The Plant Cell, Vol. 2, 973–986, October 1990 © 1990 American Society of Plant Physiologists

Functional Analysis of the *Sesbania rostrata* Leghemoglobin *glb3* Gene 5′-Upstream Region in Transgenic *Lotus corniculatus* and *Nicotiana tabacum* Plants

László Szabados,[a,1] Pascal Ratet,[a,2] Britta Grunenberg,[a] and Frans J. de Bruijn[a,b,3]

[a] Max-Planck-Institut für Züchtungsforschung, Carl-von-Linne-Weg 10, 5000 Köln 30, Federal Republic of Germany
[b] Michigan State University-Department of Energy Plant Research Laboratory, Michigan State University, East Lansing, Michigan 48824

Expression of the *Sesbania rostrata* leghemoglobin *glb3* gene was analyzed in transgenic *Lotus corniculatus* and tobacco plants harboring chimeric *glb3-uidA* (*gus*) gene fusions to identify *cis*-acting elements involved in nodule-specific gene expression and general transcriptional control. A 1.9-kilobase fragment of the *glb3* 5′-upstream region was found to direct a high level of nodule-specific β-glucuronidase (GUS) activity in *L. corniculatus*, restricted to the *Rhizobium*-infected cells of the nodules. The same fragment directed a low level of GUS activity in tobacco, restricted primarily to the roots and to phloem cells of the stem and petiole vascular system. A deletion analysis revealed that the region between coordinates −429 and −48 relative to the ATG was sufficient for nodule-specific expression. Replacement of the −161 to −48 region, containing the *glb3* CAAT and TATA boxes, with the heterologous truncated promoters Δ-p35S and Δ-p*nos* resulted in a loss of nodule specificity and reduction of GUS activity in *L. corniculatus* but a significant increase in tobacco, primarily in the roots. The same fragment could not direct nodule-specific expression when fused to a heterologous enhancer in *cis*. This region contains DNA sequences required, but not sufficient, for nodule-specific expression in *L. corniculatus* that function poorly or may be involved in promoter silencing in tobacco. By fusing further upstream fragments to the Δ-p35S and Δ-p*nos* promoters, two positive regulatory regions were delimited between coordinates −1601 and −670, as well as −429 and −162. The former region appears to function as a general enhancer because it significantly increased promoter activity in both orientations in *L. corniculatus* and tobacco. The latter region could enhance gene expression in both orientations in tobacco, but only in the correct orientation in *L. corniculatus*. These results show that efficient expression of the *S. rostrata glb3* gene in nodules is mediated by an ATG-proximal, tissue-specific element, as well as further 5′-upstream positive elements; that the *S. rostrata glb3* promoter is induced in a nodule-specific fashion in the heterologous legume *L. corniculatus*, suggesting a high degree of conservation of the relevant regulatory signals; and that the *S. rostrata lb* promoter is not silent in the nonlegume tobacco, but is expressed primarily in the roots.

INTRODUCTION

Regulated plant genes can be either inducible or tissue specific or both inducible and tissue specific. They may respond to environmental stimuli such as light, heat, anaerobic stress, wounding, fungal elicitors, circadian rhythms, hormones, or signals from pathogenic or symbiotic bacteria. Alternatively, or in addition, they can be activated in green tissues, in tubers, in seeds, in flowers, in pollen, or in nitrogen-fixing nodules induced by rhizobia. Members of both categories may also be subject to developmental control processes (see Schell et al., 1988; Weising et al., 1988; Benfey and Chua, 1989).

We are interested in studying the regulation of a class of plant genes that are specifically expressed during the process of rhizobial infection and nodule ontogeny or in mature nitrogen-fixing nodules. These genes have been termed nodulin genes (van Kammen, 1984) and appear to be induced directly or indirectly by rhizobial signals in a tissue-specific or nodule-specific fashion. In addition, their expression is developmentally controlled (see Govers et al., 1987; Verma and Delauney, 1988; Gloudemans and Bisseling, 1989; Long, 1989; de Bruijn et al., 1989, 1990).

[1] Current address: Biological Research Center Institute of Plant Physiology, P.O. Box 521, 6701 Szeged, Hungary.
[2] Current address: Institut des Sciences Vegetales, Centre National de la Recherche Scientifique, 91198 Gif-sur-Yvette Cedex, France.
[3] To whom correspondence should be addressed at Michigan State University .

Nodulins are commonly divided into early and late, reflecting their timepoint of induction. Early nodulins are involved in root hair deformation, the rhizobial infection process, and structural aspects of nodule ontogeny (see Gloudemans and Bisseling, 1989).

Late nodulins are induced around the onset of bacteroid nitrogen fixation. They participate in various aspects of nodule functioning and include enzymes involved in nitrogen assimilation (uricase II, glutamine synthetase) and carbon metabolism (sucrose synthase), peribacteroid membrane polypeptides, and oxygen transport proteins (leghemoglobins; see Appleby, 1984; Delauney and Verma, 1988; Verma and Delauney, 1988). The structure of the soybean leghemoglobins (Lbs) and the organization and expression of the corresponding *lb* genes have been examined in detail (see Appleby, 1984; Marcker et al., 1984a; Verma et al., 1986). Four expressed *lb* genes have been characterized that are induced at slightly different timepoints after nodule initiation (Marcker et al., 1984b), suggesting developmental control of these genes in addition to nodule-specific expression.

The regulation of soybean *lb* gene expression has been studied using transgenic *Lotus corniculatus* (Stougaard et al., 1986, 1987, 1990) and alfalfa (Schell et al., 1988; de Bruijn et al., 1989) plants harboring chimeric *lbc*3-*cat* (chloramphenicol acetyltransferase) gene fusions. Nodule-specific expression was found to be mediated by an approximately 2-kb 5′-upstream region in both heterologous systems (Stougaard et al., 1986; de Bruijn et al., 1989). Similar observations were made for the nodule-enhanced expression of the French bean glutamine synthetase *gln*-γ (Forde et al., 1989) and the soybean late nodulin N23 genes (Joergensen et al., 1988) in transgenic *L. corniculatus* plants. A deletion analysis of the soybean *lbc*3 gene 5′-upstream region delimited a Strong Positive Element (SPE; −1090 to −947 relative to the startpoint of transcription), a Weak Positive Element (WPE; −230 to −170), an Organ (nodule)-Specific Element (OSE; −139 to −102), and a Negative Element (NE; −102 to −49; Stougaard et al., 1987), whereby the SPE and OSE elements were found to be interdependent for maximal nodule-specific expression (Stougaard et al., 1990).

To determine which of these elements represent conserved *cis*-acting sequences responsible for *lb* gene expression and how these and other elements may be organized in the 5′-upstream region of *lb* genes from other legumes, we have initiated an analysis of the structure, organization, and regulation of *lb* genes from the distantly related legume *Sesbania rostrata* (Berhaut, 1976). This tropical legume was selected for the following reasons: *S. rostrata* is efficiently nodulated by its symbiont *Azorhizobium caulinodans* not only on the roots but also profusely on the stem (see de Bruijn, 1989). Stem nodules contain cortical cells that harbor chloroplasts, which are located immediately adjacent to bacteroid-filled cells (de Bruijn, 1989). The juxtaposition of energy source and energy sink may be beneficial for the energy-intensive nitrogen fixation process. On the other hand, the O_2 evolution associated with the former cells could represent a further stress on the O_2-sensitive nitrogenase enzyme complex in the infected cells. Therefore, it was reasonable to postulate that additional, modified, or higher levels of Lb might exist in *S. rostrata* stem nodules to provide the additional O_2 buffer capacity. In fact, seven distinct Lb forms have been identified in *S. rostrata* nodules, of which one form, Lb VI, is approximately fivefold more abundant in stem nodules (Bogusz et al., 1987), suggesting differential expression of the corresponding gene in green versus white nodule tissue (see de Bruijn, 1989). To test this hypothesis, we isolated and partially characterized the *S. rostrata* genes encoding Lb VI and Lb VII (the *glb*3 and *glb*2 genes, respectively; Metz et al., 1988; Welters et al., 1989).

A comparison of the 5′-upstream regions of the *S. rostrata* *glb*2 and *glb*3 genes revealed a high degree of homology between these two genes and between *glb*2, *glb*3, and the corresponding soybean *lbc*3 gene for a stretch of approximately 400 bp (Metz et al., 1988). The most highly conserved region covers the startpoint of transcription, the CAAT and TATA box region, and DNA sequences highly homologous with the OSE in the *lbc*3 5′-upstream region, including the typical AAAGAT/G and CTCTT sequence motifs (Sandal et al., 1987; Stougaard et al., 1987; Metz et al., 1988). Moreover, it covers highly conserved AT-rich elements, known to constitute specific binding sites for a nodule *trans*-acting factor in both the case of the *lbc*3 (Jensen et al., 1988) and in the *glb*2 and *glb*3 genes (de Bruijn et al., 1988; Metz et al., 1988). This conserved AT-rich binding site interacts specifically not only with a nodule factor but also with distinct factors from leaves and roots (Metz et al., 1988; Jacobsen et al., 1990). In fact, one of the factors has been shown to resemble the human HMG I protein, which may play a role in nucleosome positioning or nuclear scaffold-DNA interactions (Jacobsen et al., 1990).

To continue our comparative analysis to delimit general as well as nodule-specific *cis*-acting regulatory elements in the *S. rostrata* *glb*3 gene 5′-upstream region and to study the relationship between nuclear factor binding sites and these elements, we constructed transcriptional fusions of *glb*3 fragments, truncated promoters (Δ-p35S, Δ-p*nos*), an enhancer element (E-*nos*), and the *uidA* (*gus*) reporter gene in various combinations, and assayed the expression of the resulting chimeric genes in distinct tissues of transgenic *L. corniculatus* and tobacco plants by GUS enzyme activity quantification and in situ staining (Jefferson et al., 1987). Here we describe the identification of two positive regulatory elements and a promoter-proximal region essential for nodule-specific expression and compare the organization of these *cis*-acting elements with that of the soybean *lbc*3 gene. We also show that the *S. rostrata* *glb*3 promoter is functional in tobacco and is expressed predominantly in the roots, and we relate these results to

recently published observations regarding root-enhanced expression of other plant hemoglobin genes (Bogusz et al., 1988; Landsmann et al., 1988). Preliminary reports of these results have been presented at the EMBO Symposium on Signal Transduction in Higher Plants (Heidelberg, Federal Republic of Germany, September 1989) and the Lunteren Lectures (de Bruijn et al., 1989).

RESULTS

Nodule-Specific Expression of Chimeric *glb3-uidA* Genes in Transgenic *L. corniculatus* Plants

To examine the role of the *S. rostrata glb3* 5′-upstream region in tissue-specific expression, fragments containing DNA sequences from coordinates −1914 to −48, −653 to −48, −429 to −48, and −161 to −48 (relative to the *glb3* ATG) were inserted into the promoterless *uidA* plasmid pLP17, as shown in Figure 1, to form plasmids pLP32, pLP29, pLP31, and pLP63, respectively (Methods; Figure 2A). These plasmids were conjugally mobilized to *Agrobacterium rhizogenes* AR10, and the resulting transconjugants were used to infect *L. corniculatus* seedlings. Shoots were regenerated from the hairy roots, transgenic plants were obtained and nodulated with *Rhizobium loti* NZP2037, and nodules, roots, stems, and leaves were assayed for GUS activity, as described in Methods.

In these and all other experiments using *L. corniculatus*, a wide variability of GUS activity was observed in individual transgenic plants harboring the same chimeric construct, but the relative level of activity in different tissues was found to be constant (de Bruijn et al., 1990). Therefore, six to 10 individual transgenic *L. corniculatus* plants obtained from independent hairy roots were assayed for each construct and the results were averaged. Construct LP32, as shown in Figure 2A, clearly directed a high level of nodule-specific GUS activity in *L. corniculatus*; no activity was found in roots, stems, and leaves of transgenic plants (Figure 2B). Constructs LP29 and LP31 (Figure 2A) retained the ability to direct nodule-specific expression, although the levels of GUS activity were reduced twofold to threefold, respectively (Figure 2B). Construct LP63 (Figure 2A) was found to be silent (Figure 2B). Thus, DNA sequences essential for nodule-specific expression are contained within the first 429 bp of the *glb3* 5′-upstream region, and additional positive elements are located further upstream between coordinates −1914 to −654 and −429 to −162.

A histochemical analysis (see Methods) was performed on tissues of transgenic *L. corniculatus* plants harboring construct LP32. No GUS staining was observed in leaf or stem sections (data not shown), a very low level of staining was observed in the roots, as shown in Figure 3A, and very intense staining was observed in the nodules, in the bacteroid-infected, leghemoglobin-containing cells, but not in the nodule cortex or vascular tissue (Figures 3B and 3C). These results contrast sharply with those obtained with the cauliflower mosaic virus (CaMV) p35S-*uidA* construct (ROK2275, see below). Transgenic *L. corniculatus* plants harboring this construct showed intense GUS staining in leaves, stems, and roots (data not shown), as well as in infected cells, the cortex, and vascular tissue of the nodules, as shown in Figure 3D. Thus, the correct spatial expression pattern of the *glb3* gene also appears to be mediated by DNA elements present within 1.9 kb of its 5′-upstream region.

Expression of Chimeric *glb3-uidA* Genes in Transgenic Tobacco Plants

To identify general positive regulatory elements (enhancers) or negative elements (silencers), constructs LP32,

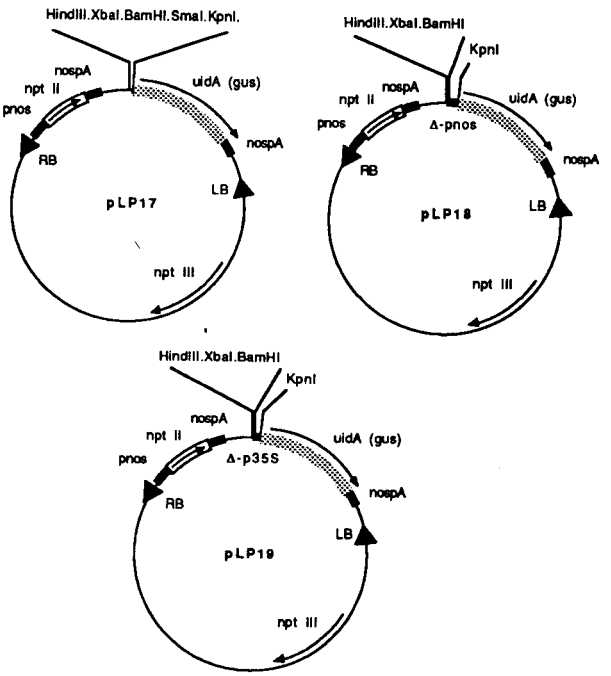


Figure 1. Structure of the Binary Vectors pLP17, pLP18, and pLP19.

The positions of the left and right borders of the T-DNA are indicated by filled triangles (LB and RB, respectively). The extent of β-glucuronidase (*uidA*) coding sequences are indicated by stippled boxes (⁘) and the direction of transcription by arrows. The location of the nopaline synthase (*nos*) gene promoter (*pnos*) and termination (polyadenylation; *nos*pA) sequences are indicated by black boxes (■). The neomycin phosphotransferase II and III (*npt* II, *npt*III) genes and their direction of transcription are indicated by open boxes (□) and with arrows. Only unique restriction sites used for cloning are shown. Δ-*pnos* indicates truncated *nos* promoter sequences, and Δ-p35S indicates truncated CaMV35S promoter sequences (see text).

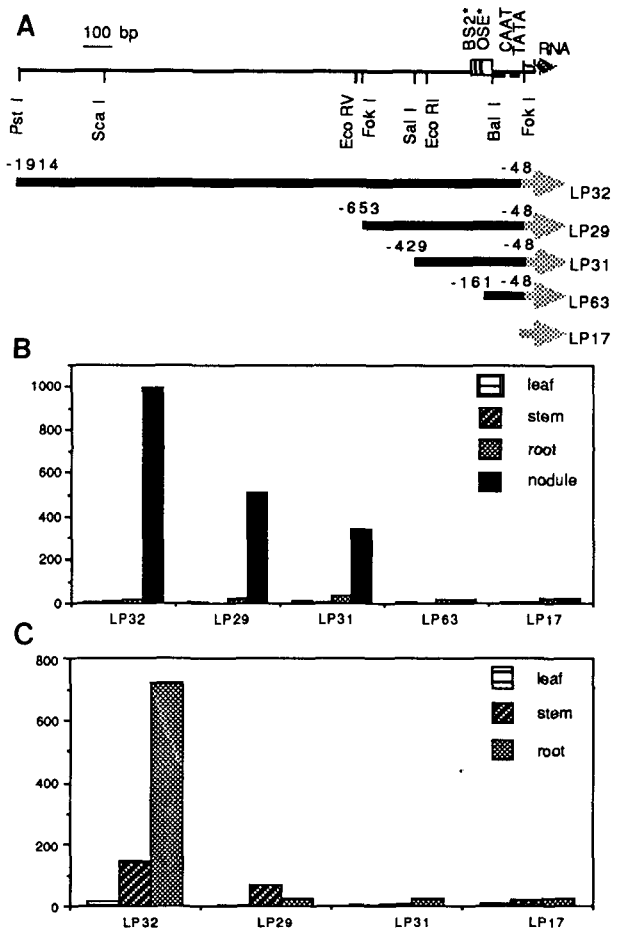


Figure 2. Structure and Expression of *glb3-uidA* 5′ Deletion Constructs in Transgenic *L. corniculatus* and tobacco plants.

(A) Structure of the *glb3* 5′-upstream region and unique restriction sites used for subcloning the respective fragments in pLP17. The positions of the CAAT and TATA boxes are indicated by black boxes (■) and the startpoint of transcription (1 bp downstream of the FokI site) is designated by an arrow labeled RNA (↠). The *uidA* structural gene sequences are indicated by stippled arrows (⇢). The striped element labeled BS2* (▥) indicates the conserved AT-rich motif that binds a nodule *trans*-acting factor (Metz et al., 1988). The open box next to the CAAT box (OSE*; □) indicates the DNA element homologous to the OSE of the soybean *lbc3* promoter (Stougaard et al., 1987; Metz et al., 1988). A scale is provided in 100 bp, and the numbers on the arrows indicate base pairs relative to the startpoint of translation. LP32 contains the PstI-FokI; LP29, the FokI-FokI; LP31, the SalI-FokI, and LP63, the BalI-FokI fragments cloned in pLP17.

(B) GUS activity levels in different tissues of *L. corniculatus* plants (average values of five to 10 independent transgenic plants) expressed in picamoles of methylumbelliferone per minute per milligram of protein in the extract.

(C) GUS activity in different tissues of tobacco plants (average of four to eight plants), expressed in the same units.

LP29, LP31, and LP17 (Figure 2A) were also introduced into tobacco plants, using a leaf-disc transformation protocol with *A. tumefaciens* 3850 (see Methods). As observed with the transgenic *L. corniculatus* plants, a high degree of variability of GUS activity was also found in transgenic tobacco plants (data not shown). Therefore, four to eight independent transgenic plants were routinely assayed. Plants harboring construct LP32 showed GUS activity primarily in the roots, some activity in the stems, but none in the leaves (Figure 2C). The level of GUS activity directed by the LP32 construct in tobacco plants corresponded to about 70% of the activity found in *L. corniculatus* (Figures 2B and 2C). However, direct comparison of the expression levels of the chimeric gene construct in different transgenic plant species is misleading. Relative promoter strengths can be deduced more accurately by comparing expression levels of different constructs within one species. The level of GUS activity directed by the complete *glb3* promoter (construct LP32) in *L. corniculatus* represented approximately 25% of that directed by the CaMV35S promoter (construct ROK2275, see below). However, in tobacco, the level of expression of LP32 was comparable with that observed with the truncated promoter constructs LP18 (Δ-p35S) and LP19 (Δ-p*nos*) and approximately 10% of the activity of the full 35S promoter construct ROK2275 (see below). Thus, we conclude that the activity of the *glb3* promoter in tobacco is approximately 40% of that found in *L. corniculatus* and that the expression is rather root specific.

Histochemical analysis revealed a very low GUS staining activity in the leaves of transgenic tobacco plants harboring LP32 around the vascular tissue (data not shown). In petiole and stem sections, more intense staining could be observed around the vascular system, particularly in the phloem, as shown in Figures 4A and 4B. In contrast, tissues of transgenic tobacco plants harboring the ROK2275 (p35S-*uidA*) construct were much more evenly stained (Figure 4C). Roots of tobacco plants harboring LP32 showed substantial staining around the root tip and the vascular bundle (data not shown). Constructs LP29 and LP31 (Figure 2A) only directed a very low to background level of GUS activity in tobacco (Figure 2C), suggesting the presence of a positive regulatory element between coordinates −1914 and −654.

Effect of ATG Proximal Deletions and CAAT/TATA Box Replacement on *glb3-uidA* Expression

To ascertain the effect of ATG proximal deletions in the 5′-upstream region on *glb3* promoter activity, several unique restriction sites (BalI, SalI, and EcoRV) were used to subclone promoter segments. These were fused to truncated promoters (including the CAAT and TATA boxes) of the nopaline synthase gene (Δ-p*nos*, −150 to

+1; Ebert et al., 1987) or the CaMV35S gene (Δ-p35S, −90 to +9, Domain A; Odell et al., 1985) and inserted in front of the *uidA* gene on plasmids pLP18 and pLP19 (Figure 1; see Methods) to form plasmids pLP44, pLP26, pLP28, and pLP27, respectively, as shown in Figure 5A. These constructs were used to generate transgenic *L. corniculatus* and tobacco plants. Substitution of the *glb*3 CAAT and TATA box region (−161 to −48) with the heterologous Δ-p*nos* (LP44) or Δ-p35S (LP26) promoters resulted in a loss of nodule-specific GUS activity and simultaneously in a 30-fold increase of root-specific GUS activity in *L. corniculatus* (Figure 5B). The latter is probably due to the presence of *cis*-acting elements in these truncated promoters that direct root-specific expression (see Discussion). The loss of nodule-specific expression of these constructs suggests that DNA sequences essential for tissue-specific induction of the *glb*3 promoter are located in the −161 to −48 (CAAT and TATA box) region. Progressive deletions to further upstream sites in constructs LP28 and LP27 resulted in a twofold to threefold decrease of total GUS activity, while maintaining root specificity (Figure 5B). The overall GUS activity level directed by the *glb*3 promoter region in construct LP32 and by the chimeric *glb*3-Δ-p*nos* or *glb*3-Δ-p35S constructs in LP44 and LP26 remained twofold to fourfold lower than that obtained with the full 35S promoter (ROK2275; Figure 5B). A histochemical analysis of transgenic *L. corniculatus* plants harboring the LP44 or LP26 constructs revealed intense GUS staining in the roots (not shown) and in the cortex and vascular system of the nodules, but not in the central bacteroid-infected cells (Figure 3E).

Constructs LP44, LP26, LP28, and LP27 showed a different expression pattern in tobacco. Replacement of the *glb*3 CAAT and TATA box region with the Δ-p*nos* promoter (LP44; Figure 5A) resulted in a 15-fold increase in promoter activity primarily in the roots and to a lesser extent in leaves and stems (Figure 5C). The pattern observed with construct LP26 (Figure 5A) was similar, except that there was no increase of GUS activity in leaves and stems and the total level of GUS activity was about 50% of that observed with LP44 (Figure 5C). The total level of GUS activity directed by construct LP44 now exceeded that directed by the full 35S promoter (ROK2275; Figure 5C). The progressive deletions in constructs LP28 and LP27 (Figure 5A) reduced overall GUS activity twofold to threefold relative to pLP26, while maintaining root-specific expression (Figure 5C). Their GUS activity was still greater than that found with the original 1.9-kb *glb*3 fragment (LP32; Figure 5A), which barely exceeded the level directed by the minimal promoters (Δ-p*nos* and Δ-p35S) themselves (LP18 and LP19; Figure 5C). These results suggest that the *glb*3 −161 to −48 (CAAT/TATA box region) does not function efficiently in tobacco and/or carries DNA sequences involved in silencing promoter activity in a nonlegume background. Moreover, the presence of positive elements between coordinates −1914 to −670 and −429 to −162 capable of functioning in heterologous systems was confirmed.

Analysis of Positive Regulatory Elements in the *glb*3 5′-Upstream Region

To examine further the nature of the positive regulatory elements in the *glb*3 5′-upstream region defined by the deletion/substitution analysis, restriction fragments covering the −1601 to −670 and −669 to −162 regions were cloned in both orientations in front of the Δ-p*nos* and Δ-p35S-*uidA* chimeric genes on pLP18 and pLP19. The resulting constructs (LP40, LP41, LP22, LP23, LP42, LP43, LP24, LP25) are shown in Figure 6A. These constructs were introduced into *L. corniculatus* and tobacco and assayed for GUS activity. The −1914 to −1602 region was ignored in these experiments because no difference in GUS activity between constructs extending upstream to −1601 or −1914 could be detected (data not shown). The −1601 to −670 fragment was found to be capable of stimulating expression of both minimal promoters threefold to 15-fold in both orientations in *L. corniculatus* and tobacco (LP40, LP41, LP22, and LP23 versus LP18 and LP19; Figures 6B and 6C). This fragment, therefore, has typical enhancer-like properties. In the proper orientation, the −669 to −162 fragment greatly stimulated the expression of both Δ-p*nos* and Δ-p35S promoters in *L. corniculatus* (LP42 and LP24 versus LP18 and LP19; Figure 6B) to a higher level than the −1601 to −670 fragment (LP42 versus LP40 and LP24 versus LP22; Figure 6B). In the opposite orientation, it did not stimulate expression in *L. corniculatus* at all (LP43 and LP25 versus LP18 and LP19; Figure 6B). In tobacco, however, the −670 to −162 fragment enhanced expression of both minimal promoters in either orientation (LP42 and LP43, LP24 and LP25 versus LP18 and LP19). Some differences between the interaction of both the −1601 to −670 and −669 to −162 fragments with either the Δ-p*nos* or Δ-p35S minimal promoters were observed because combinations with the Δ-p35S promoter were more sensitive to the orientation than those with the Δ-p*nos* promoter (LP22 versus LP23 and LP24 versus LP25). None of these constructs conferred any nodule-specific expression on the minimal promoters in *L. corniculatus*. In fact, in *L. corniculatus*, the stimulation of expression was primarily root specific (Figure 6B). In tobacco, stimulation of expression of the truncated Δ-p35S promoter followed the same pattern of root specificity (LP22, LP23, LP24, LP25; Figure 6C), whereas the Δ-p*nos* constructs also showed significant activity in stems and leaves (LP40 and LP41; Figure 6C), especially in combination with the −669 to −162 fragment (LP42 and LP43; Figure 6C). Thus, although positive elements capable of stimulating gene expression in heterologous plant hosts are clearly

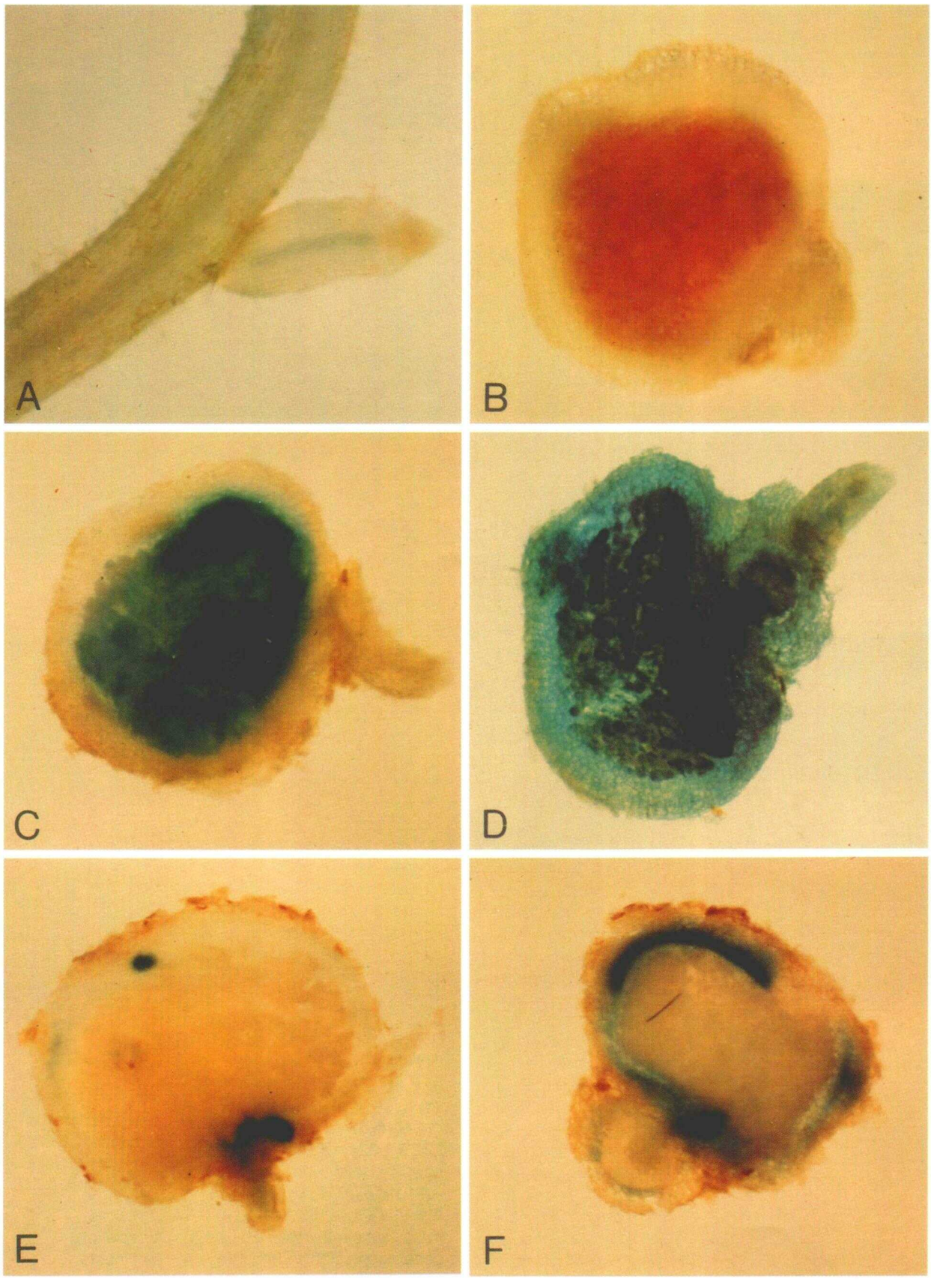


Figure 3. Histochemical Analysis of GUS Activity in Transgenic *L. corniculatus* Plants.

present in the −1601 to −162 region, none of these elements is capable of directing nodule-specific expression on its own.

Analysis of the *glb3* ATG Proximal Nodule Specificity Region

The deletion analysis had delimited a region involved in nodule-specific expression of the *glb3* gene from coordinates −429 to −48 (Figure 2B), of which the −161 to −48 DNA sequences were found to be absolutely essential (Figure 5B). To analyze this region further, a number of additional chimeric constructs were made and introduced into *L. corniculatus* and tobacco plants. Replacement of the −161 to −48 fragment of LP31 by the Δ-p*nos* (LP45) or the Δ-p35S (LP30) promoters resulted in loss of nodule-specific expression in *L. corniculatus*, as shown in Figure 7. Histochemical staining revealed GUS activity in the vascular tissue of nodules carrying construct LP45 but not in the infected cells (Figure 3F), as had been observed with the −1914 to −162 construct (LP44; Figures 3E and 5B). The LP45 and LP30 constructs showed a high level of expression in tobacco, whereas the LP31 construct was silent (Figure 7C). Thus, although clearly exhibiting the capacity to stimulate gene expression in both *L. corniculatus* and *N. tabacum*, the −429 to −162 region alone cannot direct nodule-specific expression in *L. corniculatus*. Moreover, this region can only function in tobacco when fused to a heterologous minimal promoter.

To examine whether a heterologous enhancer element could restore nodule-specific expression to the −161 to −48 region (LP63; Figure 7A), a 91-bp DNA fragment of the *nos* promoter (−150 to −60, E-*nos*; Ebert et al., 1987) was fused in *cis* to the −161 to −48 region (LP64; Figure 7A), and its activity was assayed in *L. corniculatus* plants. This fragment has enhancer-like properties when fused to the Δ-p35S promoter in both orientations (LP56 and LP57; Figures 7A and 7B). The LP64 construct directed substantially increased levels of GUS activity in *L. corniculatus*. However, this expression was not nodule specific but rather was most predominant in roots and to a lesser extent increased in leaves, stems, and nodule (Figure 7B). Histochemical analysis of the LP64-transformed *L. corniculatus* plants revealed a lack of GUS staining in the infected cells of the nodules but significant activity in the vascular system and in the cortex, analogous to the pattern obtained with construct LP45 (data not shown). To examine the possibility that the E-*nos* enhancer region used was capable of changing the tissue specificity of the *glb3* promoter when added in *cis*, it was inserted in front of the −429 to −48 *glb3* fragment (LP65; Figure 7A). Although GUS activity was increased in the roots of *L. corniculatus* plants harboring this construct, expression was not reduced in the nodules (LP65 versus LP31, Figure 7B).

DISCUSSION

We have shown that high-level nodule-specific expression of the *S. rostrata* leghemoglobin *glb3* gene is mediated by DNA elements located within approximately 2 kb of its 5′-upstream region that function efficiently in a heterologous background. This result confirms and extends observations made examining the 5′-upstream region of the soybean *lbc3* gene in transgenic *L. corniculatus* (Stougaard et al., 1986) and alfalfa (de Bruijn et al., 1989), the soybean N23 gene in *L. corniculatus* and clover (Joergensen et al., 1988), as well as the French bean glutamine synthetase *gln*γ gene in *L. corniculatus* plants (Forde et al., 1989). Thus, a conserved mechanism of late nodulin gene induction appears to exist in diverse legume species developing determinate-type (*L. corniculatus*) or indeterminate-type (alfalfa, clover) nodules. The histochemical analysis further revealed that *glb3-uidA* expression in *L. corniculatus* is confined to the central core of the nodule, which is consistent with the role of Lbs in oxygen transport to the actively respiring, nitrogen-fixing bacteroids within the infected cells. Analogous observations regarding the expression of a chimeric French bean *gln*γ*-uidA* construct in the central core of transgenic *L. corniculatus* nodules have been reported by Forde et al. (1989) that in turn are consistent with the role of the nodule-specific form of glutamine synthetase (GSn1) in rapid assimilation of the ammonia produced by the nitrogen-fixing bacteroids in the infected cells (Atkins, 1987).

The level of GUS activity directed by the *glb3* promoter in *L. corniculatus* nodules is approximately 20% to 25% of that directed by the CaMV35S promoter in the same tissue. This result contrasts with results obtained with the

Figure 3. (continued).

Different tissues were stained as described in Methods and photographed in a bright-field microscope. Blue precipitate indicates the location of GUS activity.

(A) LP32 expression in *L. corniculatus* roots. Faint GUS staining in the vascular tissue.

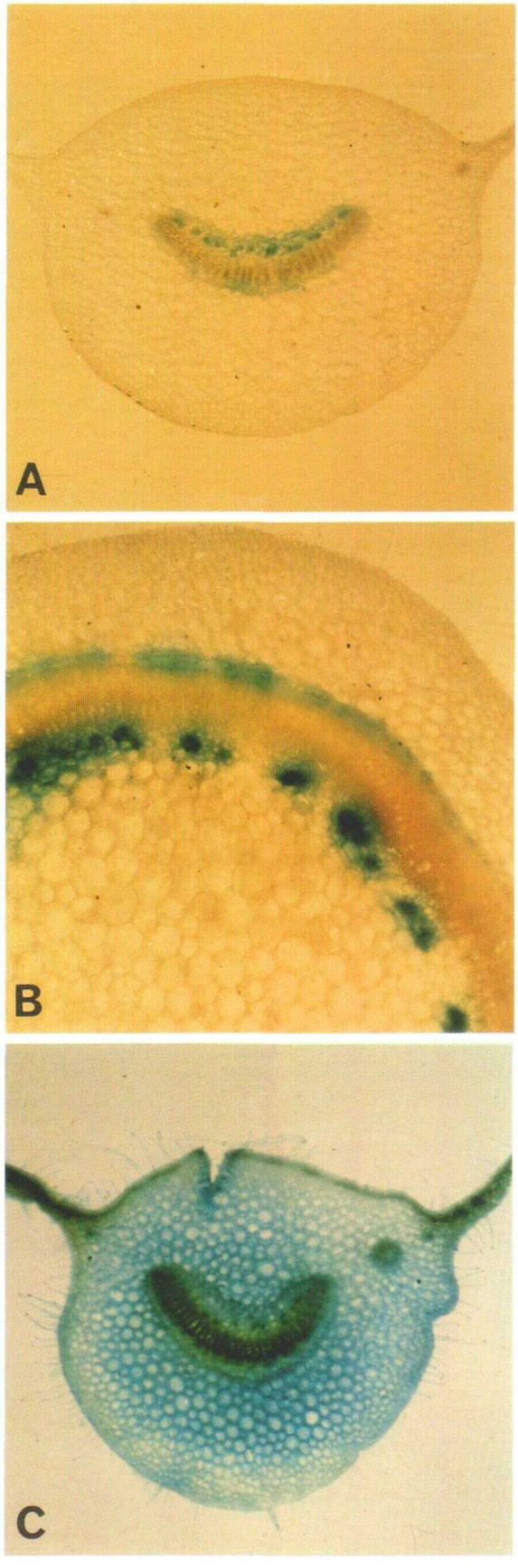
(B) Accumulation of leghemoglobin in the central infected zone of unstained *L. corniculatus* nodules.

(C) LP32 expression in *L. corniculatus* nodules. GUS staining in the central region of the nodule, no activity in nodule cortex.

(D) ROK2275 expression in *L. corniculatus* nodules. Intense GUS staining in the central zone, cortex, and vascular tissue of the nodule and in the root.

(E) LP26 expression in *L. corniculatus* nodules. GUS staining in the vascular tissue of the nodule and the root.

(F) LP45 expression in *L. corniculatus* nodules. Staining in the vascular tissue and the cortex.



soybean *lbc*3 promoter, which appears to be approximately 10-fold more active than the CaMV35S promoter in transgenic *L. corniculatus* nodules (Stougaard et al., 1990). The g*lb*3 gene product (LbVI; Bogusz et al., 1987; Welters et al., 1989) constitutes only a minor component (about 2%) of the total Lb proteins from root nodules of *S. rostrata*, which may be a direct reflection of the relatively poor g*lb*3 promoter strength.

The g*lb*3 promoter deletion and substitution analysis has shown that an approximately 380-bp region (−429 to −48) immediately upstream of the transcription start site is sufficient for nodule-specific expression. The region between coordinates −161 and −48, which includes the presumptive CAAT and TATA boxes, is clearly essential for nodule-specific expression, although it cannot achieve this by itself in the presence or absence of a heterologous enhancer element added in *cis*. The −429 to −162 region alone cannot confer nodule specificity either, although it carries (a) positive regulatory element(s). This suggests that an interaction between the −431 to −162 and −161 to −48 regions is essential for nodule-specific expression and/or that the restriction site used to subdivide this region (Ball) happens to bisect an essential element. The immediate 5′-upstream region of the soybean *lbc*3 gene, including the CAAT and TATA box region, which shares a high degree of homology with the *glb3 region* (Metz et al., 1988), has also been shown to be essential for efficient nodule-specific expression (Stougaard et al., 1990).

It should be pointed out that although both the −90 Δ-p35S and −150 Δ-p*nos* constructs have been used extensively to characterize *cis*-acting elements of a variety of genes (e.g., Stockhaus et al., 1987; Castresana et al., 1988; Chen et al., 1988; Poulsen and Chua, 1988; Stougaard et al., 1990), they have been recently shown not to be true minimal promoters. In fact, both of these truncated promoters carry root-specific elements (Benfey et al., 1989; Bouchez et al., 1989). This may explain the substantial increase in GUS activity in the roots of transgenic *L. corniculatus* plants harboring constructs LP26 or LP44. It cannot be excluded completely that interaction of these root-specific elements with g*lb*3 promoter elements added in *cis* could have led to an altered or anomalous pattern of gene expression, although the expression pattern of the E-*nos*-(−429 to −48) g*lb*3-*uidA* construct (LP65 in Figure 7) strongly suggests that the presence of the root-specific *nos* promoter elements is not capable of suppressing nodule-specific expression.

It is interesting to note that in the case of the soybean *lbc*3 promoter a DNA element has been delimited that is

Figure 4. Histochemical Analysis of GUS Activity in Transgenic Tobacco Plants.

(A) LP32 expression in tobacco petioles. GUS staining around the vascular tissue.

(B) LP32 expression in tobacco stems. GUS staining in the inner and outer phloem of the vascular tissue.

(C) ROK2275 expression in tobacco petioles. GUS staining in several different tissues: vascular system, parenchyma, and epidermis.

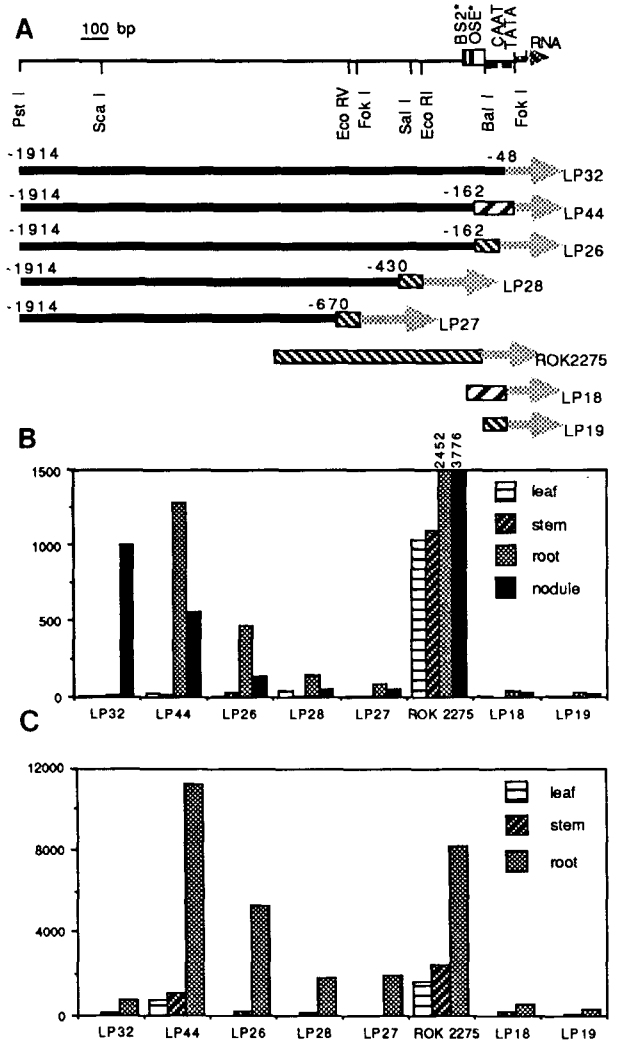


Figure 5. Effect of CAAT/TATA Box Replacement and Downstream Deletions on g*lb*3-*uidA* Expression.

(A) Structure of the g*lb*3 5′-upstream region and chimeric constructs. For details, see legend of Figure 2. LP32 contains the PstI-FokI fragment in pLP17; LP44, the PstI-BalI fragment in pLP18 and LP26 in pLP19; LP28, the PstI-SalI fragment in pLP19; and LP27, the PstI-EcoRV fragment in pLP19. The truncated promoter (Δ-p*nos* and Δ-p35S) elements are indicated by ▨ and ▧ boxes, respectively, and the full 35S promoter in ROK2275 (see Methods) by the long box (▭).
(B) and **(C)** GUS activity of these constructs in *L. corniculatus* and tobacco, respectively (for details, see legend of Figure 2). GUS activity is expressed in picamoles of methylumbelliferone per minute per milligram of protein.

essential for nodule-specific expression (OSE) and shows interdependence on an upstream positive regulatory element (SPE) for its maximal activity (Stougaard et al., 1987, 1990). This *lbc*3 OSE is highly homologous to a 36-bp region at the analogous position in the g*lb*3 promoter (−200 to −164, Metz et al., 1988; OSE*) and contains DNA sequence motifs (AAAGAT/G, CTCTT) that have been found to be conserved in various late nodulin gene 5′-upstream regions (Sandal et al., 1987; Stougaard et al.,

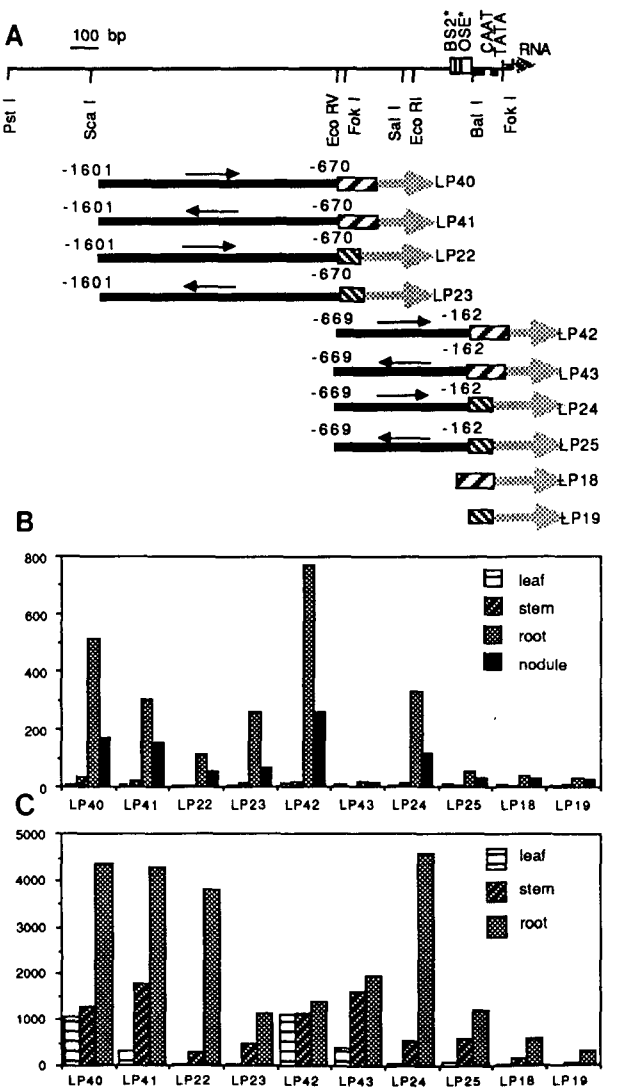


Figure 6. Identification of Positive Regulatory Elements in the g*lb*3 5′-Upstream Region.

(A) Structure of the chimeric constructs. For details, see legends of Figures 2 and 5. LP40 and LP41 contain the ScaI-EcoRV fragment in pLP18 in both orientations, and LP22 and LP23 in pLP19. LP42 and LP43 contain the EcoRV-BalI fragment in pLP18 in both orientations, and LP24 and LP25 in pLP19.
(B) and **(C)** GUS activity levels in *L. corniculatus* and tobacco, respectively (for details, see legend of Figure 2). GUS activity is expressed in picamoles of methylumbelliferone per minute per milligram of protein.

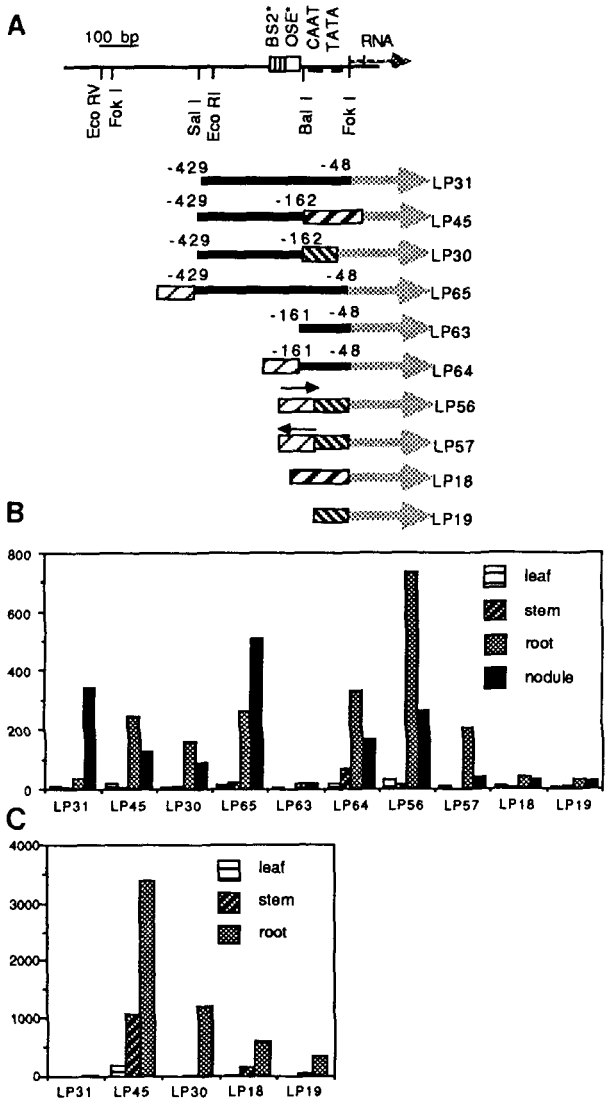


Figure 7. Characterization of the *glb3* Promoter Proximal Nodule Specificity Region.

(A) Structure of the chimeric constructs. For details, see legends of Figures 2 and 5. The *nos* enhancer (E-*nos*) segment (see text) is indicated by a striped box (▨). LP31 contains the SalI-FokI fragment in pLP17. LP45 and LP30 contain the SalI-BalI fragment in pLP18 and pLP19, respectively. LP63 contains the BalI-FokI fragment in pLP17, which is fused to the E-*nos* in *cis* in construct LP64. LP65 contains the E-*nos* enhancer fused upstream to the LP31 construct. LP56 and LP57 contain the E-*nos* in both orientations in pLP19.
(B) and **(C)** Expression of these constructs in *L. corniculatus* and tobacco (see legend of Figure 2). GUS activity is expressed in picamoles of methylumbelliferone per minute per milligram of protein.

1987; Metz et al., 1988). However, although the OSE* is contained within the *glb3* −429 to −162 region, this fragment (or the entire −1914 to −162 fragment) is not capable of conferring nodule-specific expression on chimeric *glb3-uidA* genes in *L. corniculatus*. AAAGAT/G and CTCTT motifs are also present in the −161 to −48 region of the *glb3* promoter (Metz et al., 1988). We cannot exclude the possibility that the presence of these repeated sequences in both these subfragments of the −431 to −48 region is important for nodule-specific and infected cell-specific expression of the *glb3* gene. A comparison of the soybean *lbc3* and *S. rostrata glb3* 5′-upstream regions is schematically diagrammed in Figure 8.

Immediately upstream of the *glb3* OSE-homologous element, a DNA sequence motif has been found (ATTTTTAAATTATTAATTAA, BS2*; see Figure 8) that specifically interacts with a *trans*-acting DNA-binding factor from nodules of different legume plants, as well as other distinct factors from leaves and roots (de Bruijn et al., 1988; Metz et al., 1988). This element has been shown to be repeated severalfold in the *glb3* 5′-upstream region (de Bruijn et al., 1989). Two highly homologous binding sites have been identified in the soybean *lbc3* 5′-upstream region (BS1 and BS2; Jensen et al., 1988; see Figure 8) that specifically interact with the same nodule factor, as well as (a) factor(s) from leaves (Jacobsen et al., 1990). An internal deletion of both BS1 and BS2 has been found to have a relatively minor effect on *lbc3* expression (Stougaard et al., 1990). However, the *glb3* BS2* is located in a region (−429 to −162) that carries at least one positive element and is essential for nodule specificity. Therefore, the functional significance of this binding site and the factor(s) interacting with it remain an open question.

The analysis of the expression of chimeric *glb3-uidA* constructs in tobacco has yielded some interesting insights

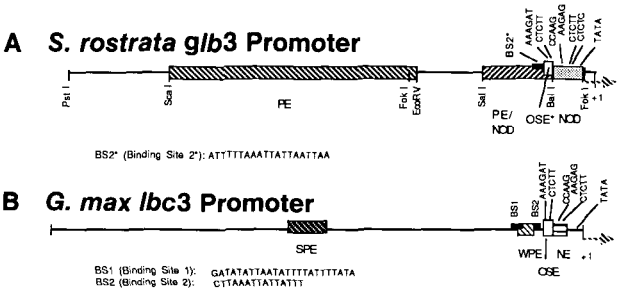


Figure 8. Functional Elements of the 5′-Upstream Region of the *S. rostrata glb3* Gene and the Soybean *lbc3* Gene.

(A) *S. rostrata glb3* promoter.
(B) Soybean *lbc3* promoter. The data in **(B)** are modified from Stougaard et al. (1987, 1990). For further details, see text.
PE, Positive Element; NOD, element(s) involved in Nodule specific Determination; OSE, Organ Specific Element (or homologous region); SPE, Strong Positive Element; WPE, Weak Positive Element; NE, Negative Element; BS1 and BS2, Binding Sites for nuclear factors. The startpoint (direction) of transcription is indicated by an arrow.

into the functioning of the *glb*3 promoter in a nonlegume heterologous background. The full (1.9-kb) *glb*3 promoter, with its own CAAT and TATA box region (−161 to −48), does direct a low level of GUS activity in different tobacco tissues. The highest level is in the roots. The hemoglobin gene (promoter) of the nonlegume *Parasponia andersonii* has also been shown to direct root-enhanced expression in transgenic tobacco plants (Landsmann et al., 1988). This observation and the identification of a functioning hemoglobin gene in the non-nodulating plant *Trema tomentosa*, which is also expressed predominantly in the roots (Bogusz et al., 1988), support the notion that plants may have functional nonsymbiotic hemoglobin genes that exhibit a root-specific expression pattern (Landsmann et al., 1988). Leghemoglobin genes, which are expressed specifically in the nodules, may have evolved from these root-specific genes by acquiring a secondary control pathway, restricting their expression to the infected cells of nitrogen-fixing nodules.

Replacing the *glb*3 CAAT and TATA box region (−161 to −48), essential for nodule-specific expression in *L. corniculatus*, with the Δ-p35S or Δ-p*nos* promoters (including CAAT and TATA boxes; Odell et al., 1985; Ebert et al., 1987) results in a very high level of predominantly root-specific gene expression in tobacco. This suggests that either the *glb*3 CAAT and TATA box region does not function efficiently in tobacco and/or carries DNA sequences that function as a negative regulatory element in tobacco (nonlegume). A negative element (NE) has been shown to be present near the CAAT and TATA boxes of the soybean *lbc*3 promoter (Stougaard et al., 1987), but this element has not been studied in tobacco. Moreover, an internal deletion in the *lbc*3 NE region does not result in constitutive expression in *L. corniculatus* (Stougaard et al., 1990). Except for the CAAT (CCAAG) and AAAGAT/G or CTCTT motifs (see above; Sandal et al., 1987), no significant homology exists between the soybean *lbc*3 NE element and the corresponding *S. rostrata* *glb*3 region (Metz et al., 1988; see Figure 8). A more precise analysis of this region using smaller deletions and/or point mutations will be needed to clarify its function.

At least two efficient positive regulatory regions are located in the *glb*3 5′-upstream region between coordinates −1601 to −670 and −669 (probably −429) to −162. The distal region acts as a general enhancer element, capable of increasing gene expression in both orientations in *L. corniculatus*, as well as in tobacco, resembling the transcriptional enhancers of the *rbcS* or CaMV35S promoters (Timko et al., 1985; Ow et al., 1987; Fang et al., 1989). The proximal region functions as an orientation-independent enhancer in tobacco, but increases gene expression in *L. corniculatus* only in the correct orientation. DNA sequences required for nodule-specific expression are also located in the −429 to −162 region (see above). Whether the positive and nodule-specific DNA elements coincide or can be separated is presently being examined. Two positive elements (SPE, WPE) have also been delimited in the soybean *lbc*3 5′-upstream region (Stougaard et al., 1987). Deleting the WPE region from the *lbc*3 promoter does not influence the expression of *lbc*3-*cat* chimeric genes in a significant fashion, and fusing this region in *cis* to the same Δ-p35S promoter used in this study does not lead to enhanced expression of the corresponding *cat* gene fusion in transgenic *L. corniculatus* plants (Stougaard et al., 1990). The *lbc*3 SPE can confer a low level of nodule-specific expression when fused directly to the *lbc*3 TATA box (Stougaard et al., 1990). The *glb*3 positive elements identified in this study (PE; Figure 8) seem to differ from their *lbc*3 counterparts (SPE, WPE; Figure 8) in their ability to serve as efficient enhancers. Neither the upstream nor downstream *glb*3 PEs can direct nodule-specific expression when fused to the Δ-p*nos* or Δ-p35S. Whether this reflects a real difference between the two promoters or an artifact due to complex interactions between additional root-specific elements present in the Δ-p*nos* and Δ-p35S constructs used in our study remains to be determined.

In conclusion, we have identified several *cis*-acting elements in the *S. rostrata* *glb*3 gene 5′-upstream region that are responsible for general transcriptional activation as well as nodule-specific expression. The use of transgenic legume and nonlegume plants to test the activity of our *glb*3-*uidA* chimeric constructs has given us a better insight into the functional components of the *glb*3 promoter region, and the use of the *uidA* reporter gene has allowed us to examine expression quantitatively as well as spatially. A comparison with the promoter region of the soybean *lbc*3 and other late nodulin genes (see also de Bruijn et al., 1990) has revealed both conserved and nonconserved elements (see Figure 8). The precise role of these elements, the DNA motifs carried by them, and their interaction with *trans*-acting regulatory factors are presently under investigation, and the results are expected to yield further insights into the signal transduction pathway responsible for tissue-specific gene regulation in this unique form of prokaryote-eukaryote interaction.

METHODS

Construction of Vectors and Gene Fusions

Plasmid pLP17 is a derivative of pROK2275 (kindly provided by Dr. R.A. Jefferson), which is based on the pBIN19 binary vector of Bevan (1984). pROK2275 carries a chimeric gene consisting of the cauliflower mosaic virus 35S promoter, β-glucuronidase (*uidA*, *gus*) coding sequences and the nopaline synthase (*nos*) gene 3′-downstream polyadenylation site (p35S-*uidA*-*nos*pA), in which the *uidA* translation site has been modified to fit the Kozak (1981) plant gene consensus sequence. To construct pLP17, the p35S was deleted and replaced by polylinker sequences of pUC18 (Yanisch-Perron et al., 1985) flanked by the HindIII and KpnI sites

(Figure 1). Plasmids pLP18 and pLP19 were derived from pLP17 by inserting a fragment of plasmid pPCV802 (C. Koncz, unpublished results) carrying the truncated *nos* promoter (Δ-p*nos*, −150 to +1; Ebert et al., 1987) or a BamHI-KpnI fragment of plasmid pMP35 (Stockhaus et al., 1987) carrying the truncated CaMVp35S (Δ-p35S −90 to +9; Domain A; Odell et al., 1985) into the polylinker in front of the *uidA* coding sequences (Figure 1). The *glb*3 5′-upstream region was kindly provided by P. Welters from our laboratory. The restriction sites used for subcloning were based on data from Metz et al. (1988), Welters et al. (1989), and P. Welters (unpublished data). The *nos* enhancer element (E-*nos*) was isolated as a SstII to SspI fragment of the *nos* promoter between −150 and −60, containing the CAAT box plus activator elements involved in *nos* promoter activity (Ebert et al., 1987). All DNA manipulations were carried out using standard protocols (Maniatis et al., 1982).

Plant Transformation Protocols

Plasmids (binary constructs) based on pLP17, pLP18, and pLP19 were mobilized from *Escherichia coli* to *Agrobacterium rhizogenes* strain AR10 (Hansen et al., 1989) or *A. tumefaciens* strain 3850 (Zambryski et al., 1983), using pRK2013 as a helper plasmid, as described previously (Ditta et al., 1980). Transgenic *Lotus corniculatus* cv Rodeo plants were generated by way of *A. rhizogenes*-mediated transformation using a modification of the protocol developed by Petit et al. (1987), as described by Tabaeizadeh (1989) and Forde et al. (1989). Seven-day-old to 10-day-old *L. corniculatus* seedlings, which had been germinated in the dark, were wounded and inoculated with overnight cultures of *A. rhizogenes* harboring the binary plasmids and were incubated in light (16-hr photoperiod; 25°C) for 3 weeks. Hairy roots that had formed at the inoculation site were excised and transferred to half-strength Murashige and Skoog medium (½ MS) (Murashige and Skoog, 1962), containing 100 mg/L claforan and 25 mg/L kanamycin (Km). Shoots that had spontaneously formed on these transgenic roots were rooted on ½ MS medium containing 10 mg/L Km. Rooted plantlets were transferred to pots containing a mixture of sand and vermiculite, inoculated with an overnight culture of *Rhizobium loti* strain NZP2037, and incubated in a Conviron growth chamber (26°C; 8-hr/16-hr photoperiod). Five weeks after inoculation, leaf, stem, root, and nodule samples were collected in liquid nitrogen and stored at −70°C. Plants derived from five to 10 independent transgenic roots were tested for each construct. Transgenic *Nicotiana tabacum* SR1 (tobacco) plants were generated using *A. tumefaciens* strain 3850 using the leaf disc transformation protocol described by Horsch et al. (1985). Km-resistant transgenic plants were grown in vitro on ½ MS medium, and leaf, stem, and root tissues were harvested, frozen in liquid N_2, and stored at −70°C. Four to eight independently obtained transgenic plants were tested for each construct.

GUS Enzymatic Assays

GUS activity of tissue samples was determined as described by Jefferson (1987). Enzymatic activity was expressed as picamoles of methylumbelliferone produced per minute per milligram of protein in the extract, as measured using the Bradford assay (Bradford, 1976).

GUS Histochemical Staining

Histochemical GUS activity analyses were performed as described (Jefferson, 1987). Thin sections of plant tissues were cut manually using a razor blade and incubated for 3 hr to 12 hr in 1 mg/mL 5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid (X-glucuronide) at 37°C. After staining, the sections were fixed and destained for chlorophyll in 70% ethanol.

ACKNOWLEDGMENTS

We thank Gisela Kobert for expert assistance in preparing the manuscript and Csaba Koncz for helpful comments. We also thank Peter Welters for providing the cloned *glb*3 5′-upstream region, for unpublished DNA sequence and restriction mapping data, and for helpful discussions. We are grateful to Richard Jefferson, Lothar Willmitzer, and Csaba Koncz for providing vectors and truncated promoter constructs. We thank our colleagues Jens Stougaard, Erik O. Jensen, and Kjeld Marcker at the Department of Molecular Biology and Plant Physiology at Aarhus University for helpful discussions and sharing data before publication. We thank Jeff Schell for scientific suggestions, for his encouragement and support, and for his inspiring interest in this area of research. L.S. and P.R. were supported by postdoctoral stipends from the Max-Planck Gesellschaft.

Received May 15, 1990; accepted August 9, 1990.

REFERENCES

Appleby, C.A. (1984). Leghemoglobin and *Rhizobium* respiration. Annu. Rev. Plant Physiol. **35,** 443–478.

Atkins, C.A. (1987). Metabolism and translocation of fixed nitrogen in the nodulated legume. Plant Soil **100,** 157–169.

Benfey, P.N., and Chua, N.-H. (1989). Regulated genes in transgenic plants. Science **244,** 174–181.

Benfey, P.N., Ren, L., and Chua, N.-H. (1989). The CaMV35S enhancer contains at least two domains which can confer developmental and tissue specific expression patterns. EMBO J. **8,** 2195–2202.

Berhaut, J. (1976). Flore Illustree du Senegal, Vol. 5 (Dakar, Senegal: Clair Afrique), p. 515.

Bevan, M. (1984). Binary agrobacterium vectors for plant transformation. Nucl. Acids. Res. **12,** 8711–8721.

Bogusz, D., Kortt, A.A., and Appleby, C.A. (1987). *Sesbania rostrata* root and stem nodule leghemoglobins: Purification and relationship among the seven major components. Arch. Biochem. Biophys. **254,** 263–271.

Bogusz, D., Appleby, C.A., Landsmann, J., Dennis, E.S., Trinick, N.J., and Peacock, W.J. (1988). Functioning hemoglobin genes in non-nodulating plants. Nature **331,** 178–180.

Bouchez, D., Tokuhisa, J.G., Llewellyn, D.J., Dennis, E.S., and

Ellis, J.G. (1989). The *ocs* element is a component of the promoters of several T-DNA and plant viral genes. EMBO J. **8,** 4197–4204.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye-binding. Anal. Biochem. **72,** 248–254.

de Bruijn, F.J. (1989). The unusual symbiosis between the diazotrophic stem-nodulating bacterium *Azorhizobium caulinodans* ORS571 and its host, the tropical legume *Sesbania rostrata*. In Plant-Microbe Interactions, Molecular and Genetic Perspectives, Volume III, T. Kosuge and E. Nester eds (New York: McGraw Hill), pp. 457–504.

de Bruijn, F.J., Jensen, E.O., Metz, B.A., Hoffmann, H.J., Welters, P., Marcker, K.A., and Schell, J. (1988). Interaction of *trans*-acting factors with distinct DNA elements in the promoter regions of the *Glycine max* and *Sesbania rostrata* leghemoglobin genes. In Molecular Genetics of Plant-Microbe Interactions, R. Palacios and D.P.S. Verma, eds (St Paul: APS Press), pp. 333–339.

de Bruijn, F.J., Felix, G., Grunenberg, B., Hoffmann, H.J., Metz, B., Ratet, P., Simons-Schreier, A., Szabados, L., Welters, P., and Schell, J. (1989). Regulation of plant genes specifically induced in nitrogen-fixing nodules: Role of *cis*-acting elements and *trans*-acting factors in leghemoglobin gene expression. Plant Mol. Biol. **13,** 319–325.

de Bruijn, F.J., Szabados, L., and Schell, J. (1990). Chimeric genes and transgenic plants to study the regulation of genes involved in symbiotic plant-microbe interactions (nodulin genes). Dev. Genet., in press.

Castresana, C., Garcia-Luque, I., Alonso, E., Malik, V.S., and Cashmore, A.R. (1988). Both positive and negative regulatory elements mediate expression of a photoregulated CAB gene from *Nicotiana plumbaginifolia*. EMBO J. **7,** 1929–1936.

Chen, Z.-L., Pan, N.-S., and Beachy, R.N. (1988). A DNA sequence element that confers seed-specific enhancement to a constitutive promoter. EMBO J. **7,** 297–302.

Delauney, A.J., and Verma, D.P.S. (1988). Cloned nodulin genes for symbiotic nitrogen fixation. Plant Mol. Biol. Rep. **6,** 279–285.

Ditta, G., Stanfield, S., Corbin, D., and Helinski, D.R. (1980). Broad host range DNA cloning system for Gram-negative bacteria: Construction of a gene bank of *Rhizobium meliloti*. Proc. Natl. Acad. Sci. USA **77,** 7347–7351.

Ebert, P., Ha, S.B., and An, G. (1987). Identification of an essential upstream element in the nopaline synthase promoter by stable and transient assays. Proc. Natl. Acad. Sci. USA **84,** 5745–5749.

Fang, R.-X., Nagy, F., Sivasubramaniam, S., and Chua, N.-H. (1989). Multiple *cis* regulatory elements for maximal expression of the cauliflower mosaic virus 35S promoter in transgenic plants. Plant Cell **1,** 141–150.

Forde, B.G., Day, H.M., Turton, J.F., Wen-jun, S., Cullimore, J.V., and Oliver, J.E. (1989). Two glutamine synthetase genes from *Phaseolus vulgaris* L. display contrasting developmental and spatial patterns of expression in transgenic *Lotus corniculatus* plants. Plant Cell **1,** 391–401.

Gloudemans, T., and Bisseling, T. (1989). Plant gene expression in early stages of *Rhizobium*-legume symbiosis. Plant Sci. **65,** 1–14.

Govers, F., Nap, J.P., van Kammen, A., and Bisseling, T. (1987). Nodulins in the developing root nodule. Plant Physiol. Biochem. **25,** 309–322.

Hansen, J., Joergensen, J.-E., Stougaard, J., and Marcker, K.A. (1989). Hairy roots—a short cut to transgenic root nodules. Plant Cell Rep. **8,** 12–15.

Horsch, R.B., Fry, J.E., Hoffmann, N.L., Eichholtz, D., Rogers, S.D., and Fraley, R.T. (1985). A simple and general method for transferring genes into plants. Science **227,** 1229–1231.

Jacobsen, K., Laursen, N.B., Jensen, E.O., Marcker, A., Poulsen, C., and Marcker, K.A. (1990). HMG I-like proteins from leaf and nodule nuclei interact with different AT motifs in soybean nodulin promoters. Plant Cell **2,** 85–94.

Jefferson, R.A. (1987). Assaying chimeric genes in plants: The GUS gene fusion system. Plant Mol. Biol. Rep. **5,** 387–405.

Jefferson, R.A., Kavanagh, T.A., and Bevan, M.W. (1987). GUS fusions: β-Glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. **6,** 3901–3907.

Jensen, E.O., Marcker, K.A., Schell, J., and de Bruijn, F.J. (1988). Interaction of a nodule specific, *trans*-acting factor with distinct DNA elements in the soybean leghemoglobin *lbc*3 5′ upstream region. EMBO J. **7,** 1265–1271.

Joergensen, J.-E., Stougaard, J., Marcker, A., and Marcker, K.A. (1988). Root nodule specific gene regulation: Analysis of the soybean nodulin N23 gene promoter in a heterologous symbiotic system. Nucl. Acids. Res. **16,** 39–50.

Kozak, M. (1981). Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. Nucl. Acids Res. **9,** 5233–5252.

Landsmann, J., Llewellyn, D., Dennis, E.S., and Peacock, W.J. (1988). Organ regulated expression of the *Parasponia andersonii* haemoglobin gene in transgenic tobacco plants. Mol. Gen. Genet. **214,** 68–73.

Long, S.R. (1989). *Rhizobium*-legume nodulation: Life together in the underground. Cell **56,** 203–214.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory).

Marcker, A., Bojsen, K., Jensen, E.O., and Paludan, K. (1984a). The soybean leghemoglobin genes. In Advances in Nitrogen Fixation Research, C. Veeger and W.E. Newton, eds (Wageningen Nijhoff/Junk, The Hague, Netherlands) pp. 573–578.

Marcker, A., Lund, M., Jensen, E.O., and Marcker, K.A. (1984b). Transcription of the soybean leghemoglobin genes during nodule development. EMBO J. **3,** 1691–1695.

Metz, B.A., Welters, P., Hoffmann, H.J., Jensen, E.O., Schell, J., and de Bruijn, F.J. (1988). Primary structure and promoter analysis of leghemoglobin genes of the stem-nodulated tropical legume *Sesbania rostrata*: Conserved coding sequences, *cis*-elements and *trans*-acting factors. Mol. Gen. Genet. **214,** 181–191.

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiol. Plant **15,** 473–497.

Odell, J.T., Nagy, F., and Chua, N.-H. (1985). Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. Nature **313,** 810–812.

Ow, D.W., Jacobs, J.D., and Howell, S.H. (1987). Functional

regions of the cauliflower mosaic virus 35S RNA promoter determined by the use of the firefly luciferase gene as a reporter of promoter activity. Proc. Natl. Acad. Sci. USA **84,** 4870–4874.

Petit, A., Stougaard, J., Kühle, A., Marcker, K.A., and Tempé, J. (1987). Transformation and regeneration of the legume *Lotus corniculatus*: A system for molecular studies of symbiotic nitrogen fixation. Mol. Gen. Genet. **207,** 245–250.

Poulsen, C., and Chua, N.-H. (1988). Dissection of 5′ upstream sequences for selective expression of the *Nicotiana plumbaginifolia rbcS*-8B gene. Mol. Gen. Genet. **214,** 16–23.

Sandal, N.N., Bojsen, K., and Marcker, K.A. (1987). A small family of nodule specific genes from soybean. Nucl. Acids Res. **15,** 1507–1519.

Schell, J., John, M., Schmidt, J., Wingender-Drissen, R., Simons, A., Metz, B., Jensen, E.O., Hoffmann, H.J., Welters, P., and de Bruijn, F.J. (1988). Regulation of gene expression in plants with special emphasis on the nodulation process. In Nitrogen Fixation: Hundred Years After. H. Bothe, F.J. de Bruijn, and W.E. Newton eds (New York: Gustav Fischer), pp. 591–598.

Stockhaus, J., Eckes, P., Rocha-Sosa, M., Schell, J., and Willmitzer, L. (1987). Analysis of *cis*-active sequences involved in the leaf-specific expression of a potato gene in transgenic plants. Proc. Natl. Acad. Sci. USA **84,** 7943–7947.

Stougaard, J., Marcker, K.A., Otten, L., and Schell, J. (1986). Nodule-specific expression of a chimeric soybean leghaemoglobin gene in transgenic *Lotus corniculatus*. Nature **321,** 669–674.

Stougaard, J., Sandal, N.N., Gron, A., Kühle, A., and Marcker, K.A. (1987). 5′-Analysis of the soybean leghaemoglobin *lbc*3 gene: Regulatory elements required for promoter activity and organ specificity. EMBO J. **6,** 3565–3569.

Stougaard, J., Joergensen, J.E., Christensen, T., Kühle, A., and Marcker, K.A. (1990). Interdependence and specificity of *cis*-acting regulatory elements in the soybean leghemoglobin *lbc*3 and N23 gene promoters. Mol. Gen. Genet. **220,** 353–360.

Tabaeizadeh, Z. (1989). Genetic transformation of a pasture legume *Lotus corniculatus* L. (Birdsfoot trefoil). Biotechnol. Lett. **11,** 411–416.

Timko, M.P., Kausch, A.P., Castresana, C., Fassler, J., Herrera-Estrella, L., Van den Broeck, G., Van Montagu, M., Schell, J., and Cashmore, A.R. (1985). Light regulation of plant gene expression by an upstream enhancer-like element. Nature **318,** 579–582.

van Kammen, A. (1984). Suggested nomenclature for plant genes involved in nodulation and symbiosis. Plant Mol. Biol. Rep. **2,** 43–45.

Verma, D.P.S., and Delauney, A.J. (1988). Root nodule symbiosis: Nodulins and nodulin genes. In Plant Gene Research: Temporal and Spatial Regulation of Plant Genes, D.P.S. Verma and R.B. Goldberg, eds (New York: Springer-Verlag), pp. 169–199.

Verma, D.P.S., Fortin, M.G., Stanley, J., Mauro, V.P., Purohit, S., and Morrison, N. (1986). Nodulin and nodulin genes of *Glycine max* . Plant Mol. Biol. **7,** 51–61.

Weising, K., Kahl, G., and Schell, J. (1988). Transfer, structure and expression of foreign genes in plants. In Architecture of Eukaryotic Genes. G. Kahl, ed (New York: VCH publishers), pp. 57–87.

Welters, P., Metz, B.A., Schell, J., and de Bruijn, F.J. (1989). Nucleotide sequence of the *Sesbania rostrata* leghemoglobin (*Srglb*3) gene. Nucl. Acids Res. **17,**1253.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985). Improved M13 cloning vectors and host strains: Nucleotide sequence of the M13mp18 and pUC19 vectors. Gene **33,** 103–110.

Zambryski, P., Joos, H., Genetello, C., Leemans, J., Van Montagu, M., and Schell, J. (1983). Ti plasmid vector for the introduction of DNA into plant cells without alteration of their normal regeneration capacity. EMBO J. **2,** 2143–2150.