Nucleic Acids Research, 2003, Vol. 31, No. 12 ***3143–3156***
DOI: 10.1093/nar/gkg411

Inactivation of the *nod* box distal half-site allows tetrameric NodD to activate *nodA* transcription in an inducer-independent manner

Jie Feng, Qiang Li, Hai-Liang Hu, Xiao-Chun Chen and Guo-Fan Hong*

State Key Laboratory of Molecular Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 320 Yue-Yang Road, Shanghai, 200031, P.R. China

Received February 18, 2003; Revised April 2, 2003; Accepted April 14, 2003

ABSTRACT

In *Rhizobium leguminosarum*, NodD can activate *nodA* transcription in response to inducer flavonoids. Here, we show that the inducible *nodA* promoter contains an intrinsic part through which NodD can activate *nodA* transcription in an inducer-independent manner. Evidence was provided that NodD binds to target DNA through anchoring the two half-sites of the *nod* box as a tetramer. An imperfect inverted repeat AT-N_{10}-GAT was found in each half-site and is critical for NodD binding. Mutation of the inverted repeat of the *nod* box distal half-site allowed NodD to activate *nodA* transcription in an inducer-independent manner *in vivo*, and to modulate the DNA bending of the NodD–*nod* box complex in the absence of inducer *in vitro*.

INTRODUCTION

In nitrate-poor soils, strains of *Azorhizobium*, *Bradyrhizobium*, *Mesorhizobium* and *Rhizobium* (collectively known as rhizobia), form nitrogen-fixing symbiosis with leguminous plants in a host-specific way (1,2). In compatible interactions, invading rhizobia and infected roots differentiate into nitrogen-fixing bacteroid in the nodules. Flavonoids released by legume roots are amongst the first signals exchanged in the molecular dialog between the two symbionts. Through rhizobial regulators of the NodD family, inducer flavonoids trigger the expression of nodulation genes (*nod*, *noe* and *nol*). In turn, most nodulation genes participate in the synthesis or secretion of a family of lipochito-oligosaccharide molecules, the Nod factors, which are required for bacterial entry into root hairs. Thus, NodD is of central importance during the nodulation process between rhizobia and host plants.

Some rhizobia species, e.g. *Bradyrhizobium japonicum*, *Rhizobium* sp. NGR234, *Rhizobium meliloti* and *Rhizobium tropici* harbor two to five copies of *nodD* genes, whereas *Rhizobium leguminosarum* bv. *viciae* and *R.leguminosarum* bv. *trifolii* have only one *nodD* gene (3). In the symbiotic plasmid pRL1JI of *R.leguminosarum* bv. *viciae*, the *nodD* gene is transcribed divergently from the *nodABCIJ* operon (4). Besides the *nodABCIJ* operon, NodD also activates the transcription of all three other identified nodulation operons (*nodFEL*, *nodMNT* and *nodO*) in response to inducer (5–7). NodD autoregulates its own *nodD* operon (8). The minimal inducible *nod* promoters contain two highly conserved DNA elements. One is a highly conserved DNA motif (47 bp *nod* box), and the other is a conserved –10 sequence, which is ~15 bp downstream of the *nod* box consensus (Fig. 1A) (9). The 47-bp *nod* boxes share 32 highly conserved consensus nucleotides. From the comparison of various *nod* box sequences of (brady)rhizobia and *Azorhizobium caulinodans*, Holsters and Goethals have proposed that a so-called NodD box with the panlindromic structure ATC-N_9-GAT is the primary NodD-binding element (10).

NodD shows extensive amino acid similarity with the LysR-type transcriptional regulators (LTTRs) and has properties similar to many LTTRs (11). These proteins have attracted a considerable amount of attention for a number of reasons. First, the LysR family is likely to be the largest family of prokaryotic regulatory proteins (a recent database search revealed over 200 members). Secondly, these proteins appear to function in ways that are somewhat atypical for classical prokaryotic regulators. Although most LTTRs are activated by small specific molecules, these small molecules generally do not greatly affect the binding affinity of the proteins to their target promoters. The regions of DNA protected by these proteins from DNase I digestion are usually 50–60 bp long, much longer than those protected by most other known transcriptional regulators. Mutational studies and analysis of amino acid sequence similarity between LTTRs have identified three basic functional domains: an N-terminal DNA-binding domain, a coinducer recognition/response domain and a conserved C-terminal domain (11,12). The LTTRs bear no obvious activation domains. DNase I footprints show that NodD specifically protects the *nod* box region (13,14). Based on analysis of the symmetry and phasing of the *cis* elements, NodD is predicted to bind to target DNA as a dimer or tetramer (10,14). Genetic evidence indicates that NodD interacts directly with inducer flavonoids to activate transcription (15,16), but direct proof of this is lacking (17). Several lines of evidence support the model that the NodD–*nod* box binding alters upon addition of appropriate flavonoids: these are increased binding to *nod* box sequences in *Sinorhizobium meliloti* and *A.caulinodans* (10,18) and changes in DNase I

*To whom correspondence should be addressed. Tel: +86 21 54921223; Fax: +86 21 54921011; Email: gfhong@sunm.shcnc.ac.cn

Nucleic Acids Research, Vol. 31 No. 12 © Oxford University Press 2003; all rights reserved

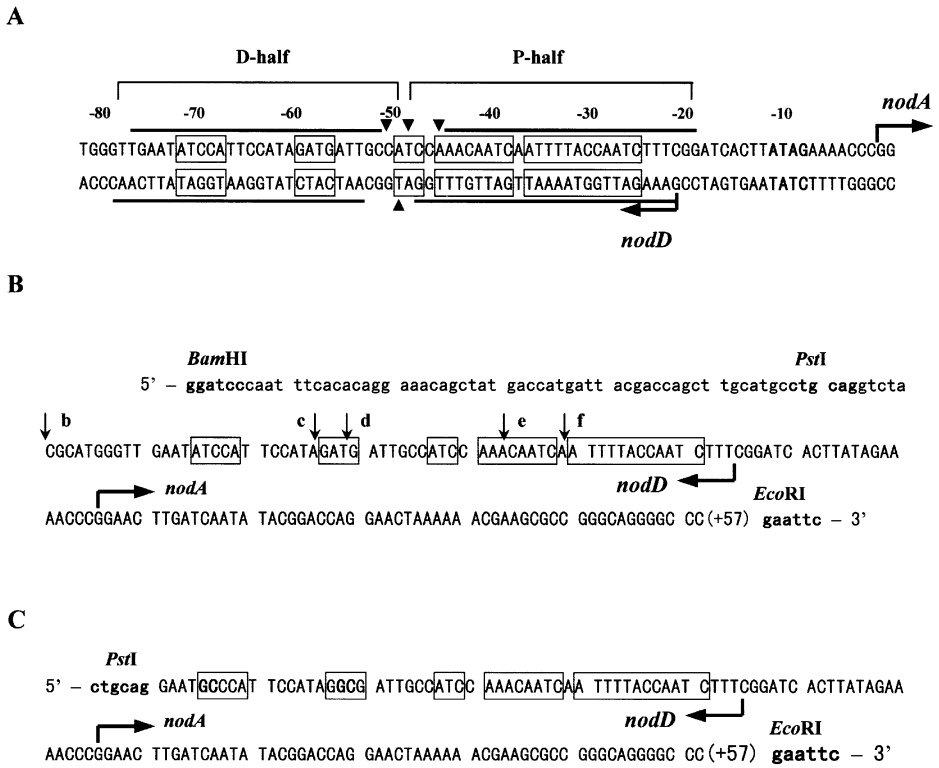

Figure 1. The *nodA* promoter of the symbiotic plasmid pRL1JI. The transcription start sites for the *nodA* and *nodD* genes are indicated by arrows (8,9). The *nodA* transcription start site is numbered +1. The consensus sequence of the canonical *nod* box is boxed (9,14). (**A**) The minimal inducible *nodA* promoter. The conserved sequence ATAG, which is part of the *nodA* promoter, is in bold (9). Lines indicate the DNase I footprints of NodD, and triangular arrowheads denote hypersensitive sites (8). The two protected half-sites of the *nod* box separated by the hypersensitive sites are named D-half and P-half, respectively. (**B**) The *nodA* promoter deletions cloned between the BamHI–EcoRI sites of pBluescriptKSII-derived plasmids pBSb, pBSc, pBSd, pBSe and pBSf (see Materials and Methods). The 5′ ends of the deletions are different and marked by the vertical arrows, and the 3′ ends are all at position +57. The upper case letters represent the *nodA* promoter region, and the lower case letters are related to relevant plasmids. The BamHI, PstI and EcoRI sites are in bold. (**C**) The *nodA* promoter (–75 to +57) with an O13 mutant D-half cloned in the PstI–EcoRI sites of pUO13 (see Materials and Methods). The four mutated residues in the O13 D-half are in bold. The PstI and EcoRI sites are also in bold.

footprints (18). However, in other cases, the affinity and/or binding of NodD for *nod* boxes seem not to be affected (19).

The mechanism by which LTTRs switch on and off their target promoters in response to coinducer is not fully understood. Recently, several clues indicate that LTTRs may regulate transcription through modulating the DNA structure (20–22). For example, *in vitro*, the ligand of OccR octopine can relax the OccR-incited DNA bend (20). The wild-type OxyR-wt causes a sharper DNA bend on the oxySRS promoter than the positive mutant OxyR-C199S, which is locked in the activated conformation (22). The DNA bending by CatR is also inducer CCM-responsive (23). These LTTRs protect their regulated operators in a similar position (approximately –75 to –25), and incite a DNA bend. The protected region is long and can be divided into two halves or subsites. *Sinorhizobium meliloti* NodD also binds to two distinct sites on the same face of the *nod* box helix and induces a bend in the DNA (14). In this work, we demonstrated in *R.leguminosarum* how tetrameric NodD binds to the *nod* box through anchoring its two tandem half-sites. In an effort to understand the role of such double-site interactions between tetrameric NodD and the two *nod* box half-sites in the NodD–*nod* box-mediated transcriptional control, the *nod* box distal half-site (D-half) was mutated and inactivated.

MATERIALS AND METHODS

Microbiological techniques

Bacterial strains and plasmids are listed in Table 1 or in the text. Media and general growth conditions were as described by Hu *et al.* (8). Diparental conjugation was performed to mobilize broad host range plasmids from *Escherichia coli* to *R.leguminosarum* as described by Simon *et al.* (24).

Enzymes and chemicals

Restriction endonucleases and DNA ligase were purchased from Promega. Bio-Rad Protein Assay Kit I based on the Bradford method was from Bio-Rad; [α-^{32}P]dATP was from Amersham; SDS–PAGE low molecular weight protein marker was from Shanghai Lizhu Dongfeng Biotech. Co.; HiFi-Bst DNA polymerase was produced in our own laboratory (25); other chemical reagents were above analytical grade.

Plasmids and DNA fragments

Three short DNA fragments TnodboxT, TDhalfT and TPhalfT were prepared by directly annealing complementary synthesized single-strand oligonucleotides (Table 1). We had previously constructed deletions of the *nodA–nodD* promoter

Table 1. Bacterial strains, plasmids and DNA fragments

	Relevant characteristics	Reference or source
Rhizobium		
8401	*Rhizobium leguminosarum* bv. *viciae* cured of its symbiotic plasmid, Str^r	(44)
Escherichia coli		
S17-1	*294 recA, chrom*, RP4 derivative	(24)
Plasmids		
pKT230	IncQ broad-host-range plasmid; Str^r, Kan^r	(45)
pIJ1518	1.8 kb BclI fragment with *nodD* in pKT230; Kan^r	(4)
pMP220	*lacZ* downstream of a multi-cloning site; IncP broad-host-range plasmid; Tc^r	(9)
pMP221	Opposite multi-cloning site in pMP220	Stored in this laboratory
pUW	*nodA* promoter (–75 to +57) cloned into the EcoRI–PstI sites of pUC119	This study
pUO13	*nodA* promoter (–75 to +57) with the O13 D-half cloned into the EcoRI–PstI sites of pUC119	This study
pMP221O13	EcoRI–PstI fragment cloned from pUO13 into pMP221	This study
Fragments		
TnodboxT	5′-**TT**GAATATCCATTCCATAGATGATTGCCATCCAAACAATCAATTTTACCAATCTTTCGG-3′ 3′-CTTATAGGTAAGGTATCTACTAACGGTAGGTTTGTTAGTTAAAATGGTTAGAAAGCC**TT**-5′	This study
TDhalfT	5′-**TT**GAATATCCATTCCATAGATGATTGCCATACTGACTGACTGACTGACTGACTGACTGA-3′ 3′-CTTATAGGTAAGGTATCTACTAACGGTATGACTGACTGACTGACTGACTGACTGACT**TT**-5′	This study
TPhalfT	5′-**TT**TGCATGCATGCATGCATGCATGCATGATCCAAACAATCAATTTTACCAATCTTTCGG-3′ 3′-ACGTACGTACGTACGTACGTACGTACTAGGTTTGTTAGTTAAAATGGTTAGAAAGCCTT-3′	This study

region from two ends and obtained two series of clones (stored in this laboratory). From five clones of one series, *nodA* promoter deletions (~200 bp) were cloned into pBluescriptIIKS, resulting in plasmids pBSb, pBSc, pBSd, pBSe and pBSf (Fig. 1B). DNA fragments EbB, EcB, EdB, EeB and EfB were the EcoRI–BamHI fragments digested from plasmids pBSb, pBSc, pBSd and pBSe, respectively. Plasmids pMP221b, pMP221c, pMP221d and pMP221e were constructed by cloning the EcoRI–PstI fragments from plasmids pBSb, pBSc, pBSd and pBSe to pMP221, respectively. Plasmids pMP220b, pMP220c, pMP220d and pMP220e were constructed by cloning the EcoRI–PstI fragments from plasmids pBSb, pBSc, pBSd and pBSe to pMP220, respectively. Three primers were synthesized: P1, AGTGAATTCG GGCCCCTGCCCG GCGCTTCG; P2, GCACTGCAGG AATATCCATT CCATAGATG; P3, GCACTGCAGG AAT**GC**CCATT CCATAG**GC**GA TTGCCATCC. With pBSb as template and P1/P2 and P1/P as primer, PCR was used to produce two *nodA* promoter fragments with a wild-type D-half and an O13 mutant D-half. pUO13 was constructed by cloning the *nodA* promoter (–75 to +57) with the O13 D-half into the EcoRI–PstI sites of pUC119. In contrast, pUW was constructed by cloning the *nodA* promoter (–75 to +57) with a wild-type D-half into the EcoRI–PstI sites of pUC119 (Fig. 1C). pMP221O13 and pMP220O13 were constructed by cloning the same *nodA* promoter (–75 to +57) with the O13 D-half into the EcoRI–PstI sites of pMP221 and pMP220, respectively.

Purification of NodD protein and determination of the active NodD concentration

NodD was prepared as described before (26). The concentration of active NodD was determined by measuring the stoichiometry of DNA bound when DNA was in high molar excess over NodD. The first titration test in gel mobility shift assays was used to choose an appropriate concentration of NodD, which was 100-fold that which could shift ~95% of the labeled TnodboxT (10^{-12} M). In the second titration test, the chosen NodD concentration and the labeled TnodboxT concentration (10^{-12} M) were fixed, and unlabeled TnodboxT was added and increased gradually in the reaction system of gel mobility shift assays. When the ratio of shifted fragment to free fragment was decreased to 1/1, the molar concentration of TnodboxT was increased to 2-fold that of the active tetrameric NodD. The molar concentration of active NodD subunit was 4-fold that of tetrameric NodD on this condition.

Gel mobility shift assay and DNase I footprinting

Gel mobility shift assays were performed as described (27). HiFi-Bst DNA polymerase was used to blunt the ends of the annealed fragments or the EcoRI-digested fragments with [α-^{32}P]dATP. In the assays to determine apparent equilibrium association constants, no non-specific competitor DNA was added. In other assays, 200 ng of calf thymus DNA (ctDNA) was added as competitor in the 10 µl reaction volume.

DNase I footprinting reactions were performed as described (27). The A+G ladder was produced by the method of Liu and Hong (28). Plasmids pBSb, pBSc, pBSd and pBSe were digested by EcoRI, the EcoRI end blunted with [α-^{32}P]dATP, and then digested by BamHI. NodD (0.4 ng/µl), radiolabeled DNA fragment (1×10^{-11} M) and ctDNA (20 ng/µl) were incubated at 28°C for 30 min in a 30 µl volume, then 2 µl of DNase I (2 µg/ml, 4 units/µg; Boehringer Mannheim) was added to digest for 30 s. The reaction was terminated, and then extracted with phenol–chloroform. The aqueous phase was precipitated with ethanol, and the samples were analyzed on 6% denaturing polyacrylamide gel.

Binding site selection

The following procedure was according to the SELEX procedure (29). The random oligonucleotides were synthesized: a 72 base oligonucleotide 5′-TTC GAG CTC CAC CGC GGT (A/G)(T/C)C CAT TCC (A/G)(T/C)A (A/G)(A/G)(T/C) GAT TGC C(A/G)(T/C) CCT GCA GGA TCC GGC CGG TAC CCA GCT TGA-3′ (D-N9) with nine residues mutated

and its corresponding primers 5′-GGC GAA TTC GAG CTC CAC CGC GG-3′ and 5′-AGG GAA TTC AAG CTG GGT ACC GGC CG-3′. The second strand of the oligonucleotides D-N9 was synthesized by primer extension in the first PCR cycle, and then the double-stranded product was amplified by PCR (five cycles). The PCR product was digested with EcoRI and radiolabeled with [α-^{32}P]dATP by HiFi-Bst DNA polymerase. Purified NodD (1 ng/μl) was used in gel mobility shift assays to shift or select target oligonucleotides from the others. The position of the shifted fragments was measured by autoradiography. Gel slices were excised from such position of the gel run on the same conditions, and ground in 1.5-ml Eppendorf tubes. TE buffer was added to rinse the selected probe out of the macerated gel by incubating the tube at 37°C for 4 h in a hybridization incubator. The extracted DNA was then amplified by PCR (30 cycles). Gel slices from the same position of the parallel lanes (two lanes apart) without NodD added were taken as the negative control. In the second and third cycles of selection, unlabeled DNA probe (200 ng) was added to make DNA probe molar excessive. The final PCR product was directly cloned into pUCm-T-vector. Twenty-nine such clones sequenced were finished (pUT-DI and pUT-DII). The control oligonucleotides from the preselection pools of D-N9 were also cloned into pUCm-T vector, and the resulting clones were named pUT-O. Twenty such clones were also sequenced (pUT-O).

β-Galactosidase assays

Rhizobium was incubated at 28°C under aeration in TY medium (8) until the A_{600} value increased to 0.5. The β-galactosidase assays were performed as described (30). Naringenin was added to a final concentration of 10 μM as the induction condition. Assays of β-galactosidase activities were performed in triplicate and were reproducible within 15% from experiment to experiment.

DNA bending by circular permutation assay

The EcoRI site of pUW and pUO13 was removed by digesting the plasmids with EcoRI, blunting the EcoRI site with *Pfu* DNA polymerase and then ligating the resulting product. With the resulting plasmids as templates, PCR was used to generate five pairs of 226-bp fragments with the following five pairs of primers: BP1, 5′-TTG AAT ATC CAT TCC ATA GAT GAT TGC CAT CCA AAC-3′ and BP2, 5′-GAA GAA TTC ATG TGC TGC AAG GCG ATT AAG-3′; BP3, 5′-CAG GAA TTC GCT ATG ACC ATG ATT AC-3′ and BP4, 5′-GTT TTC CCA GTC ACG ACG TTG-3′; BP5, 5′-TTG TGT GGA ATT GTG AGC GGA-3′ and BP6, 5′-TCT GAA TTC ATT CGG GCC CCT GCC CGG C-3′; BP7, 5′-TAG GCA CCC CAG GCT TTA CAC TTT A-3′ and BP8, 5′-TTC GAA TTC CTG ATA TTG ATC AAG TTC-3′; BP9, 5′-CGG GAA TTC AGC GCA ACG CAA TTA ATG TG-3′ and BP10, 5′-ATC CGA AAG ATT GGT AAA ATT GAT TGT TTG GAT GGC. The NodD-binding sites are distributed from the ends to the middle of the 226-bp fragments. These fragments were digested by EcoRI and radiolabeled by blunting the ends with [α-^{32}P]dATP to produce five pairs of 222-bp probes. Gel mobility shift assays were performed to detect the mobility shift of the free fragments and the NodD–fragment complexes.

RESULTS

Oligomeric form of NodD at target DNA

The stoichiometry of NodD and target DNA is an important detail to reveal the mechanism underlying the transcriptional regulation of inducible *nod* genes. Fisher and Long have reported that *R.meliloti* NodD binds to the *nodH nod* box through two distinct subsites on the same face of the DNA helix (14). We wished to establish the oligomerization degree of NodD on the intact *R.leguminosarum nodA nod* box and two mutant *nod* boxes with only one half-site.

Three 60-bp DNA fragments, named TnodboxT, TDhalfT and TPhalfT, were prepared by directly annealing synthetic complementary single strain oligonucleotides (Table 1). TnodboxT contains a wild-type *nodA nod* box. TDhalfT contains only the D-half of the *nod* box. TPhalfT contains only the proximal half (P-half) of the *nod* box. The DNA sequence of the other half in TDhalfT and TPhalfT is randomly given.

A method described by Ferguson is often used to determine the molecular weight of proteins through non-denaturing polyacryamide gel electrophoresis. Such a method is also valid to determine the molecular weight of protein–DNA complexes (31). In order to calculate the oligomerization degree of NodD, this method was used to determine the molecular weight of the NodD–DNA complexes. Several different concentrations of gel were used for gel mobility shift assays as well as the electrophoresis of molecular weight standards. After electrophoresis, the standards were located by Coomassie blue staining while the protein–DNA complex was detected by autoradiography. Rf values were determined and the molecular weight of the complex was estimated according to a calibration curve obtained through plotting twice, as in Figure 2. Then, the molecular weight of the protein component of the complex could be determined by subtracting the contribution of DNA from the total value.

A value of 170 kDa was obtained for the NodD–TnodboxT complex (Fig. 2B), 174 kDa for the NodD–TDhalfT (Fig. 2C) and 171 kDa for the NodD–TPhalfT (Fig. 2D). Subtracting 36 kDa to allow for the contribution of the DNA, yields an estimate of 134, 138 and 135 kDa for the molecular weight of the protein component of the NodD–TnodboxT complex, the NodD–TDhalfT complex and the NodD–TPhalfT complex, respectively. The calculated molecular weight is in good agreement with the value expected for binding of the homotetrameric NodD, the molecular weight of the monomer being reported as 34 kDa.

Thus, NodD appears to bind to the two ‘half-boxes’, with one intact half, as well as the intact *nod* box as a tetramer. This suggests that tetramer is the DNA-bound form of NodD at least for the uninduced NodD because the mobility shift assay patterns to generate the data in Figure 2 lack an intermediate band corresponding to dimeric NodD bound to DNA fragments TDhalfT and TPhalfT (data not shown).

Binding affinity of NodD for the intact *nod* box and the two half-boxes

To quantify NodD binding to the intact *nod* box and the two half-boxes, the purified NodD protein was titrated against DNA fragments TnodboxT, TDhalfT and TPhalfT in gel mobility shift assays. These assays were conducted using

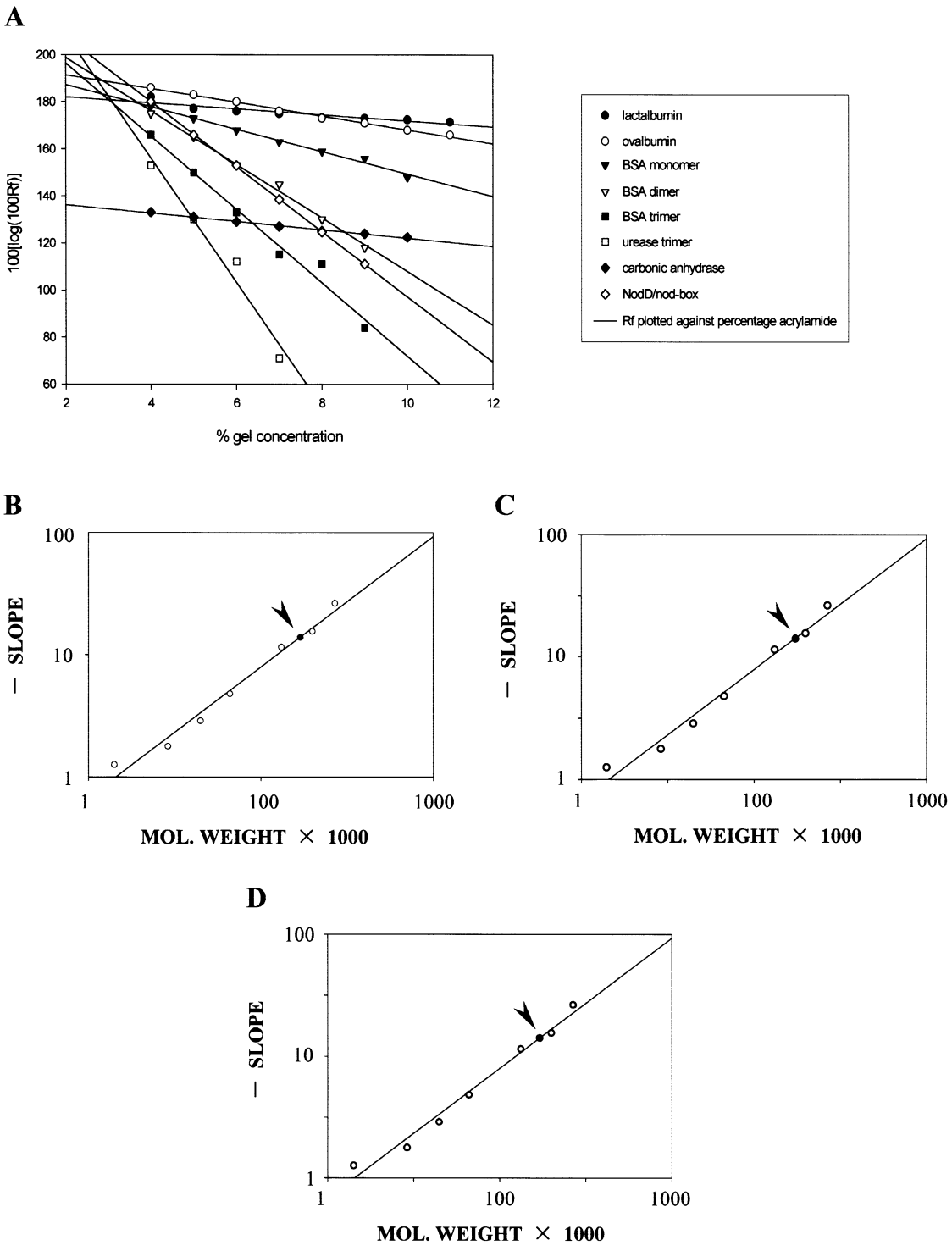

Figure 2. The molecular weight of the NodD–DNA complex NodD–TnodboxT, NodD–TDhalfT and NodD–TPhalfT. (**A**) Calibration curves. Logarithm of relative mobility (Rf) plotted against percentage acrylamide, showing the relationship between each species' mobility and the gel concentration: lactalbumin, carbonic anhydrase, ovalbumin, BSA monomer, BSA dimer, BSA trimer, urease trimer, NodD–TnodboxT. (**B**) Ferguson plot. The gradient of each line in (A) (K_f) plotted against the molecular weight of the standards (open circles), generating a standard curve from which the molecular weight of the NodD–TnodboxT complex (closed circle, also indicated by arrowhead) can be determined. (**C**) Ferguson plot to determine the molecular weight of the NodD–TDhalfT complex. (**D**) Ferguson plot to determine the molecular weight of the NodD–TPhalfT complex.

concentrations of DNA probes (1×10^{-13} M) well below the effective binding dissociation constant for the intact *nod* box to evaluate the apparent equilibrium association constant governing the assembly of NodD–DNA complex. The apparent equilibrium association constant (K_a) was determined from the averages of at least three separate experiments with each

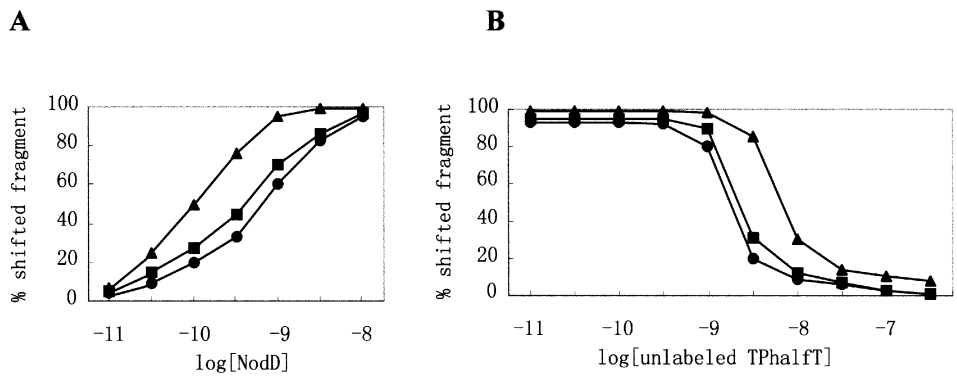

Figure 3. Binding of NodD to DNA fragments TnodboxT, TDhalfT and TPhalfT. (**A**) NodD–DNA binding isotherm curves as a function of the increasing concentration of NodD. The labeled DNA fragments TnodboxT (triangle), TDhalfT (square) and TPhalfT (circle) were added in at a final concentration of 1×10^{-13} M. (**B**) Competition titration of NodD–TnodboxT (triangle), NodD–TDhalfT (square) and NodD–TPhalfT (circle) with added unlabeled TPhalfT. NodD was added at a final concentration of 0.4 ng/μl, and the labeled DNA fragments at 1×10^{-12} M.

fragment. As shown in Figure 3A, K_a for the NodD–TnodboxT interaction, the NodD–TDhalfT interaction and the NodD–TPhalfT interaction was calculated to be 1×10^{10}, 2.4×10^9 and 1.5×10^9, respectively. Binding competition experiments with unlabeled DNA fragment TPhalfT were also performed (Fig. 3B). The competition curves are consistent with the binding curves. In Figure 3 and other cases of this study, the concentration of NodD refers to that of the active NodD subunit. We determined the concentration of the active NodD subunit by measuring the stoichiometry of DNA bound when DNA was in high molar excess over NodD (see Materials and Methods). In Figure 3, the percent shifted fragment refers to the ratio of shifted fragment to the sum of shifted and free fragments.

Specific NodD-binding determinant of the D-half and the P-half

To localize the specific NodD-binding determinant of the *nod* box more precisely, we used *nod* box deletions in gel mobility shift assays. The fragments, named EbB, EcB, EdB, EeB and EfB, were the EcoRI–BamHI fragments digested from plasmids pBSb, pBSc, pBSd, pBSe and pBSf, respectively (Fig. 1B). In gel mobility shift assays, the deletions from the D-half to the P-half twice obviously impaired the NodD-binding affinity (Fig. 4A). The first interval at which there was a significant step down occurred between –85 and –60 (Fig. 4A, lanes 1 and 2), and the second between –42 and –37 (Fig. 4A, lanes 4 and 5). DNA fragment EeB bound to NodD (Fig. 4A, lane 4), indicating that the minimal 18-bp sequence (–42 to –25) contains a basal NodD-binding determinant. Further deletion (–42 to –37) made the shifted band disappear (Fig. 4A, lane 5). DNase I footprinting showed the corresponding specific NodD protection on these fragments. The deletion up to –60 made NodD lose all its specific footprints on the left sequence of the D-half (Fig. 4B, lane 11). The 18-bp sequence (–42 to –25) allowed NodD to form a similar protection pattern on the P-half region (Fig. 4B, lane 9).

We aligned the D-half and the inverted P-half sequences. The consensus sequence is $ATN_3TN_3ATNGATN_2TTN_3A$, and the conserved consensus sequence AT-N_{10}-GAT-N_7-A. In order to test whether these nucleotides are critical for NodD binding, we selected synthetic binding sites from pools of oligonucleotides. We synthesized oligonucleotides carrying nine mutated bases (D-N9) in the context of the D-half sequence (Fig. 5). Synthetic sites bound by NodD were sequestered away from the pool of oligonucleotides by a shift in gel mobility. The bound sequences were then amplified by PCR, and the selection was repeated. In the selection assays, the purified NodD used is predicted to assume uninduced conformation. In D-N9, A is randomly mutated to A or G, T to T or C, and G to A or G.

Two independent selections (DI and DII), each with three rounds of binding and amplification, were carried out with D-N9. In order to select the strongest binding sites, the DNA fragment was molar excessive in the last two rounds of binding. The selected oligonucleotides were cloned into plasmid pUCm-T vector, and the resulting clones were named pUT-DI and pUT-DII. The control oligonucleotides from the preselection pools of D-N9 were also cloned into pUCm-T vector, and the resulting clones were named pUT-O. pUT-DI(1–15), pUT-DII(16–25), pUT-DII(27–30) and pUT-O(1–29) were successfully sequenced. The 29 sequenced selected oligonucleotides were aligned (Fig. 5A). The frequencies of such site-directed mutated nucleotides for the selected oligonucleotides were then compared with those for the oligonucleotides from the D-N9 preselection pool (Fig. 5B). All the six nucleotides of the conserved consensus AT-N_{10}-GAT-N_7-A were clearly biased for the specific NodD binding. Even the 'control' nucleotides A and T of the non-conserved consensus sequence were also biased. The last T belonging to the other half was not biased. In 21 out of 29 sequenced binding sites, AT-N_{10}-GAT was selected, indicating that these nucleotides might be cooperative and more important for NodD binding. pUT-O7 and pUT-O13 were two clones of the sequenced 'control' clones pUT-O(1–29). pUT-O7 contained a D-half sequence T**AT**CCATTCC**GC**A**GAT**-GATTGCC**GC**CC, and pUT-O13 T**GC**CCATTCC**AT**-A**GGC**GATTGCC**AT**CC. In order to testify if the nucleotides at the AT-N_{10}-GAT are critical for NodD binding, the two EcoRI–HindIII fragments of pUT-O7 and pUT-O13 were further used in the gel mobility shift assay. As expected, when the AT at positions 2/3 and 15/16 were both changed to GC, the specific NodD binding was abolished (Fig. 4C, lane

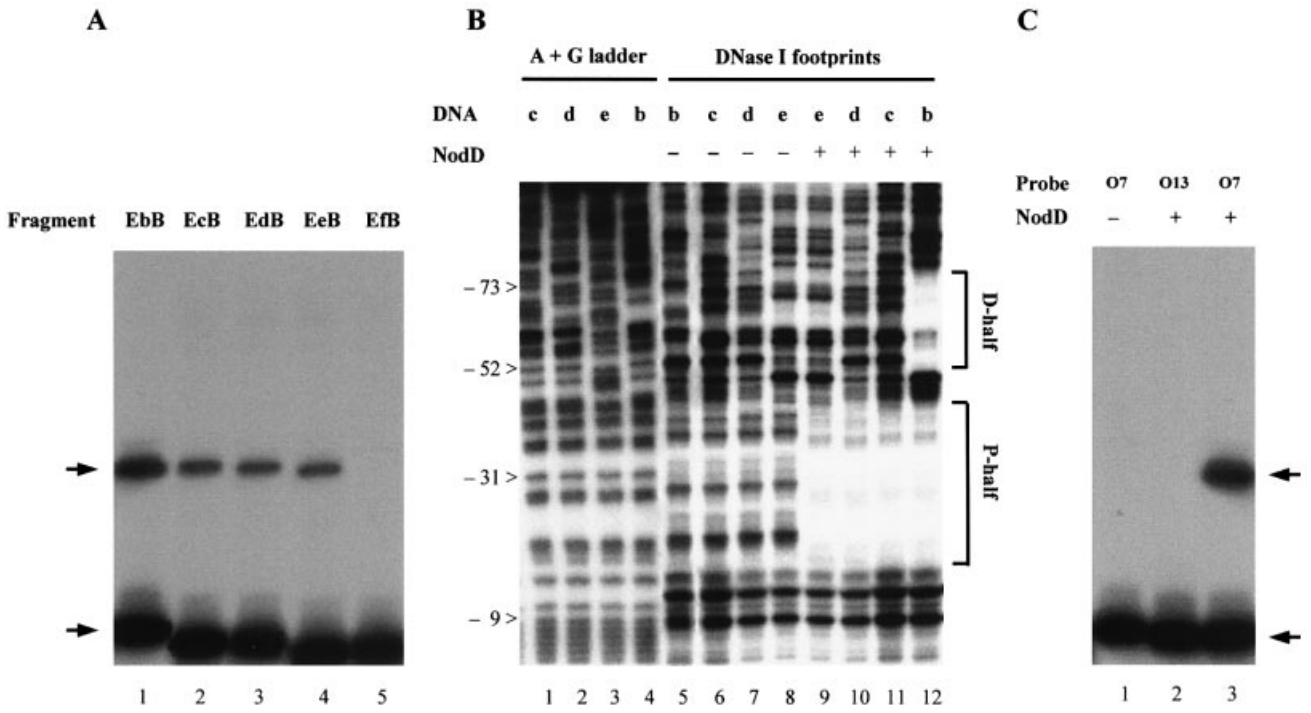

Figure 4. NodD binding to the wild-type *nodA nod* box and the mutants. (**A**) Gel mobility shift assay shows how the deletions impair the NodD binding affinity. Radiolabeled fragments EbB, EcB, EdB, EeB and EfB were used as probes in lanes 1–5, respectively. NodD was added at a final concentration of 4 pg/μl, probe at 1×10^{-12} M and ctDNA at 20 ng/μl. Arrows indicate the free probe band and the shifted band. (**B**) DNase I footprinting shows the 'footprints' of NodD on fragments EbB, EcB, EdB and EeB. The assays were performed on the sense strain of the *nodA* promoter. The sequences are numbered relative to the *nodA* transcriptional start site. The D-half and P-half regions are marked. (**C**) The radiolabeled EcoRI–HindIII fragments from pUT-O7 and pUT-O13 were used as probes in gel mobility shift assay. pUT-O7 contains T**AT**CCATTCC**GC**A**GAT**GATTGCC**GC**CC, and pUT-O13 contains T**GC**CCATTCC**ATAGGC**GATTGCC**AT**CC. NodD was added at a final concentration of 40 pg/μl, probe at 1×10^{-12} M and ctDNA at 20 ng/μl.

2), whereas when the AT at positions 11/12 and 24/25 were both changed to GC, the specific NodD binding was hardly impaired (Fig. 4C, lane 3).

Effects of the D-half inactivation on *nodA* transcription

NodD, as well as many other LTTRs, binds unusually long DNA sequences (8,14,19,20,22,23,32–36). These long DNA sequences almost always contain two protein-binding subsites or half-sites. Through anchoring the two half-sites, these LTTRs can cause a DNA bend on the DNA target. It is an emerging theme that such DNA bending is involved in the transcriptional control by LTTRs. Here, in an effort to understand the role of the double-site interaction between tetrameric NodD and the two half-sites of the *nod* box in the transcriptional regulation and the DNA structure modulation, we chose to inactivate the D-half-site of the *nod* box.

Plasmids pMP221e, pMP221d, pMP221c and pMP221b were constructed by fusing *nodA* promoter deletions to the *lacZ* gene in plasmid pMP221 (see Materials and Methods). In pMP221b, the *nodA* promoter has a wild-type D-half. In pMP221c and pMP221d, the D-half is partially deleted. In pMP221e, the *nodA* promoter is further deleted to position –43. From the above results, we had known that the O13 mutant D-half was inactive for the specific NodD binding (Fig. 4C, lane 2). Thus, a *nodA* promoter derivative (–75 to +57) with an O13 mutant D-half was also cloned into pMP221 to construct pMP221O13 (see Materials and Methods). These *nodA* promoter derivatives, named pr.*nodA*e, pr.*nodA*d pr.*nodA*c, pr.*nodA*b and pr.*nodA*O13, were tested in 8401(pKT230) and 8401(pIJ1518) for induction. *Rhizobium leguminosarum* strain 8401(pKT230) is a *nodD*$^-$ strain, while 8401(pIJ1518) is a *nodD*$^+$ strain (Table 1).

The results are shown in Figure 6. In 8401(pKT230), all of the *nodA* promoter derivatives had very low promoter activity (~35 units) in the absence and presence of inducer 10 μM naringenin. In 8401(pIJ1518), in the presence of inducer, pr.*nodA*b was activated (~1600 units), and pr.*nodA*c, pr.*nodA*d and pr.*nodA*O13 were also activated (~285 units). pr.*nodA*c, pr.*nodA*d and pr.*nodA*O13 were only partially activated by NodD when compared to pr.*nodA*b. In 8401(pIJ1518), in the absence of inducer, it is intriguing that pr.*nodA*c, pr.*nodA*d and pr.*nodA*O13 were still similarly activated (~285 units). In contrast, pr.*nodA*b was switched off (~35 units). On all the conditions, pr.*nodA*e had very low promoter activity (~35 units), which indicated that the 5′ deletion to position –43 might start to destroy the RNA polymerase recruiting site of the 'real' deficient *nodA* promoter.

Considering the above deletions are from position +65 to position +23 relative to the *nodD* transcriptional start site, we also determined their effects on *nodD* transcription. The same PstI–EcoRI fragments were cloned from plasmids pBSb, pBSc, pBSd and pBSe into pMP220, which has a reverse multi-cloning site to that of pMP221. In 8401(pKT230), the β-galactosidase level of pr.*nodD*b, pr.*nodD*c, pr.*nodD*d, pr.*nodD*e and pr.*nodD*O13 varied between 760 and 990 units. The results indicate that 5′ deletions up to position +23 have

A

D-half	T[A T]CCATTCC A T A[G A T]GATTGCC[A T]CC
Pools of oligonucleotides	T[(A/G) (C/T)]CCATTCC(A/G) (C/T) A[(A/G) (A/G) (C/T)]GATTGCC[(A/G) (C/T)]CC

DI1	TACCCATTCCATAGATGATTGCCACCC	**DII16**	TGTCCATTCCATAGGCGATTGCCGCCC
DI2	TATCCATTCCATAGATGATTGCCACCC	**DII17**	TATCCATTCCATAGATGATTGCCACCC
DI3	TATCCATTCCGTAGATGATTGCCACCC	**DII18**	TACCCATTCCATAAGTGATTGCCACCC
DI4	TATCCATTCCGTAGATGATTGCCACCC	**DII19**	TATCCATTCCGTAGATGATTGCCGCCC
DI5	TATCCATTCCGCAGATGATTGCCATCC	**DII20**	TATCCATTCCACAGATGATTGCCGCCC
DI6	TATCCATTCCACAGATGATTGCCATCC	**DII21**	TGTCCATTCCGCAAGCGATTGCCATCC
DI7	TATCCATTCCATAGATGATTGCCCGCC	**DII22**	TATCCATTCCATAGATGATTGCCATCC
DI8	TATCCATTCCGTAGATGATTGCCATCC	**DII23**	TATCCATTCCATAGATGATTGCCATCC
DI9	TATCCATTCCATAGATGATTGCCACCC	**DII24**	TGTCCATTCCGTAGATGATTGCCGTCC
DI10	TATCCATTCCGTAGATGATTGCCATCC	**DII25**	TATCCATTCCGTAGATGATTGCCGCCC
DI11	TACCCATTCCACAGACGATTGCCGTCC		
DI12	TACCCATTCCATAAGCGATTGCCATCC	**DII27**	TATCCATTCCATAGATGATTGCCGTCC
DI13	TATCCATTCCATAGATGATTGCCATCC	**DII28**	TATCCATTCCATAGATGATTGCCACCC
DI14	TATCCATTCCATAGATGATTGCCGTCC	**DII29**	TATCCATTCCGTAGATGATTGCCACCC
DI15	TATCCATTCCGTAGATGATTGCCATCC	**DII30**	TACCCATTCCATAAATGATTGCCATCC

B

```
A     26 /          18 /   3 25 /          20 /
G      3 /          11 /  26  4 /           9 /
T      / 24          / 24  /  / 25           / 15
C      /  5          /  5  /  /  4           / 14

    T A T C C A T T C C A T A G A T G A T T G C C A T C C

A     14 /          16 /  15 15 /          13 /
G     15 /          13 /  14 14 /          16 /
T      / 16          / 14  /  / 13           / 14
C      / 13          / 15  /  / 17           / 15
```

Figure 5. NodD-binding sites selected from pools of oligonucleotides carrying nine mutated nucleotides in the context of the D-half sequence. (**A**) Sequences of 29 selected probes bound specifically by NodD. The wild-type D-half sequence context is given in the first line, with the consensus conserved-sequence boxed. The mutated nucleotides are in bold: nucleotide A was randomly mutated to A or G, nucleotide T was mutated to T or C, and nucleotide G to A or G in a 1:1 ratio. Pools of 512 species of mutant oligonucleotides are denoted in the second line. The oligonucleotides selected from two independent selections (DI and DII) were aligned, with the mutant nucleotides in bold. (**B**) Summary of frequencies of nucleotide residues for selected (top bases) and unselected (bottom bases) oligonucleotides. The nucleotides of the consensus AT-N10-GAT-N7-A are clearly biased for NodD binding. Even the 'control' nucleotides A and T of the non-conserved consensus sequence are also biased. Out of 29 selected oligonucleotides, there are 21 with the sequence AT-N10-GAT selected.

not destroyed the constitutive *nodD* promoter. In 8401(pIJ1518), the β-galactosidase levels of pr.*nodD*c, pr.*nodD*d, pr.*nodD*e and pr.*nodD*O13 did not obviously change, and that of pr.*nodD*b was reduced to 120 units. The results indicate that the *nodD* negative autoregulation needs an intact *nod* box.

In conclusion, the inducible *nodA* promoter in essence contains an intrinsic part (approximately +1 to –56) through which NodD can activate or at least partially activate *nodA* transcription in an inducer-independent manner. Such an intrinsic part contains only one intact half-site of the *nod* box, the P-half. It is worth highlighting that on this condition the uninduced NodD is also able to activate *nodA* transcription. In the absence of inducer, a functional D-half appears essential for NodD to intrinsically repress such NodD-mediated activation. Inactivation of the D-half through partial deletion or site-directed mutation of the AT-N_{10}-GAT can abolish such intrinsic repression.

Effect of the O13 mutant D-half on the tetrameric NodD-induced DNA bending

From the above results, we knew that the O13 mutant D-half was inactive for its specific NodD binding, and that the O13 mutant D-half allowed tetrameric NodD to activate *nodA* transcription constitutively. Then, we were interested to see if the D-half inactivation and the resulting NodD-mediated transcriptional activation occurred with a DNA structural modulation. Circular permutation assays were performed to detect the modulation of DNA bend. These assays are based upon the observation that a bend at the middle of a DNA

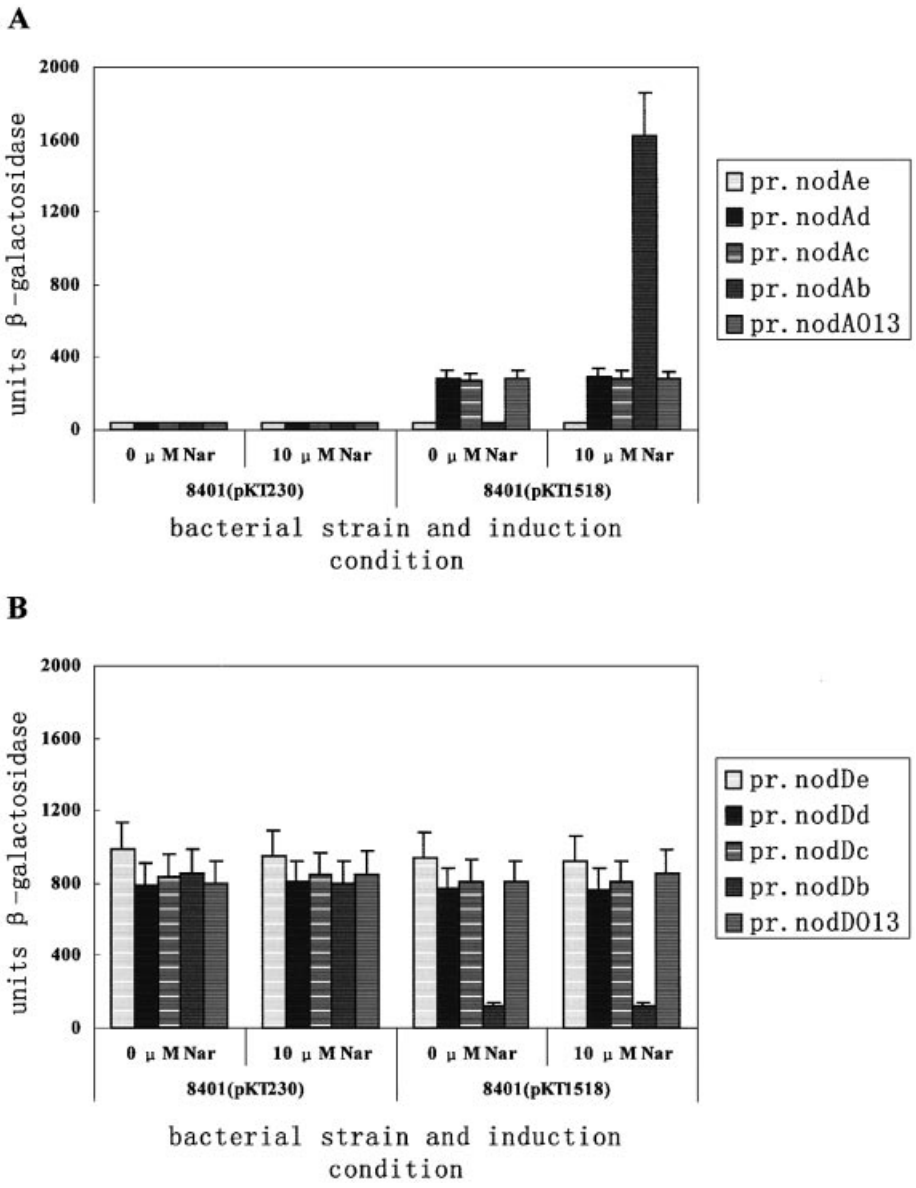

Figure 6. Inactivation of the *nod* box D-half allows NodD to partially activate *nodA* transcription in a naringenin-independent manner. The height of the bars indicates the measured β-galactosidase level of various promoter derivatives. (**A**) pr.*nodA*e, pr.*nodA*d, pr.*nodA*c, pr.*nodA*b and pr.*nodA*O13 were assayed for *nodA* promoter activity in *R.leguminosarum* 8401(pKT230) and 8401(pIJ1518) in the presence and absence of 10 μM naringenin. (**B**) pr.*nodD*e, pr.*nodD*d, pr.*nodD*c, pr.*nodD*b and pr.*nodD*O13 were assayed for *nodD* promoter activity in *R.leguminosarum* strain 8401(pKT230) and 8401(pIJ1518) in the presence and absence of 10 μM naringenin.

molecule affects the mobility of that fragment more severely than the same bend at the end of the molecule (37). Five pairs of 226-bp fragments were produced by PCR. The *nod* box region was distributed from the ends to the middle of these DNA fragments. The only divergence between every pair of DNA fragments was that four base pairs of the wild-type D-half are mutated as those of the O13 mutant D-half.

The NodD–DNA complexes that were formed exhibited a position-dependent mobility (Fig. 7). Such large alternations in mobility are generally interpreted as being due to a DNA bend (37). Complexes formed between NodD and wild-type *nod* box fragments showed a weaker position-dependent mobility (compare lanes 1–5 to lanes 6–10, Fig. 7A), indicating the O13 mutant D-half allows NodD to cause a sharper DNA bend at the *nod* box region. The free 222-bp DNA fragments did not exhibit any position-dependent mobility (data not shown), indicating that all the created sequences do not have a significant intrinsic sequence-directed curvature. The position of the bend can be found by identifying the slowest migrating complex since such a complex will have the bend at the middle of its DNA fragment (37). Therefore, the distance migrated was plotted against the number of bases between the fragment left end and the –49 A of the *nod* box. The mobility distances of the complexes were determined by measuring the distance traveled from the well during the electrophoresis. The slowest migrating band was found to have its center between the D-half and the P-half (Fig. 7B). The bend angles can be estimated by using the formula

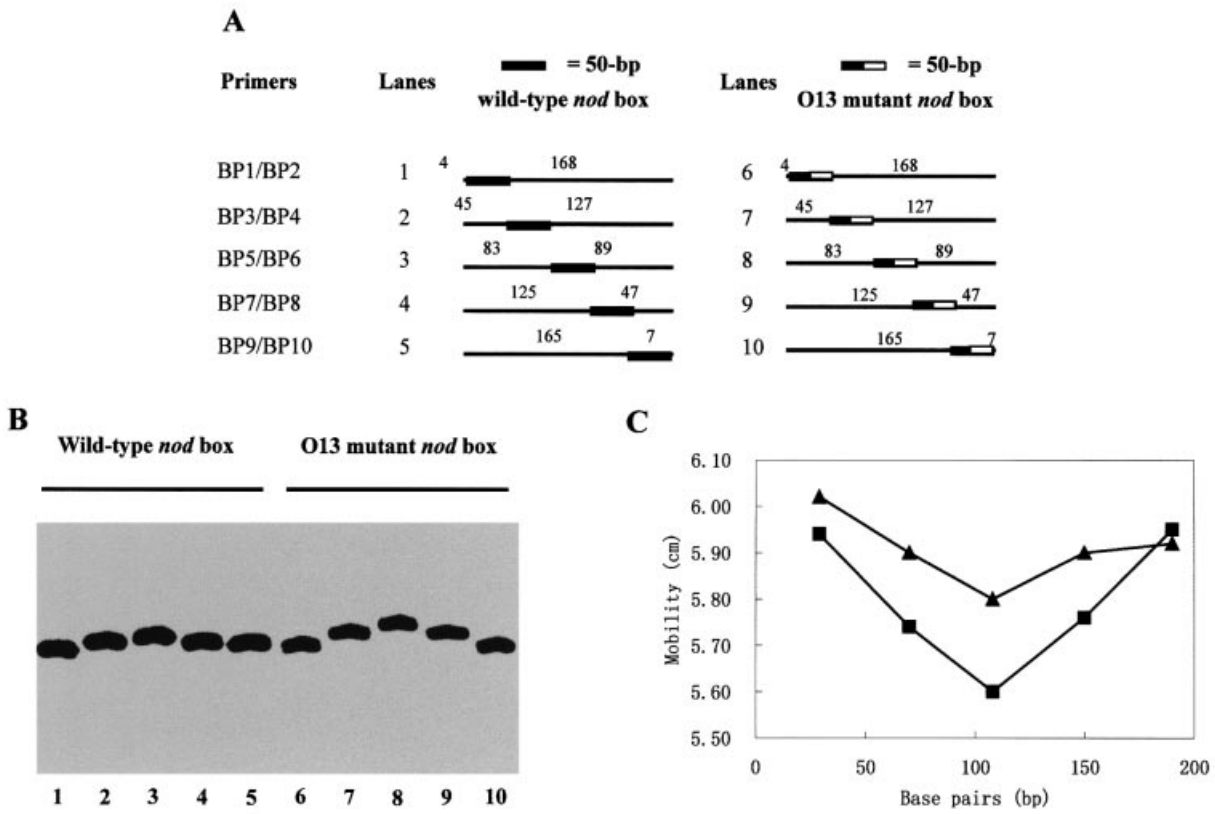

Figure 7. Determination of NodD-induced DNA bending on the wild-type *nod* box and the O13 mutant *nod* box by circular permutation. (**A**) Five pairs of 222-bp radiolabeled DNA fragments used in (B) indicating the relative position of the NodD-binding site in relation to the ends of the fragments. (**B**) The mobility of NodD–DNA complexes using the DNA fragments listed in (A). Gel mobility shift assays were done with 1 pM DNA (lanes 1–10), 20 ng/μl ctDNA (lanes 1–10), 4 pg/μl (lanes 1–5) and 40 pg/μl (lanes 6–10) NodD. (**C**) Graphical representation of NodD-induced DNA bending on the wild-type *nodA nod* box (triangle) and the O13 mutant *nod* box (square). The mobility (in cm) of the NodD–DNA complexes is plotted against the number of bases between the left end of each fragment and the midpoint of the 50 bp *nod* box region. The bending angle by which DNA deflects from linearity is measured as described in Materials and Methods. The complexes with the intact D-half are calculated to have a 30° bend, while the complexes with the mutant D-half have a 44° bend according to an experiential formula.

$\mu_m/\mu_e = \cos(\alpha/2)$, where μ_m is the mobility of the protein–DNA complex with a bend at the center of the DNA fragment, and μ_e is the mobility of the protein–DNA complex with a bend at the end of the DNA fragment (38). This formula would predict that complexes with the intact D-half have a 30° bend, while complexes with the mutant D-half have a 44° bend (Fig. 7). However, the fastest migrating fragment has its bend site rather far away from either end (~30 bp). Therefore, assuming that this formula is valid for the gel system used in this study, these calculated bend angles probably underestimate the true values.

DISCUSSION

Tetrameric NodD binds to promoter through anchoring the two half-sites of *nod* box

We propose that the most important nucleotides for the specific NodD binding are located in a 2-fold inverted repeat AT-N10-GAT-N16-ATC-N10-AT. This proposition has derived from the analysis of the symmetry of the NodD footprints on the *nod* box (Fig. 8A), the NodD affinity of various *nod* box deletions (Fig. 4) and the *in vitro* selection experiments (Fig. 5). Similarly, Toledano *et al.* have reported that the oxidized form of another LTTR OxyR also recognizes a 2-fold inverted repeat ATAGntnnnanCTATnnnnnnn-ATAGntnnnanCTAT *in vitro* (22). Such a proposed binding motif gives a clue to the general binding motif of the LysR family. In fact, the unusually long DNA targets of many LTTRs share common features in length, sequence and even location (19,20,22,23,32–36). These target sites can usually be divided into two binding subsites or half-sites, and are often located at a similar position, from approximately –75 to –25 relative to the transcriptional start site of LTTR-activated promoters. The D-half-site (–75 to –50) almost always consists of an imperfect inverted repeat, which often shares the common LysR motif T-N11-A (10). The P-half (–50 to –25) is superimposed on the –35 region for RNA polymerase (33–36). The potential inverted repeat in this half-site is usually not obvious, likely due to possible dual roles of this region as target sites for both NodD and RNA polymerase. Besides the conserved nucleotides in the proposed motif, the other nucleotides of the *nod* box region also appear to be required for the natural optimal NodD binding, but not critical to a basal NodD binding. Supporting evidence comes from our *in vitro* selection experiments. The four mutated A and T nucleotides, which do not belong to the inverted repeat AT-N10-GAT, are also biased (Fig. 5). However, mutations of

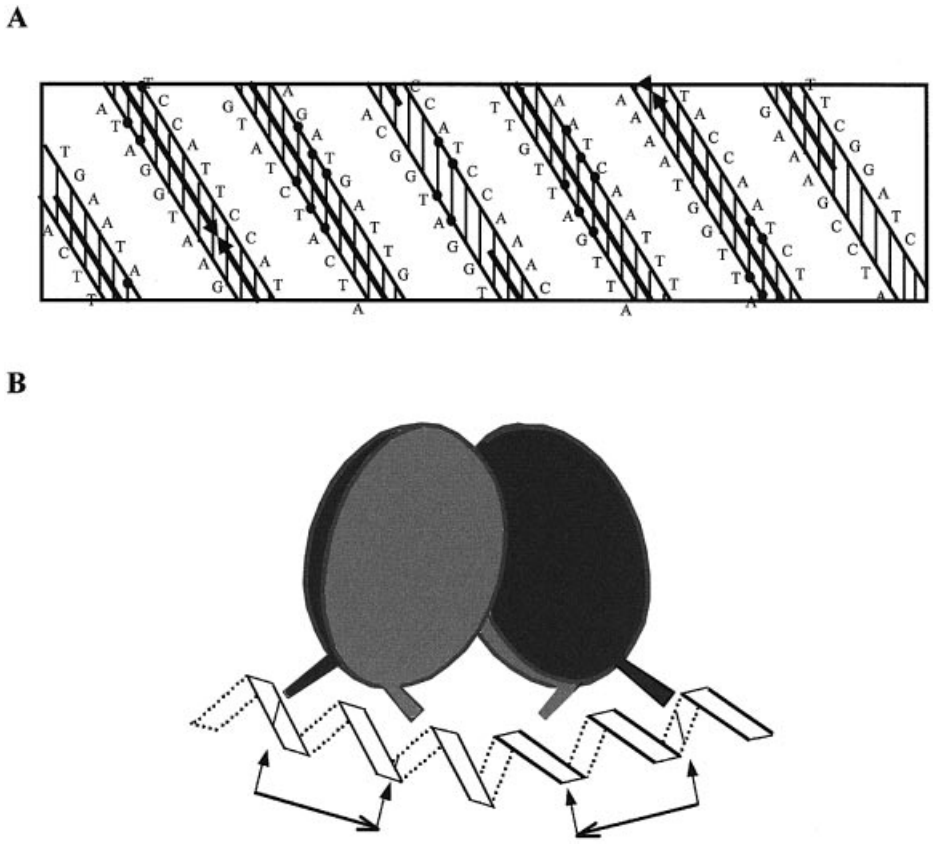

Figure 8. Model for tetrameric NodD binding on the *nod* box region. (**A**) Planar representation of the *nodA nod* box sequence. The positions of the residues are projected onto the surface of a cylinder that is then unrolled onto a flat surface. The DNA is assumed to adopt the B-form conformation, 10.4 bp per helical turn. The solid arrows indicate the regions protected from DNase I digestion. Closed circles denote the consensus sequence of the D-half (left) and P-half (right). (**B**) Tetrameric NodD assumes a cyclically symmetric dimer of NodD dimers. Each dimer is also cyclically symmetric and binds to one half-site with the basic motif AT-N_{10}-GAT. Arrows and short lines indicate the exact position of the binding motif T-N_{11}-A-N18-T-N_{11}-A.

the three out of four nucleotides at the same time do not obviously impair NodD binding (Fig. 4C, lane 3).

In Figure 8B, we propose a model for the arrangement of the NodD–*nod* box complex. In this model, tetrameric NodD is postulated to be a cyclically symmetric dimer of NodD dimers. Each NodD dimer is also suggested to be cyclically symmetric and bind to one inverted half-site with the central major groove unprotected. Many other LTTRs such as CysB, OxyR, TrpI and NahR are also reported to be tetrameric (12,22,39). The crystal structure of a dimeric CysB (88–324) has been solved (39). The dimer itself is cyclically symmetric, and the tetramer is proposed to be a dimer of dimers. We failed to detect the hypothetical NodD dimer band either by adding naringenin *in vitro* to the purified NodD or using NodD crude extract from induced *Rhizobium* cells (data not shown). This indicates that tetramer may be a functional unit of NodD at target DNA. The proposition that tetrameric NodD assumes a cyclic symmetric dimer of NodD dimers is consistent with the suggested 2-fold inverted repeat of the DNA binding motif.

In addition, we find that the two sub-halves of each half-site AT-N10-GAT bind to NodD in a highly cooperative manner. In DNase I digestion, NodD loses its specific protection on the left D-half sequence when one half of the D-half is deleted (Fig. 5B). Consistently, *nod* box deletions and its concomitant NodD affinity changes also support such a conclusion (Fig. 5A).

The inducible *nodA* promoter contains a NodD-dependent constitutive 'promoter'

It is usually thought that only the activated form or coinducer-bound form of positive transcriptional factors can rescue the deficiency of those inducible promoters (40). However, it is intriguing that we find that the inducible *nodA* promoter contains an intrinsic part through which NodD can activate *nodA* transcription in an inducer-independent manner (Fig. 7). Rather, the uninduced tetrameric NodD can also activate *nodA* transcription at least partially. This is a challenge to the classical point of view. The reason why such activation is partial may lie in that pr.*nodA*c, pr.*nodA*d and pr.*nodA*O13 have weaker NodD affinity than pr.*nodA*b. Figure 3 shows that without a specific D-half, the *nod* box region has an ~7-fold lower NodD affinity. However, the possibility still cannot be excluded that the NodD–pr.*nodA* complex with the mutant D-half is only partially competent for *nodA* transcription.

In another perspective, we suggest that in the absence of coinducer, an intact D-half is required for NodD to intrinsically repress the NodD-mediated partial activation (Fig. 6). Such intrinsic transcriptional repression requires the intact D-half or perhaps only the specific NodD binding ability of the D-half. Most recently, we have found that such intrinsic repression is dependent on the distance between the D-half and P-half. When the –54 nucleotide T was resected, NodD,

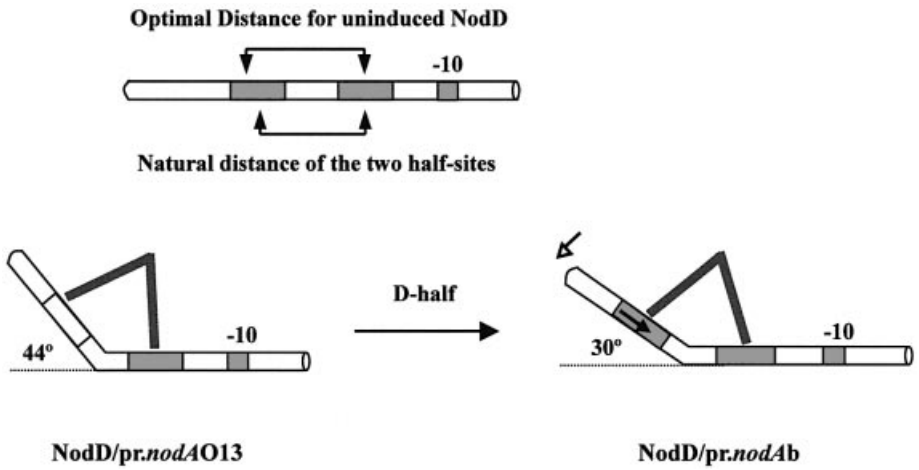

Figure 9. Model of D-half-relevant NodD-induced DNA bending. The –10 conserved sequence and the two half-sites are indicated by shadowed rectangles. Tetrameric NodD is simply represented by two converging lines. Combination of steric effect and the O13 D-half inactivation allows the NodD–*nod* box complex to adopt an optimal DNA bend. The preferential binding to the specific wild-type D-half forces NodD dimer to shift its contact position (marked by the solid arrow), resulting in a DNA bend modulation.

whether induced or not, completely repressed the transcription. When a nucleotide G was inserted to position –54, NodD could not completely repress the transcription in the absence of inducer. In the control experiment, in which the D-half (–75 to –54) was entirely resected, NodD, whether induced or not, partially activated the transcription (J. Feng, X.-C. Chen, B.-H. Hou, F. Q. Li and G.-F. Hong, unpublished data). Insight into the mechanism underlying such intrinsic D-half-relevant transcriptional repression cannot directly explain but does give a clue to the induction mechanism for the wild-type *nodA* promoter.

DNA-bend modulation in the D-half-relevant transcriptional repression

In the absence of inducer, tetrameric NodD induces a sharper DNA bend on pr.*nodA*O13 than on the wild-type *nodA* promoter (Fig. 7). On pr.*nodA*O13 and pr.*nodA*b, tetrameric NodD-induced DNA bending is similar in position, which is near the midpoint of the *nod* box region (Figs 2 and 7). Since the D-half in pr.*nodA*O13 is inactive for its specific NodD anchoring (Fig. 5), the P-half is expected to decide the orientation of tetrameric NodD alone. Wang and Winans have demonstrated that the D-half-site of a natural OccR binding site can itself also decide the orientation of tetrameric OccR (21). Though the O13 mutant D-half does not change the primary structure of the NodD–*nod* box complex and may also not destroy the 'real' deficient *nodA* promoter, it does cause a large DNA conformational change, which may account for the NodD-mediated activation.

Several lines of evidence indicate that activator proteins can act as switch factors as a result of their DNA bending properties. The only effect cAMP·CRP is known to have on the DNA binding site is inducing a bend from linearity (37). An intrinsic DNA curvature is universally found in the 5′-upstream regions of the *psbA* family, and is important for basal transcription (42). Transcriptional factor MerR can relax or twist DNA upon ligand mercuric ion binding (43,44). Recent progress on the mechanism of transcriptional regulation controlled by LTTRs argues that the induction mechanism correlates with DNA bending (20,22). IlvY mediates transcriptional regulation in a DNA supercoiling-dependent manner (41).

Fisher and Long have shown that *in vitro* addition of flavonoids does not change the DNase I footprints of NodD (19). In all our attempts, we have failed to detect the hypothetical binding of NodD dimer. Therefore, it is less likely that NodD regulates *nodA* transcription through oligomerization. Most likely, tetrameric NodD only changes its conformation in response to inducer. However, since both induced and uninduced NodD can similarly activate the *nodA* promoter derivatives pr.*nodA*c, pr.*nodA*d and pr.*nodA*O13, it is hard to imagine that NodD conformational change itself serves as the induction trigger.

In Figure 9, we suggest a model to understand how the D-half inactivation allows uninduced NodD to induce a sharper DNA bend. It is necessary first to highlight the importance of the non-sequence-specific affinity between NodD and DNA in the NodD binding to target DNA (21). Though such non-sequence-specific binding does not leave 'footprints' on the additional vector DNA in place of the missing half region (Fig. 4), it does allow NodD to form a weak protection on a weakened digestion condition with a lower concentration of DNase I and especially with a shorter digestion time (15 s) (data not shown). Such non-sequence-specific binding for OccR has been reported to cause obvious protection on the vector DNA in DNase I digestion (21). In our model, DNA is hypothesized to have an optimal distance to interact with uninduced NodD. The natural distance between the two NodD-binding half-sites is slightly shorter than the optimal distance. For the O13 mutant *nod* box, one dimer of NodD is fixed to the P-half while the other can move freely on DNA. However, the steric effect eventually forces that dimer to move to its optimal position, allowing the NodD–*nod* box complex to assume an optimal DNA bend ($\Delta G_P = -RT \ln K_P = -52.4$ kJ/mol). For the wild-type *nod* box, the preferential binding to the specific D-half allows NodD dimer to shift its

binding site to further reduce the Gibbs free energy of the whole complex ($\Delta G_W = -RT \ln K_W = -57.1$ kJ/mol). Such a shift will result in a DNA bend modulation considering the rigidity of NodD protein. Part of the Gibbs free energy might be stored in the DNA bend modulation. We suggest that the conformational alternation of the NodD–*nod* box complex, especially the promoter DNA part, accounts for the intrinsic D-half-relevant transcriptional repression. The altered conformation is proposed to be not competent for RNA polymerase any more. Our model indicates that the physiological advantage of the long target site of LTTRs might be that LTTRs could transfer the conformational change from protein to DNA through bending or twisting DNA in response to inducer. Tetrameric NodD might decrease the converging angle of its two dimers to shorten its optimal distance in response to inducer.

ACKNOWLEDGEMENTS

We are grateful to Dr Feng-Qing Li and Bo Zhou for helpful discussion in this work. This work was supported by Pan-Deng Plan of China to G.-F.H.

REFERENCES

1. Schultze,M. and Kondorosi,A. (1998) Regulation of symbiotic root nodule development. *Annu. Rev. Genet.*, **32**, 33–57.
2. Broughton,W.J., Jabbouri,S. and Perret,X. (2000) Keys to symbiotic harmony. *J. Bacteriol.*, **182**, 5641–5652.
3. van Rhijn,P.J., Feys,B., Verreth,C. and Vanderleyden,J. (1993) Multiple copies of nodD in *Rhizobium tropici* CIAT899 and BR816. *J. Bacteriol.*, **175**, 438–447.
4. Rossen,L., Shearman,A.W.B. and Downie,J.A. (1985) The *nodD* gene of *Rhizobium leguminosarum* is autoregulatory and in the presence of plant exudate induces the *nodA*, *B*, *C* genes. *EMBO J.*, **4**, 3369–3373.
5. Shearman,C.A., Rossen,L., Johnston,A.W.B. and Downie,J.A. (1986) The *Rhizobium leguminosarum* nodulation gene *nodF* encodes a polypeptide similar to acyl-carrier protein and is regulated by NodD plus a factor in pea root exudate. *EMBO J.*, **5**, 647–652.
6. Surin,B.P. and Downie,J.A. (1988) Characterization of the *Rhizobium leguminosarum* genes *nodLMN* involved in efficient host-specific nodulation. *Mol. Microbiol.*, **2**, 173–183.
7. Maagd,R.A.D., Wijfjies,A.H.M. and Lugtenberg,B.J.J. (1989) *nodO*, a new *nod* gene of the *Rhizobium leguminosarum Biovar viciae* Sym plasmid pRL1JI, encodes a secreted protein. *J. Bacteriol.*, **171**, 6764–6770.
8. Hu,H.L., Liu,S.T., Yang,Y., Chang,W.Z. and Hong,G.F. (2000) In *Rhizobium leguminosarum*, NodD represses its own transcription by competing with RNA polymerase for binding sites. *Nucleic Acids Res.*, **28**, 2784–2793.
9. Spaink,H.P., Okker,R.J.H., Wijffelman,C.A., Pees,E. and Lugtenberg,B.J.J. (1987) Promoters in the nodulation region of the *Rhizobium leguminosarum* Sym plasmid pRL1JI. *Plant Mol. Biol.*, **9**, 27–39.
10. Goethals,K., van Montagu,M. and Holsters,M. (1992) Conserved motifs in a divergent *nod* box of *Azorhizobiuim caulinodans* ORS571 reveal a common structure in promoters regulated by the LysR-type protein. *Proc. Natl Acad. Sci. USA*, **89**, 1646–1650.
11. Schell,M.A. (1993) Molecular biology of the LysR family of transcriptional regulators. *Annu. Rev. Microbiol.*, **47**, 597–626.
12. Lochowska,A. and Hryniewicz,M.M. (2001) Functional dissection of the LysR-type CysB transcriptional regulator. *J. Biol. Chem.*, **276**, 2098–2107.
13. Hong,G.F., Burn,J.E. and Johnston,A.W.B. (1987) Evidence that DNA involved in the expression of nodulation (*nod*) genes in *Rhizobium leguminosarum* binds to the product of the regulatory gene *nodD*. *Nucleic Acids Res.*, **15**, 9677–9690.
14. Fisher,R.F. and Long,S.R. (1993) Interaction 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.
15. Spaink,H.P., Wijffelman,C.A., Pees,E., Okker,R.J.H. and Lugtenberg,B.J.J. (1987) *Rhizobium* nodulation gene *nodD* as a determinant of host specificity. *Nature*, **328**, 337–339.
16. Burn,J.E., Wootton,H.J.C. and Johnston,A.W.B. (1989) Single and multiple mutations affecting properties of the regulatory gene *nodD* of *Rhizobium*. *Mol. Microbiol.*, **3**, 1567–1577.
17. Schlaman,H.R.M., Phillips,D.A. and Kondorosi,E. (1998) Genetic organization and transcriptional regulation of rhizobial nodulation genes. In Spaink,H.P., Kondorosi,A. and Hooykaas,P.J.J. (eds), *The Rhizobiaceae*. Kluwer Academic Publishers, Dordrecht, Boston, London, pp. 361–386.
18. Kondorosi,E., Gyuris,J., Schmidt,J., John,M., Duda,E., Hoffman,B., Schell,J. and Kondorosi,A. (1989) Positive and negative control of *nod* gene expression in *Rhizobium meliloti* is required for optimal nodulation. *EMBO J.*, **8**, 1331–1340.
19. Fisher,R.F. and Long,S.R. (1989) DNA footprinting analysis of the transcriptional activator proteins NodD1 and NodD3 on inducible *nod* gene promoters. *J. Bacteriol.*, **171**, 5492–5502.
20. Wang,L., Helmann,J.D. and Winans,S.C. (1992) The *A. tumefaciens* transcriptional activator OccR causes a bend at a target promoter, which is partially relaxed by a plant tumor metabolite. *Cell*, **69**, 659–667.
21. Wang,L. and Winans,S.C. (1995) The sixty nucleotide OccR operator contains a subsite essential and sufficient for OccR binding and a second subsite required for ligand-responsive DNA bending. *J. Mol. Biol.*, **253**, 691–702.
22. Toledano,M.B., Kullik,I., Trinh,F., Baird,P.T., Schneider,T.D. and Storz,G. (1994) Redox-dependent shift of OxyR-DNA contacts along an extended DNA-binding site: a mechanism for differential promoter selection. *Cell*, **78**, 897–909.
23. Mcfall,S.M., Klem,T.J., Fujita,N., Ishihama,A. and Chakrabarty,A.M. (1997) DNase I footprinting, DNA bending and *in vitro* transcription analyses of ClcR and CatR interactions with the clcABD promoter: evidence of a conserved transcriptional activation mechanism. *Mol. Microbiol.*, **24**, 965–976.
24. Simon,R., Priefer,V. and Puhler,A. (1983) A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in Gram-negative bacteria. *Biotechnology*, **1**, 784.
25. Huang,W.H., Zhai,F. and Hong,G.F. (1999) Cloning, expression, characterization and application of Bst DNA polymerase Large Fragment. *Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai)*, **31**, 379–384.
26. Feng,J., Li,F.Q., Hu,H.L., Li,Q. and Hong,G.F. (2002) Expression and purification of *Rhizobium leguminosarum* NodD. *Protein Exp. Purif.*, **26**, 321–328.
27. Liu,S.T., Chang,W.Z., Cao,H.M., Hu,H.L., Chen,Z.H., Ni,F.D. and Hong,G.F. (1998) A HU-like protein binds to specific sites within nod promoters of *Rhizobium leguminosarum*. *J. Biol. Chem.*, **273**, 20568–20574.
28. Liu,S.T. and Hong,G.F. (1998) Three-minute G+A specific reaction for DNA sequencing. *Anal. Biochem.*, **255**, 158–159.
29. Oliphant,A.R., Brandl,C.L. and Struhl,K. (1989) Defining the sequence specificity of DNA-binding proteins by selecting binding sites from random-sequence oligonucleotides: analysis of yeast GCN4 protein. *Mol. Cell. Biol.*, **9**, 2944–2949.
30. Sambrook,J. Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
31. Orchard,K. and May,G.E. (1993) An EMSA-based method for determining the molecular weight of a protein–DNA complex. *Nucleic Acids Res.*, **21**, 3335–3336.
32. Rhee,K.Y., Senear,D.F. and Hatfield,G.W. (1998) Activation of gene expression by a ligand-induced conformational change of a protein–DNA complex. *J. Biol. Chem.*, **273**, 11257–11266.
33. Coco,W.M., Parsek,M.R. and Chakrabarty,A.M. (1994) Purification of the LysR family regulator, ClcR and its interaction with the *Pseudomonas putida clcABD* chlorocatechol operon promoter. *J. Bacteriol.*, **176**, 5530–5533.
34. Jourlin-Castelli,C., Mani,N., Nakano,M.M. and Sonenshein,A.L. (2000) CcpC, a novel regulator of the LysR family required for glucose repression of the citB gene in *Bacillus subtilis*. *J. Mol. Biol.*, **295**, 865–878.

35. Lamblin,A.J. and Fuchs,J.A. (1994) Funtional analysis of the *Eshcherichia coli* K-12 *cyn* operon transcriptional regulation. *J. Bacteriol.*, **176**, 6613–6622.
36. Chang,M. and Crawford,I. (1990) The roles of indoleglycerol phosphate and the TrpI protein in the expression of *trpBA* from *Pseudomonas aeruginosa*. *Nucleic Acids Res.*, **18**, 979–988.
37. Wu,H.-M. and Crothers,D.M. (1984) The locus of sequence-directed and protein induced DNA bending. *Nature*, **308**, 509–513.
38. Thompson,J.F. and Landy,A. (1988) Empirical estimation of protein-induced DNA bending angles: applications to lambda site-specific recombination complexes. *Nucleic Acids Res.*, **16**, 9687–9705.
39. Tyrrell,R., Verschueren,K.H., Dodson,E., Murshudov,G.N., Addy,C. and Wilkinson,A.J. (1997) The structure of the cofactor-binding fragment of the LysR family member, CysB: a family fold with a surpring subunit arrangement. *Structure*, **5**, 1017–1031.
40. Schell,M.A., Brown,P.H. and Raju,S. (1990) Mini review: positive control. *J. Biol. Chem.*, **265**, 3844–3850.
41. Opel,M.L. and Hatfield,G.W. (2001) DNA supercoiling-dependent transcriptional coupling between the divergently transcribed promoters of the *ilvYC* operon of *Escherichia coli* is proportional to promoter strengths and transcript lengths. *Mol. Microbiol.*, **39**, 191–198.
42. Ansari,A.Z., Bradner,J.E. and O'Halloran,T.V. (1995) DNA-bend modulation in a repressor-to-activator switching mechanism. *Nature*, **374**, 371–375.
43. Ansari,A.Z., Chael,M.L. and O'Halloran,T.V. (1992) Allosteric underwinding of DNA is a critical step in positive control of transcription by Hg-MerR. *Nature*, **355**, 87–89.
44. Lamb,J.W., Hombrecher,G. and Johnston,A.W.B. (1982) Plasmid-determined nodulation and nitrogen-fixation abilities in *Rhizobium phaseoli*. *Mol. Gen. Genet.*, **186**, 449–452.
45. Bagdasarian,M., Lurz,R., Ruckert,B., Franklin,F.C.H., Bagdasarian,M.M., Frey,J. and Timmis,K.N. (1981) Specific-purpose plasmid cloning vectors. II. Broad host range, high copy number, RSF1010-derived vectors and a host-vector system for gene cloning in *Pseudomonas*. *Gene*, **16**, 237–247.