The Plant Cell, Vol. 5, 781–794, July 1993 © 1993 American Society of Plant Physiologists

Soybean Nodulin-26 Gene Encoding a Channel Protein Is Expressed Only in the Infected Cells of Nodules and Is Regulated Differently in Roots of Homologous and Heterologous Plants

Guo-Hua Miao and Desh Pal S. Verma[1]

Department of Molecular Genetics and Biotechnology Center, The Ohio State University, Columbus, Ohio 43210

Nodulin-26 (N-26) is a major peribacteroid membrane protein in soybean root nodules. The gene encoding this protein is a member of an ancient gene family conserved from bacteria to humans. N-26 is specifically expressed in root nodules, while its homolog, soybean putative channel protein, is expressed in vegetative parts of the plant, with its highest level in the root elongation zone. Analysis of the soybean N-26 gene showed that its four introns mark the boundaries between transmembrane domains and the surface peptides, suggesting that individual transmembrane domains encoded by a single exon act as functional units. The number and arrangement of introns between N-26 and its homologs differ, however. Promoter analysis of N-26 was conducted in both homologous and heterologous transgenic plants. The *cis*-acting elements of the N-26 gene are different from those of the other nodulin genes, and no nodule-specific *cis*-acting element was found in this gene. In transgenic nodules, the expression of N-26 was detected only in the infected cells; no activity was found in nodule parenchyma and uninfected cells of the symbiotic zone. The N-26 gene is expressed in root meristem of transgenic *Lotus corniculatus* and tobacco but not in untransformed and transgenic soybean roots, suggesting the possibility that this nodulin gene is controlled by a *trans*-negative regulatory mechanism in homologous plants. This study demonstrates how a preexisting gene in the root may have been recruited for symbiotic function and brought under nodule-specific developmental control.

INTRODUCTION

Root nodules, an organ sui generis, are inhabited by compatible rhizobia that fix dinitrogen symbiotically. In an effective symbiosis, several thousand bacteria invade a single host cell by a process resembling endocytosis. "Internalization" of bacteria requires formation of a peribacteroid membrane (PBM; Robertson et al., 1978; Verma et al., 1978) in the infected cells. This novel membrane keeps the bacteria compartmentalized from the cytoplasm of the host and serves as a primary interface for both nutrient and signal exchanges between the two partners (Brewin, 1991; Verma, 1992*a*). Mutations in bacteria affect synthesis and stability of the PBM, as shown by *Bradyrhizobium* mutants that form empty PBM vesicles (Regensburger et al., 1986; Roth and Stacey, 1989) and degenerate the PBM (Werner et al., 1985), resulting in ineffective nodules. These observations suggest that formation of the PBM is pivotal for symbiosis.

Prior to or during the formation of the PBM, a number of nodulin genes (Legocki and Verma, 1980) are activated, some of which encode proteins specifically targeted to the PBM (Fortin et al., 1985). Nodulin-26 (N-26) is a major 26-kD PBM protein in soybean root nodules (Fortin et al., 1987; Miao et al., 1992). The topology of N-26 in the PBM has recently been determined. This protein spans the PBM six times with both the amino and carboxyl termini facing the host cytoplasm (Miao et al., 1992). It is phosphorylated by a Ca^{2+}-dependent and calmodulin-independent protein kinase located in the PBM (Weaver et al., 1991; Miao et al., 1992). N-26 resembles other recently identified channel proteins, including an *Escherichia coli* pore-type channel, glycerol facilitator (GlpF; Muramatsu and Mizuno, 1989; Sweet et al., 1990), a yeast membrane protein affecting sugar metabolism (FPS1; Van Aelst et al., 1991), a mammalian eye lens major intrinsic protein (MIP-26; Gorin et al., 1984), a Drosophila neurogenic protein (BIB; Rao et al., 1990), an erythrocyte integral membrane protein (CHIP-28; Preston and Agre, 1991), a tobacco root meristem–specific protein (TobRB7; Yamamoto et al., 1990), a tonoplast integral membrane protein (TIP; Johnson et al., 1990; Höfte et al., 1992), and a turgor-regulated protein of pea (7*a*; Guerrero et al., 1990). Whereas GlpF facilitates translocation of glycerol and other small molecules (Sweet et al., 1990), CHIP-28 (Preston et al., 1992), and TIP (Maurel et al., 1993) have been shown to be

[1] To whom correspondence should be addressed at: The Ohio State Biotechnology Center, 1060 Carmack Road, Columbus, OH 43210.

water channels. This suggests that N-26 may also be a channel protein that facilitates transport of small molecules across the PBM. Recently, based on the correlation between phosphorylation status of the N-26 and malate uptake, this protein has been suggested to be involved in dicarboxylate transport across the PBM of soybean nodules (Ouyang et al., 1991). Direct demonstration of the biochemical function of this channel protein, however, remains to be ascertained.

Several nodulin genes have been isolated and characterized. The expression of these genes is primarily controlled at the transcriptional level. Detailed analyses of promoter regions of the two late nodulin genes encoding leghemoglobin (Lbc3; Stougaard et al., 1987; Szabados et al., 1990) and nodulin-23 (Jørgensen et al., 1988, 1991), a PBM-specific protein (Jacobs et al., 1987), have identified organ-specific positive and negative *cis*-acting elements. Moreover, a number of consensus sequences in the 5′ promoter regions of these genes have been identified (Mauro et al., 1985; Sandal et al., 1987). On the other hand, the nodulin-35 gene (Nguyen et al., 1985), expressed in uninfected nodule cells of the ureide-producing legumes, lacks these conserved sequences, suggesting that different *cis*-acting sequence motifs may be involved in controlling the expression of different nodulin genes.

Expression of the soybean N-26 gene was detected in nodules elicited by a *Bradyrhizobium* mutant T8-1 (Morrison and Verma, 1987) in which bacteria were unable to release from the infection thread, indicating that the N-26 gene is induced prior to endocytosis of bacteria. Although N-26 is only expressed in soybean nodules and not in roots (Fortin et al., 1987; Miao et al., 1992), the widespread presence of N-26 homologs in other plants (Pao et al., 1991; see also Verma, 1992b) raised the possibility that soybean N-26 or some of its homologs may also be expressed in other tissues of the soybean plant.

We report here the isolation, characterization, and expression of the N-26 gene and its two homologs in different tissues of soybean. The significant sequence diversion, intron–exon organization, and developmental expression patterns of the N-26 gene compared to its homologs provides compelling molecular evidence that the N-26 gene is derived from an ancient gene family encoding channel proteins to perform specific function in root nodules. Soybean N-26 is expressed in both nodules and root meristem of heterologous transgenic plants, while in homologous plants, this gene is only expressed in the infected cells of nodules. The promoter region of the N-26 gene was analyzed to identify sequences responsible for conferring nodule-specific and root meristem–specific expression in transgenic plants.

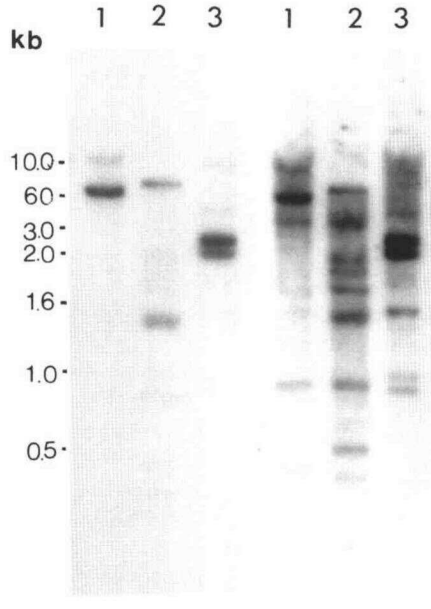

Figure 1. Genomic DNA Gel Blots of the Soybean N-26 Gene under High- and Low-Stringency Hybridization Conditions.

At left is soybean genomic DNA digested with BamHI (lane 1), HindIII (lane 2), and KpnI (lane 3) hybridized under high-stringency conditions. At right is the same gel blot as shown at left, except that hybridization was performed at a lower stringency. The 1-kb DNA molecular length markers were used as the standard. Hybridization conditions are as given in Methods.

RESULTS

Cloning of N-26 Homologs from Soybean

Hybridization of soybean genomic DNA with ^{32}P-labeled N-26 cDNA (Miao et al., 1992) revealed the complexity of the N-26 gene family in soybean. Under high-stringency hybridization conditions, one major BamHI fragment (6.2 kb), two HindIII fragments (7.0 and 1.4 kb), and two KpnI fragments (2.3 and 2.0 kb) were observed, as shown in Figure 1, which is consistent with the restriction map of an N-26 genomic clone (λGM16; see below). However, when the hybridization was performed at a lower stringency, several bands were detected (Figure 1), suggesting that N-26 is a member of a gene family. Although a single gene encodes a nodule-specific sequence, some of the other members of this family may be expressed in different parts of the plant.

To isolate N-26 homologs from soybean, two degenerate primers (see Methods) corresponding to the conserved sequence motif, as shown in Figure 2, among known MIP gene family members were synthesized and used for polymerase chain reaction (PCR) with cDNAs reverse transcribed from total RNAs from soybean root and leaf. The amplified products of expected size were confirmed to be related to N-26 by sequencing, and they were used as probes to screen a soybean nodule cDNA library (Delauney and Verma, 1990). Two distinct homologs of N-26 cDNA (encoding soybean putative channel proteins; SPCP1 and SPCP2) were thus isolated and corresponded to the sequences initially amplified by PCR from the root and leaf cDNAs, respectively (Figure 2). The SPCP1

Figure 2. Amino Acid Sequence Comparison of Soybean N-26, SPCP1, and SPCP2, and Arabidopsis γ-TIP.

The shaded areas mark the regions of transmembrane domains (A to F). The box represents the most conserved sequence motif in the MIP family. Arrows indicate the regions of the sequence used for designing the degenerate oligonucleotides for PCR to clone soybean SPCP1 and SPCP2. Dashes represent the conserved amino acid residues among N-26, SPCP1, SPCP2, and Arabidopsis γ-TIP.

and SPCP2 cDNAs encode primary translation products of 249 and 255 amino acids with calculated molecular masses of 25 and 26 kD, respectively. Both peptides have the putative six-transmembrane domains characteristic of the MIP family (Figure 2). Sequence homology comparison revealed that SPCP1 and SPCP2 share only 32 and 35% amino acid sequence identity with the N-26 peptide (Figure 2). However, SPCP1 shares 80 and 59% amino acid identity with Arabidopsis γ-TIP and α-TIP, respectively (Höfte et al., 1992). Hence, it is likely that SPCP1 is the homolog of Arabidopsis γ-TIP. Similarly, SPCP2 displays 78 and 57% amino acid identity with the Arabidopsis γ-TIP and α-TIP (Höfte et al., 1992). In spite of the significant sequence homology with γ-TIP and α-TIP, SPCP2 appears to represent a new member of this gene family because the developmental expression pattern of SPCP2 is different from that of γ-TIP or α-TIP genes (see below). The significant sequence differences between N-26 and its homologs indicate that N-26 has diverged from the other members of the family during early evolution, although they all share the basic six-transmembrane structural feature.

Differential Expression of the N-26 Gene and Its Homologs in Soybean

To determine the developmental expression patterns of N-26, SPCP1, and SPCP2 in different tissues, total RNA from different parts of the soybean plant was analyzed by primer extension using sequence-specific oligonucleotides corresponding to the 5′ nontranslated regions of N-26, SPCP1, and SPCP2 cDNAs (see Methods). This allowed us to eliminate the possibility of cross-hybridization by RNA gel blot analysis between the related sequences of this gene family. Figure 3A shows that a major extension product was made from a primer specific to the N-26 sequence using RNA from nodules, but no extension product could be detected using RNAs from root meristem, root elongation zone, leaf meristem, mature leaf, or the flower. This is consistent with the protein gel blot analysis showing the presence of N-26 only in the PBM (Miao et al., 1992). One major and two minor extension products were detected using a primer specific to SPCP1 with RNA from the root elongation zone, although lower levels of expression were also detected in all vegetative tissues (Figure 3B). This further supports the contention that SPCP1 is a soybean homolog of Arabidopsis γ-TIP, presumably involved in cell elongation in root (Ludevid et al., 1992). A moderate level of SPCP1 expression was also found in root tips of soybean. Under the same experimental conditions, SPCP2 was not detected by primer extension or RNA gel blot analysis in RNAs from vegetative tissues, dry seeds, or germinating cotyledons (data not shown). Thus, we reason that the expression level of SPCP2 in the plant may be extremely low and beyond the sensitivity of detection by these methods. The difference in the expression pattern of SPCP2 from that of TIP-type genes (Höfte et al., 1992; Ludevid et al., 1992) suggests that SPCP2 may represent a new member of this gene family. The integrity of total RNA from different tissues was confirmed by gel blotting with

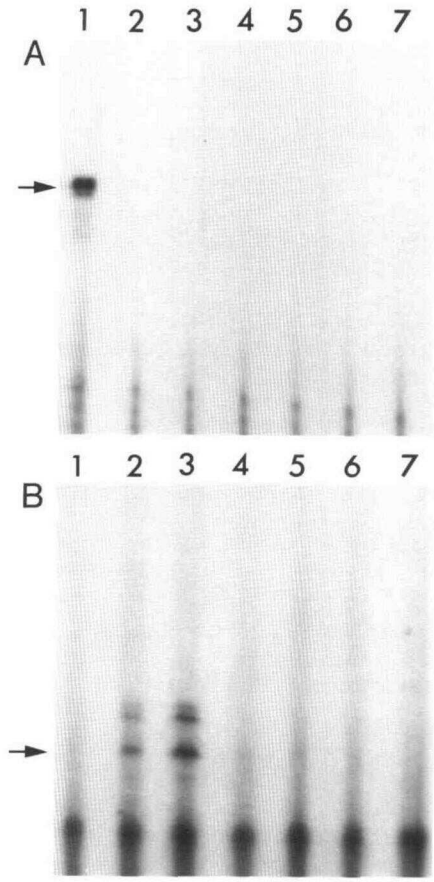

Figure 3. Expression Levels of the Soybean N-26 and SPCP1 Transcripts in Different Tissues as Determined by Primer Extension Analysis.

(A) Soybean N-26.
(B) Soybean SPCP1.
Lanes 1 contain total RNA from nodules; lanes 2, root meristem; lanes 3, root elongation zone; lanes 4, leaf; lanes 5, leaf meristem; lanes 6, flower; lanes 7, control without the addition of RNA. The size of the major primer extension product (arrow; **[A]**, lane 1) corresponds to the initiation site at position −85 from the translation start codon, as shown in Figure 6. The sizes of the SPCP1 primer extension products **(B)** were not determined. Arrow **(B)** indicates the major extension product.

different cDNA probes, including soybean glutamine synthetase and cdc2 protein kinase (data not shown). These results demonstrate that N-26 is indeed specifically induced during the development of root nodules. Because the N-26 transcript was not detected in soybean root meristem or in the root elongation zone, it appears that the regulation of the N-26 gene is different from that of the root meristem–specific gene TobRB7 in tobacco (Conkling et al., 1990; Yamamoto et al., 1991) and γ-TIP in Arabidopsis (Höfte et al., 1992; Ludevid et al., 1992).

Structure of the Soybean N-26 Gene and the Correlation of Intron-Exon Organization with the Topology of the Encoded Peptide Domains

Two of the six positive clones isolated by screening a soybean genomic library with a full-length N-26 cDNA probe (Miao et al., 1992) were found to contain the complete N-26 gene based on restriction enzyme digestion patterns and their ability to hybridize to specific oligonucleotides from both the 5′ and 3′ ends of the N-26 cDNA. The restriction map of one of the two clones, λGM16, is shown in Figure 4A. The other four clones encode two different sequences, which cross-hybridize to the full-length N-26 cDNA but have different restriction patterns and fail to hybridize either to the 5′- or 3′-specific probes. These clones may represent different members of the N-26 gene family of soybean.

The complete N-26 gene (Figure 4A) encompasses 4467 bp, including both the 5′ and 3′ noncoding regions, and is contained within a 6.2-kb BamHI fragment of λGM16. The sequence of the N-26 gene shows 61 nucleotide differences in the coding region in comparison to the cDNA, resulting in 36 amino acid differences over 271 amino acids (Figure 4B). There are only two nucleotide differences in the 77 bp of the 5′ noncoding region and 10 nucleotides in the 182 bp of the 3′ noncoding region of the two sequences (Figure 4B). The high percentage of similarity at both the 5′ and 3′ noncoding regions strongly suggests that λGM16 indeed encodes N-26. This was further supported by the genomic DNA gel blot analysis (Figure 1) showing a single 6.2-kb BamHI fragment that strongly hybridized to the N-26 cDNA. The differences in the coding region found in the N-26 gene and cDNA may be due to the soybean variety differences between the genomic clone (soybean cultivar Dare) and cDNA clone (soybean cultivar Prize), as observed previously (Suzuki and Verma, 1991). They may also represent two allelic genes (Miao et al., 1991).

A correlation between the location of introns in the N-26 gene and topology of the peptide in the membrane (Miao et al., 1992) was observed and is presented schematically in Figure 5A. Intron I marks the boundary between the N-terminal hydrophilic peptide and the first transmembrane domain, A. The second and third introns are similarly located between the junction of hydrophilic connecting loops (surface peptides 2 and 4) and transmembrane domains C and E, respectively. The smallest intron interrupted transmembrane domain E close to the surface peptide 5. The tobacco TobRB7 (Yamamoto et al., 1991) and Arabidopsis α-TIP (Höfte et al., 1992) genes both contain only two introns. The locations of the introns in the latter two genes are precisely conserved and are located at the sequences corresponding to the ends of the first and third transmembrane domains of the two peptides (Figure 5B). Arabidopsis γ-TIP contains only one intron, which is located at the same position as the second intron of the TobRB7 and α-TIP genes (Figure 5B). The differences in intron organization between the N-26 gene and its homologs further support the hypothesis that the N-26 gene has significantly diverged

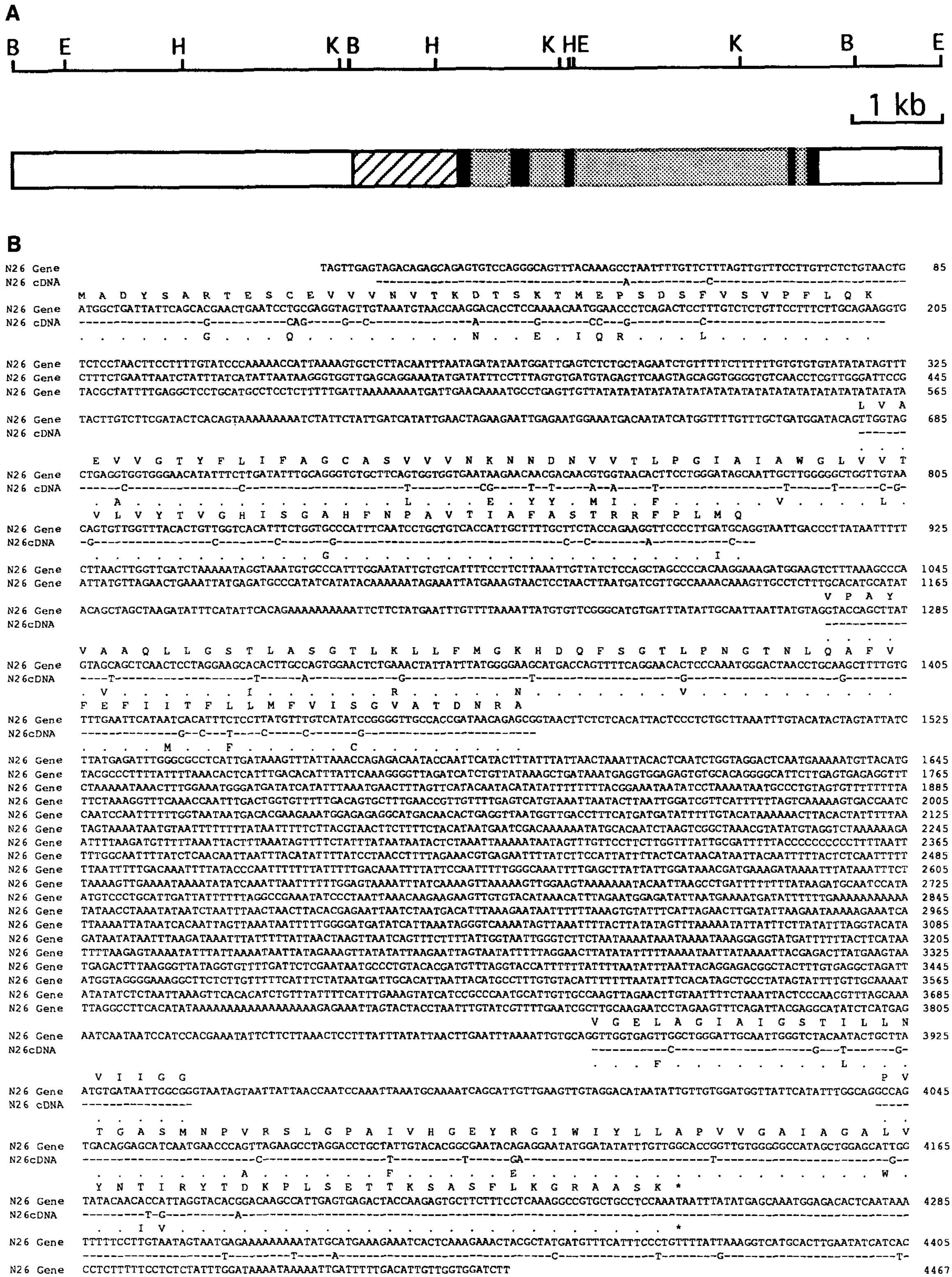

Figure 4. Organization and Structure of the Soybean N-26 Gene.

(A) Restriction map of λGM16 containing the soybean N-26 gene. Solid areas denote exons; stippled areas, introns; hatched area, 1344-bp 5′ flanking region containing promoter activity. Restriction enzyme sites are as follows: B, BamHI; E, EcoRI; H, HindIII; K, KpnI.

(B) Comparison of the nucleotide sequences of the N-26 gene starting from the transcription start site to the poly(A) addition site and the soybean N-26 cDNA (Miao et al., 1992). Amino acid sequences deduced from the N-26 gene and cDNA are given above and below the nucleotide sequences, respectively. Dashes indicate identical nucleotides; periods designate identical amino acids in the genomic and cDNA sequences.

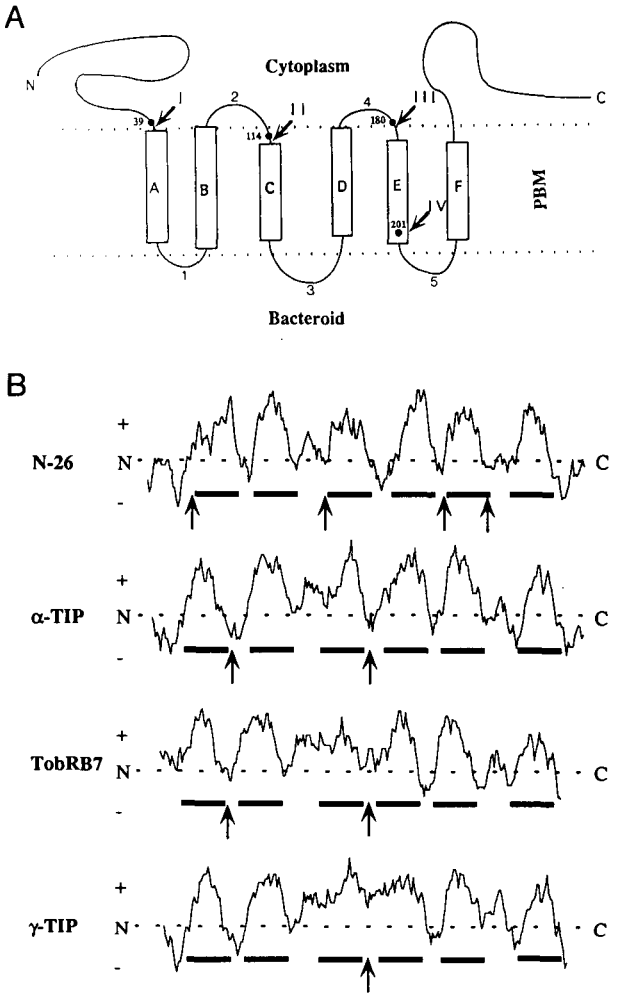

Figure 5. Topology of N-26 in the PBM and Comparison of Intron Organization with Other Plant Members of This Gene Family.

(A) The positions of introns marking the boundaries of transmembrane domains (A to F) and connecting loops (1 to 5) (see Miao et al., 1992). Dots indicate the positions of amino acid residues, and arrows mark the positions of introns.

(B) Comparison of intron organization in relation to transmembrane domains among different members of plant N-26 homologs, including α-TIP, γ-TIP (Höfte et al., 1992), and TobRB7 (Yamamoto et al., 1991). Arrows indicate the positions of introns, and bars show the putative transmembrane domains. N and C denote amino and carboxyl termini, respectively.

from the other members of the family (Figure 2). On the other hand, a correlation between the intron position and transmembrane domains of the peptides is observed in all members of this gene family for which the genomic DNA sequences have been determined (Figure 5B). This indicates that introns marking the boundaries of transmembrane domains is a common feature of this group of membrane proteins.

Transcription Initiation and Putative *cis*-Regulatory Elements of the N-26 Gene Involved in Root- and Nodule-Specific Expression

The transcription initiation site of the N-26 gene was determined by primer extension analysis of total nodule RNA using a 34-nucleotide primer corresponding to the sequence adjacent to the translation initiation site, as shown in Figure 6A (see also Methods). A major primer extension product was observed corresponding to the initiation start site at position −85 (indicated by an open triangle; Figures 6A and 6B) from the translation initiation codon (Figure 6A).

The nodule-specific elements (5′-AAAGAT-3′ and 5′-CTCTT-3′; Stougaard et al., 1987, 1990) located upstream of the CAAT box in Lbc3 and nodulin-23 genes were not found in the N-26 promoter. Instead, one of the elements (5′-CTCTT-3′) is present between the TATA and CAAT boxes. This is consistent with the observation that the expression pattern of the N-26 gene is different from that of Lb and nodulin-23 genes (Morrison and Verma, 1987). Three 8-bp direct repeats (5′-ATTGTATT-3′) were found in the region of the N-26 promoter between −206 and −157. The space between these direct repeats is exactly the same (13 bp) and comprised of mostly A residues. An inverted repeat sequence (5′-TTATGAACGATCGTTCCTAA-3′), with only one mismatch as indicated by underlining, was also found at position −598 (Figure 6A). Both sequence motifs have been shown to play roles in controlling the expression of this gene, as revealed by deletion analysis (see below).

The Soybean N-26 Promoter Confers Both Root Meristem- and Nodule-Specific Expression in Heterologous Transgenic Plants but Is only Expressed in Infected Cells of Soybean Nodules

As diagrammed in Figure 7, a transcriptional fusion construct of 1259 bp of the N-26 promoter with the GUS (*E. coli uidA*) reporter gene (D-0) was created and was strongly expressed in nodules of transgenic *L. corniculatus*, as shown in Table 1 and Figure 8. Surprisingly, the GUS activity was also found in the roots of transgenic *Lotus* (Table 1) and tobacco plants (data not shown), although activity in the entire root is much less as compared to that in nodules (Figure 8). However, in the root meristem, GUS activity is comparable to that in nodules (see below). No GUS activity was detected in the leaves. When the same construct was introduced into soybean, no GUS activity was detected in transgenic roots and GUS activity was detected only in nodules, as illustrated in Figure 9. Extended incubation of transgenic soybean roots with the GUS substrate did not yield any GUS activity, and neither was GUS activity detected by fluorometric assay. This is consistent with the primer extension experiment (Figure 3A) and protein gel blot analysis (Miao et al., 1992) showing expression of the N-26 gene only in the nodules. The difference in the expression of the N-26 gene in homologous and heterologous transgenic plants suggests that a control mechanism may exist in soybean

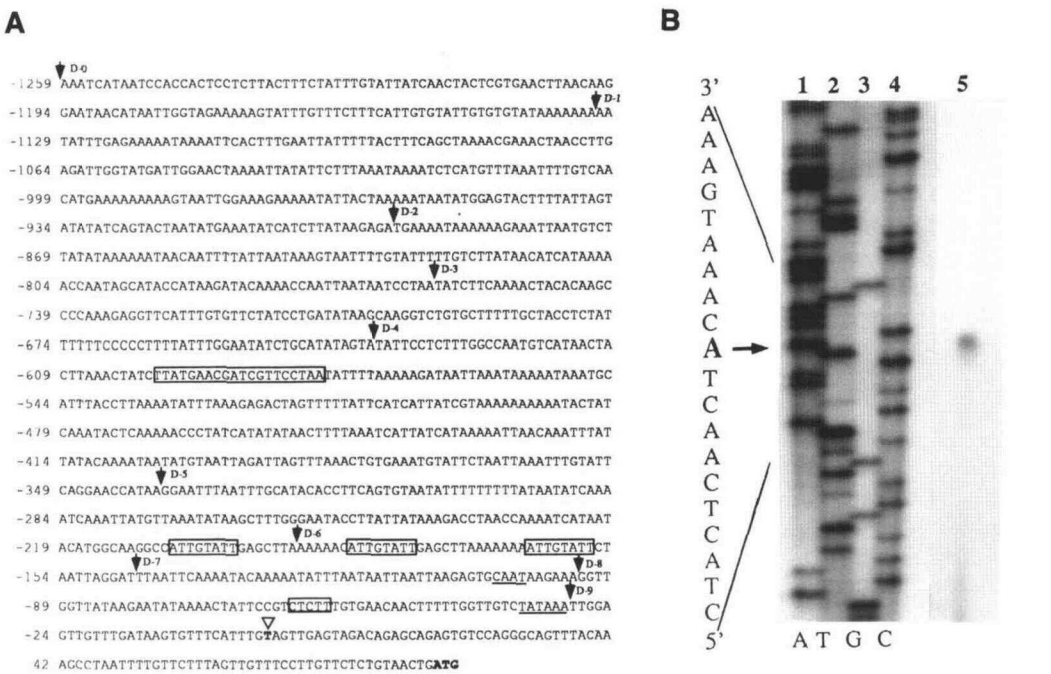

Figure 6. The 5′ Flanking Sequence of the Soybean N-26 Gene and Transcription Initiation Site.

(A) The 1344-bp upstream region from ATG containing TATA and CAAT boxes (underlined). The translation start site is in bold letters, and the major transcription initiation site is marked by an open triangle. Three perfect direct repeats between −206 and −156 and an inverted repeat at position −598 are boxed. In addition, a conserved nodule-specific sequence (5′-CTCTT-3′) that is present between TATA and CAAT is also boxed. Arrows indicate the 5′ end sequence of different constructs (D-0 to D-9), which are shown in Figure 7.

(B) Primer extension analysis showing the transcription start site of the N-26 gene. A sequencing reaction (lanes 1 to 4) was run next to the primer extension (lane 5) using the same primer to determine the position of the transcription start site indicated by the arrow (−85 from the translation start codon).

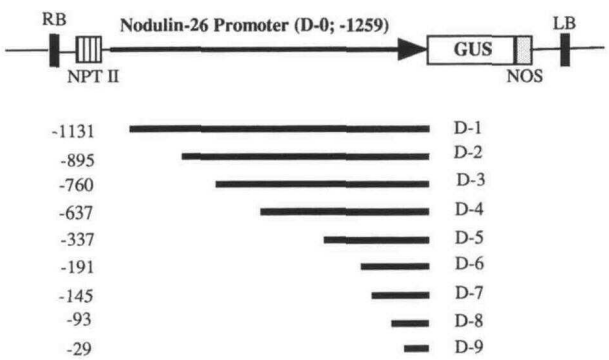

Figure 7. Construction of N-26–GUS Reporter Genes and Different Deletion Derivatives.

The 1259-bp 5′ flanking region starting from the N-26 transcription initiation site was fused to the GUS reporter gene and the nopaline synthetase (NOS) 3′ end sequences and placed downstream of the neomycin phosphotransferase II (NPT II) gene in pBI101.2. RB and LB are right and left T-DNA borders, respectively. Different deletions (D-1 to D-9) were constructed as described in Methods, and their sequences were determined (see Figure 6A); the sizes of the promoter fragments are indicated on the left.

that suppresses the expression of this gene in the roots of homologous plants.

In roots of transgenic *Lotus*, GUS activity was confined to the root meristem (Figures 8A and 8B). Moreover, this activity can be detected prior to the emergence of lateral roots (Figure 8A). Localization of N-26 in the root meristem corresponds to the expression pattern of the tobacco N-26 homolog TobRB7 (Yamamoto et al., 1991). This suggests a common control mechanism shared between these two genes. In nodules, GUS activity was localized in the infected cells only (Figures 8C, 8D, and 9E), consistent with the fact that N-26 is a PBM protein (Fortin et al., 1987; Miao et al., 1992). This activity was detected in very young nodules (Figure 9C). Thus, several independent lines of experiments demonstrated that the N-26 gene is one of the truly symbiotically induced host genes.

cis-Regulatory Elements of the N-26 Gene Required for Root Meristem- and Nodule-Specific Gene Expression in Transgenic *Lotus*

To further dissect the essential regulatory sequence motifs controlling the tissue-specific expression of the N-26 gene, a

Table 1. GUS Activity in Leaves, Roots, and Nodules of Transgenic *L. corniculatus* Plants Expressing the Soybean N-26 Promoter (D-0; 1259 bp) Fused with the GUS Reporter Gene

	GUS Activity (pmol MU/mg protein/min)		
Construct	Leaf	Root	Nodule
N-26–GUS (D-0)	15.8 ± 22.2	197 ± 61	1574 ± 651
pBI101	1.5 ± 2.5	6.5 ± 0.5	5 ± 6

GUS activities of transgenic plants carrying the N-26–GUS construct and the pBI101 represent an average value (mean ± SD) of seven and two independent transgenic plants, respectively. MU, methylumbelliferone.

number of 5′ end deletions were constructed and fused to the GUS reporter gene, as shown in Figure 7. The breakpoint for each construct, indicated in Figure 6A, was verified by sequencing. GUS activity was detected in both roots and nodules in transgenic *Lotus* with all deletion constructs containing sequences longer than 276 bp (−191 to +85, D-6), although GUS activity differed between roots and nodules, as shown in Figure 10. GUS activity in transgenic *Lotus* roots was confined to the root meristem, as shown in Figure 8. Similar expression patterns were found in transgenic tobacco (data not shown). In all cases, the deletion construct (D-6), containing −191 bp of the 5′ flanking sequence, was still capable of directing both root meristem– and nodule-specific expression, whereas further deletion of 46 bp (D-7) resulted in the complete loss of GUS activity in both tissues. This suggests that the sequence between −145 and −191 contains one or more indispensable *cis*-regulatory elements conferring both root meristem and nodule specificity. Further analysis revealed the presence of three perfect direct repeats in this region (Figure 6A). In construct D-7, where both proximal sequence elements were deleted, no GUS activity was detected. Thus, we conclude that the distal repeat element is dispensable for the expression of this gene in the root meristem and nodule, and the 5′ flanking sequence, including two proximal repeat elements, is essential for the root meristem– and nodule-specific expression.

An enhancement in GUS activity was observed when the region between the −191 and −337 was included; the activity of D-6 dropped almost threefold in comparison to that of D-5. This may suggest a role of the distal direct repeat sequence in controlling the level of expression, although it is clearly dispensable for both root meristem– and nodule-specific expression. A comparison of GUS activity in transgenic *Lotus* expressing different deletion constructs also allowed identification of a weak negative regulatory sequence located between positions −337 and −637. In this region, an inverted repeat element was found. Sequential deletions from the 5′ end of the promoter did not allow us to separate sequence element(s) controlling root meristem– and nodule-specific expression. Thus, the *cis*-regulatory elements controlling expression of the N-26 gene in root meristem and nodules most likely overlap and are not separable.

DISCUSSION

The Soybean N-26 Gene Is Diverged from an Ancient Gene Family Encoding Channel Proteins and Has Been Brought under Nodule Development Control

N-26 is a member of a gene family conserved from bacteria to mammals (Pao et al., 1991; Verma, 1992b). Although most of these proteins (MIP, CHIP-28, and GlpF) are located in the plasma membrane, TIP (Johnson et al., 1990; Höfte et al., 1992) and N-26 (Fortin et al., 1987; Miao et al., 1992) are present in the vacuole membrane and PBM, respectively. A number

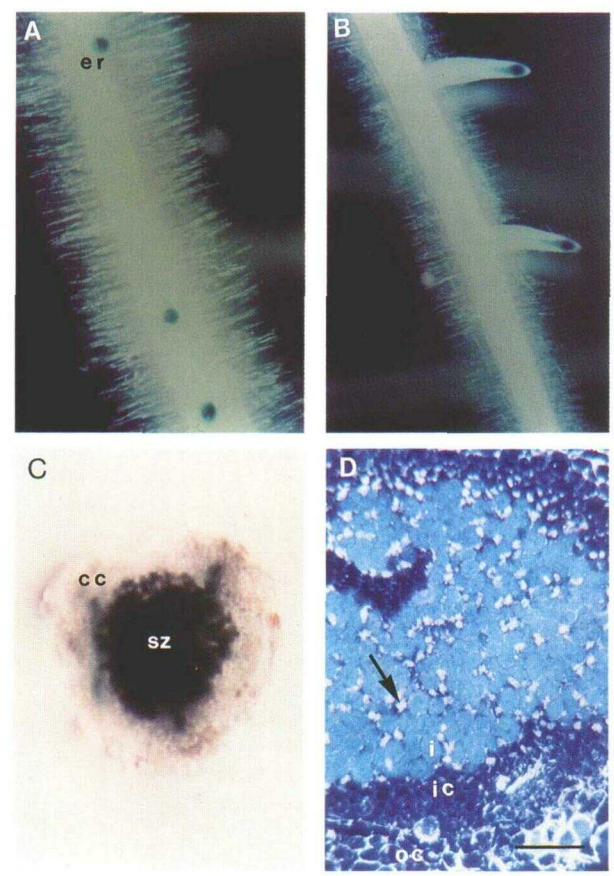

Figure 8. Expression of the N-26–GUS Gene Fusion in Roots and Nodules of Transgenic *L. corniculatus*.

(A) Histochemical localization of GUS activity in the lateral root meristem prior to any physical demarcation on the root surface.
(B) Localization of GUS activity in the lateral root meristem.
(C) Localization of GUS activity in the symbiotic zone of nodules.
(D) Thin section of a nodule showing the presence of GUS activity in the infected cells of the symbiotic zone of root nodules.
cc, cortex cells including both the inner and outer cortex; er, emerging root; i, infected cells; ic, inner cortex; oc, outer cortex; sz, symbiotic zone; arrow in **(D)** points to the uninfected cells. Bar = 200 μm.

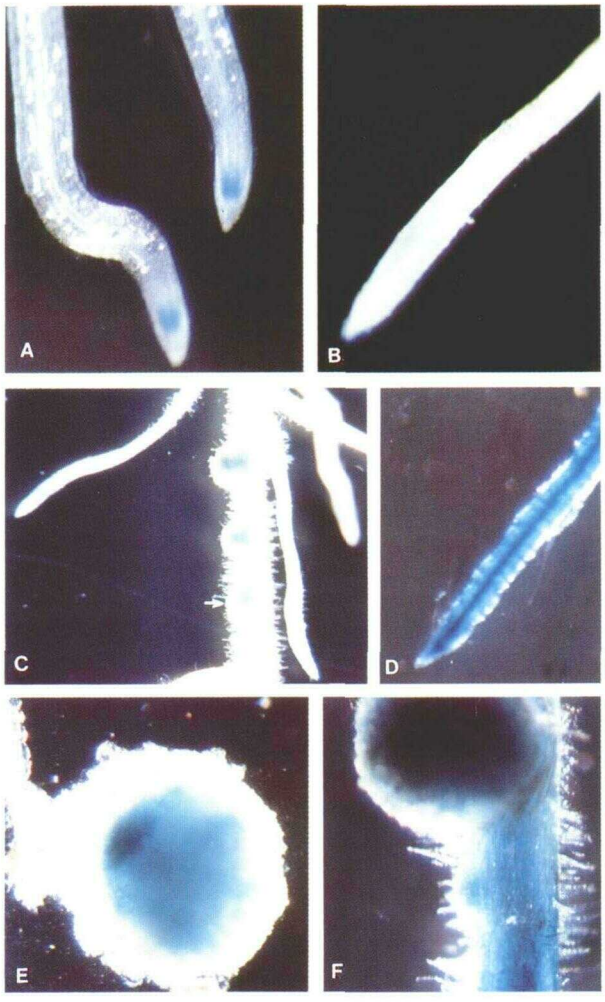

Figure 9. N-26–GUS Expression in Tissues of Transgenic Plants.

(A) to **(C)** and **(E)** show localization of the soybean N-26–GUS gene fusion in homologous and heterologous transgenic plants. **(D)** and **(F)** show GUS activity in transgenic soybean roots and mature nodules, respectively, expressing the cauliflower mosaic virus (CaMV) 35S–GUS construct.

(A) Root meristem of transgenic *L. corniculatus* expressing N-26–GUS.

(B) Transgenic soybean root expressing N-26–GUS.

(C) Young soybean nodules (including emerging nodules, arrow) expressing N-26–GUS.

(D) Soybean roots expressing CaMV 35S–GUS.

(E) Mature soybean nodules expressing N-26–GUS.

(F) Mature soybean nodule expressing CaMV 35S–GUS.

Note the lack of N-26–GUS activity in transgenic primary **(B)** and lateral roots **(C)** of soybean.

of N-26 homologs have been isolated from different tissues of plants (see Verma, 1992b). The presence of distinct homologs of N-26 in the same plant suggests that these genes are derived from a common ancestor. These homologs are differentially expressed in different parts of an organism. This wide distribution and evolutionary conservation suggest a fundamental role for these proteins in cell function.

Primer extension experiments using RNA from different tissues of soybean plant demonstrated that N-26 is specifically expressed in nodules (Figure 3A). Thus, N-26 is one of the truly symbiotically induced genes that is not expressed in any other part of the plant (cf. early nodulin genes; Nap and Bisseling, 1990). In contrast, its homolog, SPCP1, is primarily expressed in the root elongation zone with lower levels of expression in other vegetative tissue (Figure 3B). This is consistent with the expression pattern of Arabidopsis γ-TIP, which is confined to a narrow zone just behind the apical meristem (Höfte et al., 1992; Ludevid et al., 1992). The expression of SPCP2 is very low, and its tissue specificity is not clear. Phylogenetic analysis (Höfte et al., 1992) of the MIP gene family showed that the N-26 gene is evolutionarily distant from all other members of this family. The intron organization of N-26 is also quite different from the other members (Figure 5B). The N-26 gene apparently evolved from a preexisting host gene and has been brought under nodule developmental control.

The presence of different members of the N-26 gene family in soybean suggests that the different homologs may perform related but distinct functions in different tissues. It has been suggested that N-26 is involved in dicarboxylic acid transport across the PBM of soybean nodules (Ouyang et al., 1991). The *E. coli* GlpF and red blood cell CHIP-28 perform quite different transport functions. CHIP-28 acts as a water channel in red blood cells (Preston et al., 1992), whereas GlpF facilitates

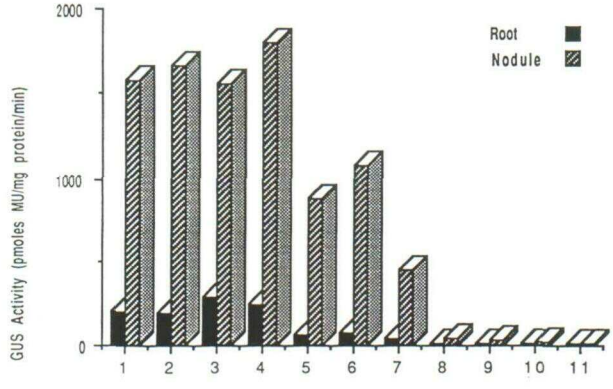

Figure 10. Deletion Analysis of the Soybean N-26 Promoter–GUS Reporter Gene Fusion.

Each construct, as shown in Figure 7, and pBI101 (used as control) were introduced into *L. corniculatus* (see Methods), and transgenic roots and nodules were assayed for the level of GUS expression. GUS activity was determined fluorometrically as described in Methods. Deletion constructs shown on the x-axis are as follows: 1, D-0; 2, D-1; 3, D-2; 4, D-3; 5, D-4; 6, D-5; 7, D-6; 8, D-7; 9, D-8; 10, D-9; 11, pBI101. GUS activity on the y-axis is given in picomoles of methylumbelliferone (MU) per minute per milligram of protein. Data points are the means of five to seven individual transformants, except for pBI101, which is the mean of two independent transformants.

transport of glycerol and other small molecules across the plasma membrane (Sweet et al., 1990). Arabidopsis α- and γ-TIP (Höfte et al., 1992; Ludevid et al., 1992) are both located in the tonoplast and may encode water channels (Maurel et al., 1993). However, these two genes display completely different expression patterns, and they may be involved in different developmental processes of the plant. Similarly, SPCP1 and SPCP2 are highly homologous at the amino acid sequence level (74% identity); however, the expression patterns of these two genes are clearly different.

Correlation of Introns in the N-26 Gene and Its Homologs with the Topology of the Proteins in the Membrane

The intron locations are precisely conserved among TobRB7, α-TIP, and γ-TIP, whereas introns of the N-26 gene are located at different positions (Figure 5). Sequence comparison of N-26 with the other members of this gene family shows only 30 to 35% amino acid identity (Figure 2); however, a homology of 55 to 80% was found between TIP and SPCP. These data provide molecular evidence that N-26 has distantly diverged from the other members of this gene family.

A correlation between the location of introns and topology of the N-26 peptide was observed (Figure 5). The introns I, II, and III interrupt the amino acid residues located in the hydrophobic connector loops, while the intron IV interrupts transmembrane domain E close to surface peptide 3 (Figure 5A). No intron was found in the regions conserved among members of the MIP family (Figure 2). Despite the differences in the locations of introns in the N-26 gene, a correlation between intron position and transmembrane domains of the peptides exists in all members of this gene family (Figure 5B). That the locations of introns generally mark the boundary of transmembrane regions suggests the functional unity of the domains encoded by individual exons. A correlation between intron location and peptide domains has also been observed in other proteins (Nathans and Hogness, 1983; Liscum et al., 1985; Kopito et al., 1987).

Two Direct Repeat Sequence Elements Confer Both Root Meristem- and Nodule-Specific Expression of the N-26 Gene in Transgenic Legume Plants

In transgenic nodules, N-26 GUS activity was localized in the infected cells of both heterologous (Figures 8C and 8D) and homologous (Figure 9E) plants. On the other hand, soybean nodulin-35, encoding uricase, has been shown to be primarily localized in the uninfected cells of soybean nodules (Nguyen et al., 1985). Thus, specific regulatory mechanisms may exist for controlling the cell-type–specific expression of nodulin genes.

Several *cis*-acting sequence elements controlling nodulin gene expression have been identified in late nodulin genes (Stougaard et al., 1987, 1990; Jørgensen et al., 1988, 1991; Szabados et al., 1990), but no *trans*-acting factor specific to nodulin genes has yet been identified. The sequence motifs 5′-AAAGAT-3′ and 5′-CTCTT-3′ located upstream of the CAAT boxes of Lbc3 and nodulin-23 genes have been shown to be organ-specific elements essential for gene expression in nodules (Stougaard et al., 1987, 1990; Jørgensen et al., 1988; Szabados et al., 1990). These sequences are not present upstream of the CAAT box in the N-26 promoter, suggesting that different *cis*-acting sequence elements may be involved in controlling expression of this gene. Three perfect direct repeats with constant spacing between each repeat (Figure 6A) were identified in the N-26 promoter region, of which two proximal direct repeat elements are essential for the expression of this gene.

Differential Expression of the N-26 Gene in Roots of Homologous and Heterologous Plants Suggests Possible Involvement of a Negative Control Mechanism for Expression of the N-26 Gene in the Soybean Root

Similar to hemoglobin genes of both legumes (Szabados et al., 1990) and nonlegumes (Bogusz et al., 1988), the N-26 gene was expressed in roots and nodules of transgenic *Lotus* and roots of tobacco plants. However, expression of Lb genes was only determined in heterologous transgenic plants. We demonstrate here that the expression of the N-26 gene differs between homologous and heterologous systems, suggesting that tissue-specific genes may need to be assayed in homologous systems, if possible.

What mechanism is involved in silencing expression of the N-26 gene in soybean roots? Many specific *trans*-acting DNA binding proteins interact with general transcription machinery and bring about the regulation of gene expression (Benfey and Chua, 1990). Both positive and negative *cis* elements have been found to be involved in regulation of gene expression (Castresana et al., 1988; Martini et al., 1988; Barrett et al., 1992). The apparent difference in tissue-specific expression of the soybean N-26 gene in homologous and heterologous transgenic plants is unlikely to be due to the presence of negative *cis*-acting elements (Castresana et al., 1988; Gilgenkrantz et al., 1992) located upstream of the promoter region analyzed (1259 bp; D-0) because the construct containing 5 kb of the N-26 promoter still confers GUS activity in the root meristem of transgenic *Lotus* (data not shown).

The lack of a positive or the presence of a negative *trans*-acting factor(s) in roots of soybean might be responsible for silencing the expression of the N-26 gene in roots. Because 5′ deletion analysis did not allow us to separate *cis*-acting element(s) essential for root meristem– or nodule-specific gene expression (Figure 10), the same *cis*-acting element(s) that interacts with the corresponding *trans*-acting factor(s) in roots may be responsible for activating gene expression in nodules. Because N-26 is also active in tobacco root meristem (data

not shown), it further suggests that no nodule-specific *trans*-acting factor is involved in its activation. Although not proven, our results are consistent with the hypothesis that while one factor encoded by the same gene is capable of interacting with the *cis*-acting element(s) of the N-26 gene and activating it in nodules, a modified factor (Montmayeur and Borrelli, 1991; Mäkelä et al., 1992) may be predominant in soybean roots. The latter binds to the same site but is unable to activate the expression of this gene.

That a single gene encodes both activator(s) and repressor(s) controlling gene expression seems to be a common feature for several transcription factors (see Foulkes and Sassone-Corsi, 1992). Similar regulatory mechanisms have recently been observed in controlling the expression of mammalian genes (Thayer and Weintraub, 1990; Mäkelä et al., 1992). Such a mechanism may explain the evolution of a nodule-specific gene from a preexisting plant gene by silencing its expression in roots and allowing it to be expressed in nodules. Expression of hemoglobin genes of legumes (Szabados et al., 1990) and nonlegumes (Bogusz et al., 1988) has also been detected in roots of transgenic tobacco and roots and nodules of transgenic *Lotus*, which may follow the same regulatory mechanism as described here. All of these observations support the notion that nodule-specific gene expression evolved from a preexisting root-specific expression mechanism. Isolation of the genes encoding factor(s) responsible for this regulatory mechanism and study of the interaction of these factors with the *cis*-acting elements may provide further insight into the molecular mechanism of control of nodulin gene expression.

METHODS

Growth of Plants and Bacterial Strains

Plants were grown and maintained as described previously (Miao et al., 1991). *Agrobacterium tumefaciens* (LBA4404) and *A. rhizogenes* (A4) were used for tobacco and *Lotus corniculatus* transformations (Miao et al., 1991), and *A. rhizogenes* K599, kindly provided by P. M. Gresshoff (University of Tennessee, Knoxville), was used for the induction of hairy roots on soybean.

Isolation of the cDNA Clones of N-26 Gene Homologs

Two degenerate primers (P-A: 5′-CATT(C)TGAAT(C)CCT(C)GCTGTG-(C)AC-3′; P-B: 5′-GCT(A)GGGTTCATT(G)C(G)ATG(C)CT(A)CC-3′) corresponding to the conserved domains of the major intrinsic protein (MIP) gene family (Figure 2) were synthesized to amplify nodulin-26 (N-26) homologous sequences from soybean. Total RNAs from root and leaf were reverse transcribed into cDNAs that were used as templates for polymerase chain reaction (PCR) amplification using an RNA-PCR kit (Perkin-Elmer) according to the manufacturer's instructions. The amplified fragments of ~350 bp were cloned into pUC18 for sequencing. The confirmed DNA fragments were labeled with α-^{32}P-dATP and used for screening a soybean nodule cDNA library made in λ ZAPII (Delauney and Verma, 1990). The positive clones were isolated and sequenced by dideoxy nucleotide chain termination sequencing with Sequenase 2.0 (U.S. Biochemicals). Sequences of SPCP1 and SPCP2 cDNAs have GenBank accession numbers of L12257 and L12258, respectively.

RNA Preparation and Primer Extension Analysis

Soybean nodules, roots, root meristems, leaves, leaf meristems, and flowers were harvested and immediately frozen in liquid nitrogen for RNA preparation. Three-week-old nodules were used for total nodule RNA preparation. Root RNA was isolated from roots of 3-day-old germinating soybean seeds, including the root meristem, elongation zone, and lateral root–forming zone, and root meristem RNA was prepared from the meristematic tissues (2 mm from the root tip). Unopened flowers were used for flower RNA isolation. Total RNA preparation was made by the procedure described by Verwoerd et al. (1989).

Primer extension experiments were performed using a primer extension system (Promega) according to the manufacturer's instructions. Two micrograms of total RNA from different tissues was annealed with the γ-^{32}P-ATP–labeled 5′-specific primer for N-26 (5′-CGCGGATCCA-GTTACAGAGAACAAGGAAACAACT-3′; see also Figure 6A) adjacent to the translation initiation codon of the N-26 cDNA or a 5′-specific primer for SPCP1 (5′-GACTTATGAAGGATCAAGTGAACACCG-3′) or SPCP2 (5′-CTTCTGCTTCTTTCTCTAAGATTGTTTCG-5′) corresponding to the 5′ untranslated region of SPCP1 and SPCP2 at 55°C for 15 min. A reverse transcription reaction was performed with 50 units of avian myeloblastosis virus reverse transcriptase containing 50 mM Tris-HCl, pH 8.3, 10 mM $MgCl_2$, 50 mM KCl, 10 μM DTT at 42°C for 30 min. The extension products were resolved on an 8% acrylamide gel following denaturation of samples at 90°C for 5 min in the presence of 40% formamide. A sequencing reaction was performed using the same primer and the N-26 genomic clone (pGM26-6.2) as template, and the product was electrophoresed next to the primer extension reaction to determine the size of the extension products.

Screening of a Genomic DNA Library and Characterization and Sequencing of the N-26 Gene

A soybean genomic library constructed in the λEMBL3 vector was screened by plaque hybridization at 55°C using N-26 cDNA (Miao et al., 1992) as probe. Several positive clones were recovered, and phage DNAs were isolated. Restriction maps of positive clones were constructed with the Lambda Map System according to the manufacturer's instructions (Promega). The approximate locations of the 5′ and 3′ ends of the N-26 gene were mapped by hybridization with oligonucleotides corresponding to the 5′ and 3′ ends of the N-26 cDNA in a 6 × SSC (1 × SSC is 150 mM NaCl, 15 mM sodium citrate, pH 7.0) hybridization solution at 42°C (Kinnaird et al., 1982). The 6.2-kb BamHI fragment containing the entire N-26 gene (λGM16) was subcloned into the pUC19 plasmid (pGM26-6.2), and nested deletions were made by exonuclease III digestion (Henikoff, 1984) for DNA sequencing. DNAs from overlapping deletions covering the N-26 gene were purified and directly sequenced by dideoxy nucleotide chain termination (Sanger et al., 1977). In certain regions, specific oligonucleotide primers were synthesized and used for double-stranded DNA sequencing. The DNA sequence data were compiled and analyzed with MacVector (International Biotechnologies, Inc., New Haven, CT) and the University of Wisconsin Genetics Computer Group (Madison) programs. The

soybean N-26 gene sequence has a GenBank accession number of M94764.

Genomic DNA Gel Blot Analysis

Genomic DNA was isolated from young soybean leaves according to the method of Dellaporta et al. (1983). Ten micrograms of soybean genomic DNA was digested with BamHI, HindIII, and KpnI and subjected to agarose gel electrophoresis and blotted on GeneScreen membranes (DuPont, New England Nuclear). The hybridization was performed, as described by Mahmoudi and Lin (1989), with ^{32}P-dCTP–labeled N-26 cDNA for 24 hr at 65°C for high stringency or 55°C for low stringency. The filters were washed twice in 40 mM phosphate buffer, pH 7.2, 5% SDS, 1 mM EDTA for 15 min, followed by washing in 40 mM phosphate buffer, pH 7.2, 1% SDS, 1 mM EDTA for 30 min at the hybridization temperature and exposed to x-ray film.

5′ Nested Deletions of the N-26 Promoter and Reporter Gene Fusions

A 1.34-kb fragment from the 5′ flanking region of the N-26 gene was obtained by PCR. An oligonucleotide (same as described for primer extension, see above) with a BamHI linker added at the 5′ end was used as the first primer, and the reverse primer that annealed to the pUC vector was used as the second primer in the PCR, using pGM26-6.2 as a template. The resulting 1.34-kb fragment was digested with BamHI and cloned into the BamHI site of pBI101.2. The different deletions from the 5′ end of the promoter were constructed by exonuclease III digestion (see Figure 6), and each end point was sequenced, as shown in Figure 4A. The deletions of the N-26 promoter were subcloned into the BamHI-digested pBI101.2, and the orientation of each clone was confirmed before mobilization into *Agrobacterium* for plant transformations.

Plant Transformation

Transgenic tobacco and *L. corniculatus* plants were obtained as described by Miao et al. (1991). Five to eleven independent transgenic *L. corniculatus* plants were regenerated from each construct. The transformed plants regenerated from hairy roots were inoculated with the wild-type *Rhizobium loti* (USDA347), and nodules were collected over a period of 4 to 8 weeks after inoculation. New plants were propagated and maintained from cuttings. Transgenic soybean roots were obtained by infecting soybean hypocotyls with *A. rhizogenes* (K599) containing pBI101, pBI121, or the N-26 promoter–GUS (*E. coli uidA*) construct (D-0). Approximately 50 to 80% of the hairy roots induced by *A. rhizogenes* containing these constructs were found to be GUS positive (data not shown). The plants containing transgenic hairy roots were maintained in vermiculite and inoculated with *Bradyrhizobium japonicum* (strain 61A76). More than 20 individual transgenic hairy roots bearing nodules, representing independent transformation events, were collected and assayed for GUS activity after 3 to 4 weeks of inoculation.

Histochemical and Fluorometric GUS Assay

Fluorometric GUS assay was performed as described by Jefferson et al. (1987). Protein concentration was determined by the method of Bradford (1976). Histochemical localization of GUS activity was performed using X-gluc (5-bromo-4-chloro-3-indolyl β-D-glucuronide) as a chromogenic substrate (Jefferson et al., 1987) for 12 hr at 37°C in a humidified chamber. Sections for light microscopy were processed as described previously (Miao et al., 1991).

ACKNOWLEDGMENTS

We wish to thank Dr. Zonglie Hong for his help in soybean transformation, Dr. Peter Gresshoff for providing *A. rhizogenes* K599, and Dr. Anthony Young for critical comments on this manuscript. This study was supported by National Science Foundation Grants Nos. DCB 88-19399 and DCB 89-04101.

Received March 15, 1993; accepted May 13, 1993.

REFERENCES

Barrett, G., Horiuchi, M., Paul, M., Pratt, R.E., Nakamura, N., and Dzau, V.J. (1992). Identification of a negative regulatory element involved in tissue-specific expression of mouse renin genes. Proc. Natl. Acad. Sci. USA **89,** 885–889.

Benfey, P.N., and Chua, N.-H. (1990). The cauliflower mosaic virus 35S promoter: Combinatorial regulation of transcription in plants. Science **250,** 959–966.

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.

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.

Brewin, N.J. (1991). Development of the legume root nodule. Annu. Rev. Cell Biol. **7,** 191–226.

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.

Conkling, M.A., Cheng, C.-L., Yamamoto, Y.T., and Goodman, H.M. (1990). Isolation of transcriptionally regulated root-specific genes from tobacco. Plant Physiol. **93,** 1203–1211.

Delauney, A.J., and Verma, D.P.S. (1990). A soybean gene encoding Δ^1-pyrroline-5-carboxylate reductase was isolated by functional complementation in *Escherichia coli* and found to be osmoregulated. Mol. Gen. Genet. **221,** 299–305.

Dellaporta, S.L., Wood J., and Hicks, J.B. (1983). A plant DNA minipreparation: Version II. Plant Mol. Biol. Rep. **1,** 19–21.

Fortin, M.G., Zelechowska, M., and Verma, D.P.S. (1985). Specific targeting of membrane nodulins to the bacteroid-enclosing compartment in soybean nodules. EMBO J. **4,** 3041–3046.

Fortin, M.G., Morrison, N.A., and Verma, D.P.S. (1987). Nodulin-26, a peribacteroid membrane nodulin, is expressed independently of the development of the peribacteroid compartment. Nucl. Acids Res. **15,** 813–824.

Foulkes, N.S., and Sassone-Corsi, P. (1992). More is better: Activators and repressors from the same gene. Cell **68,** 411–414.

Gilgenkrantz, H., Hugnot, J.-P., Lambert, M., Chafey, P., Kaplan, J.-C., and Kahn, A. (1992). Positive and negative regulatory DNA elements including a CCArGG box are involved in the cell type-specific expression of the human muscle dystrophin gene. J. Biol. Chem. **267,** 10823–10830.

Gorin, M.B., Yancey, S.B., Cline, J., Revel, J.-P., and Horwitz, J. (1984). The major intrinsic protein (MIP) of the bovine lens fiber membrane: Characterization and structure based on cDNA cloning. Cell **39,** 49–59.

Guerrero, F.D., Jones, J.T., and Mullet, J.E. (1990). Turgor-responsive gene transcription and RNA levels increase rapidly when pea shoots are wilted: Sequence and expression of three inducible genes. Plant. Mol. Biol. **15,** 11–26.

Henikoff, S. (1984). Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28,** 351–359.

Höfte, H., Hubbard, L., Reizer, J., Ludevid, D., Herman, E.M., and Chrispeels, M.J. (1992). Vegetative and seed-specific forms of tonoplast intrinsic protein in the vacuolar membrane of *Arabidopsis thaliana.* Plant Physiol. **99,** 561–570.

Jacobs, F.A., Zhang, M., Fortin, M.G., and Verma, D.P.S. (1987). Several nodulins of soybean share structural domains but differ in their subcellular location. Nucl. Acids Res. **15,** 1271–1280.

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.

Johnson, K.D., Höfte, H., and Chrispeels, M.J. (1990). An intrinsic tonoplast protein of protein storage vacuoles in seeds is structurally related to a bacterial solute transporter (GlpF). Plant Cell **2,** 525–532.

Jørgensen, 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 heterologous symbiotic systems. Nucl. Acids Res. **16,** 39–50.

Jørgensen, J.-E., Stougaard, J., and Marcker, K.A. (1991). A two-component nodule-specific enhancer in the soybean N23 gene promoter. Plant Cell **3,** 819–827.

Kinnaird, J.H., Keighren M.A., Kinsey, J.A., Eaton, M., and Fincham, J.R.S. (1982). Cloning of the *am* (glutamate dehydrogenase) gene of *Neurospora crassa* through the use of a synthetic DNA probe. Gene **20,** 387–390.

Kopito, R.R., Andersson, M., and Lodish, H.R. (1987). Structure and organization of the murine band 3 gene. J. Biol. Chem. **262,** 8035–8040.

Legocki, R.P., and Verma, D.P.S. (1980). Identification of nodule-specific host proteins (nodulins) involved in the development of *Rhizobium*-legume symbiosis. Cell **20,** 153–163.

Liscum, L., Finer-Moore, J., Stroud, R.M., Luskey, K., Brown, M., and Goldstein, J. (1985). Domain structure of 3-hydroxy-3-methylglutaryl coenzyme-A reductase. J. Biol. Chem. **260,** 522–530.

Ludevid, D., Höfte, H., Himelblau, E. and Chrispeels, M.J. (1992). The expression pattern of the tonoplast intrinsic protein γ-TIP in *Arabidopsis thaliana* is correlated with cell enlargement. Plant Physiol. **100,** 1633–1639.

Mahmoudi, M., and Lin, V.K. (1989). Comparison of two different hybridization systems in northern transfer analysis. Biotechniques **7,** 331–333.

Mäkelä, T.P., Koskinen, P.J., Västrik, I., and Alitalo, K. (1992). Alternative forms of Max as enhancers or suppressors of *myc-ras* cotransformation. Science **256,** 373–377.

Martini, B.D., Lingappa, J.R., and Kafatos, F.C. (1988). Temporal regulation in development: Negative and positive *cis* regulators dictate the precise timing of expression of a *Drosophila* chorion gene. Proc. Natl. Acad. Sci. USA **85,** 3029–3033.

Maurel, C., Reizer, J., Schroeder, J.I., and Chrispeels, M.J. (1993). The vacuolar membrane protein g-TIP creates water specific channels in Xenopus oocytes. EMBO J., in press.

Mauro, V.P., Nguyen, T., Katinakis, P., and Verma, D.P.S. (1985). Primary structure of the soybean nodulin-23 gene and potential regulatory elements in the 5′-flanking regions of nodulin and leghemoglobin genes. Nucl. Acids Res. **13,** 239–249.

Miao, G.-H., Hirel, B., Marsolier, M.C., Ridge, R.W., and Verma, D.P.S. (1991). Ammonia-regulated expression of a soybean gene encoding cytosolic glutamine synthetase in transgenic *Lotus corniculatus.* Plant Cell **3,** 11–22.

Miao, G.-H., Hong, Z., and Verma, D.P.S. (1992). Topology and phosphorylation of soybean N-26, an intrinsic protein of the peribacteroid membrane. J. Cell Biol. **118,** 481–490.

Montmayeur, J.-P., and Borrelli, E. (1991). Transcription mediated by a cAMP-responsive promoter element is reduced upon activation of dopamine D_2 receptors. Proc. Natl. Acad. Sci. USA **88,** 3135–3139.

Morrison, N., and Verma, D.P.S. (1987). A block in the endocytosis of *Rhizobium* allows cellular differentiation in nodules but affects the expression of some peribacteroid membrane nodulins. Plant Mol. Biol. **7,** 51–61.

Muramatsu, S., and Mizuno, T. (1989). Nucleotide sequence of the region encompassing the *glpKF* operon and its upstream region containing a bend DNA sequence of *Escherichia coli.* Nucl. Acids Res. **17,** 4378.

Nap, J.P., and Bisseling, T. (1990). Developmental biology of plant-procaryotes symbiosis: The legume root nodule. Science **250,** 948–954.

Nathans, J., and Hogness, D.S. (1983). Isolation, sequence analysis, and intron-exon arrangement of the gene encoding bovine rhodopsin. Cell **34,** 807–814.

Nguyen, T., Zelechowska, M., Foster, V., Bergmann, H., and Verma, D.P.S. (1985). Primary structure of the soybean nodulin-35 gene encoding uricase II localized in the peroxisomes of uninfected cells of nodules. Proc. Natl. Acad. Sci. USA **82,** 5040–5044.

Ouyang, L.-J., Whelan, J., Weaver, C.D., Roberts, D.M., and Day, D.A. (1991). Protein phosphorylation stimulates the rate of malate uptake across the peribacteroid membrane of soybean nodules. FEBS Lett. **293,** 188–190.

Pao, G.M., Wu, L.-F., Johnson, K.D., Höfte, H., Chrispeels, M.J., Sweet, G., Sandal, N.N., and Saier, M.H., Jr. (1991). Evolution of the MIP family of integral membrane transport proteins. Mol. Microbiol. **5,** 33–37.

Preston, G.M., and Agre, P. (1991). Isolation of the cDNA for erythrocyte integral membrane protein of 28 kilodaltons: Member of an ancient channel family. Proc. Natl. Acad. Sci. USA **88,** 11110–11114.

Preston, G.M., Carroll, T.P., Guggino, W.B., and Agre, P. (1992). Appearance of water channels in Xenopus oocytes expressing red cell CHIP28 protein. Science **256,** 385–387.

Rao, Y., Jan, L.Y., and Jan, Y.N. (1990). Similarity of the product of the *Drosophila* neurogenic gene *big* brain to transmembrane channel proteins. Nature **345,** 163–167.

Regensburger, B., Meyer, L., Filser, M., Webber, J., Studer, D., Lamb, J.W., Fischer, H.-M., Hahn, M., and Hennecke, H. (1986). *Bradyrhizobium japonicum* mutants defective in root-nodule bacteroid development and nitrogen fixation. Arch. Microbiol. **144,** 355–366.

Robertson, J.G., Lyttleton, P., Bullivant, S., and Grayston, G.F. (1978). Membrane in lupin root nodules I. The role of Golgi bodies in the biogenesis of infection threads and peribacteroid membranes. J. Cell Sci. **30,** 129–149.

Roth, L.E., and Stacey, G. (1989). Cytoplasmic membrane systems involved in bacterium release into soybean nodule cells as studied with two *Bradyrhizobium japonicum* mutant strains. Eur. J. Cell Biol. **49,** 24–32.

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.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74,** 5463–5467.

Stougaard, J., Sandal, N.N., Grøn, 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., Jørgensen, J.-E., Christensen, T., Kühle, A., and Marcker, K.A. (1990). Interdependence and nodule specificity of *cis*-acting regulatory elements in the soybean leghemoglobin lbc_3 and N23 gene promoters. Mol. Gen. Genet. **220,** 353–360.

Suzuki, H., and Verma, D.P.S. (1991). Soybean nodule-specific uricase (nodulin-35) is expressed and assembled into functional tetrameric holoenzyme in *Escherichia coli*. Plant Physiol. **95,** 384–389.

Sweet, G., Gandor, C., Voegele, R., Wittekindt, N.A., Beuerle, J., Truniger, V., Lin, E.C.C., and Winfried, B. (1990). Glycerol facilitator of *Escherichia coli*: Cloning of *glp*F and identification of the *glp*F product. J. Bacteriol. **172,** 424–430.

Szabados, L., Ratet, P., Grunenberg, B., and de Bruijn, F.J. (1990). Functional analysis of the *Sesbania rostrata* leghemoglobin *glb*3 gene 5′-upstream region in transgenic *Lotus corniculatus* and *Nicotiana tabacum* plants. Plant Cell **2,** 973–986.

Thayer, M.J., and Weintraub, H. (1990). Activation and repression of myogenesis in somatic cell hybrids: Evidence for *Trans*-negative regulation of MyoD in primary fibroblasts. Cell **63,** 23–32.

Van Aelst, L., Hohmann, S., Zimmermann, F.K., Jans, A.W.H., and Thevelein, J.M. (1991). A yeast homologue of the bovine lens fibre MIP gene family complements the growth defect of a *Saccharomyces cerevisiae* mutant on fermentable sugars but not its defect in glucose-induced RAS-mediated cAMP signalling. EMBO J. **10,** 2095–2104.

Verma, D.P.S. (1992a). Signals in root nodule organogenesis and endocytosis of *Rhizobium*. Plant Cell **4,** 373–382.

Verma, D.P.S. (1992b). Soybean nodulin-26: A channel protein conserved from bacteria to mammals. In Transport and Receptor Proteins of Plant Membranes, D.T. Clarkson and D.T. Cooke, eds (New York: Plenum Press), pp. 113–117.

Verma, D.P.S., Kazazian, V., Zogbi, V., and Bal, A.K. (1978). Isolation and characterization of the membrane envelope enclosing the bacteroids in soybean root nodules. J. Cell Biol. **78,** 919–939.

Verwoerd, T.C., Dekker, B.M.M., and Hoekema, A. (1989). A small-scale procedure for the rapid isolation of plant RNAs. Nucl. Acids Res. **17,** 2362.

Weaver, C.D., Crombie, B., Stacey, G., and Roberts, D.M. (1991). Calcium-dependent phosphorylation of symbiosome membrane proteins from nitrogen-fixing soybean nodules: Evidence for phosphorylation of nodulin-26. Plant Physiol. **95,** 222–227.

Werner, D., Mellor, R.B., Hahn, M.G., and Grisebach, H. (1985). Soybean root response to symbiotic infection glyceollin I accumulation in an ineffective type of soybean nodules with an early loss of the peribacteroid membrane. Z. Naturforsch. **40c,** 179–181.

Yamamoto, Y.T., Cheng, C.-L., and Conkling, M.A. (1990). Root-specific genes from tobacco and *Arabidopsis* homologous to an evolutionarily conserved gene family of membrane channel proteins. Nucl. Acid Res. **18,** 7449.

Yamamoto, Y.T., Taylor, C.G., Acedo, G.N., Cheng, C.-L., and Conkling, M.A. (1991). Characterization of *cis*-acting sequences regulating root-specific gene expression in tobacco. Plant Cell **3,** 371–382.