The Plant Cell, Vol. 3, 819–827, August 1991

A Two-Component Nodule-Specific Enhancer in the Soybean N23 Gene Promoter

Jan-Elo Jørgensen, Jens Stougaard, and Kjeld A. Marcker[1]

Department of Molecular Biology, Laboratory of Gene Expression, University of Aarhus, Gustav Wieds Vej 10, DK-8000 Aarhus C, Denmark

The two positive *cis* elements in the soybean nodulin N23 gene promoter were investigated in transgenic *Lotus corniculatus* plants and shown to constitute a two-component nodule-specific enhancer. Equal quantitative contributions from the two components were suggested by the similar expression level of chimeric N23-chloramphenicol acetyltransferase genes after deletion of either the distal positive element (PE-A, −320 to −298) or the proximal positive element (PE-B, −257 to −165). A combined effect of the two elements was indicated by orientation-dependent effects in the N23 promoter, and by the observation that neither PE-A nor PE-B separately was able to confer any activity to the cauliflower mosaic virus 35S minimal promoter. Reactivation of the minimal N23 and the minimal cauliflower mosaic virus 35S promoters by the inverted complete element (PE-AB) further suggested that PE-AB is a nodule-specific enhancer containing two equally strong enhancer components. Two 12-bp sequence motifs, InvA and InvB, constituting an inverted repeat, were identified as the core of the enhancer components PE-A and PE-B, respectively. Point mutations in InvA or InvB resulted in lower expression levels and mutations in both abolished enhancer activity. Point mutations in two nodulin consensus sequences, 5′-CTCTT and 5′-AAAGAT located downstream of PE-AB, resulted in a decreased level of expression, confirming the involvement of these two motifs in nodulin gene expression. The binding site for the nodule-specific *trans*-acting factor, NAT2, present in the PE-A segment, was removed without affecting expression significantly. This interaction is, therefore, dispensable for enhancer activity.

INTRODUCTION

In eukaryotes, initiation of transcription is regulated by protein factors that recognize and bind to *cis* elements. The widely conserved motifs, the TATA box and the CAAT box, constitute a basic promoter that ensures correct transcription initiation on nuclear genes by RNA polymerase II. In addition, many eukaryotic promoters contain *cis* elements responding to environmental and developmental changes, thereby directing selective expression of the relevant gene (see Kuhlemeier et al., 1987; Maniatis et al., 1987; Benfey and Chua, 1989, for reviews). In legumes, formation of nitrogen-fixing nodules after invasion by bacteria of the genus *Rhizobium* or *Bradyrhizobium* initiates a developmental process where specific *cis* elements direct patterns of temporal and spatial expression of specific host genes (see de Bruijn et al., 1990; Nap and Bisseling 1990, for reviews). Several genes encoding polypeptide nodulins involved in nodule morphogenesis and function are selectively expressed. Genes for "early nodulins" involved in the infection process and nodule formation are induced soon after invasion. Later in development, "late nodulins" like the oxygen carrier leghemoglobin are synthesized.

In soybean, the late nodulin N23 is associated with the peribacteroid membrane surrounding bacteroids in the invaded plant cells (Jacobs et al., 1987). The N23 gene is a member of a gene family including genes N44, N20, N22 (Sandal et al., 1987), and N26b (Jacobs et al., 1987). The function of the corresponding proteins is not known, but all contain a conserved putative metal binding domain that seems to be important for their function (Sandal et al., 1987). Both the N23 and the leghemoglobin gene families in soybean are expressed at a high level in root nodules. In the growing nodule, a low level of transcripts from the leghemoglobin genes is detectable 4 days before the N23, N44, and N20 transcripts (Marcker et al., 1984; Sengupta-Gopalan et al., 1986; Jørgensen et al., 1988). A large increment in leghemoglobin transcription at about day 12 postinfection coincides with the induction of the N23 family of genes. Other late nodulins are also induced at this developmental stage, indicating a common regulatory mechanism for late nodulin expression. To investigate this mechanism, promoter regions from the soybean leghemoglobin c3 (*lbc3*) and the N23 genes were analyzed for

[1] To whom correspondence should be addressed.

cis-acting regulatory elements (Stougaard et al., 1986, 1987, 1990; Jørgensen et al., 1988) and interacting *trans*-acting factors (Jensen et al., 1988) (for review, see de Bruijn et al., 1990).

cis analysis revealed upstream positive elements (Stougaard et al., 1987; Jørgensen et al., 1988), promoter proximal organ-specific elements (Stougaard et al., 1987, 1990), and, in the case of *lbc3*, a negative element overlapping the basic promoter (Stougaard et al., 1987). Currently, the only well-defined conserved sequences in these gene promoters are the nodulin consensus motifs 5′-CTCTT and 5′-AAAGAT located in the organ-specific element (Sandal et al., 1987; Stougaard et al., 1987). A similar promoter organization with upstream positive elements and organ-specific control elements overlapping or in proximity of the basic promoter was also found in the *Sesbania rostrata* leghemoglobin g*lb*3 gene promoter (Szabados et al., 1990).

trans-Acting factors interacting with described *cis* elements were identified in gel retardation experiments using restriction fragments from the promoter regions. Interaction of a nodule-specific factor, NAT2, a root/nodule factor, NAT1, and a leaf high-mobility group protein, LAT1, was demonstrated at several AT-rich sequences in upstream regions of the soybean N23 (Jacobsen et al., 1990) and *lbc3* genes (Jensen et al., 1988). NAT2 also interacts with the *S. rostrata* g*lb*3 (de Bruijn et al., 1989) and the bean nodule-specific glutamine synthetase (Forde et al., 1990) gene promoters. The weak positive element of the soybean *lbc3* promoter and a positive element in the *S. rostrata* g*lb*3 promoter includes binding sites for the NAT2 factor, whereas one of the positive elements in the soybean N23 promoter includes binding sites for NAT2 and LAT1.

In the N23 promoter, two positive elements are located at positions −344 to −298 (PE-A) and −257 to −165 (PE-B) (Jørgensen et al., 1988). PE-A contains three putative sequence motifs for positive regulation. First, there is the AT-rich stretch between −344 and −329 that binds the nodule-specific protein factor NAT2 (Jacobsen et al., 1990). Second, there is a 5′-CTCTT motif located at position −325 to −321. Previous experiments suggested that this motif is significant for high-level organ-specific expression from the N23 promoter (Stougaard et al., 1990). Third, PE-A contains one part of an almost perfect 12-bp inverted repeat InvA (5′-ACATCCCTTTGT), which has only one mismatch compared with InvB (5′-ACAAAGGGTTGT), present in PE-B. Because inverted and direct repeats play an essential role for quantitative as well as qualitative regulation of several other plant genes (Ellis et al., 1987; Chen et al., 1988; Bustos et al., 1989; Mitra and An, 1989), the InvA/InvB inverted repeat constitutes a putative positive element and/or organ-specific element in the N23 gene promoter.

We have used transgenic *Lotus corniculatus* plants containing chimeric N23-chloramphenicol acetyltransferase (CAT) genes to investigate the qualitative and quantitative contributions of PE-A and PE-B to the nodule-specific expression. Point mutation experiments were designed to determine the functions of InvA and InvB as well as the two putative nodulin consensus sequence motifs 5′-CTCTT and 5′-AAAGAT present in the defined organ-specific box (Stougaard et al., 1990) at positions −126 to −122 and −93 to −88, respectively (Sandal et al., 1987). The results suggest that the nodule-specific expression of the N23 gene is regulated positively by the inverted repeat InvA/InvB in combination with the two nodulin consensus motifs 5′-CTCTT and 5′-AAAGAT.

RESULTS

The chimeric N23-CAT reporter genes were transferred into *L. corniculatus* plants using *Agrobacterium rhizogenes* AR19 (see Methods), which contains a cauliflower mosaic virus (CaMV) 35S-β-glucuronidase (GUS) construct as a constitutively expressed marker gene. CAT activity was monitored in roots and nodules from six to 10 GUS^+ plants of each line. The uncoordinated expression of reporter and marker genes did not allow the use of GUS activity for normalization of CAT gene expression.

Definition of PE-A

The distal positive element, PE-A (−344 to −298), was previously described as the strong positive element of the N23 promoter (Jørgensen et al., 1988). Removal of this region in a 5′ deletion reduced expression from full level to about 15%. Within its primary sequence, three potential motifs for regulation of expression level and/or specificity can be recognized, as shown in Figure 1. First, a binding site for the nodule-specific protein factor NAT2 (Jacobsen et al., 1990) is found from −344 to −329, then a nodulin consensus 5′-CTCTT motif follows at position −320 to −325. Further downstream, one part of an inverted repeat, InvA (5′-ACATCCCTTTGT), is located at position −308 to −297, and is therefore almost completely included in PE-A. The presence of the corresponding inverted part, InvB, in the proximal positive element, PE-B, suggested that InvA might constitute the core of PE-A.

To test the significance of the NAT2 binding site as well as the 5′-CTCTT motif, a deletion mutant, Δ(−320), was created at the unique FokI site in PE-A. The Δ(−320) plant lines showed no significant decrease in CAT activity compared with the full-length promoter shown in Figure 2A. This observation strongly suggests that neither the 5′-CTCTT motif nor the NAT2 binding site contributes significantly to the effect of PE-A and delimited PE-A to position −320 to −298, leaving InvA as a major candidate for a positive regulatory element.

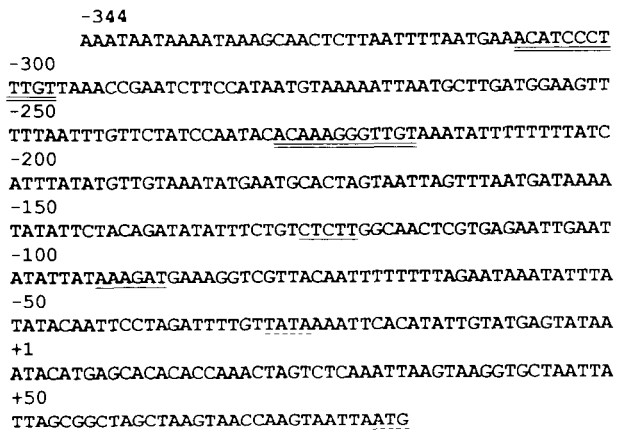

Figure 1. DNA Sequence of the 5′ Flanking Region of the N23 Gene Containing All Necessary *cis* Elements for High-Level Nodule-Specific Expression.

The start codon ATG and TATA box are indicated by the dashed lines. The nodulin consensus motifs 5′-CTCTT and 5′-AAAGAT are underlined with one bar; InvA and InvB are underlined with two bars. Positions are indicated relative to the major cap site of the original N23 promoter (Wong and Verma, 1985).

Quantitative Contribution of PE-B

A 5′ deletion analysis previously defined the proximal positive element, PE-B (−257 to −165), as a weak positive element, conferring only 15% of full expression to the almost silent minimal (−165) N23 promoter (Jørgensen et al., 1988). The presence of InvB in PE-B indicated that this interpretation might not be correct and that removal of PE-B by an internal deletion would show an effect identical to the removal of PE-A. To test this hypothesis, an internal deletion Δ(−286 to −165) (Figure 2A) was constructed to remove PE-B entirely from a fully active −344 promoter. The subsequent decrease in CAT activity to 15% of full expression in Δ(−286 to −165) plant lines was similar to the decrease in Δ(−298) lines containing the 5′ deletion of PE-A (Figure 2A), indicating that PE-A and PE-B contributed equally to the level of expression.

Qualitative Contribution and Relative Interdependence of PE-A and PE-B

The qualitative contributions of PE-A and PE-B were analyzed by combining the two elements in front of the CaMV 35S minimal promoter (Figure 2C). The fragments (−344 to −286) that included PE-A and (−257 to −175) PE-B were, in the correct order and orientation, able to confer nodule-specific expression to the CaMV 35S minimal promoter (2E12). Accordingly, PE-AB can be regarded as a positive nodule-specific element. The level of expression from PE-AB was comparable with the activity from a control construct containing the complete CaMV 35S promoter sequence. This result was confirmed by a similar expression level observed from constructions D1-3 and D1-5 containing the uninterrupted PE-AB sequence in opposite orientations. The relative interdependence of PE-A and PE-B was investigated by placing the two elements in different combinations in front of the N23 and CaMV 35S minimal promoters (Figures 2B and 2C). None of the elements was able separately to confer any expression to the CaMV 35S minimal promoter (E218 and E404), confirming the high degree of interdependence already demonstrated by the deletion mutant experiments. Inversion of PE-A relative to PE-B (2E3 and C8-7) reduced CAT activity approximately fivefold compared with constructions containing PE-A in the correct orientation (2E12 and C8-4). This lower expression level may reflect the loss of a stem and loop structure that might be formed by InvA and InvB placed in proper orientation. Otherwise, the simple repositioning of InvA relative to InvB could cause this effect.

Enhancer Properties of PE-AB

According to the definition of an enhancer element, it is possible to invert the enhancer sequence without changing its effect on the expression level. The N23 positive element PE-AB was tested for this particular enhancer property (Figures 2B and 2C) by inverting the complete PE-AB region (−344 to −175) in front of the minimal (−165) N23 promoter (C1-4) and the CaMV 35S minimal promoter (D1-5). Both promoter constructions preserved the nodule-specific CAT activity. It is notable that the inverted PE-AB directed a somewhat lower expression level than did the correctly oriented PE-AB fragment in front of the N23 minimal promoter (C1-3). The opposite effect was obvious in front of the CaMV 35S minimal promoter, where inversion of PE-AB caused a significantly higher expression level than a correctly oriented PE-AB fragment (D1-3). These opposite results might have been caused by position effects because C1-3, which resulted in a full level of expression, was the only promoter construction that contained PE-AB in its original position. Furthermore, the changed surroundings in the hybrid promoters of D1-3 and D1-5 might have influenced the enhancement of expression. Nevertheless, the significant observation in these experiments is that the inversion of PE-AB preserved the nodule-specific enhancer activity. One of the enhancer properties, therefore, can be assigned to PE-AB. It remains to be investigated whether this element will be able to direct high-level gene expression when moved further away from the transcription start or even downstream of the coding region.

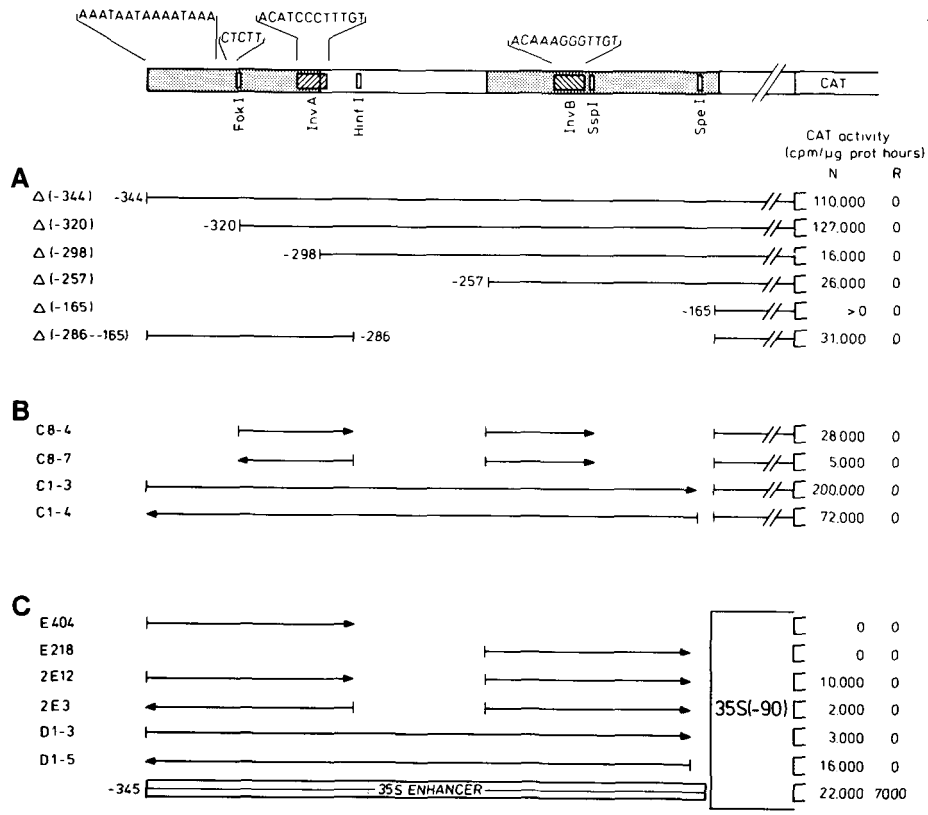

Figure 2. CAT Activity in Transgenic Roots and Nodules from *L. corniculatus* Plants Containing Various N23 Promoter Constructions.

(A) Delimitation of PE-A and PE-B by 5′ deletions and one internal deletion in the N23 promoter region
(B) Fragments containing PE-A and PE-B in different combinations in front of the N23 (−165) minimal promoter reveal mutual interdependency of PE-A and PE-B and enhancer property of PE-AB.
(C) Hybrid promoter experiments combining PE-A and PE-B in front of the CaMV 35S (−90) minimal promoter demonstrate qualitative contributions and mutual interdependency of PE-A and PE-B.
In all cases, the remaining sequences are represented with bars. Arrows indicate orientations with respect to the original promoter arrangement. CAT activity is given as average values for nodules (N) and roots (R) from approximately 10 plant lines representing each construction. The CaMV 35S enhancer (−343 to −90) is boxed below.
The diagram at the top is a representation of a part of the N23 promoter containing PE-A and PE-B (shaded). The positions and sequences of the NAT2 binding site, a 5′-CTCTT motif, and InvA and InvB are indicated together with the positions of some relevant restriction endonuclease recognition sites.

Additional Delimitation of PE-A and PE-B

Attempts to reduce PE-A and PE-B toward the InvA and InvB elements resulted in significantly lower nodule-specific CAT activity (Figure 2B). As an example, a fragment containing the intact sequences of InvA and InvB in correct order in front of the N23 minimal promoter (C8-4) lowered the expression level to approximately 25% of full activity. This observation either indicates requirement for additional *cis* elements present in the deleted sequences or reflects an effect of changes in the mutual positions of PE-A and PE-B. To determine further the importance of the inverted repeat, point mutations were introduced into the InvA and InvB sequences of the complete −915 promoter.

Point Mutations in InvA and InvB

To introduce mutations into the full-length promoter of N23, an oligonucleotide-directed mutagenesis was performed (as described in Methods) using oligonucleotides with base substitutions in the sequences corresponding to InvA and InvB.

Five point mutations were introduced into InvA and InvB, as shown in Figures 3 and 4. Because of the relatively high number of guanines and cytosines in InvA and InvB compared with the rest of the N23 promoter, the selected base substitutions were from guanine and cytosine to adenine or thymidine, as shown in Figure 3. Three mutant promoters, M967 (containing Mut-InvA), M966 (containing Mut-InvB), and M9676 (containing both Mut-InvA and Mut-InvB), were constructed. Expression of CAT activity from mutant promoters was compared with activities from comparable deletion mutants. M9676 plants barely expressed detectable CAT activity, as was the case with Δ(−165) lines without PE-AB (Figures 2A and 4). The M966 promoter-mutant decreased CAT activity to 15% of normal activity, as observed with the internal deletion Δ(−286 to −165) without PE-B. Exactly half the analyzed M967 plant lines showed the 15% of full nodule-specific expression that could be expected from previous investigations of the 5′ deletion Δ(−298) without PE-A. The remaining lines showed a surprisingly high expression rate. Currently, there is no obvious explanation for the high expression rate in these lines. This observation questioned the significance of InvA in nodule-specific gene expression. The effect of InvA, however, was very clear in the M9676 lines because the only difference between M966 and M9676 is the mutated InvA motif in M9676. Redundancy of InvA would have resulted in similar expression levels of the M9676 and M966 lines.

These observations strongly indicate that InvA and InvB can be regarded as the core of a two-component nodule specific enhancer element.

InvA	-308	5′ACATCCCTTTGT 3′	-297
Mut.InvA		5′AAATAAATTTCT 3′	
InvB	-228	5′ACAAAGGGTTGT 3′	-217
Mut.InvB		5′AGAAATTTTTCT 3′	
	-325	5′CTCTT 3′	-321
Mut.G		5′CTGTT 3′	
Mut.Tot1		5′AGGAC 3′	
	-392	5′AAAGAT 3′	-387
Mut.T		5′AAATAT 3′	
Mut.Tot2		5′TTCATA 3′	

Figure 3. Base Mutations Introduced into the Putative Regulatory Elements of the N23 Promoter.

Mutated sequences are shown below the original sequences. Positions are indicated with respect to the major cap site of the original N23 promoter (Wong and Verma, 1985). Underlining indicates the single mismatch between InvA and InvB.

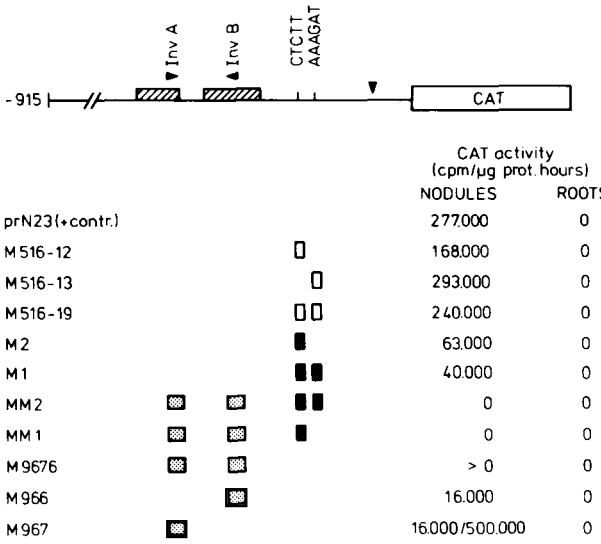

Figure 4. Activity from the CAT Reporter Gene after Introduction of Base Mutations into the Full-Length (−915) N23 Promoter.

Boxes indicate presence of MutG/MutT (open), MutTot1/MutTot2 (black), and Mut.InvA/Mut.InvB (shaded) (see Figure 3). PE-A and PE-B are marked with hatched boxes. The positions of InvA and InvB (horizontal arrows) as well as the 5′-CTCTT and 5′-AAAGAT motifs (vertical lines) are indicated. A vertical arrow indicates the proposed major cap site of the N23 gene (Wong and Verma, 1985). Activities from the CAT reporter gene are given as average values for nodules and roots from approximately 10 plant lines representing each construction.

Contribution of the Putative Nodulin Consensus Motifs

Two short sequence motifs, 5′-CTCTT and 5′-AAAGAT, are present in the immediate 5′ upstream region of described nodulin gene sequences (Sandal et al., 1987). Accordingly, these motifs might be involved in the nodule-specific expression from the N23 promoter as well as from the *lbc3* promoter. In both cases, nodule-specific expression was regained when silent deletion mutants, including the nodulin consensus motifs, were reactivated with the (−343 to −90) CaMV 35S enhancer (Stougaard et al., 1987, 1990).

To confirm their critical role in nodule-specific gene regulation, mutations were introduced into the 5′-CTCTT and 5′-AAAGAT sequences in the N23 promoter (Figures 3 and 4). Single base mutations Mut-G and Mut-T created the sequences 5′-CTGTT and 5′-AAATAT, respectively, whereas the totally mutated sequences Mut-Tot1 and Mut-Tot2 completely substituted the two consensus motifs (Figure 3). None of the single base mutations led to significant changes in nodule-specific expression level (M516-12, M516-13, and M516-19), whereas the promoters of M2 (containing Mut-Tot1) and M1 (containing both Mut-Tot1 and Mut-Tot2) resulted in a significant drop in CAT activity in nodules. In agreement with previous reactivation experiments (Stougaard et al., 1990), elimination of the 5′-CTCTT sequence alone was enough to cause a reduced expression level. Although a further minor decrease in CAT activity in the M1 plant lines was observed, it remains an open question whether the 5′-AAAGAT sequence and the 5′-CTCTT sequence are of equal importance for expression. The conservation of both these motifs in nodulin gene promoters, however, suggests that this is indeed the case.

Removal of both 5′-CTCTT and 5′-AAAGAT in M1 plant lines did not completely eliminate nodule-specific expression, and this effect was not as drastic as the mutation of InvA and InvB in the M9676 construct. This confirmed the ability of InvA and InvB to direct nodule specificity and assigned this inverted repeat as a major positive element compared with the effect of the 5′-CTCTT/5′-AAAGAT arrangement.

The very low expression level observed in plant lines harboring the M9676 construct was not detectable in lines containing the MM1 and MM2 constructs, which contained, in addition to Mut-InvA and Mut-InvB, the highly mutated Mut-Tot1 and Mut-Tot1 and Mut-Tot2, respectively. This confirmed that the 5′-CTCTT and the 5′-AAAGAT sequences constitute an additional positive nodule-specific element.

In summary, the mutation experiments indicate that the nodule-specific expression from the N23 promoter is directed by InvA/InvB in combination with 5′-CTCTT/5′-AAAGAT, where the latter motifs seem to be of less importance than the combination of InvA/InvB.

DISCUSSION

Investigation of the soybean nodulin N23 gene promoter revealed that four sequence motifs, located fairly close to one another, direct the full-level nodule-specific expression of a CAT reporter gene. All four motifs were qualitative elements, directing organ-specific expression. The inverted repeat InvA/InvB constituted a strong positive element that exhibited some characteristics of an enhancer element. The nodulin consensus arrangement, 5′-CTCTT/5′-AAAGAT, which in itself is not able to express the reporter gene, was dependent on InvA/InvB and acted to assure the full nodule-specific expression level. This is in agreement with previous investigations demonstrating interdependent regulatory elements in other nodulin genes (Stougaard et al., 1990; Szabados et al., 1990). The experimental approach did not allow identification of negative elements. Therefore, the presence of one or more negative elements in the N23 promoter cannot be excluded.

The N23 gene is a member of a small nodulin gene family that includes the N44, N20, N22 (Sandal et al., 1987), and N26b (Jacobs et al., 1987) genes. InvA and InvB are present in the 5′ flanking regions of the N23 and N44 genes, whereas the N20 and N22 genes do not have these motifs in the sequenced part of the 5′ regions (positions −532 and −463, respectively). As previously noted, all genes contain the 5′-CTCTT/5′-AAAGAT arrangement in appropriate positions. The 5′ flanking region of N26b has not been sequenced, but we predict that this gene also contains the InvA and InvB arrangement because further analysis of the gene family suggested that the N23, N44, and N26b genes belong to the same evolutionary subfamily, whereas the N20 and N22 genes are members of another subfamily (Sandal et al., 1987; Richter et al., 1991). The positive elements in the *glb3* gene promoter diverge from the N23 PE-AB and the *lbc3* strong positive element (SPE) by their ability to direct positive regulation in roots (Szabados et al., 1990). The inverted repeat InvA/InvB is also not present in this promoter. Some less conserved InvA and InvB sequences are located in the far upstream region of the *lbc3* promoter, which does not seem to influence the nodule-specific expression level. Although PE-AB from the N23 gene promoter directs organ specificity like the strong positive element from the soybean *lbc3* promoter (Stougaard et al., 1990), no obvious sequence homologies are noticeable between *lbc3*-SPE and N23-PE-AB, and it is not possible to locate any obvious inverted repeats in the *lbc3*-SPE. The PE-AB enhancer, therefore, cannot be compared directly with the previously assigned *lbc3*-SPE (Stougaard et al., 1987). However, it is still possible that the *lbc3*-SPE is also composed of multiple components. The lack of significant InvA/InvB arrangements in the promoters of the *lb* genes and even in the closely N23-related N20 and N22 genes suggests that the elements regulating the nodule-specific expression of the late nodulin genes exhibit a larger diversity than suggested by their similar induction kinetics.

The delimitation of PE-A (−320 to −298) and the subsequent mutation-based assignment of InvA as the core of PE-A excludes that the NAT2 binding at −344 to −330 (Jacobsen et al., 1990) is responsible for the high-level expression directed by PE-AB, as was previously proposed based on the presence of similar binding sites in the soybean *lbc3* (Jensen et al., 1988) and bean glutamine synthetase (Forde et al., 1990) promoters. Furthermore, removal of binding sites 1 and 2 from the soybean *lbc3*

promoter did not lead to a changed expression level (Stougaard et al., 1990). Jacobsen and co-workers (1990) argue that the NAT2 binding site can be replaced by several putative NAT2 binding sites located further upstream in the N23 promoter and that the importance of NAT2 binding activity for nodule specific expression, therefore, cannot be excluded. However, the 5′ deletion Δ(−320) does not leave any obvious NAT2 binding sites in the N23 promoter region; therefore, this suggestion can only be supported if a binding site elsewhere in the surrounding regions can replace the deleted site.

The mechanism by which InvA and InvB regulate expression is not immediately obvious. Their inverted correspondence and the mutual orientation dependency suggest that a stem and loop formation might occur to connect InvA and InvB. The decrease in expression level caused by inverting PE-A could also be a consequence of the reposition rather than the inversion of InvA. Thus, the possibility of one or more protein factors binding separately to InvA and InvB without causing loop formation cannot be excluded.

Protein-DNA interaction at an inverted repeat was observed previously by the binding of the GBF factor to the G-box present in a number of different plant gene promoters, including the alcohol dehydrogenase gene from Arabidopsis (McKendree et al., 1990), the chalcone synthase gene from snapdragon (Staiger et al., 1989), and several of the ribulose-1,5-bisphosphate carboxylase small subunit (RBCS) genes (Gilmartin et al., 1990; Schindler and Cashmore, 1990). Unlike InvA and InvB, the inverted repeat in the G-box constitutes a palindrome. This is also the case for the octopine synthase (*ocs*) enhancer, which constitutes a 16-bp palindrome in the *ocs* promoter (Ellis et al., 1987). The *ocs trans*-acting factor has two binding sites, one in each half of the palindrome. To make this complex function as an enhancer, both binding sites must be occupied (Tokuhisa et al., 1990). The strong mutual interdependence of InvA and InvB indicates that the enhancement of the *ocs* and the N23 gene expression could operate by analogous mechanisms. In contrast, a lesser degree of mutual interdependence between two positive elements was found in the nopaline synthase promoter (Mitra and An, 1989), where the downstream element was more important than the distant upstream element of an 8-bp inverted repeat.

The importance of guanines and cytosines for enhancement of expression by the relatively GC-rich InvA and InvB elements was confirmed by base mutation experiments (Figures 3 and 4). It is noteworthy that no enhancement occurred in the M9676 plant lines, although the mutated InvA and InvB motifs still constitute an inverted repeat in these lines. Another 12-bp GC-rich stretch is directly as well as invertedly repeated several times in two positive elements, PR-1 and PR-2, within the distal upstream region of the chlorophyll *a/b*-binding protein (*Cab-E*) gene promoter from *Nicotiana plumbaginifolia* (Castresana et al., 1988). The regulatory activity of PR-1 and PR-2 shares some similarities with the N23 PE-AB because they are strongly dependent on proximally located light regulatory elements in the *Cab-E* promoter. Although proximal qualitative elements might exhibit some positive effects on the gene expression level, their regulatory activity often seems to depend upon stronger positive elements with which they interact to potentiate expression (Walker et al., 1987; Fang et al., 1989). Likewise, the nodulin consensus motifs 5′-CTCTT and 5′-AAAGAT show nearly no independent ability to direct nodule-specific transcription. However, they do play a significant role by increasing the transcriptional activity of the distal positive element, PE-AB.

Whereas the PR-1 and PR-2 elements from the *Cab-E* promoter seem to act only by stimulating the expression level, the positive elements of the β-conglycinin α-subunit gene, containing four GC-rich 6-bp direct repeats, constitute a tissue-specific and temporally regulating enhancer (Chen et al., 1986, 1988). This enhancer as well as the enhancer from the light-regulated *rbc*S-3A gene promoter, which confers both light-regulated and organ-specific expression to a reporter gene (Aoyagi et al., 1988), resemble PE-AB in being multiqualitative plant gene enhancers.

The interplay of several *cis* elements occurs in various plant gene promoters (Benfey and Chua, 1989; Goldberg et al., 1989; Bruce and Quail, 1990; Gilmartin et al., 1990). Whereas in most cases it is possible to distinguish between qualitative and quantitative elements, the results presented here suggest that the individual elements in the N23 promoter exhibit a combined effect on the reporter gene. A similar arrangement is found in the *lb* gene promoters (Stougaard et al., 1990; Szabados et al., 1990). Consequently, the *cis* elements of the late nodulin gene promoters are part of complex promoter structures in which the characteristics of the individual *cis* elements are strongly influenced by the neighboring elements.

METHODS

Nucleic Acid Manipulation

Standard techniques for DNA manipulations were used as described in Maniatis et al. (1982). DNA-modifying enzymes were purchased from Boehringer Mannheim, Mannheim, Germany, Stratagene, Heidelberg, Germany, and New England Biolabs, Beverly, MA, and used according to the manufacturers' instructions.

Construction of Chimeric N23-CAT Genes

5′ deletion mutants in the 5′-N23-CAT-3′ *lbc3* gene were produced by Bal31 digestion as described by Jørgensen et al. (1988) or by cleavage at endonuclease restriction sites as indicated in Figure 2. Fragments with cohesive ends were refilled with DNA

polymerase I (Klenow fragment). All constructions were subcloned into the pIV10 integration vector (J. Stougaard, unpublished data) that carries the spectinomycin and streptomycin gene for selection after transfer to *Agrobacterium rhizogenes*. The CaMV minimal (−90) 35S-CAT fusion (Stougaard et al., 1990) was used for the hybrid promoter constructions. All constructions were checked by sequencing using the United States Biochemical Corporation (Cleveland, OH) Sequenase kit, version 2.0, on denatured double-stranded plasmids (Hattori and Sakaki, 1986).

Oligonucleotide-Directed in Vitro Mutagenesis

Chimeric 5′-N23-CAT-3′ *lbc3* genes carrying mutations in the N23 promoter region were created using the second version of the Amersham oligonucleotide-directed in vitro mutagenesis system following the manufacturer's instructions. A fragment containing the (−915) N23 promoter region was subcloned into an M13-mp18 vector. Oligonucleotides were annealed onto the M13 single-stranded template, and DNA polymerase I (Klenow fragment) completed duplication. The original DNA strand was nicked at a unique NciI site and subsequently removed by exonuclease III digestion. DNA polymerase I duplicated the mutated strand, and the double-stranded plasmids were transformed into TG1 cells to produce M13 phages carrying the mutated genes in a single-stranded form. All mutants were checked by sequencing of the single-stranded M13 DNA. The mutated N23 promoter fragments were subcloned into a CAT-3′ *lbc3* cassette in the pIV10 integration vector to reconstitute a chimeric 5′-N23-CAT-3′ *lbc3* gene.

Transformation of Plants

Chimeric N23-CAT genes were transferred to *A. rhizogenes* as described by Van Haute et al. (1983). Transgenic *Lotus corniculatus* plants with transformed hairy roots on untransformed shoots were produced according to Hansen et al. (1989) using strain AR19 of *A. rhizogenes*. AR19 carries 35S-GUS in the left T-DNA region (TL segment) and is analogous to AR12 (Hansen et al., 1989) but supplied with the *sacB* gene, which causes the bacteria to degenerate in 6% sucrose. Roots of Agrobacterium-free plants were inoculated with *Rhizobium loti* NZP2037 to produce transgenic root nodules. Crude extracts from roots and nodules were prepared according to Hansen et al. (1989).

Biochemical Assays

Activity from the CAT reporter gene was measured in eight to 12 plants as previously described (Stougaard et al., 1986, 1987). GUS activity from the GUS marker gene was measured by fluorometric assay (Jefferson et al., 1987).

ACKNOWLEDGMENTS

This research was supported by the Danish Biotechnology Programme.

Received April 22, 1991; accepted June 25, 1991.

REFERENCES

Aoyagi, K., Kuhlemeier, C., and Chua, N.-H. (1988). The pea *rbcS-3A* enhancer-like element directs cell-specific expression in transgenic tobacco. Mol. Gen. Genet. **213,** 179–185.

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

Bruce, W.B., and Quail, P.H. (1990). *cis*-Acting elements involved in photoregulation of an oat phytochrome promoter in rice. Plant Cell **2,** 1081–1089.

Bustos, M.M., Guiltinan, M.J., Jordano, J., Begum, D., Kalkan, F.A., and Hall, T.C. (1989). Regulation of β-glucuronidase expression in transgenic tobacco plants by an A/T-rich, *cis*-acting sequence found upstream of a French bean β-phaseolin gene. Plant Cell **1,** 839–853.

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., Schuler, M.A., and Beachy, R.N. (1986). Functional analysis of regulatory elements in a plant embryo-specific gene. Proc. Natl. Acad. Sci. USA **83,** 8560–8564.

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.

de Bruijn, F.J., Grunenberg, F.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 are used to study the regulation of genes involved in symbiotic plant-microbe interactions (nodulin genes). Dev. Genet. **11,** 182–196.

Ellis, J.G., Llewellyn, D.J., Walker, J.C., Dennis, E.S., and Peacock, W.J. (1987). The *ocs* element: A 16 base pair palindrome essential for activity of the octopine synthase enhancer. EMBO J. **6,** 3203–3208.

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., Freeman, J., Oliver, J.E., and Pineda, M. (1990). Nuclear factors interact with conserved A/T-rich elements upstream of a nodule-enhanced glutamine synthetase gene from French bean. Plant Cell **2,** 925–939.

Gilmartin, P.M., Sarokin, L., Memelink, J., and Chua, N.-H. (1990). Molecular light switches for plant genes. Plant Cell **2,** 369–378.

Goldberg, R.B., Barker, S.J., and Perez-Grau, L. (1989). Regulation of gene expression during plant embryogenesis. Cell **56,** 149–160.

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

Hattori, M., and Sakaki, Y. (1986). Dideoxy sequencing using denatured plasmid templates. Anal. Biochem. **152,** 232–238.

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 locations. Nucl. Acids Res. **15,** 1271–1280.

Jacobsen, K., Laursen, N.B., Jensen, E.Ø., 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., 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.Ø., 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 leghaemoglobin *lbc*3 5′ upstream region. EMBO J. **7,** 1265–1271.

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.

Kuhlemeier, C., Green, P.J., and Chua, N.-H. (1987). Regulation of gene expression in higher plants. Annu. Rev. Plant Physiol. **38,** 221–257.

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

Maniatis, T., Goodbourn, S., and Fischer, J.A. (1987). Regulation of inducible and tissue-specific gene expression. Science **236,** 1237–1245.

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

McKendree, W.L., Paul, A.-L., DeLisle, A.J., and Ferl, R.J. (1990). In vivo and in vitro characterization of protein interactions with the dyad G-box of the *Arabidopsis Adh* gene. Plant Cell **2,** 207–214.

Mitra, A., and An, G. (1989). Three distinct regulatory elements comprise the upstream promoter region of the nopaline synthase gene. Mol. Gen. Genet. **215,** 294–299.

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

Richter, H.E., Sandal, N.N., Marcker, K.A., and Sengupta-Gopalan, C. (1991). Characterization and genomic organization of a highly expressed late nodulin gene subfamily in soybeans. Mol. Gen. Genet., in press.

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.

Schindler, U., and Cashmore, A.R. (1990). Photoregulated gene expression may involve ubiquitous DNA binding proteins. EMBO J. **9,** 3415–3427.

Sengupta-Gopalan, C., Pitas, J.W., Thompson, D.V., and Hoffman, L.M. (1986). Expression of host genes during nodule development in soybeans. Mol. Gen. Genet. **203,** 410–420.

Staiger, D., Kaulen, H., and Schell, J. (1989). A CACGTG motif of the *Antirrhinum majus* chalcone synthase promoter is recognized by an evolutionary conserved nuclear protein. Proc. Natl. Acad. Sci. USA **86,** 6930–6934.

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

Stougaard, J., Sandal, N., Grøn, A., Kühle, A., and Marcker, K.A. (1987). 5′ Analysis of the soybean leghemoglobin *lbc3* gene: Regulatory elements required for promoter activity and organ specificity. EMBO J. **6,** 3565–3369.

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.

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

Tokuhisa, J.G., Singh, K., Dennis, E.S., and Peacock, W.J. (1990). A DNA-binding protein factor recognizes two binding domains within the octopine synthase enhancer element. Plant Cell **2,** 215–224.

Van Haute, E., Joos, H., Maes, M., Warren, G., Van Montagu, M., and Schell, J. (1983). Intergenetic transfer and exchange recombination of restriction fragments cloned into pBR322: A novel strategy for reversed genetics of the Ti-plasmids of *Agrobacterium tumefaciens*. EMBO J. **2,** 411–417.

Walker, J.C., Howard, E.A., Dennis, E.S., and Peacock, W.J. (1987). DNA sequences required for anaerobic expression of the maize alcohol dehydrogenase 1 gene. Proc. Natl. Acad. Sci. USA **84,** 6624–6628.

Wong, S.L., and Verma, D.P.S. (1985). Promoter analysis of a soybean nuclear gene coding for nodulin-23, a nodule-specific polypeptide involved in symbiosis with *Rhizobium*. EMBO J. **4,** 2431–2438.