JOURNAL OF BACTERIOLOGY, Oct. 1989, p. 5492–5502
0021-9193/89/105492-11$02.00/0
Copyright © 1989, American Society for Microbiology
Vol. 171, No. 10

DNA Footprint Analysis of the Transcriptional Activator Proteins NodD1 and NodD3 on Inducible *nod* Gene Promoters

ROBERT F. FISHER AND SHARON R. LONG*

Department of Biological Sciences, Stanford University, Stanford, California 94305-5020

Received 29 June 1989/Accepted 5 July 1989

The *Rhizobium meliloti nodD1* and *nodD3* gene products (NodD1 and NodD3) are members of the *lysR-nodD* gene regulator family. They are functionally distinct in that NodD1 transcriptionally activates other *nod* genes in the presence of a flavonoid inducer such as luteolin, while NodD3 is capable of activating *nod* gene expression at high levels in the absence of inducer. NodD1 and NodD3 are DNA-binding proteins which interact with DNA sequences situated upstream of the transcription initiation sites of at least three sets of inducible *nod* genes. We report the footprinting of NodD1- and NodD3-DNA complexes with both DNase I and the 1,10-phenanthroline-copper ion reagent. NodD1 and NodD3 both interacted with the *nodABC*, *nodFE*, and *nodH* promoters and protected from cleavage an extensive piece of DNA, including the *nod* box, from approximately −20 to −75 from the transcription start site for each of the three promoters. The constitutively activating protein NodD3 displayed an additional hypersensitive cleavage site in its footprint compared with NodD1.

Nitrogen fixation in alfalfa occurs following root infection by *Rhizobium meliloti*. A complex interaction, requiring the function of both plant and bacterial genes, is required in order to establish a productive symbiosis. We have focused on several sets of *Rhizobium* genes involved in the formation of root nodules (*nod* genes), which harbor the nitrogen-fixing *Rhizobium* bacteroids (28). The common *nod* genes, *nodABC*, found in all *Rhizobium* species examined to date (9, 24, 41, 45, 46, 52), are required for the initial stages of nodule development: epidermal root hair deformation, infection thread formation, cortical cell division, and nodule morphogenesis (5, 36). Genes which are apparently involved in nodulation efficiency and the specification of host range, i.e., the range of plants which a given *Rhizobium* species is able to infect, include the divergently transcribed *nodFE* and *nodH* in *R. meliloti* (6, 7, 14, 22, 23, 45, 48, 50).

While *nodABC*, *nodFE*, and *nodH* are poorly expressed under free-living conditions, they are induced over 30-fold in the presence of alfalfa or alfalfa exudates (10, 14, 19, 35, 44; J. T. Mulligan, Ph.D. thesis, Stanford University, Stanford, Calif., 1987). The most active inducing compound isolated from alfalfa seed exudates is luteolin (3′,4′,5,7-tetrahydroxyflavone) (39). This induction requires the expression of *nodD1* (19, 21, 34), which is transcribed divergently from *nodABC* (9, 11) (Fig. 1). Both mutagenesis and genetic transfer studies have indicated that the *nodD* gene product functions in transcriptional activation (19, 34, 42, 49). *R. meliloti* harbors two additional alleles of *nodD1* (17, 21); the positions of the *nodD1*, *nodD2*, and *nodD3* homologs are shown in Fig. 1. NodD1 is activated when cells are supplied with a complex plant seed extract or one purified inducer, luteolin (35). Overexpressed NodD2 is activated when cells are supplied with the complex extract, but not with purified luteolin (35). Overexpressed NodD3 causes high basal (uninduced) levels of *nodC-lacZ* expression; NodD3 activation is unaffected by seed extract or luteolin (35). Evidence has accumulated in recent years that *nodD*, the transcription activator, also plays a role in host specificity by determining which flavonoid compounds are able to serve as *nod* gene inducers and inhibitors (3, 22, 49).

* Corresponding author.

We have recently demonstrated by gel mobility shift assays that NodD1 and NodD3 are DNA-binding proteins which interact specifically with DNA sequences found upstream of the inducible *nod* genes *nodABC*, *nodFE*, and *nodH* (12). Examination of these promoter sequences shows that the only element shared by these upstream regions is a highly conserved 47-base-pair (bp) segment known as the *nod* box (7, 11, 14, 43, 45, 46, 48, 49). The specific NodD-*nod* promoter binding can be inhibited with a double-stranded DNA oligomer homologous to a portion of the *nod* box (12). Specific binding can also be inhibited by clearing NodD1 and NodD3 from extracts with an antibody directed against a LacZ-NodD1 fusion protein (13). The simplest model of *nod* gene positive activation predicts that NodD1 and NodD3 function by binding to the *nod* box and, by some as yet unknown mechanism, directing RNA polymerase to initiate transcription from an adjacent site (12, 43, 48). In this report we use DNase I (16) and 1,10-phenanthroline-copper ion (*o*Phen-Cu) (26) footprinting of NodD1- and NodD3-promoter DNA complexes to show that both bind to the promoters at the *nod* box. Approximately 55 bp of DNA are protected, extending ~4 bp on either side of the *nod* box, while the center of the *nod* box is highly prone to DNase I cleavage, suggesting that the DNA is bent so that this central portion is more accessible to DNase I. The NodD3 footprint displays a slightly larger set of hypersensitive cleavage sites. The *o*Phen-Cu chemical nuclease produces a footprint slightly smaller than the DNase I footprint.

MATERIALS AND METHODS

Recombinant plasmid construction and labeling of *nod* box fragments. pRmE36 (10) was used as a source of a 210-bp *Hpa*II-*Rsa*I fragment containing the *nod* box upstream of *nodA* (9). This fragment was ligated with *Acc*I- and *Sma*I-digested pUC119 (53) to generate pRmF528. A gel-purified 1.1-kilobase (kb) *Eco*RI-*Hin*dIII fragment from pRmF58 (12) was digested with *Sau*3A and *Hpa*II, and the 186-bp *Sau*3A-*Hpa*II fragment containing the *nod* box upstream of *nodF* (14) was cloned into pUC119 to generate pRmF536. Similarly, the gel-purified 0.5-kb *Pvu*II fragment from pRmF59 (12) was digested with *Hpa*II and *Hae*III, and the 140-bp *Hpa*II-*Hae*III fragment containing the *nod* box upstream of

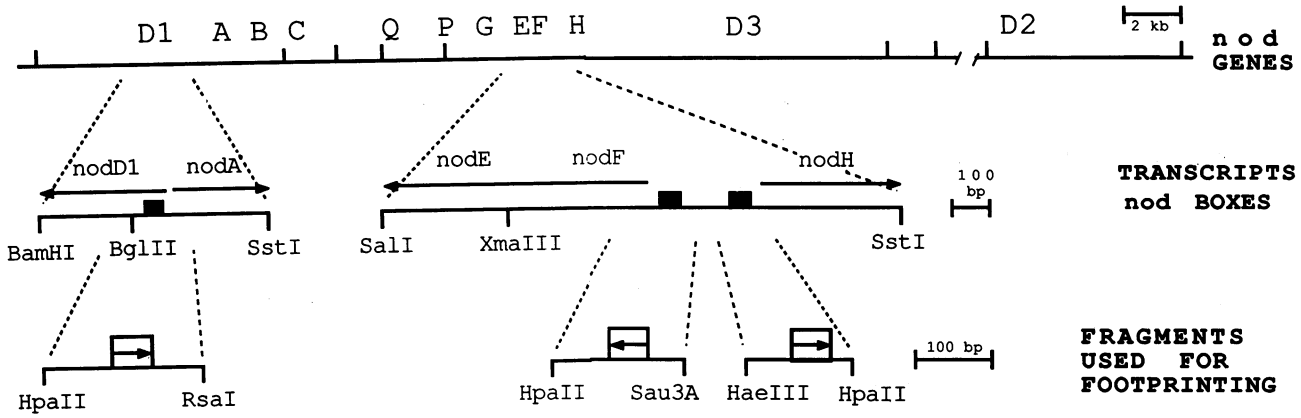


FIG. 1. Physical map of *R. meliloti* 1021 (SU47 Str^r) *nod* gene region. (*Top*) Map of *nod* gene regions on pSym*a*. Vertical lines represent *Eco*RI sites. The gap in the map between the other *nod* genes and *nodD2* is 42 kb. (*Middle*) Expanded scale, indicating transcripts (arrows) and their relationships to the *nod* boxes (black boxes), which lie 26 to 28 bp upstream from the transcription initiation sites. Appropriate restriction sites are indicated. Note that the *nodA nod* box also lies within the *nodD* transcript leader. (*Bottom*) Restriction fragments used to make subclones for footprinting. Small boxes with an arrow inside indicate position and relative orientation of *nod* boxes.

nodH (14) was cloned into pUC119 to generate pRmF569. This resulted in construction of a set of plasmids containing the *nodA*, *nodF*, and *nodH nod* boxes located approximately midway between the vector polylinker *Eco*RI and *Hin*dIII sites, which were subsequently used to generate specifically end-labeled substrates for footprinting analysis. Each strand of each *nod* box fragment was end labeled at the 3′ end following an initial digestion with either *Eco*RI or *Hin*dIII, by filling in with [α-^{32}P]dATP and unlabeled dCTP, dGTP, and dTTP with the Klenow fragment of DNA polymerase I (30) and then secondarily digesting with *Hin*dIII or *Eco*RI (whichever enzyme was not used during the primary digestion) after heat inactivation of the Klenow fragment. The appropriate 200- to 300-bp fragments were subsequently purified by polyacrylamide gel electrophoresis (31) for use in footprinting experiments.

Purification of NodD3. *R. meliloti* JM96 is a *nodD1-lacZ nodD2-uidA* fusion strain whose only intact *nodD* allele is *nodD3*; pRmE65 is a broad-host-range plasmid which overexpresses *nodD3* under control of the *Salmonella typhimurium trp* promoter (12). *R. meliloti* JM96(pRmE65) was grown in Luria broth (32) with 0.2% sucrose to an OD_{595} of 4.1 in a Biogen 200-liter fermentor and used as a source for the purification of NodD3. Cells were harvested by centrifugation, flash-frozen in liquid N_2, and stored at −80°C.

After suspension in TED (50 mM Tris hydrochloride [pH 8.0], 0.5 mM EDTA, 0.5 mM dithiothreitol) plus 250 mM NaCl to an OD_{595} of 175, cells were lysed in a French pressure cell at 10,000 to 14,000 lb/in^2. A mixture of protease inhibitors (final concentrations: leupeptin, 8 μg/ml; chymostatin, 2 μg/ml; pepstatin, 10 μg/ml; and 1 mM phenylmethylsulfonyl fluoride) was added to the lysate (fraction I), which was cleared by centrifugation at 30,000 rpm in a Beckman 45Ti rotor for 1 h at 4°C. Ammonium sulfate was slowly added to 0.26 g/ml with constant stirring at 4°C. The precipitated protein pellet was collected by centrifugation at 27,000 × *g* for 20 min. The pellet was washed once in a Dounce homogenizer with 0.25 volume of TED–0.1 M NaCl–0.22 g of ammonium sulfate per ml and twice with 0.1 volume of the same buffer. The remaining insoluble pellet was redissolved in TED–0.1 M NaCl and dialyzed for 70 min at 4°C against TED–0.1 M NaCl to give fraction II. Fraction II (70 mg of protein per ml) was adjusted to 5 mM $MgCl_2$ and 0.5 mM ATP and diluted to 7 mg of protein per ml with buffer A (50 mM Tris hydrochloride [pH 7.4], 25% glycerol, 1 mM EDTA, 5 mM dithiothreitol, 5 mM $MgCl_2$, and 0.5 mM ATP). The sample was applied to a 45-ml column of BioRex 70 which was equilibrated with buffer A plus 50 mM NaCl. The column was washed with 125 ml of buffer A–50 mM NaCl, and a linear 500-ml gradient of buffer A with 50 mM to 0.6 M NaCl was applied.

NodD3 activity was monitored by assaying for the specific shift in gel mobility of *nod* box-containing DNA fragments during polyacrylamide gel electrophoresis (12, 15, 33). The peak of NodD3 activity eluted at approximately 250 mM NaCl. Peak fractions were pooled and dialyzed against buffer A until a conductivity equivalent to that of buffer A plus 50 mM NaCl was achieved and then applied to a previously equilibrated 5-ml heparin-agarose (BioRad Laboratories) column. The column was washed with 2 column volumes of buffer A–50 mM NaCl, and a 15-column-volume gradient of buffer A plus 50 mM to 1 M NaCl was applied. Active fractions were flash-frozen in liquid N_2 and stored at −80°C in portions. In the NodD3 preparation used in the experiments reported here, NodD3 constituted approximately 25% of the protein in the fraction, as judged by Coomassie blue staining of sodium dodecyl sulfate (SDS)-polyacrylamide gels (27). NodD1 was purified by immunoaffinity chromatography as described previously (12).

DNase I cleavage within the polyacrylamide gel slice. Foot-

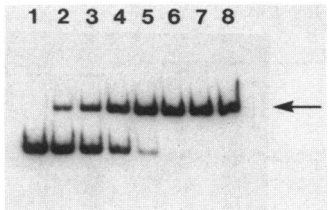


FIG. 2. Interaction of NodD3 with a *nod* box fragment. Increasing amounts of the partially purified NodD3 preparation were mixed with the end-labeled *nodF nod* box fragment (Fig. 1) as described previously (12). Lane 1 displays the electrophoretic migration pattern of labeled restriction fragment in the absence of any added NodD3 material. Lanes 2 to 8 display the same but in the presence of 16, 32, 64, 80, 160, 320, and 400 ng of the NodD3 preparation, respectively. The arrow designates the single electrophoretically retarded complex.

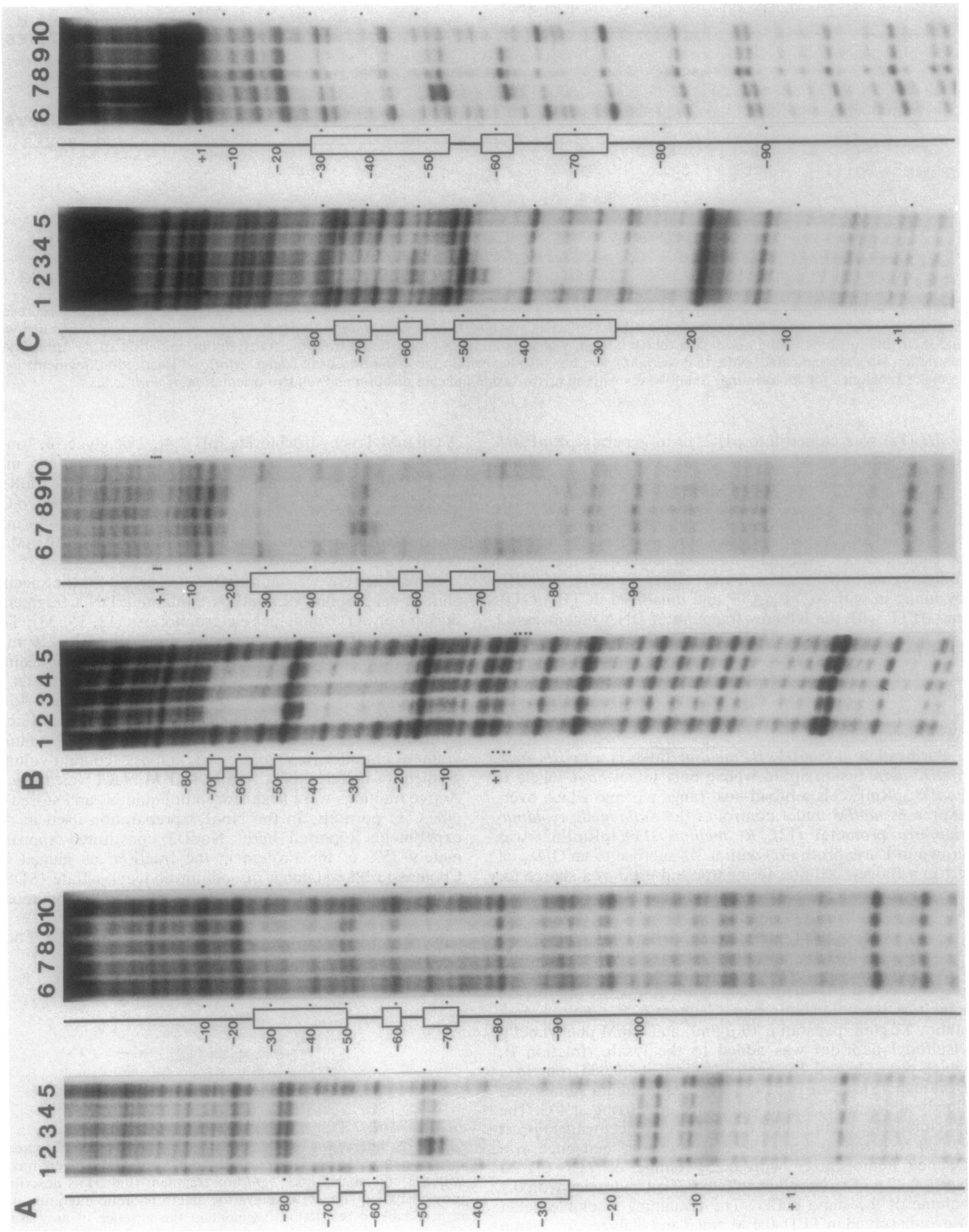

A
1 2 3 4 5
6 7 8 9 10
B
1 2 3 4 5
6 7 8 9 10
C
1 2 3 4 5
6 7 8 9 10

FIG. 3. DNase I footprint analysis of NodD protection of *nod* box sequences. Free DNA fragments and protein-DNA complexes (see Fig. 2) were subjected to DNase I footprinting within the polyacrylamide gel slice as described in Materials and Methods. (A) The DNA is a 246-bp fragment that contains the *nodA nod* box, whose position is indicated by the boxes to the left of each panel. The left panel (lanes 1 to 5) shows the results when the top strand was labeled at the right end, according to the orientation of the sequence in Fig. 4, and the right panel (lanes 6 to 10) shows the results when the bottom strand of Fig. 4 is labeled at the left end. The +1 indicates the transcription start site, and the number of base pairs upstream from that site is designated every 10 bp by the small dots. Lanes 1, 5, 6, and 10 present the DNase I cleavage pattern of the free (i.e., unshifted) DNA fragment. Lanes 2 and 7 show the DNase I cleavage pattern obtained following complex formation of the DNA fragments and NodD3. Lanes 3 and 8 and lanes 4 and 9 display the cleavage patterns obtained after DNA complexing with NodD1 in the absence and presence of 1 μM luteolin, respectively. (B) As in panel A, except the DNA fragment used contains the *nodF nod* box. (C) As in panel A, except the DNA fragment used contains the *nodH nod* box.

printing with DNase I was performed on protein-DNA complexes essentially as described by David C. Straney, Susan B. Straney, and Donald M. Crothers (personal communication) as follows. All steps were conducted at room temperature. Following brief autoradiography, the free or shifted bands were excised in an approximately 30-μl gel fragment volume and placed in microfuge tubes. Then, 3 μl of a DNase I solution (10 mM Tris hydrochloride [pH 8.0], 2 mM dithiothreitol, 5% glycerol, 0.5 mg of bovine serum albumin per ml, and 0.4 μg of DNase I per ml) was spread on the surface of the gel slices and allowed to incubate for 45 min. Then, 3 μl of a solution containing 50 mM $MgCl_2$ and 50 mM $CaCl_2$ was spread over the gel slices and allowed to incubate for 4 min before 15 μl of 0.1 M EDTA was added to stop the DNase I. After an additional 4 min, 2.5 μl of 1% SDS was added. DNA was electroeluted in 150 μl of TBE (89 mM Tris, 89 mM borate, 2.8 mM EDTA) for 1 h at 150 V (30). Samples were recovered from the dialysis tubing, and the DNA was ethanol precipitated and dissolved in formamide loading dye (80% [vol/vol] formamide, 10 mM NaOH, 1 mM EDTA, 0.1% bromophenol blue, 0.1% xylene cyanol) before being run on sequencing gels. DNA sequencing ladders of the identical fragments were generated by the method of Maxam and Gilbert (31).

Chemical nuclease footprinting within the polyacrylamide matrix. Chemical cleavage of protein-DNA complexes or free DNA fragments took advantage of the nuclease activity of *o*Phen-Cu (26). Footprinting was conducted exactly as described before (26) with the following modifications. After the digestion was quenched with 2,9-dimethyl-1,10-phenanthroline, gel slices and liquid were placed in dialysis tubing, 70 μl of TBE was added, and the digestion products were electroeluted at 150 V for 45 min (30). The liquid was recovered from the tubing, and the DNA was ethanol precipitated and suspended in formamide loading dye before samples were run on sequencing gels adjacent to Maxam-Gilbert (31) sequencing ladders.

RESULTS

Positive transcriptional activators bind to unique target DNA sequences, forming stable, biochemically detectable protein-DNA complexes which are thought to function by directing precise transcriptional initiation by RNA polymerase (8, 18, 29, 38, 40). To determine the precise site of action of NodD1 and NodD3 in the *nod* gene regulatory regions, we used both enzymatic and chemical nuclease footprinting of restriction fragments containing the *nodA*, *nodF*, and *nodH* promoters. The patterns of cleavage obtained in the presence and absence of NodD1 or NodD3 are compared on DNA sequencing gels. A decrease in the intensity of a cleavage product band results from protection of the cleavage site by NodD1 or NodD3, the DNA-binding proteins. We used both NodD1 and NodD3 to study these interactions and expected that their footprints would be similar but not identical, because while NodD3 is able to activate *nod* gene transcription in the absence of other known factors, NodD1 requires an inducer from the plant to achieve transcriptional activation (34, 35). We used the gel mobility shift assay to enrich for NodD1- and NodD3-promoter DNA complexes. Figure 2 shows a typical gel mobility shift assay, in which increasing amounts of partially purified NodD3 are mixed with an end-labeled restriction fragment that contains the *nodF nod* box. In all such assays, we observed only a single shifted (reduced mobility) band. We subjected the free DNA fragment (Fig. 2, lower band) and protein-DNA complexes

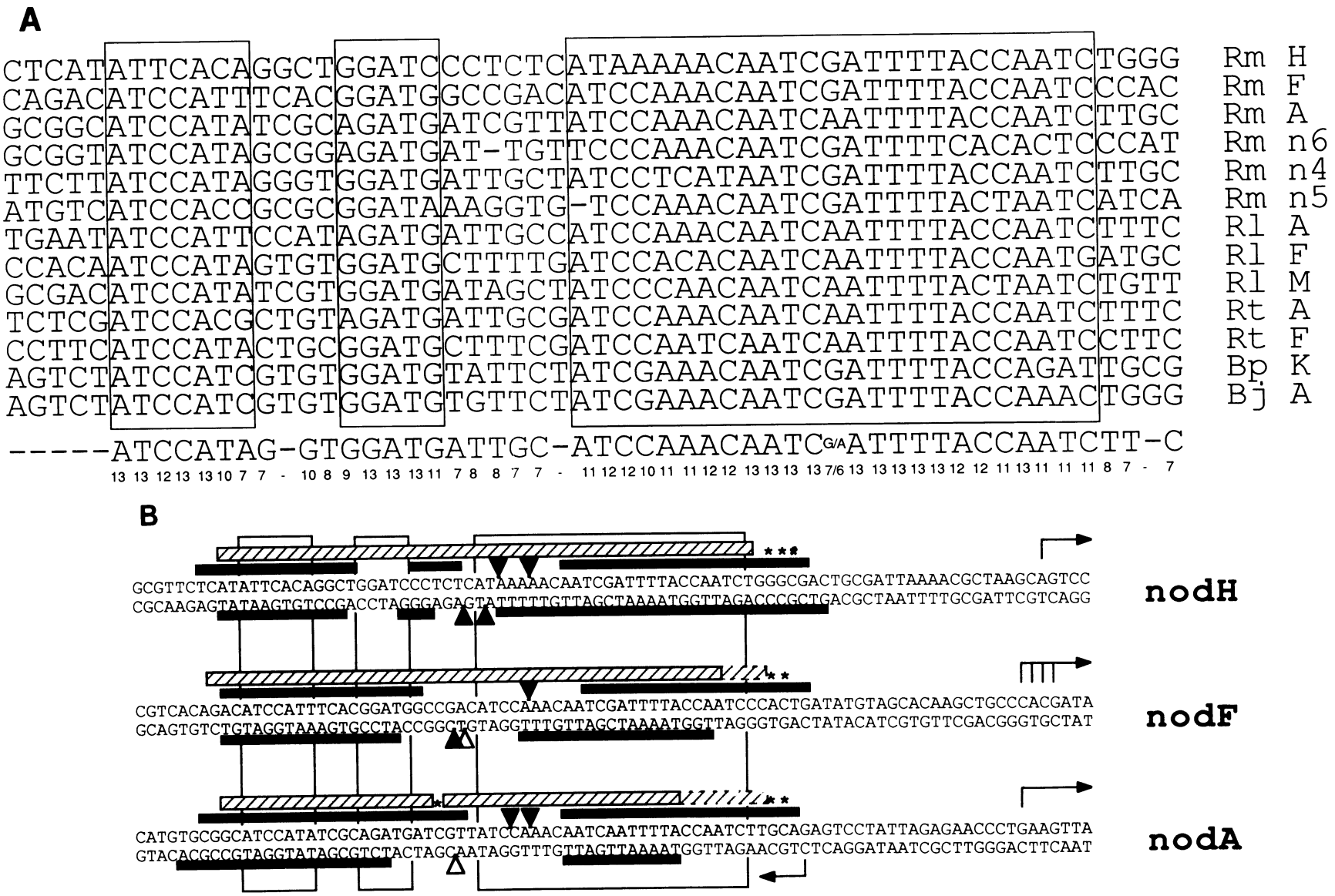

A
CTCATATTCACAGGCTGGATCCCTCTCATAAAAACAATCGATTTTACCAATCTGGG Rm H
CAGACATCCATTTCACGGATGGCCGACATCCAAACAATCGATTTTACCAATCCCAC Rm F
GCGGCATCCATATCGCAGATGATCGTTATCCAAACAATCAATTTTACCAATCTTGC Rm A
GCGGTATCCATAGCGGAGATGAT-TGTTCCCAAACAATCGATTTTCACACTCCCAT Rm n6
TTCTTATCCATAGGGTGGATGATTGCTATCCTCATAATCGATTTTACCAATCTTGC Rm n4
ATGTCATCCACCGCGCGGATAAAGGTG-TCCAAACAATCGATTTTACTAATCATCA Rm n5
TGAATATCCATTCCATAGATGATTGCCATCCAAACAATCAATTTTACCAATCTTTC Rl A
CCACAATCCATAGTGTGGATGCTTTTGATCCACACAATCAATTTTACCAATGATGC Rl F
GCGACATCCATATCGTGGATGATAGCTATCCCAACAATCAATTTTACTAATCTGTT Rl M
TCTCGATCCACGCTGTAGATGATTGCGATCCAAACAATCAATTTTACCAATCTTTC Rt A
CCTTCATCCATACTGCGGATGCTTTCGATCCAATCAATCAATTTTACCAATCCTTC Rt F
AGTCTATCCATCGTGTGGATGTATTCTATCGAAACAATCGATTTTACCAGATTGCG Bp K
AGTCTATCCATCGTGTGGATGTGTTCTATCGAAACAATCGATTTTACCAAACTGGG Bj A
-----ATCCATAG-GTGGATGATTGC-ATCCAAACAATCG/AATTTTACCAATCTT-C
13 13 12 13 13 10 7 7 - 10 8 9 13 13 13 11 7 8 8 7 7 - 11 12 12 10 11 11 12 12 13 13 13 13 7/6 13 13 13 13 13 12 12 11 13 11 11 11 8 7 - 7
B
GCGTTCTCATATTCACAGGCTGGATCCCTCTCATAAAAACAATCGATTTTACCAATCTGGGCGACTGCGATTAAAACGCTAAGCAGTCC
CGCAAGAGTATAAGTGTCCGACCTAGGGAGAGTATTTTTGTTAGCTAAAATGGTTAGACCCGCTGACGCTAATTTTGCGATTCGTCAGG
nodH
CGTCACAGACATCCATTTCACGGATGGCCGACATCCAAACAATCGATTTTACCAATCCCACTGATATGTAGCACAAGCTGCCCACGATA
GCAGTGTCTGTAGGTAAAGTGCCTACCGGCTGTAGGTTTGTTAGCTAAAATGGTTAGGGTGACTATACATCGTGTTCGACGGGTGCTAT
nodF
CATGTGCGGCATCCATATCGCAGATGATCGTTATCCAAACAATCAATTTTACCAATCTTGCAGAGTCCTATTAGAGAACCCTGAAGTTA
GTACACGCCGTAGGTATAGCGTCTACTAGCAATAGGTTTGTTAGTTAAAATGGTTAGAACGTCTCAGGATAATCGCTTGGGACTTCAAT
nodA

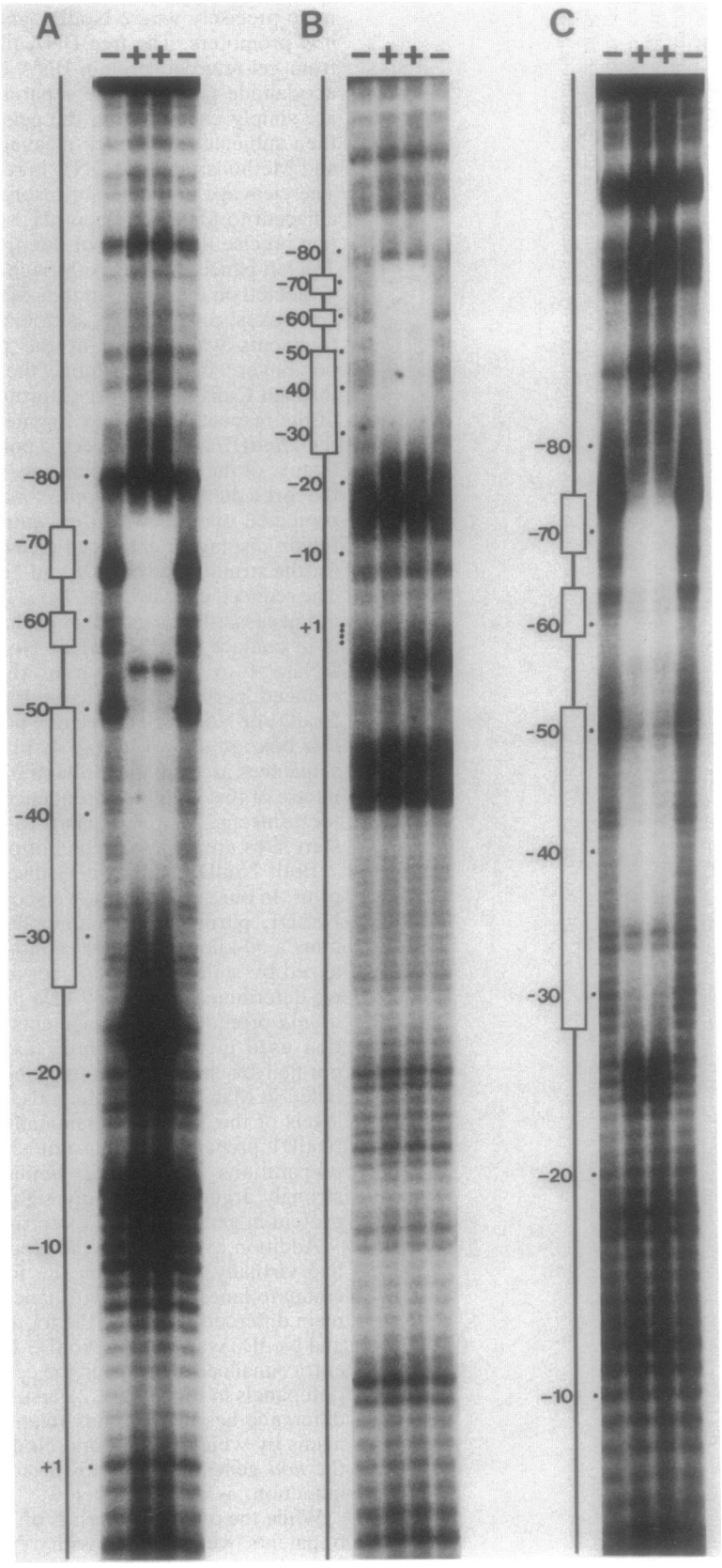
FIG. 4. *nod* box sequences and summary of footprint protection data. (A) The DNA sequence and location of 13 different *nod* boxes are presented for five different organisms: *R. meliloti* (Rm) (43), *R. leguminosorum* bv. *viciae* (Rl) (48), *R. leguminosarum* bv. *trifolii* (Rt) (45), *Bradyrhizobium* sp. (*Parasponia*) (Bp) (46), and *Bradyrhizobium japonicum* (Bj) (37). Also indicated is the letter designating the gene which lies downstream of each *nod* box (e.g., A indicates *nodA*) or the number of the *nod* box assigned by Rostas et al. (43). At the bottom, a consensus sequence is presented, along with the number of times that the specific consensus base is found in the different *nod* boxes. (B) Summary of footprint studies presented in this report. The DNA sequence of the *nodA*, *nodF*, and *nodH nod* boxes and adjacent segments is presented, and the *nod* box is boxed. Transcription start sites, including that for divergent *nodD1* on the bottom line, are indicated by the arrows. Solid bars indicate the regions protected by NodD from DNase I cleavage on each strand. Open triangles indicate hypersensitive DNase I cleavage sites due to the presence of NodD1 or NodD3, and solid triangles indicate hypersensitive sites found only on addition of NodD3. Hatched bars indicate regions protected by NodD3 from nuclease activity of *o*Phen-Cu. Asterisks indicate hypersensitive *o*Phen-Cu nuclease cleavage sites (3′ of the indicated base) in the presence of NodD3. Note that for the *nodA* and *nodF nod* boxes, the absolute limit of NodD3 protection by *o*Phen-Cuo was not determined, which is represented by the change in the hatched pattern.

(Fig. 2, upper band) to both enzymatic (DNase I) and chemical nuclease (*o*Phen-Cu) footprinting in order to determine precisely where NodD1 and/or NodD3 interacts with *nod* promoters. The free DNA fragment is readily resolved from gel-retarded protein-DNA complexes on native polyacrylamide gels, and the separated radiolabeled fragments are simply excised from the gels. The DNA fragments are then subjected to limited cleavage as detailed in Materials and Methods, and the DNA is recovered by electroelution. The cleavage products are resolved on polyacrylamide gels adjacent to Maxam-Gilbert (31) sequencing ladders, permitting precise localization of the protected segments of DNA. The left panels of Fig. 3A through 3C show which bases were protected on the upper strands of the *nodA*, *nodF*, and *nodH nod* boxes, respectively, as oriented in Fig. 4. Each of these fragments was labeled at the *Eco*RI site of the vector polylinker. In like fashion, the right panels of Fig. 3A through C display the bases protected on the lower strands of the respective *nod* box fragments, which were labeled at the *Hin*dIII site of the vector polylinker. The most striking feature of the DNase I footprints shown here is the extent of the protected region; approximately 55 bp of DNA was protected overall, with the central portion of the *nod* box region displaying enhanced cleavage. This has implications for the structure of NodD1 and NodD3, as discussed below. The reduced cleavage by DNase I in the *nod* box region does not necessarily indicate that each base in the DNA sequence is in contact with NodD1 or NodD3; rather, the access of DNase I to that portion of the DNA strand is simply reduced, perhaps by steric hindrance due to the presence of NodD1 or NodD3. The relative positions of the footprint and *nod* box are shown in Fig. 4. In the top panel, 13 *nod* box sequences are displayed, illustrating the highly conservative nature of this regulatory sequence. The position of the DNA footprint and its relationship to the known transcriptional start sites are shown in the bottom panel.

Both NodD1 and NodD3 displayed this extensive footprint. In our earlier work, we showed that the preparation of NodD1, purified by immunoaffinity chromatography, contains a 59-kilodalton (kDa) contaminant which can be visualized by staining SDS-polyacrylamide gels (12). However, we determined that the 59-kDa protein was not able to bind to *nod* promoter DNA fragments (12). The NodD3 preparation used in the footprinting experiments described here, purified by standard ion-exchange chromatography as detailed in Materials and Methods, did not contain detectable levels of the 59-kDa contaminant which was present in the NodD1 preparation (data not shown). The fact that both preparations produced a similar footprint (Fig. 3) also strongly argues against any significant role for the 59-kDa protein in generating the footprint at the *nod* box.

Addition of luteolin to the NodD1-DNA reaction mixture had virtually no effect on the footprints obtained (Fig. 3, compare lanes 3 and 8 with lanes 4 and 9, all panels). The main difference between the footprints generated by NodD1 and NodD3 was that the NodD3-DNA complex displayed an extra enhanced cleavage in the central portion of the *nod* box (left panels in Fig. 3, lane 2 versus lanes 3 and 4). This subtle difference has implications relating to the potential mechanisms by which NodD3 and NodD1 induce transcription of the *nod* genes and also to the role of luteolin in *nod* gene induction, as discussed below.

While the overall footprints on the three different *nod* box fragments were similar with regard to the extent of the footprint, they did not have identical cleavage patterns. This may partially reflect the known sequence specificity of


A
B
C
- + + -
-80
-70
-60
-50
-40
-30
-20
-10
+1

DNase I cleavage (1, 2). Thus, the segments of diverging DNA sequence that are interspersed among the different *nod* boxes probably affect cleavages by DNase I. The inherent binding affinities of NodD1 and NodD3 for the different *nod* boxes also probably contribute to the small differences in the footprints, including the relative weakness of the *nodH nod* box footprint.

Because DNase I cleaves only a subset of DNA backbone positions, we also used the chemical nuclease activity of *o*Phen-Cu to further characterize the NodD-DNA interactions (26). Such footprints are generally smaller than DNase I-directed footprints because DNase I is a bulkier molecule and is prevented from cleaving the DNA immediately adjacent to the protein-binding site due to steric hindrance (26). Again, we footprinted within the polyacrylamide gel slice following gel mobility shift enrichment of NodD3-DNA complexes. With low-molecular-weight, readily diffusible chemical reagents, *o*Phen-Cu cleavage occurs within the polyacrylamide matrix and yields footprints consistent with those obtained in solution (26). Figure 5 displays the NodD3–*o*Phen-Cu footprints of the upper (coding) strands of the *nodA*, *nodF*, and *nodH nod* boxes shown in Fig. 4B. As expected, a slightly smaller footprint was observed on all three *nod* box fragments than was observed after DNase I cleavage. A striking feature of each *o*Phen-Cu footprint was the appearance of an extremely hypersensitive cleavage product that mapped to a position at the right edge of the *nod* box, as oriented in Fig. 4. The *nod* box regions protected by NodD3 as detected by *o*Phen-Cu footprinting are also summarized in Fig. 4B (hatched bars).

DISCUSSION

Our results show that NodD1 and NodD3, positive activators required for induction of other *nod* genes, bind to the *nod* boxes located starting 26 to 28 bp upstream of the transcription start sites of *nodA*, *nodF*, and *nodH*. In an independent study, Kondorosi et al. (25) used *Rhizobium* extracts containing NodD to show protection of the *nodA nod* box from DNase I cleavage, which is consistent with the results presented here.

The two regulatory genes *nodD1* and *nodD3* differ in their activating behavior. NodD1 requires the presence of a plant factor in order to cause *nod* gene induction in vivo (34). Overexpressed NodD3, on the other hand, activates *nod* gene expression in the absence of any exogenous plant factor (35). The transcription start sites for *nod* genes activated by NodD1 and NodD3 are identical (35) which is consistent with the similarities in the footprints for NodD1 and NodD3. The results reported here, along with those of previous studies with a mobility shift gel system to assay NodD-promoter interactions (12), and genetic data showing that different *nodD* alleles interfere with each other's activities (35), all support the model that NodD-promoter binding is not affected by inducer and is essentially similar for all NodD proteins.

We studied the interaction of both NodD1 and NodD3 on three different *nod* box fragments in order to obtain a consensus view of how NodD interacts with *nod* boxes while fulfilling its role as a positive activator. It was surprising to find that such an extensive region of DNA (~50 bp) was protected by NodD1 and NodD3 from DNase I and *o*Phen-Cu cleavage. By contrast, the *Escherichia coli* catabolite activator protein (CAP), a 22.5-kDa monomeric protein which functions as a 45-kDa dimer, only protects ~25 bp of the *gal* or *lac* promoters from DNase I cleavage in the absence of RNA polymerase (47, 51). When RNA polymerase is added to the *gal* promoter, cooperative binding of CAP to a second, upstream site occurs, which lengthens the protected region (47). To protect more than 20 bp of DNA, CAP must induce a bent or kinked conformation in the DNA (47). This suggests that in order to protect such a large segment in these DNA protection assays, NodD1 and NodD3, 35-kDa monomeric proteins (9, 10, 12), function as multimeric proteins, have a very unusual tertiary structure, or induce bending or kinking of the target DNA. These properties are not necessarily mutually exclusive. The fact that the central portion of each *nod* box studied here displays hypersensitivity to DNase I (Fig. 4) is consistent with the formation of a similar kink in the DNA of each of the *nod* boxes upon NodD1 or NodD3 binding.

We have noticed elements of twofold rotational symmetry (data not shown) in the *nod* boxes and adjacent sequences reported by Rostas et al. (43). The twofold symmetry in each individual promoter sequence was not found exclusively in the conserved *nod* box sequences. Classically, it could be predicted that NodD might function as a dimer binding to symmetrical sites. This model alone, however, would be insufficient to account for all of the data; the *nodF nod* box segment had the weakest twofold rotational symmetry elements of the six *nod* boxes in *R. meliloti* yet yielded the tightest footprint (Fig. 3B and 5B) of the three *nod* boxes examined.

The *nodD* DNA sequence shows it to be a member of a newly defined group of procaryotic activator genes, designated the LysR family (20). These proteins are highly related to each other but not to other bacterial regulatory proteins. All of the members of the LysR family are between 30 and 35 kDa in size, and several regulate the expression of an overlapping promoter on the opposite strand of the template. One of these, OxyR, is required for the induction of a regulon of hydrogen peroxide-inducible genes in *E. coli* and *Salmonella typhimurium* (4). Crude extracts of cells overproducing OxyR yield extensive (~45 bp) footprints on the *S. typhimurium ahpC* and *E. coli katG* promoters, part of the regulon of hydrogen peroxide-inducible genes (L. Tartaglia, G. Storz, and B. Ames, J. Mol. Biol., in press). As we found with NodD, they observed that the central portion of the footprint contains hypersensitive cleavage sites. In addition, IlvY activates the divergent *ilvC* gene in *E. coli*, the second enzyme in the parallel isoleucine-valine biosynthetic pathway (54). Cell extracts enriched for IlvY protect two adjacent 27-bp segments upstream of *ilvC* which are separated from each other by 5 bp. Unlike the NodD-*nod* box interaction, however, the two IlvY-protected segments in the *ilvC*

FIG. 5. Binding of NodD3 to *nod* box fragments as determined by *o*Phen-Cu nuclease footprint analysis. Free and gel-retarded DNA fragments were subjected to *o*Phen-Cu footprinting within the polyacrylamide gel matrix as described in Materials and Methods. (A) The top strand, as oriented in Fig. 4, of the *nodA nod* box fragment, whose position is indicated by the boxes on the left, was subjected to *o*Phen-Cu cleavage in the presence (+) and absence (−) of NodD3. The position of the transcription start site is indicated by +1, and the number of base pairs upstream from the site is denoted every 10 bp by the small dots. (B) As in panel A, except the DNA fragment used contains the *nodF nod* box. (C) As in panel A, except the DNA fragment footprinted contains the *nodH nod* box. Duplicate gel applications of *o*Phen-Cu cleavage products are shown.

promoter have characteristics of classical procaryotic operators; the recognition sites contain a nucleotide sequence that is an inverted repeat. It will be of interest to see whether the other LysR-type proteins display similarly large footprints and to examine what patterns may be displayed by the various DNA-binding sites.

Using the footprint assay, we looked for keys to the luteolin effect and to the difference between the two *nodD* alleles *nodD1* and *nodD3*. Genetic evidence from the *R. leguminosarum* biovar *viciae nod* gene system suggests that NodD plays a direct role in mediating the response to various flavonoid compounds during *nod* gene induction. Mutation of *nodD* in *R. leguminosarum* bv. *viciae* results in an altered response to a spectrum of flavonoid inducer molecules and inhibitors (3). Transfer of native *nodD* alleles from diverse species into *nodD R. leguminosarum* (49) or *R. meliloti* (22) mutant backgrounds also alters the response to various flavonoid inducers. We have shown here that NodD1 is able to protect the *nod* box region from DNase I cleavage whether or not luteolin, the most active inducer molecule from alfalfa seed exudates, is present during the formation of protein-DNA complexes. In other experiments, we added a vast excess of luteolin to the shifted NodD1-*nod* box complexes which had been excised from shift gels prior to DNase I treatment, and still observed no effect on the footprint pattern (data not shown). Thus, we found no biochemical evidence by this assay for a direct interaction between a flavonoid inducer and NodD1. However, NodD1 may interact directly with luteolin to effect a change at the NodD1-RNA polymerase interface and not at the level of NodD1 binding to its target, the *nod* box. We are also unable to rule out the possibility that some other factor mediates an interaction between the inducer molecules and NodD1 during the induction process.

That luteolin has shown no effect on the footprint in these assays makes it even more interesting to compare the behavior of NodD1 and NodD3. When overexpressed in *R. meliloti*, *nodD3* functions as a naturally occurring constitutive variant of the more typical *nodD1*-like, inducible activator which is found as the unique *nodD* in other systems (35, 42, 48). In particular, the basal level of an overexpressed NodD3-induced *nodC-lacZ* fusion is about 100-fold higher than background, and addition of plant exudates fails to significantly elevate expression of the fusion (35). In *R. leguminosarum* bv. *viciae*, either mutation of *nodD* by as little as one codon (3) or construction of certain chimeric *nodD* genes (55) can result in a *nodD* which activates *nod* gene expression constitutively, although the overall activity is not as high as that of *R. meliloti nodD3*.

We observed a potentially significant difference between the footprints observed with NodD1 and NodD3. Use of NodD3 resulted in one or two additional hypersensitive sites in the central portion of the *nod* box on each *nod* box fragment tested (near bp −47; Fig. 3, compare lanes 2 and 3). If this extra NodD3-dependent hypersensitivity represents the positioning of the NodD3-*nod* box complex into an activated (for transcription) state, then its absence in the presence of NodD1 may reflect the need for components besides luteolin to achieve NodD1-mediated transcriptional activation. Alternatively, the difference between the NodD1- and NodD3-generated footprints may simply result from the different ways they were purified. It is also possible that the distinctiveness of the NodD3 footprint is due to other structural differences between the *nodD3* and *nodD1* gene products which are not related to the functional difference. We are currently determining the sequence of the *nodD3* gene to see how divergent it is from that of *nodD1*. In addition, we plan to distinguish between these possibilities by combined genetic and molecular analyses.

A functional demonstration of the role of NodD in transcriptional activation itself will require the pursuit of several goals: mutagenesis of both *nodD* genes and the target sequences of their gene products, to define the points of critical contact, should be carried out in parallel with further biochemical tests such as methylation protection and an analysis of in vivo promoter strength and NodD activity. Finally, faithful in vitro expression from inducible *nod* promoters will be required to confirm the identity of essential components needed for *nod* gene activation and to permit an analysis of their mechanism of transcriptional activation.

ACKNOWLEDGMENTS

We thank Tom Egelhoff for his generous gift of purified NodD1 and David Bramhill for advice and assistance in NodD3 purification; D. B. and Janette Carey for advice on footprinting procedures; D. B. and Joy Ogawa for suggestions on Maxam-Gilbert sequencing; C. Yanofsky for the use of materials and equipment; and Frank Witney (Bio-Rad Laboratories) for a gift of T4 DNA ligase. Special thanks to Marc Gartenberg (Yale University) for the tip on footprinting within the gel slice, and to M. G. and Don Crothers for providing the protocol for this procedure prior to publication. We are grateful to all the members of our research group for valuable discussions during the course of this work, to D. B., Nancy Federspiel, Melanie Yelton, and Julie Schwedock for comments on the manuscript, and to Alexandra Bloom for her care in its preparation.

This research was supported by Public Health Service grant R01-GM30962 to S.R.L. from the National Institutes of Health.

LITERATURE CITED

1. **Bernardi, A., C. Gaillard, and G. Bernardi.** 1975. The specificity of five DNAses as studied by the analysis of 5′-terminal doublets. Eur. J. Biochem. **52**:451–457.
2. **Bernardi G., D. E. Stanislav, and J.-P. Thiery.** 1973. The specificity of deoxyribonucleases and their use in nucleotide sequence studies. Nature (London) New Biol. **246**:36–40.
3. **Burn, J., L. Rossen, and A. W. B. Johnston.** 1987. Four classes of mutations in the *nodD* gene of *Rhizobium leguminosarum* biovar *viciae* that affect its ability to autoregulate and/or activate other *nod* genes in the presence of flavonoid inducers. Genes Dev. **1**:456–464.
4. **Christman, M. F., G. Storz, and B. N. Ames.** 1989. OxyR, a positive regulator of hydrogen peroxide-inducible genes in *Escherichia coli* and *Salmonella typhimurium*, is homologous to a family of bacterial regulatory proteins. Proc. Natl. Acad. Sci. USA **86**:3484–3488.
5. **Dart, P. J.** 1977. Infection and development of leguminous nodules, p. 367–472. *In* R. W. Hardy and W. S. Silver (ed.), A treatise on dinitrogen fixation, vol. III. John Wiley & Sons, Inc., New York.
6. **Debellé, F., C. Rosenberg, J. Vasse, F. Maillet, E. Martinez, J. Dénarié, and G. Truchet.** 1986. Assignment of symbiotic developmental phenotypes to common and specific nodulation (*nod*) genetic loci of *Rhizobium meliloti*. J. Bacteriol. **168**:1075–1086.
7. **Debellé, F., and S. B. Sharma.** 1986. Nucleotide sequence of *Rhizobium meliloti* RCR2011 genes involved in host specificity of nodulation. Nucleic Acids Res. **14**:7453–7472.
8. **De Crombrugghe, B., S. Busby, and H. Buc.** 1984. Activation of transcription by the cyclic AMP receptor protein. Science **224**:831–837.
9. **Egelhoff, T. T., R. F. Fisher, T. W. Jacobs, J. T. Mulligan, and S. R. Long.** 1985. Nucleotide sequence of *Rhizobium meliloti* 1021 nodulation genes: *nodD* is read divergently from *nodABC*. DNA **4**:241–248.
10. **Egelhoff, T. T., and S. R. Long.** 1985. *Rhizobium meliloti*

nodulation genes: identification of *nodDABC* gene products, purification of *nodA* protein, and expression of *nodA* in *Rhizobium meliloti*. J. Bacteriol. **164**:591–599.

11. **Fisher, R. F., H. L. Brierley, J. T. Mulligan, and S. R. Long.** 1987. Transcription of *Rhizobium meliloti* nodulation genes: identification of a *nodD* transcription initiation site *in vitro* and *in vivo*. J. Biol. Chem. **262**:6849–6855.
12. **Fisher, R. F., T. T. Egelhoff, J. T. Mulligan, and S. R. Long.** 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.
13. **Fisher, R. F., T. T. Egelhoff, J. T. Mulligan, M. M. Yelton, and S. R. Long.** 1988. *Rhizobium meliloti* NodD binds to DNA sequences upstream of inducible nodulation genes, p. 391–398. *In* H. Bothe, F. J. de Bruijn, and W. E. Newton, (ed.), Nitrogen fixation: hundred years after. Gustav Fisher Verlag, Stuttgart.
14. **Fisher, R. F., J. Swanson, J. T. Mulligan, and S. R. Long.** 1987. Extended region of nodulation genes in *Rhizobium meliloti* 1021. II. Nucleotide sequence, transcription start sites, and protein products. Genetics **117**:191–201.
15. **Fried, M., and D. M. Crothers.** 1981. Equilibria and kinetics of *lac* repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. **9**:6505–6525.
16. **Galas, D. J., and A. Schmitz.** 1978. DNAase footprinting: a simple method for the detection of protein-DNA binding specificity. Nucleic Acids Res. **5**:3157–3170.
17. **Göttfert, M., B. Horvath, E. Kondorosi, P. Putnoky, F. Rodriguez-Quinones, and A. Kondorosi.** 1986. At least two *nodD* genes are necessary for efficient nodulation of alfalfa by *Rhizobium meliloti*. J. Mol. Biol. **191**:411–420.
18. **Gussin, G. N., C. W. Ronson, and F. M. Ausubel.** 1986. Regulation of nitrogen fixation genes. Annu. Rev. Genet. **20**: 567–591.
19. **Györgypal, Z., N. Iyer, and A. Kondorosi.** 1988. Three regulatory *nodD* alleles of diverged flavonoid-specificity are involved in host-dependent nodulation by *Rhizobium meliloti*. Mol. Gen. Genet. **212**:85–92.
20. **Henikoff, S., G. W. Haughn, J. M. Calvo, and J. C. Wallace.** 1988. A large family of bacterial activator proteins. Proc. Natl. Acad. Sci. USA **85**:6602–6606.
21. **Honma, M., and F. M. Ausubel.** 1987. *Rhizobium meliloti* has three functional copies of the *nodD* regulatory gene. Proc. Natl. Acad. Sci. USA **84**:8558–8562.
22. **Horvath, B., C. W. B. Bachem, J. Schell, and A. Kondorosi.** 1987. Host-specific regulation of nodulation genes in *Rhizobium* is mediated by a plant-signal interacting with the *nodD* gene product. EMBO J. **6**:841–848.
23. **Horvath, B., E. Kondorosi, M. John, J. Schmidt, I. Torok, L. Z. Gyorgypal, I. Barabas, U. Wieneke, J. Schell, and A. Kondorosi.** 1986. Organization, structure and symbiotic function of *Rhizobium-meliloti* nodulation genes determining host specificity for alfalfa. Cell **46**:335–344.
24. **Jacobs, T. W., T. T. Egelhoff, and S. R. Long.** 1985. Physical and genetic map of a *Rhizobium meliloti* nodulation gene region and nucleotide sequence of *nodC*. J. Bacteriol. **162**:469–476.
25. **Kondorosi, E., J. Gyuris, J. Schmidt, M. John, E. Duda, B. Hoffmann, J. Schell, and A. Kondorosi.** 1989. Positive and negative control of *nod* gene expression in *Rhizobium meliloti* is required for optimal nodulation. EMBO J. **8**:1331–1340.
26. **Kuwabara, M. D., and D. S. Sigman.** 1987. Footprinting DNA-protein complexes in situ following gel retardation assays using 1,10-phenanthroline-copper ion: *Escherichia coli* RNA polymerase-*lac* promoter complexes. Biochemistry **26**:7234–7238.
27. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the heat of bacteriophage T4. Nature (London) **227**:680–685.
28. **Long, S. R.** 1989. *Rhizobium*-legume nodulation: life together in the underground. Cell **56**:203–214.
29. **Majors, J.** 1975. Specific binding of CAP factor to *lac* promoter DNA. Nature (London) **256**:672–674.
30. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. p. 545. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
31. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
32. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 466. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
33. **Miller, V. L., R. K. Taylor, and J. J. Mekalanos.** 1987. Cholera toxin transcriptional activator ToxR is a transmembrane DNA binding protein. Cell **48**:271–279.
34. **Mulligan, J. T., and S. R. Long.** 1985. Induction of *Rhizobium meliloti nodC* expression by plant exudate requires *nodD*. Proc. Natl. Acad. Sci. USA **82**:6609–6613.
35. **Mulligan, J. T., and S. R. Long.** 1989. A family of activator genes regulates expression of *Rhizobium meliloti* nodulation genes. Genetics **122**:7–18.
36. **Newcomb, W.** 1981. Nodule morphogenesis and differentiation, p. 247–298. *In* K. L. Giles and A. G. Atherly (ed.), Biology of the rhizobiaceae. Academic Press, Inc., New York.
37. **Nieuwkoop, A. J., Z. Banfalvi, N. Deshmane, D. Gerhold, M. G. Schell, K. M. Sirotkin, and G. Stacey.** 1987. A locus encoding host range is linked to the common nodulation genes of *Bradyrhizobium japonicum*. J. Bacteriol. **169**:2631–2638.
38. **Pabo, C., and R. Sauer.** 1984. Protein-DNA recognition. Annu. Rev. Biochem. **53**:293–321.
39. **Peters, N. K., J. W. Frost, and S. R. Long.** 1986. A plant flavone, luteolin, induces expression of *Rhizobium meliloti* nodulation genes. Science **233**:917–1008.
40. **Raibaud, O., and M. Schwartz.** 1984. Positive control of transcription initiation in bacteria. Annu. Rev. Genet. **18**:173–206.
41. **Rossen, L., A. W. B. Johnston, and J. A. Downie.** 1984. DNA sequence of the *Rhizobium leguminosarum* nodulation genes *nodAB* and *C* requires for root hair curling. Nucleic Acids Res. **12**:9497–9508.
42. **Rossen, L., C. A. Shearman, A. W. B. Johnston, and J. A. Downie.** 1985. The *nodD* gene of *Rhizobium leguminosarum* is autoregulatory and in the presence of plant exudate induces the *nod A, B, C* genes. EMBO J. **4**:3369–3373.
43. **Rostas, K., E. Kondorosi, B. Horvath, A. Simoncsits, and A. Kondorosi.** 1986. Conservation of extended promoter regions of nodulation genes in *Rhizobium*. Proc. Natl. Acad. Sci. USA **83**:1757–1761.
44. **Schmidt, J., M. John, U. Wieneke, H.-D. Kruessmann, and J. Schell.** 1987. Expression of the nodulation gene *nodA* in *Rhizobium meliloti* and localization of the gene product in the cytosol. Proc. Natl. Acad. Sci. USA **83**:9581–9585.
45. **Schofield, P. R., and J. M. Watson.** 1986. DNA sequence of *Rhizobium trifolii* nodulation genes reveals a reiterated and potentially regulatory sequence preceding *nodABC* and *nodFE*. Nucleic Acids Res. **14**:2891–2903.
46. **Scott, K. F.** 1986. Conserved nodulation genes from the nonlegume symbiont *Bradyrhizobium* sp. (*Parasponia*). Nucleic Acids Res. **14**:2905–2919.
47. **Shanblatt, S. H., and A. Revzin.** 1987. Interactions of the catabolite activator protein (CAP) at the galactose and lactose promoters of *Escherichia coli* probed by hydroxyl radical footprinting. J. Biol. Chem. **262**:11422–11427.
48. **Shearman, C. A., L. Rossen, A. W. B. Johnston, and J. A. Downie.** 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.
49. **Spaink, H. P., C. A. Wijffelman, E. Pees, R. J. H. Okker, and B. J. J. Lugtenberg.** 1987. *Rhizobium* nodulation gene *nodD* as a determinant of host specificity. Nature (London) **328**:337–340.
50. **Swanson, J., J. K. Tu, J. M. Ogawa, R. Sanga, R. Fisher, and S. R. Long.** 1987. Extended region of nodulation genes in *Rhizobium meliloti* 1021. I. Phenotypes of Tn*5* insertion mutants. Genetics **117**:181–189.
51. **Taniguchi, T., M. O'Neill, and B. De Crombrugghe.** 1979. Interaction site of *Escherichia coli* cyclic AMP receptor protein on DNA of galactose operon promoters. Proc. Natl. Acad. Sci. USA **76**:5090–5094.
52. **Török, I., E. Kondorosi, T. Stepkowski, J. Posfai, and A. Kondorosi.** 1984. Nucleotide sequence of *Rhizobium meliloti*

nodulation genes. Nucleic Acids Res. **12**:9509–9524.
53. **Vieira, J., and J. Messing.** 1987. Production of single-stranded plasmid DNA. Methods Enzymol. **153**:3–11.
54. **Wek, R. C., and G. W. Hatfield.** 1988. Transcriptional activation at adjacent operators in the divergent-overlapping *ilvY* and *ilvC* promoters of *Escherichia coli*. J. Mol. Biol. **203**:643–663.
55. **Wijffelman, C. A., H. P. Spaink, R. J. H. Okker, E. Pees, S. A. J. Zaat, A. A. N. Van Brussel, and B. J. J. Lugtenberg.** 1988. Regulation of *nod* gene expression and nodulation, p. 417–422. *In* H. Bothe, F. J. de Bruijn, and W. E. Newton (ed.), Nitrogen fixation: hundred years after. Gustav Fischer Verlag, Stuttgart.