Mutants in the nodFEL promoter of Rhizobium leguminosarum bv. viciae reveal a role of individual nucleotides in transcriptional activation and protein binding

Okker, R.J.H.; Spaink, H.P.; Lugtenberg, E.J.J.; Schlaman, W.R.M.; Lugtenberg, B.J.J.; Schlaman, H.R.M.

Arch Microbiol (2001) 175 : 152–160
DOI 10.1007/s002030000249

ORIGINAL PAPER

Robert J. H. Okker · Herman P. Spaink
Ben J. J. Lugtenberg · Helmi R. M. Schlaman

Mutants in the *nodFEL* promoter of *Rhizobium leguminosarum* bv. *viciae* reveal a role of individual nucleotides in transcriptional activation and protein binding

Received: 31 July 2000 / Revised: 22 November 2000 / Accepted: 13 December 2000 / Published online: 9 February 2001

Abstract The highly conserved *nod* box sequence in the promoters of the inducible nodulation genes of rhizobia is required for transcription activation together with NodD, a LysR-type transcriptional regulator, and a flavonoid as a coinducer. DNA fragments containing *nod* box sequences form two binding complexes when crude preparations of *Rhizobium leguminosarum* bv. *viciae* are used: a NodD-dependent and an additional, NodD-independent complex. The role of individual nucleotides in the conserved *nod* box sequence in complex formation and in nodulation gene expression was investigated by introducing 13 individual base-pair substitutions in the *nodF nod* box of *R. leguminosarum* bv. *viciae* and studying their effect on promoter activity and protein-DNA complex formation. Two mutants showed decreased NodD binding and decreased promoter activity. Five mutants showed a NodD-dependent complex as with the wild-type *nodF nod* box, whereas their promoter activity was severely reduced after induction. This result is in agreement with earlier observations that NodD DNA binding also occurs in the absence of inducer. Four mutants were impaired in the formation of the NodD-independent retardation complex. Three of them showed no alterations in promoter activity, meaning that no specific role for the protein forming the NodD-independent complex could be established. The two mutants in the highly conserved LysR motif of the *nod* box were unable to direct coinducer-dependent promoter activity but, unexpectedly, their retardation patterns were not altered. The remaining two mutants showed constitutive promoter activity. The results are discussed in terms of the relevance of conserved nucleotides and motifs identified in the *nod* box.

R. J. H. Okker and H. R. M. Schlaman contributed equally to this paper.

R. J. H. Okker · H. P. Spaink · B. J. J. Lugtenberg
H. R. M. Schlaman (✉)
Leiden University, Institute of Molecular Plant Sciences,
Wassenaarseweg 64, 2333 AL Leiden, The Netherlands
e-mail: schlaman@rulbim.leidenuniv.nl,
Tel.: +31-71-5274927, Fax: +31-71-5274999

Keywords *Rhizobium* · *nod* Box · *nodD* · Nodulation · Transcriptional regulation

Introduction

Bacteria of the genus *Rhizobium* can induce root nodules on leguminous plants in a host-specific way. Within the nodules, the bacteria reside in an altered form, designated as bacteroids, in which they fix atmospheric nitrogen. Both plant and bacterial signal molecules are required for the formation of nodules. A large group of bacterial genes, known as *nod*, *nol*, and *noe* (together known as nodulation) genes, is involved in the nodulation process. In fast-growing rhizobia like *Rhizobium leguminosarum* biovar (bv.) *viciae* and *Sinorhizobium meliloti*, nodulation genes are localized on a Sym(biosis) plasmid and organized in different operons. Mutations in the *nodABC* and *nodFE* genes, each comprising separate operons, abolish nodulation. Most nodulation genes are involved in the synthesis of specific lipo-chitin oligosaccharides that induce the formation of root nodules (Downie 1998).

Expression of many of the nodulation genes is positively regulated by the product of *nodD* in the presence of flavonoids or isoflavonoids that are released by the plant. NodD presumably interacts directly with the coinducer to activate transcription but direct proof of this is lacking (for a review see Schlaman et al. 1998). The protein NodD is a member of the large LysR family of transcriptional regulators which share common features at the protein level, in DNA target site recognition, and in their mode of transcription activation (reviewed in Schell 1993). The promoter regions of the nodulation operons, which are positively regulated by NodD, contain a highly conserved sequence that is known as the *nod* box (Rostas et al. 1986, and see Fig. 1) and which is essential for promoter activity (Fisher and Long 1989; Rostas et al. 1986; Spaink et al. 1987). The *nod* boxes contain three highly conserved stretches of nucleotides, separated by two, short less-well-conserved regions of three and six nucleotides, respectively (see Fig. 1). Moreover, a so-called LysR motif, which is char-

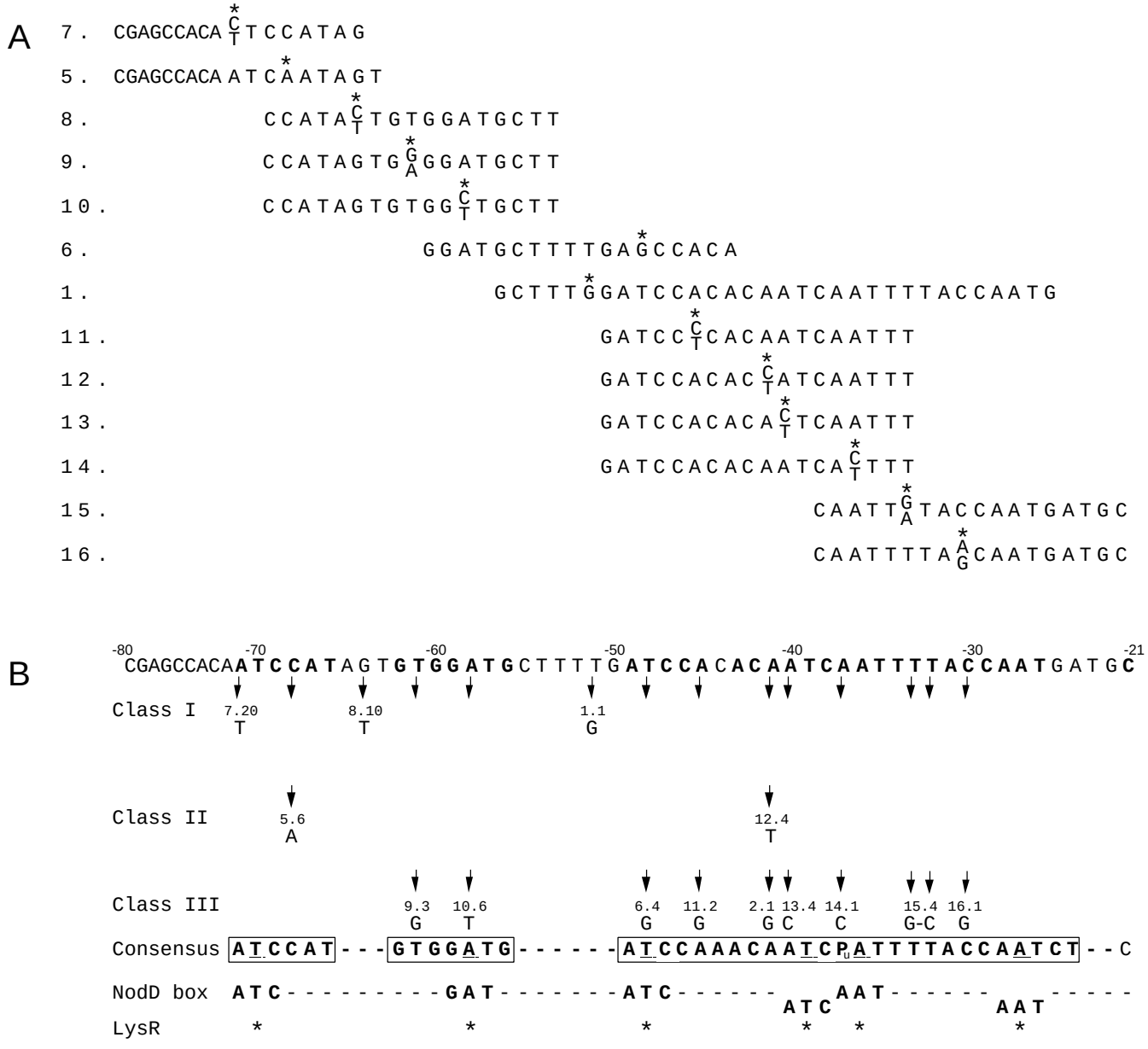

Fig. 1A, B Primers used to induce mutations in the *nodF nod* box and mutants obtained. **A** Primers are indicated with the mutant nucleotide marked with an *asterisk* (*). Type 2 primers containing two different mutant nucleotides at a certain position in a 1:1 ratio are indicated with an *asterisk* and the two mutations. **B** The sequences of the wild-type *nodF nod* box and the mutants obtained. Bases are numbered relative to the most 5′ transcription initiation site of *nodF* (Spaink et al. 1989). Mutant 2.1 is a "spontaneous" mutant that was described earlier (Schlaman 1992). Distribution of mutants in different classes is after their phenotype (see text for further details). Sequences in the *nodF nod* box homologous to the consensus sequence are in *bold*. *Boxed bases* indicate the consensus sequence of the rhizobial *nod* boxes (Fisher and Long 1989; Spaink et al. 1987). Furthermore, the NodD box motif which has the conserved A-T-C-N_9-G-A-T repeat (Goethals et al. 1992) is indicated and the LysR motif characterized as a T-N_{11}-A repeat (Goethals et al. 1992) is marked by *asterisks*. P_u is A or G

acterized by the sequence T-N_{11}-A and which is present in many promoters regulated by LysR-type transcriptional regulators, has been recognized in *nod* boxes (Goethals et al. 1992; Schell 1993) as well as a so-called NodD box, which is a discrete, inverted repeat structure with the basic structure A-T-C-N_9-G-A-T (Goethals et al. 1992). Deletions in the highly conserved regions and insertions in the central less-well-conserved region (nucleotides –55 through –50 in Fig. 1B) lead to loss of promoter activity (Fisher and Long 1993; Spaink et al. 1987; Wang and Stacey 1991). However, insertion of ten nucleotides in this region has very little effect on promoter activity (Fisher and Long 1993). The NodD protein binds in the presence and in the absence of flavonoids in vitro to the *nod* box, as shown by gel electrophoresis retardation studies (Fisher and Long 1993; Hong et al. 1988; Kondorosi et al. 1989; Schlaman et al. 1992), and it protects a large area, includ-

ing virtually the entire *nod* box, from enzymatic and chemical cleavage (Fisher and Long 1989; Kondorosi et al. 1989; Machado et al. 1998). NodD induces a bending of the *nod* box upon binding (Fisher and Long 1993) and critical points of contact between NodD and nucleotides in the *nod* box have been determined (Fisher and Long 1993; Schlaman 1992). The precise mode of action of NodD and coinducer in inducing nodulation gene transcription has not been completely revealed, although it has been demonstrated for some cases that, in the presence of flavonoids, the binding of NodD to its target DNA becomes stronger, and changes in Dnase I sensitivity of the NodD-*nod* box complex suggest an altered binding (Goethals et al. 1992; Kondorosi et al. 1989).

Besides the NodD-specific retardation complex, gel electrophoresis retardation studies, using crude protein extracts of *R. leguminosarum* bv. *viciae* and the *nodA*, *nodF* and *nodM nod* boxes but not the *nodO nod* box, have revealed a second retardation complex (Hong et al. 1988; Schlaman 1992; Schlaman et al. 1992). This complex, which migrates between the NodD-containing complex and the free DNA fragment, might be similar to or different from the NolR-containing complex observed when using extracts from *S. meliloti* (Kondorosi et al. 1989). Sequences similar to *nolR* sequences have been found in *R. leguminosarum* bv. *viciae*, but a NolR target site is only present in the *nodA nod* box of *R. leguminosarum* bv. *viciae* (Kiss et al. 1998). The second retardation complex is formed independently from NodD and has therefore been designated as a NodD-independent complex. It shows the existence of another *nod*-box-binding protein in *R. leguminosarum* bv. *viciae* whose function is not yet understood. Previously, a point mutation in the *nodF nod* box of *R. leguminosarum* bv. *viciae* was described which could not form the NodD-independent complex and which showed very low promoter activity after flavonoid induction, whereas the NodD-dependent complex was normally present (Schlaman 1992), suggesting that this NodD-independent complex contains an activator of transcription.

To assess the role of individual nucleotides in the conserved *nod* box sequence, we introduced single nucleotide substitutions in the *nod* box in front of the *nodFEL* genes of *R. leguminosarum* bv. *viciae* and transcription activation was compared with complex formation. Crude bacterial extract was used for these latter assays because we wanted to obtain insight into a possible role of the various complexes in *nod* gene transcription. The results are discussed in terms of the relevance of conserved nucleotides and motifs identified in the *nod* box.

Materials and methods

Plasmids, molecular cloning, and bacterial strains

Plasmid pMP2066 contains the *nodFEL* promoter of *R. leguminosarum* bv. *viciae* isolated on a 116-bp *Sal*I fragment using PCR (Schlaman et al. 1992). This fragment and adjoining sequences of the multi-cloning site were recloned as a *Sph*I-*Bam*HI fragment from pMP2066 in pIC20R (Marsh et al. 1984), resulting in pMP2070. DNA fragments with assumed mutations in the *nod* box, as generated using PCR with pMP2070 as template (see below), were cloned as *Bgl*II-*Eco*RI fragments into pIC20H (Marsh et al. 1984). These constructs were named after the primer from which they originated (Fig. 1), e.g., as pMP7.20, and so on. Fragments with an established mutation were cloned in front of a promoterless *lacZ* gene in the wide-host range vector pMP220 (Spaink et al. 1987). These plasmids were designated, e.g., as pMP220–7.20, and so on. Plasmid pMP2073 contains the wild-type *nodFEL* promoter in front of *lacZ* in pMP220.

Plasmids were transferred to *R. leguminosarum* bv. *viciae* wild-type strain RBL5560 (Zaat et al. 1987) using tri-parental mating (Ditta et al. 1980). Plasmid pMP220–12.4 was also transferred to rhizobial strains RBL5561, harboring *nodD*::Tn*5* (Zaat et al. 1987), and LPR5045, cured from the Sym plasmid (Hooykaas et al. 1982).

Derivatives of *Rhizobium* strain RBL1391, which is *R. leguminosarum* bv. *viciae* 248 cured of the Sym plasmid, harboring either plasmid pMP280 or pMP1070, were used to make crude protein extracts (Schlaman et al. 1992). Plasmids pMP280 and pMP1070 are IncP vectors with the promoter and structural gene of *nodD*, and with the *nodABCIJ* promoter region of *R. leguminosarum* bv. *viciae*, respectively.

Site-specific mutagenesis

To mutate the *nodF nod* box, a set of oligonucleotide primers was designed which differed in only one nucleotide from the original sequence. The primers contained either one single substitution (type 1 primers) or they were a mixture of two different primers in each of which a nucleotide was replaced by another nucleotide (type 2 primers) in a 1:1 ratio (see Fig. 1). In this latter design, a pyrimidine or purine was replaced by both possible purines and pyrimidines, respectively. Mutations were introduced in the *nodF nod* box by a two-step polymerase chain reaction (PCR) using Supertaq polymerase enzyme and pMP2070 as template. In the first step of PCR, the mutagenic primers were used together with the standard –40 M13 primer. The double-stranded reaction product was purified by electrophoresis on a 5% polyacrylamide gel (Sambrook et al. 1989) and used as primer in a subsequent PCR reaction together with the standard M13 reverse primer to elongate the product to cover the entire *nodF nod* box. In this second step of PCR, the first ten cycles were performed with only the M13 reverse primer to enrich the template strand in order to favor annealing of the desired strand of the reaction product of the first PCR. Subsequently, the first PCR product was added to the mix and 30 more PCR cycles were performed. The final PCR product was purified using native polyacrylamide gel electrophoresis (Sambrook et al. 1989), digested with *Bgl*II and *Eco*RI, and cloned in pIC20H. DNA sequences were confirmed by the dideoxy chain-termination method (Sambrook et al. 1989).

Gel electrophoresis retardation assay

Wild-type *nodF nod* box derived from pMP2070 and mutant *nodF nod* boxes from pMP7.20 and so on were isolated as 116-bp *Sal*I fragments and labeled by filling in the 3′ recessive ends using Klenow enzyme and (α^{32}P)-dCTP using standard methods (Sambrook et al. 1989). Estimated specific activity ranged between $7–17 \cdot 10^3$ dpm·(ng DNA)$^{-1}$. Crude protein preparations were prepared from rhizobia as described previously (Schlaman et al. 1992). Binding reactions were performed in volumes of 15 µl containing 37 mM Tris-HCl (pH 8.3); 3.7 mM EDTA; 100 mM KCl; 0.15 mM dithiothreitol; 3.7% sucrose; 13% glycerol; 65 µg ml^{-1} of each herring sperm DNA, polydA·polydT, and polydI·polydC (Pharmacia, Woerden, The Netherlands) as competing DNA; 13–15 µg total protein extract; and 0.3–0.7 ng *nodF nod* box DNA equivalent to 5000 dpm. This extract:DNA ratio was found in a pilot experiment to be optimal to visualize the different complexes (data not shown). The mix was incubated during 20 min at ambient temperature; subsequently, 3 µl Ficoll mix containing xylene cyanol FF and bromophenol blue dyes was added. The samples

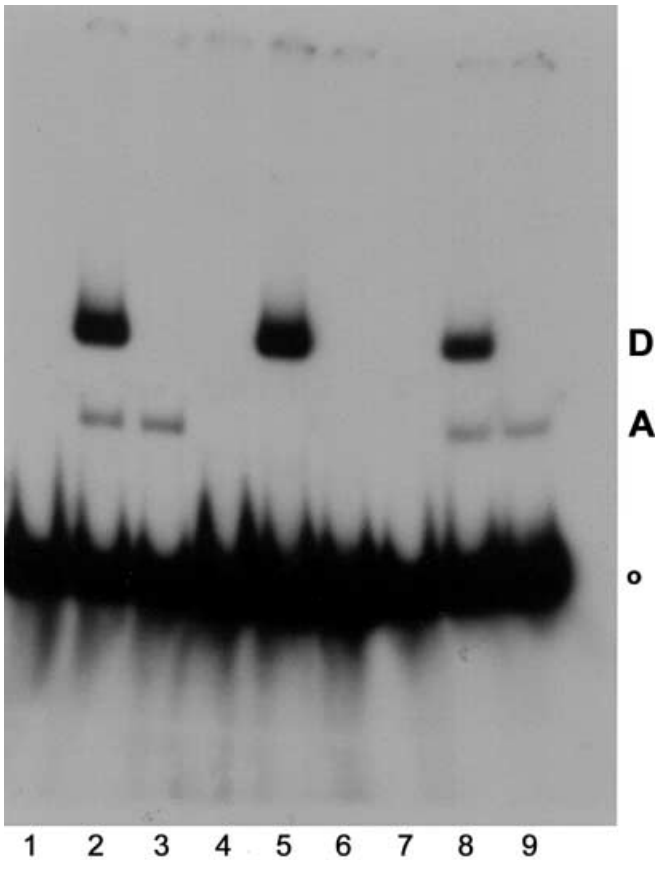

Fig. 2 Retardation complexes formed with the *nodF nod* box. DNA fragments are wild-type *nodF nod* box sequence (*lanes 1–3, 7–9*) and mutant 2.1 (*lanes 4–6*). Protein extracts added to the lanes are from: no protein extract (*lanes 1, 4, 7*), RBL1391.pMP280 (with NodD, *lanes 2, 5, 8*) and RBL1391.pMP1070 (without NodD, *lanes 3, 6, 9*). Protein extracts were obtained from naringenin-induced (*lanes 1–6*) or uninduced (*lanes 7–9*) cultures. The NodD-dependent and the NodD-independent retardation complexes are indicated with *D* and *A*, respectively. The unbound DNA fragment is indicated with ❍

were loaded on 5% polyacrylamide gels in Tris-borate-EDTA buffer and run at ambient temperature at 10 V cm^{-1}. The experiments of which the results are presented in Fig. 2 were performed under slightly different conditions as follows: binding was performed at 21 °C in the presence of 500 μg ml^{-1} herring sperm DNA as competitor and the gels were run at 4 °C. Gels were dried and exposed to Fuji X-ray films using intensifying screens.

β-Galactosidase assay

To measure promoter activity, levels of β-galactosidase were determined according to published procedures (Spaink et al. 1987). If appropriate, hesperitin (150 nM) was added to the bacterial cultures to induce nodulation gene transcription.

Results

Design of *nod* box mutagenesis

In *R. leguminosarum* bv. *viciae* the promoter of the *nodABCIJ* operon promotes stronger transcriptional activity than do promoters of other nodulation operons (Spaink et al. 1987). However, this promoter overlaps with the *nodD* promoter, which is constitutively expressed and autoregulated. To avoid complications, the *nod* box from the slightly weaker *nodFEL* promoter (indicated as *nodF nod* box) was selected for mutational analysis. The *nod* box consists of three highly conserved DNA regions, separated by two less-well-conserved regions (Fig. 1B). Sites for mutation were chosen in the highly conserved regions, including the LysR motif (Goethals et al. 1992), as these regions are the most important for both promoter activity and NodD binding (Fisher and Long 1993; Kondorosi et al. 1989; Spaink et al. 1987). Furthermore, they were chosen to be evenly dispersed over the conserved regions with two to three nucleotides in between them. In the central less-well-conserved region, only one site was mutated to introduce a *Bam*HI site (indicated as 1.1 in Fig. 1B). The desired mutations were introduced in the *nodF nod* box using PCR techniques and the mutagenic primers depicted in Fig. 1A.

Isolated *nodF nod* box mutants

Forty clones with a presumed *nodF nod* box mutation, which were isolated after PCR amplification with individual mutagenic primers, showed restriction fragments of the expected size and were further analyzed by nucleotide sequencing. Twenty out of these 40 clones showed more than one base-pair mutation and were discarded, except for one (designated as 15.4 in Fig. 1B). From the remaining 20 clones, six showed the same base-pair substitution at the same position as in other clones and were therefore discarded, leaving us 14 different *nod* box mutants.

The single *nodF nod* box mutations that were subjected to further analyses are shown in Fig. 1. Primer 8, consisting of a 1:1 mixture of primers with two possible base-pair substitution at a certain position, was the only one out of ten of this type of mixed primers that resulted in both possible mutants (mutants 8.10 and 8.11). One mutant, designated as 15.4 in Fig. 1B, contained the expected bp substitution according to primer 15 as well as a "spontaneous" mutation ($T_{-32} \rightarrow C$) next to it. Because this latter mutation was at an interesting position of the *nod* box, this mutant was included for further characterization. In weak *nod* gene promoters, for example the *nodO nod* box, a C residue is always substituted for the consensus T_{-32}.

The mutant designated as 2.1 in Fig. 1B has been isolated before and was described earlier as a Supertaq-polymerase-induced base-pair substitution in the synthesis of the *nodF nod* box as a 116-bp *Sal*I fragment (Schlaman 1992).

Promoter activity of mutant *nod* boxes

To determine the promoter activity of the mutant *nodF nod* boxes, the mutated DNA fragments were cloned in front of the promoterless *lacZ* reporter gene in pMP220

Table 1 Promoter activity of *nodF nod* boxes with base pair substitutions as determined by measuring β-galactosidase activity. Plasmids are present in *Rhizobium leguminosarum* bv. *viciae* strain RBL5560. Hesperitin (150 nM) was used as flavonoid coinducer. β-Galactosidase activity is expressed in Miller Units (U). Data are the mean of four independent experiments, except for the control, which is the mean of ten independent experiments, each performed in quadruplicate, with a deviation (standard error) less than 10%

Class	Plasmid	β-Galactosidase		Activity relative to wild-type	
		–Hesperitin	+Hesperitin	–Hesperitin	+Hesperitin
	Wild-type				
	pMP2073	220	2390		
	Mutants				
I	pMP220–7.20	310	1820	1.4	0.8
	pMP220–8.10	300	2320	1.4	1.0
	pMP220–8.11	360	2480	1.6	1.0
	pMP220–1.1	220	2380	1.0	1.0
II	pMP220–5.6	830	1710	3.8	0.7
	pMP220–12.4	3320	3630	15.1	1.5
III	pMP220–9.3	120	180	0.6	0.1
	pMP220–10.6	190	270	0.9	0.1
	pMP220–6.4	380	340	1.7	0.1
	pMP220–11.2	140	100	0.6	0.0
	pMP220–2.1	190	420	0.9	0.1
	pMP220–13.4	240	170	1.1	0.1
	pMP220–14.1	160	330	0.7	0.1
	pMP220–15.4	300	330	1.4	0.1
	pMP220–16.1	300	270	1.4	0.1

Table 2 Promoter activity of the constitutive mutant *nodF nod* box pMP220–12.4 in various bacterial backgrounds. β-Galactosidase activity is expressed in Miller Units (U). Hesperitin (150 nM) was used as flavonoid coinducer. Data are the mean of four independent experiments each performed in quadruplicate, with a deviation (standard error) less than 10%. *ND* Not determined

Bacterial strains and plasmids	β-Galactosidase		Activity relative to wild-type	
	–Hesperitin	+Hesperitin	–Hesperitin	+Hesperitin
R. leguminosarum bv. *viciae*				
RBL5560 (*nodD*$^+$)				
pMP2073	220	2390		
pMP220–12.4	3320	3630	15.1	1.5
RBL5561(*nodD*$^-$)				
pMP2073	240	310		
pMP220–12.4	4400	4400	18.3	14.2
LPR5045(pSym$^-$)				
pMP2073	240	300		
pMP220–12.4	4100	ND	17.0	ND
E. coli				
JM101				
pMP2073	0	0		
pMP220–12.4	1500	1500	>1500	>1500

(Spaink et al. 1987) and transferred to *R. leguminosarum* bv. *viciae* wild-type strain RBL5560. As a positive control, plasmid pMP2073, which contains *lacZ* under control of the wild-type *nodFEL* promoter, was used. Activity was determined as the number of units of β-galactosidase produced in the absence or presence of the nodulation gene coinducer hesperitin. The results, presented in Table 1, showed three classes of mutants. In class I mutants, the promoter activity was not, or only slightly, affected (mutants 7.20, 8.10, 8.11,and 1.1). Class II mutants had an elevated level of promoter activity in the absence of coinducer, i.e. they are constitutive mutants (mutants 5.6 and 12.4). In mutant 5.6, the promoter activity showed a fourfold increase and a 30% reduction compared to the control in the absence and presence of coinducer, respectively. Mutant 12.4 showed almost the same level of promoter activity in the presence or absence of coinducer, and the levels were substantially higher than that of the wild-type. Therefore, transcription from this mutant promoter is completely constitutive. In class III mutants, the promoter activity was very low in the presence of coinducer (mutants 9.3, 10.6, 6.4, 11.2, 2.1, 13.4, 14.1, 15.4, and 16.1). In the absence of coinducer, the three mutants 6.4, 15.4, and 16.1 had a higher level of activity than the uninduced control. Mutants 2.1, 9.3, 10.6, 11.2, 13.4, and 14.1 showed an activity equal to or slightly reduced compared to the uninduced control.

The completely constitutive mutant 12.4 was further characterized (Table 2). Surprisingly, it was found that the constitutive phenotype appeared to be completely independent of the presence of NodD protein, as shown by the high levels of promoter activity in the NodD-lacking *Rhizobium* strains RBL5561 (*nodD*::Tn*5*) and LPR5045 (pSym$^-$). Even in *Escherichia coli*, considerable promoter

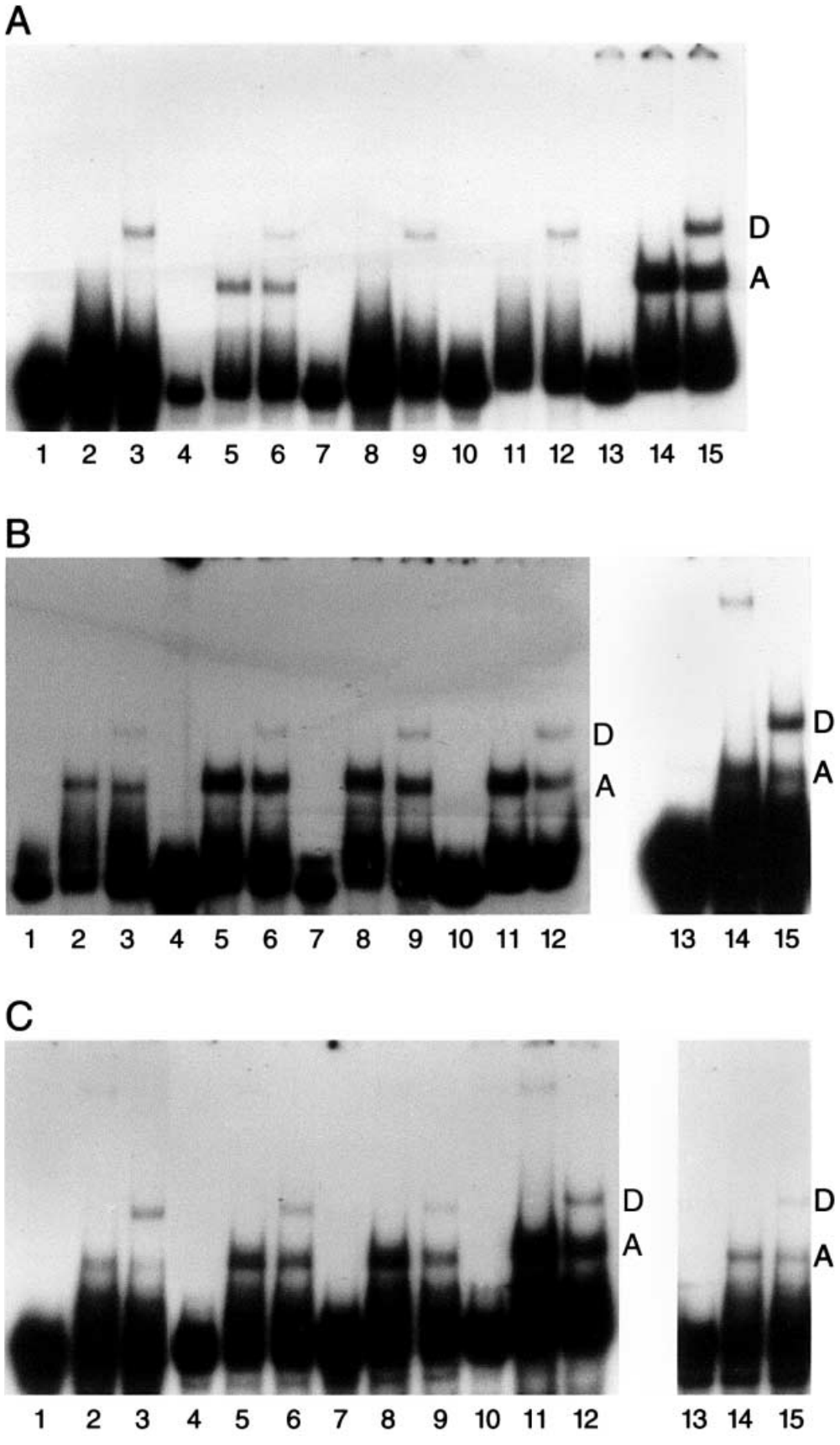

Fig. 3A–C Retardation complexes formed with mutated *nodF nod* boxes harboring a single base pair substitution. In **A–C**, protein extracts added to the labeled DNA fragment are from the following sources: *Lanes 1, 4, 7, 10, 13, 16* no protein extract added, *lanes 2, 5, 8, 11, 14, 17* RBL1391.pMP1070 (without NodD), *lanes 3, 6, 9, 12, 15, 18* RBL1391.pMP280 (with NodD). Mutant *nodF nod* box DNA fragments are: **A** *Lanes 1–3* pMP7.20, *lanes 4–6* pMP5.6, *lanes 7–9* pMP8.10, *lanes 10–12* pMP8.11, *lanes 13–15* pMP2070 (wild-type). **B** *Lanes 1–3*, pMP10.6, *lanes 4–6* pMP6.4, *lanes 7–9* pMP11.2, *lanes 10–12* pMP2070 (wild-type), *lanes 13–15* pMP9.3 (long exposure). **C** *Lanes 1–3* pMP12.4, *lanes 4–6* pMP14.1, *lanes 7–9* pMP15.4, *lanes 10–12* pMP2070, *lanes 13–15* pMP13.4 (long exposure). *D* NodD-dependent complex, *A* NodD-independent complex

Table 3 Summary of phenotypes of mutants in the *nodF nod* box

Class	Mutant numbers	Promoter activity	Retardation complexes	
			NodD Complex	NodD-independent Complex
I	7.20; 8.10; 8.11	(Almost) normal	Normal	Absent
II	5.6; 12.4	At least partially constitutive	Normal	Normal or reduced
IIIA	2.1	Decreased	Normal	Absent
IIIB	9.3; 13.4	Decreased	Reduced	Normal or reduced
IIIC	10.6; 6.4; 11.2; 14.1; 15.4	Decreased	Normal	Normal

activity was detected, whereas the control plasmid pMP2073 did not show any activity at all in this heterologous background.

Behavior of mutant *nod* boxes in electrophoretic retardation assays

To analyze whether patterns of protein complexes formed by total rhizobial extracts with the mutated *nodF nod* boxes are altered compared to the wild-type *nodF nod* box, electrophoretic gel-retardation assays were performed. The wild-type *nodF nod* box showed two retardation complexes. One of the complexes is NodD-dependent (designated as D in Fig. 2 and Fig. 3), as it only formed with protein extracts obtained from *R. leguminosarum* bv. *viciae* strains containing *nodD*. The other complex (designated as A in Fig. 2 and Fig. 3) formed with protein preparations derived from all tested *R. leguminosarum* bv. *viciae* strains, including those strains lacking *nodD*; it is therefore indicated as a NodD-independent complex (or A-complex). The observation that more than one retardation complex is formed with *nod* box DNA has been reported before for *R. leguminosarum* bv. *viciae* (Hong et al. 1988; Schlaman 1992; Schlaman et al. 1992) as well as for other rhizobia (Goethals et al. 1992; Kondorosi et al. 1989). Usually, it was observed that the NodD-independent complex with wild-type *nodF nod* box was more abundant than the nodD-dependent complex under our experimental conditions. However, the opposite was observed when higher amounts of competitor DNA were used (up to 500 $\mu g\ ml^{-1}$) suggesting a higher sequence specificity of NodD protein than of the protein forming the NodD-independent complex for *nod* box DNA. Since more A-complex seemed to be formed by lysates lacking NodD than by lysates having NodD, it may be that binding sites of NodD overlap with those of the unknown protein forming the A-complex.

Special interest in the NodD-independent retardation complex (A) arose from the observation with *nodF* box mutant 2.1, which could not form this complex and showed very low promoter activity after flavonoid induction, whereas the NodD-dependent complex was normally present (Fig. 2 and Schlaman 1992). To investigate the relationship between the NodD-independent complex (A) and promoter activation in more detail, all other *nodF nod* box mutants were subjected to gel electrophoretic retardation assays. Several independent experiments, each with a set of various mutants, were performed and in each set wild-type *nodF nod* box served as a control. The results are presented in Fig. 3; within each panel the results of the mutants should be compared with that of the wild-type in the same panel. As shown in Fig. 3, extracts of *R. leguminosarum* bv. *viciae* carrying or not carrying *nodD* did not form the NodD-independent complex with the *nodF nod* box mutants 7.20, 8.10, and 8.11 (Fig. 3A, lanes 1–3, 7–12), whereas with all other mutants the A-complex was formed.

The mutants 9.3 and 13.4 showed reduced binding with NodD, as could be concluded from the fact that, reproducibly, much longer exposure times were needed to observe the NodD-dependent retardation complex (D) in the case of both these mutants (Fig. 3B, lanes 13–15 and Fig. 3C, lanes 13–15, respectively). Moreover, mutant 13.4 showed also weak A-complex formation. All other mutants (5.6, 12.4, 10.6, 6.4, 11.2, 14.1, and 15.4) showed patterns of retardation complex formation indistinguishable from that of the wild-type *nodF nod* box control (Fig. 3A, lanes 13–15 and Fig. 3B, lanes 10–12).

The phenotypes of the *nodF nod* box mutants in transcription activation and retardation complex formation are summarized in Table 3.

Discussion

Inducible nodulation gene expression

Transcription of many nodulation genes in rhizobia requires the transcriptional activator NodD protein, an activating flavonoid (so-called coinducer) derived from the host plant, and a conserved *nod* box DNA sequence. Information on structural requirements of *nod* boxes for expression has mainly been obtained from deletion and insertion analyses. Deletions from the 5′ and 3′ ends of the *nod* box covering more or less of the conserved regions abolishes transcription (Spaink et al. 1987; Wang and Stacey 1991) as do insertions in the central non-conserved region (Fisher and Long 1993). However, insertion of exactly ten nucleotides in this region has no effect on transcription activation (Fisher and Long 1993). From in vitro assays it has become evident that NodD protein binds to *nod* boxes both in the presence and absence of coinducers. Several lines of evidence support the model that the

NodD-*nod* box binding alters upon addition of appropriate flavonoid coinducers: increased binding to *nod* box sequences in *S. meliloti*, *Azorhizobium caulinodans*, and *Sinorhizobium fredii* (Goethals et al. 1992; Kondorosi et al. 1989; Machado et al. 1998) and changes in DNase I footprints (Kondorosi et al. 1989). However, in other cases the affinity and/or binding of NodD for *nod* boxes seemed not to be affected (Fig. 2 and Fisher and Long 1989; Fisher et al. 1988; Hong et al. 1988). These apparently conflicting data may be the result of differences in the experimental set-ups. Except for NolR (Kondorosi et al. 1991), the role of additional proteins that bind to the *nod* box in transcription activation is not understood.

The present study is the first one in which single nucleotide mutants, evenly dispersed over the entire *nod* box, were investigated for their role in promoter activation and protein complex formation. All mutants showed an altered phenotype compared to the wild-type *nodF nod* box either in transcription activation, protein complex formation, or both, confirming the importance of the conserved residues in the *nod* box.

Promoter activity and NodD binding

Most of the *nodF nod* box mutants showed complete loss of promoter activity, whereas altered protein complex formation was only detected in some cases (Table 3). Surprisingly, in none of these so-called class III mutants was NodD binding completely abolished; only mutants 9.3 and 13.4 showed severely reduced NodD binding (Fig. 3B, lanes 13–15 and 3C, lanes 13–15, respectively), indicating that NodD protein binding is not sufficient for promoter activity. It should be noted that both these mutants are transversions and both are located at a similar position in either one of the inverted repeats of the NodD box. It has been proposed that the inverted repeats of the NodD box form loop structures to which the NodD protein binds as a dimer (Goethals et al. 1992). The mutations 9.3 and 13.4 might severely influence the 3-dimensional configuration of such loop structures, thereby preventing efficient binding of NodD protein.

Two mutants (10.6 and 6.4) were chosen for analysis of the LysR motif (Goethals et al. 1992; Schell 1993). These mutants showed lack of, or at least very decreased, transcription activation upon addition of coinducer. Mutations in the LysR motif constructed previously also showed complete lack of transcription activation (Goethals et al. 1992). The results show, however, that the retardation patterns of our mutants were unaltered (class IIIC mutants, Fig. 3B, lanes 1–6); this is in contrast to earlier observations in which a mutation in the LysR motif severely reduced NodD binding (Goethals et al. 1992). It cannot be excluded that our mutants lead to altered NodD binding, but the retardation assays used here are not sensitive enough to detect subtle differences.

Promoter activity and the NodD-independent binding complex

Additional NodD-independent binding complexes migrating in the retardation assay between the NodD-dependent complex and the free DNA fragment have been observed using extracts from *R. leguminosarum* bv. *viciae* (Hong et al. 1988; Schlaman 1992; Schlaman et al. 1992), *S. meliloti* (Kondorosi et al. 1989) and *A. caulinodans* (Goethals et al. 1992). The nature of such complexes may differ, depending on the *nod* box and the bacterial strain. For *S. meliloti*, it has been shown that this complex with the *nodA* promoter is formed by the repressor of nodulation gene expression NolR (Kondorosi et al. 1989, 1991). In *R. leguminosarum* bv. *viciae nolR*, homologous sequences have been detected but a NolR DNA target site is absent in the *nodF nod* box (Kiss et al. 1998). It has been shown previously (Schlaman et al. 1992) that in *R.* leguminosarum bv. *viciae* the NodD-independent complex was found to be specific for *nod* boxes of strong nodulation gene promoters, since it is not formed with, for instance, the weaker *nodO* promoter (De Maagd et al. 1989). Furthermore, the complex is specific for rhizobial protein extracts, as it is not observed using *E. coli* extracts (Schlaman 1992), indicating that it is not formed by proteins of the general transcription machinery.

It was reported previously that *nodF nod* box mutant 2.1 formed an unaltered NodD retardation complex, but that the NodD-independent binding complex was absent and transcription activation was decreased (Schlaman 1992). This prompted the hypothesis that an unknown protein may be involved in the stimulation of *nod* promoter activity. However, here it was shown that, although with class I mutants 7.20, 8.10, and 8.11 the NodD-independent retardation complex was not formed (Fig. 3A, lanes 1–3 and 8–12, respectively), these mutants nevertheless exhibit high inducible promoter activity (Table 1). The NodD-independent complex formed with *R. leguminosarum* bv. *viciae nodF nod* box has therefore no clear function in transcription activation but apparently is not required for NodD-dependent transcription activation.

Constitutive *nod* gene transcription

Surprisingly, two *nodF nod* box mutants (5.6 and 12.4) were (partially) independent from flavonoids in their ability to activate nodulation gene transcription (class II mutants, Table 1). These mutants are at different locations in the *nod* box (Fig. 1), and therefore it is unlikely that a single conformational change is responsible for this phenotype. It cannot be excluded that the mutation(s) resulted in a new recognition site for RNA polymerase, making transcription therefore independent of NodD and coinducers. Mutant 12.4 showed the most pronounced phenotype in that its transcription activation was completely independent from NodD and that it even showed significant promoter activity in *E. coli* (Table 2), an observation that, until now, has never been reported for rhizobial promoters.

In the gel-retardation assay, both class II mutants formed retardation patterns characteristic of the NodD-dependent and the NodD-independent complexes. However, mutant 12.4 distinguished itself from the wild-type in that the NodD-dependent complex was more abundant than the (A) complex, whereas in the wild-type the reverse was the case under the same experimental conditions (Fig. 3C, lanes 1–3). Interestingly, mutant 12.4 ($A_{-41}\rightarrow T$ transversion) maps at exactly the same position as mutant 2.1 ($A_{-41}\rightarrow G$ transition), which showed severely decreased transcription activation and lacked the NodD-independent retardation complex (Fig. 2) suggesting that this nucleotide is at a key position for binding of the protein in the NodD-independent retardation complex. In agreement is our result of reduced amounts of NodD-independent complex observed for mutant 13.4, which maps directly adjacent at position –40 (Fig. 3C, lanes 13–15). The NodD-dependent complex is unexpected for mutant 5.6, since the altered nucleotide in this mutant forms a base pair at a critical point of contact between the protein NodD and the *nod* box (Fisher and Long 1993; Schlaman 1992). It is likely that other interactions between NodD and the *nod* box still exist and are sufficient to form complexes visible in a retardation assay.

In summary, we conclude from the data obtained with the class II and class IIIC *nod* box mutants, both of which showed unaltered retardation complex formation but completely opposite phenotypes in promoter activity, that formation of a NodD-dependent retardation complex is an insufficient criterion to judge functionality in transcription activation. Mutants such as our class II and class IIIC types can be very helpful in developing assays to discriminate between protein–promoter binding, which does or does not lead to transcription activation. Such discriminating assays would be of great use for the study of all promoters dependent on activation by proteins of the LysR family. Furthermore, it can be concluded that the different classes of mutants we obtained are evenly dispersed over the entire *nod* box. This shows that the *nod* box is not built up of simple blocks with distinct functions, but rather that it functions as an entity in transcription activation.

References

De Maagd RA, Wijfjes AHM, Spaink HP, Ruiz-Sainz JE, Wijfelman CA, Okker RJH, Lugtenberg BJJ (1989) *nodO*, a new *nod* gene of *Rhizobium leguminosarum* biovar *viciae* Sym plasmid pRL1JI, encodes a secreted protein. J Bacteriol 171:6764–6770

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

Downie JA (1998) Function of rhizobial nodulation genes. In: Spaink HP, Kondorosi A, Hooykaas PJJ (eds) The Rhizobiaceae. Kluwer, Dordrecht, pp 387–402

Fisher RF, Long SR (1989) DNA footprint analysis of the transcriptional activator proteins NodD1 and NodD3 on inducible *nod* gene promoters. J Bacteriol 171: 5492–5502

Fisher RF, Long SR (1993) Interactions of NodD at the *nod* box: NodD binds to two distinct sites on the same face of the helix and induces a bend in the DNA. J Mol Biol 233: 336–348

Fisher RF, Egelhoff TT, Mulligan JT, Long SR (1988) Specific binding of proteins from *Rhizobium meliloti* cell-free extracts containing NodD to DNA sequences upstream of inducible nodulation genes. Genes Dev 2: 282–293

Goethals K, van Montagu M, Holsters M (1992) Conserved motifs in a divergent *nod* box of *Azorhizobium caulinodans* ORS571 reveal a common structure in promoters regulated by the LysR-type protein. Proc Natl Acad Sci USA 89: 1646–1650

Hong, G-F, Burn *H*-E, Johnston AWB (1988) Evidence that DNA involved in the expression of nodulation (*nod*) genes in *Rhizobium* binds to the product of the regulatory gene *nodD*. Nucleic Acids Res 15: 9677–9690

Hooykaas PJJ, Schnijdewindt FGM, Schilperoort RA (1982) Identification of the Sym plasmid of *Rhizobium leguminosarum* stain 1001 and its transfer to and expression in other Rhizobia and *Agrobacterium tumefaciens*. Plasmid 8: 73–82

Kiss E, Mergaert P, Olàh B, Kereszt A, Staehelin C, Davies AE, Downie JA, Kondorosi A, Kondorosi E (1998) Conservation of *nolR* in the *Sinorhizobium* and *Rhizobium* genera of the *Rhizobiaceae* family. Mol Plant-Microbe Interact 11: 1186–1195

Kondorosi E, Gyuris J, Schmidt J, John M, Duda E, Hoffman B, Schell J, Kondorosi A (1989) Positive and negative control of *nod* gene expression in *Rhizobium meliloti* is required for optimal nodulation. EMBO J 8: 1331–1341

Kondorosi E, Pierre M, Cren M, Haumann U, Buiré M, Hoffmann B, Schell J, Kondorosi A (1991) Identification of NolR, a negative transacting factor controlling the *nod* regulon in *Rhizobium meliloti*. J Mol Biol 222: 885–896

Machado D, Pueppke SG, Vinardel JM, Ruiz-Sainz JE, Krishnan HB (1998) Expression of *nodD1* and *nodD2* in *Sinorhizobium fredii*, a nitrogen-fixing symbiont of soybean and other legumes. Mol Plant-Microbe Interact 11: 375–382

Marsh JL, Erfle M, Wykes EJ (1984) The pIC plasmid and phage vectors with versatile cloning sites for recombinant selection by insertional inactivation. Gene 32: 481–485

Rostas K, Kondorosi E, Horvath B, Simoncsits A, Kondorosi A (1986) Conservation of extended promoter regions of nodulation genes in *Rhizobium*. Proc Natl Acad Sci USA 83: 1757–1761

Sambrook JE, Fritsch F, Maniatis T (1989) Molecular cloning: a laboratory manual. 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Habor, NY

Schell MA (1993) Molecular biology of the LysR family of transcriptional regulators. Annu Rev Microbiol 47: 597–626

Schlaman HRM (1992) Regulation of nodulation gene expression in *Rhizobium leguminosarum* biovar *viceae*. PhD Thesis, Leiden University, Leiden, The Netherlands

Schlaman HRM, Lugtenberg BJJ, Okker RJH (1992) The NodD protein does not bind to the promoters of inducible nodulation genes in extracts of bacteroids of *Rhizobium leguminosarum* biovar *viciae*. J Bacteriol 174: 6109–6116

Schlaman HRM, PhillipsDA, Kondorosi E (1998) Genetic organization and transcriptional regulation of rhizobial nodulation genes. In: Spaink HP, Kondorosi A, Hooykaas PJJ (eds) The Rhizobiaceae. Kluwer Acad Publ, Dordrecht Boston London, pp 361–386

Spaink HP, Okker RJH, Wijffelman CA, Pees E, Lugtenberg BJJ (1987) Promoters in the nodulation region of the *Rhizobium leguminosarum* Sym plasmid pRL1JI. Plant Mol Biol 9: 29–37

Spaink HP, Okker RJH, Wijffelman CA, Tak T, Goosen-De Roo L, Pees E, Van Brussel AAN, Lugtenberg BJJ (1989) Symbiotic properties of *Rhizobia* containing a flavonoid-independent hybrid *nodD* product. J Bacteriol 171: 4045–4053

Wang S-P, Stacey G (1991) Studies of the *Bradyrhizobium japonicum nodD1* gene promoter: a repeated structure for the *nod* box. J Bacteriol 173: 3356–3365

Zaat SAJ, Wijffelman CA, Spaink HP, Van Brussel AAN, Okker RJH, Lugtenberg BJJ (1987) Induction of the *nodA* promoter of *Rhizobium leguminosarum* sym plasmid pRL1JI by plant flavanones and flavones. J Bacteriol 169: 198–204