The Plant Cell, Vol. 6, 659–668, May 1994 © 1994 American Society of Plant Physiologists

A Protein Binding AT-Rich Sequence in the Soybean Leghemoglobin *c3* Promoter Is a General *cis* Element That Requires Proximal DNA Elements to Stimulate Transcription

Niels Bech Laursen, Knud Larsen, Jesper Yde Knudsen, Hans Jürgen Hoffmann, Carsten Poulsen, Kjeld A. Marcker,[1] and Erik Østergaard Jensen

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

A nodule nuclear factor, NAT2, interacts with two AT-rich binding sites (NAT2 BS1 and NAT2 BS2) in the soybean leghemoglobin (*lb*) *c3* promoter. In transgenic *Lotus corniculatus* nodules, an oligonucleotide containing NAT2 BS1 activated an inactive −159 *lbc3* promoter when placed immediately upstream of the promoter. The activation was independent of the orientation of NAT2 BS1 but was dependent on its position in the promoter. The abilities of different mutated binding sites to activate expression in vivo were correlated to their respective in vitro affinities for binding NAT2. This suggested that the interaction between NAT2 and NAT2 BS1 is responsible for the observed reactivation. Further activation experiments with the *lbc3* and the leaf-specific *Nicotiana plumbaginifolia* ribulose bisphosphate carboxylase/oxygenase small subunit (*rbcS-8B*) promoter suggested that another specific *cis* element(s) is required for the function of NAT2 BS1. Thus, the −102 *lbc3* promoter lacking the organ-specific element (−139 to −102) was not reactivated by the presence of the binding site, and the *rbcS-8B* promoter required sequences between −312 and −257 to be activated by NAT2 BS1. This implies that NAT2 has to work in combination with other *trans*-acting factor(s) to increase expression. The finding of NAT2-like binding activities in different plant organs and the specific expression of the hybrid NAT2 BS1/−312 *rbcS-8B* promoter in leaves suggest that NAT2 is a general activator of transcription.

INTRODUCTION

Several plant genes termed nodulin genes are specifically expressed during the formation of nitrogen-fixing root nodules on legumes. The trigger of the developmental program leading to the new organ is the infection of the root by soil bacteria of the genera *Bradyrhizobium*, *Rhizobium*, or *Azorhizobium* (reviewed by Nap and Bisseling, 1990). Most of the nodulin genes are only expressed in certain cell types and/or periods of the nodule development. Thus, soybean *lbc3*, which encodes the leghemoglobin c3 protein, is expressed in the central infected tissue of the nodule (Lauridsen et al., 1993).

Functional analyses in transgenic legumes of nodulin promoters fused to bacterial reporter genes identified several regulatory sequences in the early nodulin 2 (*Enod2*), *lba*, *lbc3*, and *N23* genes from soybean, the glutamine synthetase (*gln-γ*) gene from French bean, and the *Srglb3* gene from *Sesbania rostrata* (Stougaard et al., 1987, 1990; Jørgensen et al., 1988, 1991; Szabados et al., 1990; Shen et al., 1992; Lauridsen et al., 1993; She et al., 1993a, 1993b). One of the most intensively studied of these genes is *lbc3* from soybean. Promoter analyses in transgenic *Lotus corniculatus* identified the following major *cis* elements: a strong positive enhancer-like element (SPE; −1364 to −947) and a weak positive element (WPE; −230 to −170), capable of retaining 2 to 5% of the full promoter activity. In addition, an organ-specific element (OSE; −139 to −102) and a negative element (NE; −102 to −49) were identified (Figure 1C) (Stougaard et al., 1987; She et al., 1993b).

The WPE extends over two AT-rich sequence motifs interacting with a nuclear factor NAT2, which is found in the root nodules of soybean (Jensen et al., 1988). Although NAT2 BS1 (−246 to −223) binds NAT2 stronger than NAT2 BS2 (−176 to −161), the NAT2 BS2 is the best conserved of the two binding sites. NAT2 BS2 is found in all four soybean *lb* promoters and is also present in the promoter regions of the *lb* genes *Srglb2* and *Srglb3* from *S. rostrata* (Metz et al., 1988). Protein binding sites with sequences slightly different from NAT2 BS1 are present in the promoters of two other nodule-specific genes, namely the *N23* gene from soybean (Jacobsen et al., 1990) and the *gln-γ* gene from French bean (Forde et al., 1990).

Protein factors that bind to AT-rich binding sites, such as NAT2 BS1, NAT2 BS2, and related sequences in other late nodulin promoters, have been identified in nuclear extracts from *S. rostrata* nodules, roots, and leaves, alfalfa nodules and leaves, French bean roots and nodules, and soybean plumules.

[1] To whom correspondence should be addressed.

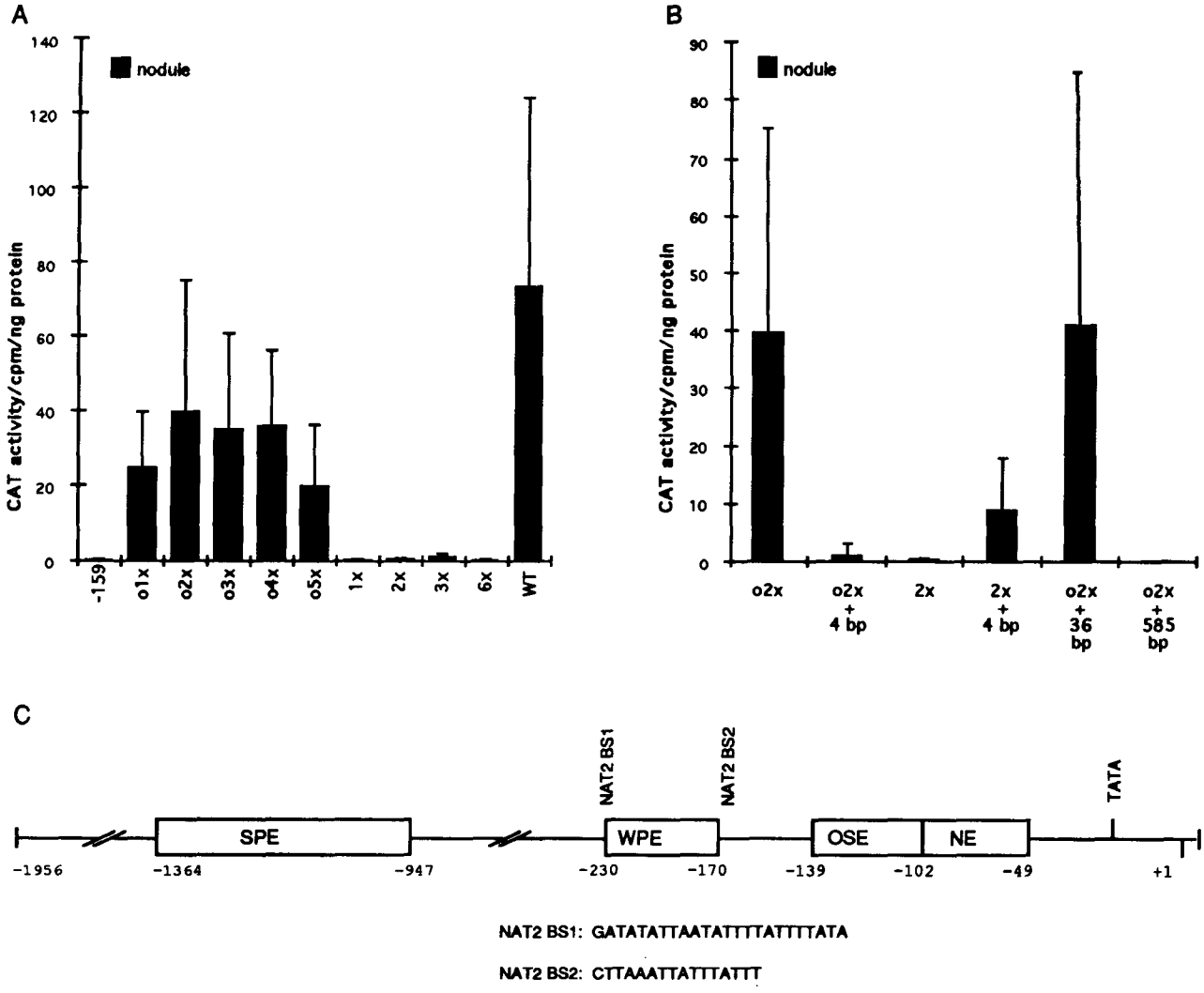

Figure 1. Reactivation of Expression from the Inactive −159 *lbc3* Minimal Promoter.

(A) CAT activity of the truncated −159 *lbc3*–CAT construct after reactivation by NAT2 BS1. Various copy numbers of the 32-bp NAT2 BS1 oligonucleotide were inserted in opposite (o1x to o5x) or normal (1x to 3x and 6x) orientation upstream of the inactive −159 *lbc3* promoter (−159). Wild type (WT) is the full-length 2-kb *lbc3* promoter (Stougaard et al., 1987). Error bars (95%) are indicated above the columns. Between seven and 20 independent transformations were measured for each construct. No activity was detected in roots or leaves with any of the constructs (data not shown). 1x and 6x were only measured in roots and nodules.

(B) Effect of insertion of DNA fragments between the NAT2 BS1 and the −159 *lbc3* promoter. Two copies of NAT2 BS1 in opposite orientation (o2x) were separated from the −159 *lbc3* promoter by a 4-bp insertion (o2x + 4 bp). 2x and 2x + 4 bp are as described above but with NAT2 BS1 inserted in its normal orientation. Insertion of a 36-bp (o2x + 36 bp) or a 585-bp (o2x + 585 bp) pUC19 fragment is also shown. Error bars (95%) are indicated above the columns. Between 11 and 27 independent transformations were measured for each construct. No activity was detectable in roots (data not shown).

(C) Organization of regulatory DNA elements in the 2-kb 5′ region of the soybean leghemoglobin *lbc3* gene. Binding sites for the NAT2 *trans*-acting factor are indicated by NAT2 BS1 and NAT2 BS2.

However, the mobilities of the complexes formed by extracts from the various organs differ slightly in gel retardation assays (Metz et al., 1988; Forde et al., 1990; Czarnecka et al., 1992).

The functional significance of the NAT2 binding sites was analyzed in the *lbc3* promoter. Removal of the WPE in a 5′ deletion series resulted in loss of activity (Stougaard et al., 1987), whereas an internal deletion removing the WPE from

a fully active promoter only had a minor effect on the level of expression (Stougaard et al., 1990).

cis-acting AT-rich sequences have previously been described in the promoters of β-phaseolin and phytohemagglutinin-L from French bean (Bustos et al., 1989; Riggs et al., 1989), *Gmhsp17.3-B* from soybean (Rieping and Schöffl, 1992), *cabE* (encoding the chlorophyll *a/b* binding protein) from *Nicotiana plumbaginifolia* (Castresana et al., 1988; Schindler and Cashmore, 1990), and ribulose bisphosphate carboxylase/oxygenase small subunit (*rbcS-3A*) from pea (Lam et al., 1990). Recently, an AT-rich *cis* element from the soybean heat shock gene *Gmhsp17.5E* has been shown to stimulate transcription when placed 5′ to a truncated (−140) maize alcohol dehydrogenase 1 (*Adh1*) promoter (Czarnecka et al., 1992).

To elucidate the functional significance of the NAT2 recognition sequences in the *lbc3* promoter, one to six copies of short NAT2 BS1 oligonucleotides (see Methods for sequence data) were linked to an inactive chimeric −159 *lbc3*–chloramphenicol acetyltransferase (CAT) gene, and the expression was subsequently analyzed in transgenic *L. corniculatus*. Here, we show that the NAT2 binding site activated the inactive truncated *lbc3* promoter in an orientation-independent manner. However, the activity of NAT2 BS1 is highly dependent on its distance relative to other downstream *cis* elements, because reactivation required the presence of a *cis* element(s) contained within the OSE region. The general effect of the presence of the NAT2 BS1 was investigated in reactivation experiments with the leaf-specific *rbcS-8B* and the constitutive cauliflower mosaic virus (CaMV) 35S promoters. Analogous to the *lbc3* promoter, the NAT2 BS1 enhanced expression of a truncated leaf-specific promoter *rbcS-8B* only when other *cis* elements between −312 and −257 were present in the construct. NAT2 BS1 is therefore capable of increasing expression from two truncated promoters derived from genes that are specifically expressed in two different kinds of cells. Therefore, this suggests that NAT2 BS1 functions as a general *cis* element requiring other proximal DNA elements to stimulate transcription.

RESULTS

NAT2 BS1 Reactivates Transcription from an Inactive −159 *lbc3* Minimal Promoter

To investigate the functional significance of the NAT2 binding sites in the *lbc3* promoter, oligonucleotides containing one of the binding sites (−246 to −223, NAT2 BS1) were fused to the silent −159 *lbc3* promoter deletion. Different copy numbers of a 32-bp oligonucleotide containing NAT2 BS1 were inserted in both orientations 20 bp upstream of the chimeric −159 *lbc3*–CAT gene. The constructs were transferred to *L. corniculatus*, and CAT activity was measured in roots, nodules, and leaves of the transgenic plants. The results in Figure 1A show that in nodules NAT2 BS1 reactivates the expression from the inactive −159 minimal *lbc3* promoter (column −159) when inserted in the opposite orientation (columns o1x, o2x, o3x, o4x, and o5x), but not when placed in its normal orientation (columns 1x, 2x, 3x, and 6x). When one to five copies of the binding site were inserted in opposite orientations (columns o1x, o2x, o3x, o4x, and o5x), the level of expression corresponded to ~50% of the 2-kb full-length wild-type promoter activity. These observations indicate that NAT2 BS1 functions as a positive *cis* element capable of stimulating transcription.

Activity of NAT2 BS1 Is Position Dependent

The failure of NAT2 BS1 placed in the normal orientation to reactivate the silent *lbc3* promoter might be due to a suboptimal distance between the binding site and other promoter elements. To investigate this possibility, the position of the binding sites was changed by inserting 4 bp between two copies of NAT2 BS1 and the −159 *lbc3* promoter. The results from transgenic *L. corniculatus* showed that NAT2 BS1 was then active in the normal orientation (Figure 1B, lane 2x + 4 bp) but nearly inactive in the opposite orientation (lane o2x + 4 bp). Thus, the activity of NAT2 BS1 was independent of its orientation in the promoter, but it was highly dependent on its position relative to other proximal promoter elements.

To investigate this effect in more detail, two copies of NAT2 BS1 in opposite orientations (o2x) were separated from the −159 *lbc3* promoter by inserting 36- and 585-bp pUC19 DNA fragments, respectively. The results shown in Figure 1B demonstrate that the small insertion (36 bp) between NAT2 BS1 and the −159 *lbc3* promoter has no effect on the expression (column o2x + 36 bp), whereas an insertion of a 585-bp fragment results in an inactive construct (column o2x + 585 bp). Two copies of NAT2 BS1 in normal orientation separated from the −159 *lbc3* promoter by the same 585 bp (2x + 585 bp) also failed to reactivate the promoter (data not shown), suggesting that NAT2 BS1 must be located close to other promoter elements to function.

lbc3 OSE Is Required for Reactivation

The −159 *lbc3* promoter contains the OSE and the NE, in addition to the basal promoter elements. To investigate whether the NAT2 BS1 is able to reactivate shorter promoters containing only the most basal *cis* elements, such as the TATA and CAAT boxes, several constructs were designed and tested for expression.

Two different minimal promoter constructs were used in this investigation: the −102 *lbc3*–CAT, which includes the NE but not the OSE, and the heterologous −90 35S–CAT gene fusion (Stougaard et al., 1987). Both minimal promoters are inactive in roots, nodules, and leaves. Although the −90 35S promoter is expressed to some extent in roots of tobacco (Poulsen and Chua, 1988; Benfey and Chua, 1989), this promoter is inactive in *L. corniculatus* (Stougaard et al., 1990; Jørgensen et al., 1991). The addition of two copies of NAT2

BS1 in either orientation had no effect on the two promoters, because no expression was detected in any organ. Similarly, the insertion of NAT2 BS1 at the −90 position between the 35S enhancer and 35S promoter in the −829 35S–CAT gene fusion did not lead to any change in specificity and expression level (data not shown).

Thus, the presence of the OSE (−139 to −102) and its 5′ flanking sequence between −159 and −139 is required for NAT2 BS2 to enhance expression from the *lbc3* basal promoter. This suggests that NAT2 works in combination with other *trans*-acting factors that associate with the OSE region and thereby mediate this effect. To see whether it would be possible to reactivate the two minimal promoters, the SPE from the *lbc3* promoter (−1364 to −947) was placed upstream of both promoters in a control experiment. Both the −102 *lbc3* promoter and the −90 35S promoters were specifically expressed in nodules when fused to the SPE (data not shown).

Correlation between in Vitro Binding of NAT2 to an *lbc3* Promoter Fusion and in Vivo Expression of the Gene

Specific point mutations in NAT2 BS1 influence the binding of NAT2 (Jacobsen et al., 1990). To study whether the reactivation is correlated to the binding of the *trans*-acting factor NAT2, two mutant NAT2 BS1 oligonucleotides were designed, one with low affinity for NAT2 and another that does not bind NAT2. The former NAT2 BS1 mutant contained six AT transversions, and the latter oligonucleotide had six A-to-G and T-to-C transitions in the same positions. The binding of these oligonucleotides to NAT2 was tested in gel retardation experiments, and the relative binding affinity compared to the wild-type NAT2 BS1 oligonucleotide was determined. The binding strength of NAT2 to the AT-mutated oligonucleotide is 40% of the wild type, whereas the CG mutant does not bind (Figure 2A).

Two copies of the mutant oligonucleotides were inserted 20 bp upstream of the −159 *lbc3*–CAT gene in the opposite orientation and the constructs transferred to *L. corniculatus*. The AT mutations resulted in an eightfold lower expression when compared to the wild-type oligonucleotide (Figure 2B, columns AT and WT), whereas the CG mutant failed to reactivate the promoter (Figure 2B, column CG). The nodule specificity of the expression was not changed in any of the new constructs. Thus, the positive effect of NAT2 BS1 on the expression from the truncated *lbc3* promoter is most likely due to its capacity to bind NAT2.

NAT2 BS1 Can Activate a Leaf-Specific Truncated Promoter

The reactivation of expression reported above was obtained with the homologous *lbc3* minimal promoter in nodule tissue. To test whether NAT2 BS1 could reactivate or enhance the expression of a heterologous promoter in organs other than

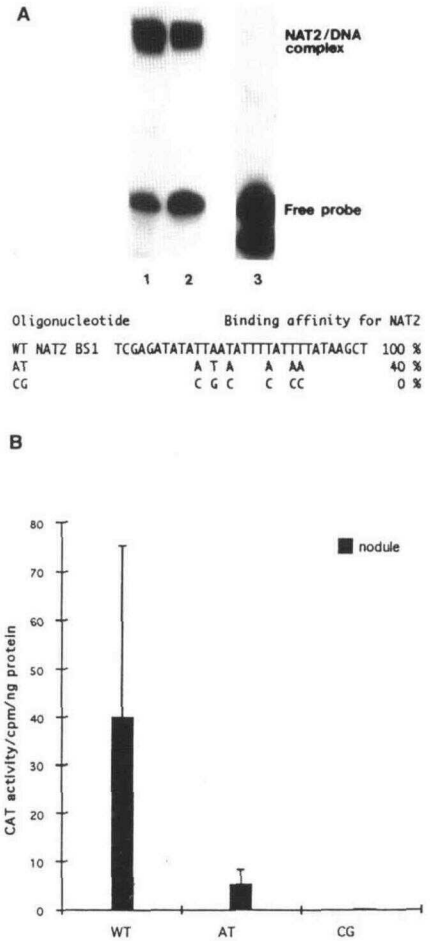

Figure 2. Reactivation by Mutant NAT2 BS1.

(A) Binding affinity of NAT2 for mutated wild-type NAT2 BS1 oligonucleotides. Nodule nuclear extract from soybean was incubated with labeled DNA fragments and salmon sperm. Lane 1 contains labeled wild-type NAT2 BS1; lane 2, the labeled AT mutant oligonucleotide; lane 3, the labeled CG mutant oligonucleotide. Only nucleotides differing from the NAT2 BS1 wild-type (WT) sequences are shown. Protein–DNA complexes were fractionated by electrophoresis on an 8% polyacrylamide gel to resolve bound and free NAT2 BS1 probe. The amount of probe present in the NAT2–DNA complexes and the unbound DNA fragments was measured by scintillation counting. The binding affinity was determined as the ratio between the amounts of probe in the complexes and total amount of probe. The binding affinity of the wild-type fragment was set to 100%. Similar analysis of a mutant NAT2 BS1 oligonucleotide with only three AT transversions shows only minor reduction in NAT2 binding capacity compared to the wild type.

(B) Reactivation of the −159 *lbc3*–CAT construct by insertion of the wild-type and mutated NAT2 BS1 oligonucleotides. Two copies of the oligonucleotides shown in **(A)** were inserted in opposite orientation upstream of the −159 *lbc3* promoter. Error bars (95%) are indicated above the columns. Between 17 and 21 independent transformations were measured for each construct. No activity was detectable in roots (data not shown).

nodules, we designed several constructs based on upstream elements from the *rbcS-8B* promoter of *N. plumbaginifolia* (Poulsen and Chua, 1988). We used the following upstream *rbcS-8B* fragments: the −1038 fragment with wild-type promoter activity in *N. plumbaginifolia*, the −312 and the −257 deletions with 30 and 25% activity, respectively (Poulsen and Chua, 1988), and the −192 deletion with very low promoter activity. The three former constructs (−1038, −312, and −254) are expressed specifically in the leaves of *N. plumbaginifolia* (Poulsen and Chua, 1988).

The results presented in Figure 3A show that the *rbcS-8B* promoter deletions exhibited the same organ specificity and the same relative promoter strength in *L. corniculatus* as in *N. plumbaginifolia* (columns −1038, −312, −257, and −192). No significant enhancement or reactivation was detected when two copies of NAT2 BS1 were placed upstream of the shortest deletion (columns 2x/−192 and o2x/−192), but some increase was observed for deletions 2x/−257 and o2x/−257. However, the addition of NAT2 BS1 to the −312 *rbcS-8B* promoter resulted in an eightfold enhancement of the expression in leaves when NAT2 BS1 was in the normal orientation (2x/−312) and a fivefold increase when placed in the opposite orientation (o2x/−312). The two copies of NAT2 BS1 placed in the normal orientation enhanced the transcription level above the level of the −1038 promoter. There was no increase of expression in roots or nodules (data not shown).

Thus, NAT2 BS1 seems to act as a general *cis* element capable of enhancing the transcription level of both nodule- and leaf-specific promoters. It does not direct or influence the tissue specificity of these promoters. Analogous to the *lbc3* minimal promoter, the protein interacting with NAT2 BS1 probably cooperates with other *trans*-acting factors binding to specific *cis* elements present between −312 and −257 of the *rbcS-8B* promoter. In control experiments in which the constitutive CaMV 35S enhancer (Benfey and Chua, 1990) was placed upstream of the four different truncated *rbcS-8B* promoters, the 35S constitutive enhancer (35SE) was found to stimulate transcription levels of all four promoters. Although transcription in roots and nodules was enhanced, the expression in leaves was still much more pronounced (Figure 3B, columns 35SE/−1038, 35SE/−312, 35SE/−257, and 35SE/−192).

NAT2 or a NAT2-like Protein Is Present in Nodules and Leaves of *L. corniculatus*

The results obtained using the *rbcS-8B* promoter fragments in combination with NAT2 BS1 showed that NAT2 BS1 is active in *L. corniculatus* leaves. This conflicts with the previous observation that NAT2 is a nodule-specific *trans*-acting factor (Jensen et al., 1988). To demonstrate the presence of proteins binding to NAT2 BS1 in *L. corniculatus*, nuclear extracts were prepared from leaves and nodules of *L. corniculatus*. Figure 4 shows that both nodules and leaves contain proteins that form complexes with NAT2 BS1. These complexes have the

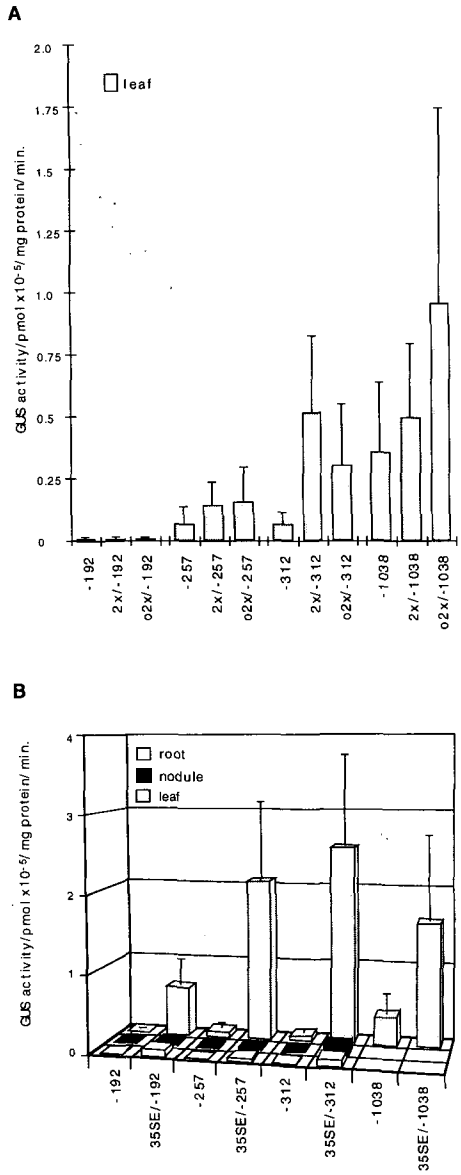

Figure 3. Expression of *rbcS-8B*–GUS Constructs Fused to Either the 35S Enhancer (SE) or NAT2 BS1.

(A) Two copies of NAT2 BS1 were inserted in the normal (2x/−192 to 2x/−1038) or opposite (o2x/−192 to o2x/−1038) orientation upstream of the various 5′ deleted *rbcS-8B* promoters (−192 to −1038). No activity was detectable in roots or nodules (data not shown).

(B) Effect of placing the 35S enhancer (35SE/−192 to −1038) upstream of 5′ deleted *rbcS-8B* promoters (−192 to −1038). Error bars (95%) are indicated above the columns. Between 10 and 19 independent transformations were measured for each construct. The GUS activity in roots and nodules of all the −1038 constructs was analyzed histochemically instead of fluorometrically (see Methods). No significant staining was observed in any of the root or nodules of the −1038 constructs, and compared to leaves only a weak staining was observed in the root and nodules of the 35SE/−1038 construct (data not shown).

same mobility as that of the complex formed between soybean NAT2 and NAT2 BS1.

The proteins present in extracts from nodules and leaves of *L. corniculatus* bind NAT2 BS1 in a sequence-specific manner, because the addition of small amounts of nonradioactive NAT2 BS1 eliminated binding, whereas binding was not affected by the addition of salmon sperm DNA (Figure 4, lanes 3 and 5). The nodule and leaf extracts from *L. corniculatus* also showed the same binding activities relative to the mutated AT and CG oligonucleotides, as demonstrated for soybean NAT2 (data not shown), indicating a close relationship between soybean NAT2 and the corresponding proteins in nodules and leaves of *L. corniculatus.*

DISCUSSION

The soybean nodule nuclear factor NAT2 recognizes two AT-rich sequences in the soybean *lbc3* promoter (Jensen et al., 1988). The functional significance of this interaction was unclear, because the results of 5′ deletions and internal deletions of the binding sites suggested only a minor role in gene activation (Stougaard et al., 1987, 1990). Here, we analyzed the in vivo expression conferred by oligonucleotides that contain NAT2 binding sites fused to either the soybean *lbc3* promoter or the heterologous *rbcS-8B* promoter in transgenic *L. corniculatus.* In both cases, the NAT2 binding site increased the level of expression from the two promoters, whereas it did not influence the specificity of expression. Inserted at the −90 position between the 35S enhancer and promoter, the NAT2 BS1 was also unable to confer nodule specificity to the −829 35S promoter. Inserted at the same −90 position, the OSE plus NE from the *lbc3* promoter conferred nodule specificity to the −829 35S promoter (Lauridsen et al., 1991). Therefore, we suggest that the NAT2 binding site is a quantitative rather than a qualitative *cis* element. The activity of NAT2 BS1 in the constructs analyzed here was dependent on the presence of *cis* elements located between −159 and −102 in the *lbc3* promoter and between −312 and −257 in the *rbcS-8B* promoter. Similarly, Lam et al. (1990) show that a tetramer of box VI (−51 to −31) from the pea *rbcS-3A* promoter is capable of enhancing transcription from the −90 CaMV 35S promoter but not from the −46 CaMV 35S promoter. An AT-rich element from the soybean *Gmhsp17.3-B* promoter (−593 to −321) is able to enhance transcription from the −56 35S promoter only when inserted together with the distal heat shock element (−321 to −254) from the *Gmhsp17.3-B* promoter (Rieping and Schöffl, 1992). Regulatory elements that must be juxtaposed with other *cis* elements to function have been defined as class B elements (Fromental et al., 1988; Czarnecka et al., 1992). Therefore, it is appropriate to classify NAT2 BS1 as a class B element.

NAT2 BS1 functions in both orientations, but it has no effect when placed far upstream from the transcription start site. Thus, NAT2 BS1 is only active when placed immediately upstream of other promoter elements. The NAT2 BS1 oligonucleotide used in the reactivation experiments is 32 bp long. Thus, multimers of NAT2 BS1 inserts are in helical-turn phase and will therefore have their corresponding NAT2 contact point(s) located on the same side of the DNA helix. The apparent position dependency demonstrated by insertion of 4 bp between the BS1 oligonucleotide and the −159 *lbc3* promoter (Figure 4B) might therefore be due in this case to a requirement for the NAT2/DNA contact point(s) to be located on a particular side of the DNA helix relative to the position of the contact point for other transcription factors.

Similar behavior was also reported for AT-rich sequences (−907 to −899 and −843 to −826) present in the soybean *Gmhsp17.5E* gene; these sequences bind soybean plumule nuclear proteins. Although this region does not appear to be functional in the *Gmhsp17.5E* promoter, a 33-bp element homologous to these AT-rich sequences acts as an orientation-independent positive upstream element when placed 5′ to a truncated −140 *Adh* promoter in sunflower tumors. However, when placed upstream of a −410 *Adh* promoter, this element has only a minor effect on expression of the reporter gene (Czarnecka et al., 1992).

AT-rich sequences that interact with proteins that may be related to NAT2 have also been identified in positive regulatory elements of the *gln-γ* promoter from French bean and the *Srglb3* promoter from *S. rostrata* (Szabados et al., 1990; Shen et al., 1992). However, these regulatory elements cover regions of 223 and 267 bp, respectively; therefore, it is not clear whether these binding sites are part of or constitute the functional sequences in the delimited regions.

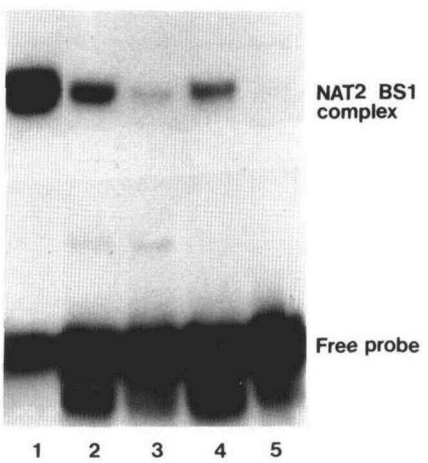

Figure 4. Binding of Nuclear Protein Extracts from Soybean and *L. corniculatus* to NAT2 BS1.

Nuclear extract from soybean nodules (lane 1), crude extract from *L. corniculatus* leaves (lanes 2 and 3), and nuclear extract from *L. corniculatus* nodules (lanes 4 and 5) were incubated with labeled NAT2 BS1 (monomer) and salmon sperm DNA. Unlabeled NAT2 BS1 (1000-fold molar excess) was added as specific competitor (lanes 3 and 5). Protein–DNA complexes were fractionated by electrophoresis on an 8% polyacrylamide gel to resolve bound and free NAT2 BS1 probe.

An important regulatory element spanning the region from −1098 to −1009 in the tomato *rbcS-3A* promoter overlaps an AT-rich domain containing a recognition sequence for the nuclear factor AT-1, which can be isolated from pea plumules (Datta and Cashmore, 1989; Ueda et al., 1989). Assuming that the AT-1 binding site is the functional *cis* element in this region, the mechanism of activation must be different from that occurring between NAT2 and the soybean plumule nuclear proteins and their respective AT-rich sequences (Czarnecka et al., 1992), due to the far upstream location of the AT-1 binding site. However, AT-1 also binds to an AT-rich negative regulatory element in the *N. plumbaginifolia cabE* gene (Castresana et al., 1988). Thus, AT-1 appears active in positive as well as negative gene regulation. A functional AT-rich sequence has also been described in the French bean β-phaseolin promoter region (−682 to −628). A 55-bp oligonucleotide containing this sequence binds a nuclear protein from cotyledons in vitro and is able to activate a −90 35S promoter construct independent of the orientation of the oligonucleotide (Bustos et al., 1989). The failure of NAT2 BS1 to activate the identical −90 35S construct suggests that interactions at the two binding sites are different.

An important question is whether the NAT2 protein is identical to or related to known AT binding proteins. Several of the nuclear factors interacting with AT-rich sequences in plant genes have been characterized to some extent. The AT-1 protein is temperature sensitive because activity is gradually lost when incubated at 37°C (Datta and Cashmore, 1989). In contrast, our experiments showed that the NAT2 protein is rather heat stable, surviving treatment at 80°C for 10 min without significant loss of activity (K. Larsen and H. J. Hoffmann, unpublished results). Furthermore, AT-1 binding is modulated by phosphorylation; however, we have no evidence that this is also the case for NAT2 activity (K. Larsen and H. J. Hoffman, unpublished results).

The 3AF1 factor interacting with an AT-rich sequence in the −45 region of the *rbcS-3A* promoter is a metal-dependent DNA binding protein because metal chelating agents completely inhibit complex formation (Lam et al., 1990). We found that NAT2 binding activity is unaffected by the addition of metal chelating compounds to the reaction mixture (K. Larsen and H. J. Hoffmann, unpublished results).

High mobility group (HMG) proteins also bind to AT-rich sequences (Grasser et al., 1990; Jacobsen et al., 1990; Pedersen et al., 1991), but we have shown previously that NAT2 cannot be classified as an HMG protein (Jacobsen et al., 1990). Jofuku et al. (1987) showed that a 60-kD protein from soybean embryo recognizes AT-rich sequences in the genes encoding soybean lectin and the Kunitz trypsin inhibitor. The relationship between this protein and NAT2 is unclear at present. Czarnecka et al. (1992) characterized the nuclear factors from soybean plumules interacting with the AT-rich *cis* elements of the *Gmhsp17.5E* promoter. There are two types of nuclear factors that bind: a high molecular mass group consisting of several proteins with molecular masses in the range of 46 to 69 kD and a low molecular mass group with molecular masses ranging from 23 to 32 kD. The latter group of proteins most likely consists of HMG proteins. In gel retardation assays, the mobility of the complex formed between NAT2 BS1 and plumule nuclear proteins from the high molecular mass group is consistent with the mobility of the complex between NAT2 BS1 and NAT2. Further characterization reveals that this binding activity is primarily associated with a protein with molecular mass of ~54 kD (Czarnecka et al., 1992).

We have purified NAT2 activity from soybean nodule nuclei and found that the activity is associated with two proteins with molecular masses of about 54 and 69 kD, respectively (K. Larsen and H. J. Hoffmann, unpublished results). Furthermore, the AT-rich sequence binding the plumule nuclear proteins can stimulate transcription from a truncated *Adh* promoter similar to the way NAT2 BS1 can activate or increase transcription from the inactive *lbc3* promoter and the truncated *rbcS* promoters (Czarnecka et al., 1992). Thus, the interaction between NAT2 and NAT2 BS1 is most likely identical to or closely related to the interaction between the plumule nuclear protein(s) and the corresponding AT-rich sequences. The presence of similar types of factors in *L. corniculatus* nodules and leaves therefore suggests that AT-rich elements interacting with these types of proteins function as general *cis*-regulatory elements.

METHODS

Nucleic Acids Manipulation

Standard techniques for DNA manipulations were used, as described in Sambrook et al. (1989). All end filling of double-stranded DNA with sticky ends was conducted with the Klenow fragment of DNA polymerase I. DNA-modifying enzymes were purchased from Boehringer Mannheim and used according to the manufacturer's instructions.

Oligonucleotides

The 32-bp synthetic NAT2 BS1 oligonucleotide containing binding site 1 that interacts with nuclear factor NAT2 has the following sequence:

5′-AAGCTTCGAGATATATTAATATTTTATTTTAT-3′
3′-AGCTCTATATAATTATAAAATAAAATATTCGA-5′.

The monomer of the NAT2 BS1 was end filled before insertion. The multimers of NAT2 BS1 were made by self-ligation followed by end filling before insertion. The 32-bp oligonucleotides of the AT mutant and CG mutant (for sequences, see Figure 2A) were treated as NAT2 BS1.

Construction of Chimeric Genes

lbc3 *Constructs*

The wild-type 2-kb *lbc3*–chloramphenicol acetyltransferase (CAT) gene fusion is similar to the pCAT-14 (2-kb leghemoglobin *lbc3* 5′-CAT-825-bp *lbc3* 3′) plasmid described in Stougaard et al. (1987). The −159

and −102 *lbc3*–CAT constructs were made by cloning the SalI fragment of the respective *lbc3* 5′ deletions from Stougaard et al. (1987) into the SalI site in the polylinker of the pIV10 plasmid (Ramlov et al., 1993). Monomers and multimers of NAT2 BS1 were inserted in the SmaI site 20 bp upstream of the −159 and −102 *lbc3*–CAT constructs. The strong positive enhancer-like element (SPE) (a NheI/XbaI fragment from the *lbc3* promoter) was inserted in the XbaI site immediately upstream of the −102 *lbc3*–CAT construct. In the 2x and o2x constructs (Figure 1), a BamHI restriction site is present between the dimers of NAT2 BS1 and the −159 *lbc3* promoter fragment. A BamHI digest of the 2x and o2x constructs followed by end filling and religation results therefore in a 4-bp insertion in all of these constructs. Two different Sau3A fragments from pUC19 (pUC19 coordinates 2346 to 2377 and 2382 to 276) resulted in the o2x + 36 bp and o2x + 585 bp constructs (Figure 1B). The latter fragment was inserted in opposite orientation.

35S Constructs

The minimal −90 35S–CAT fusion is described in Stougaard et al. (1990). Dimers of NAT2 BS1 and the end-filled SPE were inserted in an end-filled HindIII site 47 bp upstream of the −90 35S–CAT fusion. Dimers of NAT2 BS1 were inserted in a HpaI linker at the −90 position of the −829 35S–CAT fusion (derived from Jefferson et al., 1987; Lauridsen et al., 1991).

rbcS *Constructs*

The −1038, −312, −257, and −192 *rbcS*–β-glucuronidase (GUS) fusions were made by cloning the HindIII end-filled and BamHI fragments of the 5′ *rbcS* deletions (Poulsen and Chua, 1988) into a SmaI/BamHI digest of pIV20 (Ramlov et al., 1993), which carries the GUS reporter gene from pBI101.1 (Jefferson et al., 1987). Dimers of NAT2 BS1 and the 35S enhancer (−829 to −90) were inserted in an end-filled XbaI site located 8 bp upstream of the *rbcS*–GUS fusions.

All constructs were checked by sequencing using the United States Biochemical Corporation Sequenase kit, version 2.0, on denatured double-stranded plasmids (Hattori and Sakaki, 1986). The DNA sequence of the complete 2-kb 5′ *lbc3* region (Christensen et al., 1989) has accession number X15061, and the DNA sequence of *rbcS-8B* has accession number M36685 in the EMBL, GenBank, and DDB data banks.

Transformation of Plants

Gene constructions were transferred into *Agrobacterium rhizogenes* as described by Van Haute et al. (1983). The AR12 strain carrying cauliflower mosaic virus (CaMV) 35S–GUS in the left T-DNA (TL) segment was used as vector for the chimeric CAT constructs, whereas the AR14 strain carrying 35S–CAT in the TL segment (Hansen et al., 1989) was used as vector for the chimeric GUS constructs. Composite plants with untransformed shoots on transformed "hairy roots" were produced according to Hansen et al. (1989). The transformation regeneration procedure produced whole transformed plants (Petit et al., 1987).

Biochemical Assays

CAT activity was measured as previously described (Stougaard et al., 1987). GUS activity was measured by the fluorometric assay of Jefferson et al. (1987). The only exception was that the GUS activity in roots and nodules of the constructs containing the −1038 *rbcS-8B* fragment was measured by the histochemical assay of Jefferson et al. (1987). Protein levels in extracts were determined by the dye-binding assay (Spector, 1978).

Preparation of Protein Extracts

For the preparation of nuclear extracts, the methods previously described by Jacobsen et al. (1990) were used with modifications.

Nuclear Extracts from Nodules

Soybean plants (*Glycine max* var Evans) were inoculated with *Bradyrhizobium japonicum* USDA 110, and *Lotus corniculatus* plants (var Rodéo) were inoculated with *Rhizobium loti* NZP2037. Three weeks after inoculation, 25 g of nodules were harvested and frozen in liquid nitrogen. The frozen nodules were homogenized with mortar and pestle under liquid nitrogen and dissolved in 175 mL of NIB (10 mM 2-(*N*-morpholino)ethane-sulfonic acid), pH 6.0, 5 mM EDTA, 0.15 mM spermine, 0.5 mM spermidine, 20 mM β-mercaptoethanol, 0.2 mM phenylmethylsulfonyl fluoride, 0.6% Triton X-100, and 0.25 M sucrose). The homogenate was filtered twice through Miracloth (Calbiochem) and centrifuged for 20 min at 2000*g* to obtain a crude nuclear pellet. Nuclei were washed once in NIB and resuspended in NIB plus Percoll (5.5 g of NIB and 44 g of Percoll). The suspension was centrifuged for 30 min at 8000*g*, the fraction floating on the top of the gradient was resuspended in NIB plus Percoll, and the floating step was repeated. The floated fraction was washed twice in NIB (the last wash without Triton X-100) and resuspended in 2.5 mL of NXB-500 (500 mM NaCl, 20 mM Hepes, pH 7.9, 25% glycerol, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM phenylmethylsulfonyl fluoride, and 0.5 mM DTT). Following sonication, the mixture was incubated for 30 min at 0°C, and nuclear debris was removed by a 10-min centrifugation at 15,000*g*. All steps were carried out at 4°C. Extracts containing nuclear proteins were aliquoted and stored at −80°C.

Crude Extract from Leaf

Leaves (5.5 g) of different age from *L. corniculatus* were harvested, frozen, homogenized, and dissolved in 100 mL of NIB, filtered through Miracloth, and centrifuged for 15 min at 2000*g*. Nuclei were washed three times in NIB (the last wash without Triton X-100). The washing steps included only centrifugation for 10 min at 500*g*. The pellet was resuspended in 1 mL of NXB-100 (NXB-500 with 100 mM, rather than 500 mM, NaCl). Following sonication, the mixture was incubated for 30 min at 0°C and centrifugated 10 min at 7000*g*. The pellet was resuspended in 1 mL of NXB-500, sonicated, and incubated for 30 min at 0°C; nuclear debris was removed by a 20-min centrifugation at 9400*g*. All steps were performed at 4°C. Extracts containing nuclear proteins were aliquoted and stored at −80°C. The NXB-500 fraction was used for the gel retardation assays.

Gel Retardation Assay

Gel retardation assays were performed as described by Jensen et al. (1988). The probes were end-labeled with ^{32}P-dATP using the Klenow fragment. Salmon sperm DNA (200 ng) were added to each lane. Seven to 12 μg of protein extracts were used in each lane.

ACKNOWLEDGMENTS

We thank Jens Stougaard for stimulating discussions. This research was supported by the Carlsberg Foundation, the Danish Biotechnology Programme, and Human Frontier Science Program.

Received January 24, 1994; accepted March 15, 1994.

REFERENCES

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

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

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.

Christensen, T., Sandal, N.N., Stougaard, J., and Marcker, K.A. (1989). 5′ Flanking sequence of the soybean leghemoglobin *lbc3* gene. Nucl. Acids Res. **17,** 4383.

Czarnecka, E., Ingersoll, J.C., and Gurley, W.B. (1992). AT-rich promoter elements of soybean heat shock gene *Gmhsp17.5E* bind two distinct sets of nuclear proteins in vitro. Plant Mol. Biol. **19,** 985–1000.

Datta, N., and Cashmore, A.R. (1989). Binding of a pea nuclear protein to promoters of certain photoregulated genes is modulated by phosphorylation. Plant Cell **1,** 1069–1077.

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.

Fromental, C., Kanno, M., Nomiyama, H., and Chambon, P. (1988). Cooperativity and hierarchical levels of functional organization in the SV40 enhancer. Cell **54,** 943–953.

Grasser, K.D., Maier, U.G., Haass, M.M., and Feix, G. (1990). Maize high mobility group proteins bind to CCAAT and TATA boxes of a zein gene promoter. J. Biol. Chem. **265,** 4185–4188.

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 denaturated plasmid templates. Anal. Biochem. **152,** 232–238.

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 *lbc3* 5′ upstream region. EMBO J. **7,** 1265–1271.

Jofuku, K.D., Okamuro, J.K., and Goldberg, R.B. (1987). Interaction of an embryo DNA binding protein with a soybean lectin gene upstream region. Nature **328,** 734–737.

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.

Lam, E., Kano-Murakami, Y., Gilmartin, P., Niner, B., and Chua, N.-H. (1990). A metal-dependent DNA-binding protein interacts with a constitutive element of a light-responsive promoter. Plant Cell **2,** 857–866.

Lauridsen, P., Sandal, N.N., Kühle, A., Marcker, K., and Stougaard, J. (1991). Regulation of nodule-specific genes. In Plant Molecular Biology **2,** R.G. Herrman and B. Larkins, eds (New York: Plenum Press), pp. 131–137.

Lauridsen, P., Franssen, H., Stougaard, J., Bisseling, T., and Marcker, K.A. (1993). Conserved regulation of the soybean early nodulin *ENOD2* gene promoter in determinate and indeterminate transgenic root nodules. Plant J. **3,** 483–492.

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

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

Pedersen, T.J., Arwood, L.J., Spiker, S., Guiltinan, M.J., and Thompson, W.F. (1991). High mobility group chromosomal proteins bind to AT-rich tracts flanking plant genes. Plant Mol. Biol. **16,** 95–104.

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

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

Ramlov, K.B., Laursen, N.B., Stougaard, J., and Marcker, K.A. (1993). Site-directed mutagenesis of the organ-specific element in the soybean leghemoglobin *lbc3* gene promoter. Plant J. **4,** 577–580.

Rieping, M., and Schöffl, F. (1992). Synergistic effect of upstream sequences, CCAAT box elements, and HSE sequences for enhanced expression of chimaeric heat shock genes in transgenic tobacco. Mol. Gen. Genet. **231,** 226–232.

Riggs, C.D., Voelker, T.A., and Chrispeels, M.J. (1989). Cotyledon nuclear proteins bind to DNA fragments harboring regulatory elements of phytohemagglutinin genes. Plant Cell **1,** 609–621.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory).

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

She, Q., Sandal, N.N., Stougaard, J., and Marcker, K.A. (1993a). Comparative sequence analysis of *cis* elements present in *Glycine max* L. leghemoglobin *lba* and *lbc3* genes. Plant Mol. Biol. **22,** 931–935.

She, Q., Lauridsen, P., Stougaard, J., and Marcker, K.A. (1993b). Minimal enhancer elements of the leghemoglobin *lba* and *lbc3* gene promoters from *Glycine max* L. have different properties. Plant Mol. Biol. **22,** 945–956.

Shen, W.-J., Williamson, M.S., and Forde, B.G. (1992). Functional analysis of the promoter region of a nodule-enhanced glutamine synthetase gene from *Phaseolus vulgaris* L. Plant Mol. Biol. **19,** 837–846.

Spector, T. (1978). Refinement of the Coomassie blue method of protein quantitation. Anal. Biochem. **86,** 142–146.

Stougaard, J., Sandal, N.N., Grøn, A., Kühle, A., and Marcker, K.A. (1987). 5′ Analysis of the soybean leghaemoglobin *lbc3* 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 *lbc3* and N23 gene promoters. Mol. Gen. Genet. **220,** 353–360.

Szabados, L., Ratet, P., Grunenberg, B., and de Bruijn, R.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.

Ueda, T., Pichersky, E., Malik, V.S., and Cashmore, A.R. (1989). Level of expression of the tomato *rbcS-3A* gene is modulated by a far upstream promoter element in a developmentally regulated manner. Plant Cell **1,** 217–227.

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.