggs CD, Voelker TA, Chrispeels MJ: Cotyledon nuclear proteins bind to DNA fragments harboring regulatory elements of phytohemagglutinin genes. Plant Cell 1: 609–621 (1989).
28. Sambrook J, Fritsch EF, Maniatis T: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989).
29. Schuler MA, Schitt ES, Beachy RN: Closely related families of genes code for the α′- and α-subunits of the soybean 7S storage protein complex. Nucl Acids Res 10: 8225–8244 (1982).
30. Shirsat A, Wilford N, Croy R, Boulter D: Sequences responsible for the tissue specific promoter activity of a pea legumin gene in tobacco. Mol Gen Genet 215: 326–331 (1989).
31. Shirsat AH, Meakin PJ, Gatehouse JA: Sequences 5′ to the conserved 28 bp Leg box element regulate the expression of pea seed storage protein gene *legA*. Plant Mol Biol 15: 685–693 (1990).
32. Stockhaus J, Schell J, Willmitzer L: Identification of enhancer elements in the upstream region of the nuclear photosynthetic gene ST-LS1. Plant Cell 1: 805–813 (1989).
33. Stougaard J, Sandal NN, Gron A, Kühle A, Marcker KA: 5′ analysis of the soybean leghaemoglobin *lbc3* gene: regulatory elements required for promoter activity and organ specificity. EMBO J 6: 3565–3569 (1987).
34. Voelker T, Sturm A, Chrispeels MJ: Differences in expression between two seed lectin alleles obtained from normal and lectin-deficient beans are maintained in transgenic tobacco. EMBO J 6: 3571–3577 (1987).