A Family of Activator Genes Regulates Expression of *Rhizobium meliloti* Nodulation Genes

John T. Mulligan and Sharon R. Long

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

Manuscript received August 29, 1988
Accepted for publication January 27, 1989

ABSTRACT

Nodulation (nod) gene expression in *Rhizobium meliloti* requires plant inducers and the activating protein product of the nodD gene. We have examined three genes in *R. meliloti* which have nodD activity and sequence homology. These three nodD genes are designated nodD1, nodD2 and nodD3, and have distinctive properties. The nodD1 gene product activates expression of the nodABC operon, as measured by a nodC-lacZ fusion or by transcript analysis, in the presence of crude seed or plant wash or the inducer, luteolin. The nodD3 gene product can cause a high basal (uninduced) level of nodC-lacZ expression and nodABC transcripts which is relatively unaffected by inducers. The effect of nodD3 is dependent on the presence of another gene, syrM (symbiotic regulator). By primer extension analysis we determined that the transcription start site is the same for nodD1 plus luteolin or nodD3-syrM mediated expression of nodA and nodH mRNAs. syrM also enhances the expression of another symbiotically important trait, production of extracellular polysaccharide. This regulatory effect of syrM requires locus syrA, which is linked to nodD3 and syrM. The syrM-syrA mediated increase in polysaccharide production requires at least some of the previously identified exo genes and may be a parallel regulatory event to the syrM-nodD3 control of nod promoters.

BACTERIA in the genus *Rhizobium* invade specific host plants and stimulate the development of nitrogen-fixing root nodules. This complex and highly specific process occurs between a given species of *Rhizobium* and a set of plants which defines its host range. *Rhizobium meliloti* establishes a productive symbiosis with alfalfa (*Medicago sativa*) and related plants. The early events in the development of nodules include bacterial stimulation of cell divisions in the plant root cortical layer, giving rise to a new meristem, and bacterial invasion of epidermal root hairs and underlying cells (BAUER 1981; NEWCOMB 1981).

At least three sets of *R. meliloti* genes are involved in mediating the early stages of alfalfa nodule formation. One set, nodABC, is present in all *Rhizobium* species so far examined, and mutations in these genes lead to a Nod− phenotype: the bacteria neither stimulate cortical cell divisions, nor deform or invade root hairs (ROSTAS et al. 1986; DEBELLE et al. 1986; DUDLEY, JACOBS and LONG 1987; JACOBS, EGELHOFF and LONG 1985). The members of the second set of nodulation genes (in *R. meliloti*, nodFE, nodH, nodG, nodPQ and others) are not necessarily present in or functionally conserved among all *Rhizobium* species; these affect the rate and frequency of nodule formation and influence host range (HORVATH et al. 1986; DJORDJEVIC, SCHOFIELD and ROLFE 1985; DEBELLE et al. 1986; RODRIGUEZ-QUINONES et al. 1987; SWANSON et al. 1987; SCHWEDOCK and LONG 1989).

The *Rhizobium* exopolysaccharides, whose production is under the control of an additional set of symbiotic genes, the exo genes, are also required for normal nodule development (FINAN et al. 1985; CHAKRABORTY et al. 1982; CHEN et al. 1985; BORTHAKUR et al. 1986; LEIGH, SIGNEr and WALKER 1985; DYLAN et al. 1986. Most exo gene mutations of *R. meliloti* (deficient in acidic exopolysaccharides) block the normal bacterial invasion of the plant root. When alfalfa is inoculated with exo− strains, nodules develop but few or no infection threads are formed, few if any bacteria enter the nodules, and no nitrogen fixation occurs (FINAN et al. 1985).

A family of regulatory genes, which are referred to as nodD, has been identified in all *Rhizobium* species so far examined (MULLIGAN and LONG 1985; EGELHOFF and LONG 1985; ROSEN et al. 1985; GÖTTFERT et al. 1986; SPAINK et al. 1987a, b; SheARMAN et al. 1986). Although some *Rhizobium* species carry only a single copy of nodD (DOWNIE et al. 1984; SCHOFIELD and WATSON 1985), many species carry two or three similar but not identical copies of nodD (RODRIGUEZ-QUINONES et al. 1987; APPEI BaUM A et al. 1988). Expression of nodABC, nodFE and nodH is induced in the presence of an exudate of legume seeds or roots; this activation requires a nodD gene product. *R. meliloti* has three copies of nodD, designated nodD1, nodD2 and nodD3 (PUTNOKY and KONDOROSI 1986; GÖTTFERT et al. 1986; HONMA and AUSUEL 1987).

The flavone, luteolin, isolated from alfalfa seed wash, is the most active inducer of *R. meliloti* nod

Genetics 122: 7–18 (May, 1989)
genes when nodD1 is overexpressed (Peters, Frost and Long 1986). A wide variety of flavones and flavanonones are present in legume exudates, some of which are active in inducing nod gene expression, and others in blocking induction (Djordjevic et al. 1987; Redmond et al. 1986; Firmin et al. 1986; Zaat et al. 1987; Peters and Long 1988). Several lines of evidence indirectly suggest that the nodD gene products interact with the inducing molecules. First, the substitution of the naturally occurring nodD gene of one Rhizobium species for the nodD gene of another species in otherwise isogenic strains can determine the range of compounds which will induce expression of the nod genes (Spaink et al. 1987a). Second, a mutation in nodD can broaden the spectrum of plant or synthetic compounds which will induce nod gene expression (Burn, Rossen and Johnston 1987). Third, the species source of nodD in a transgenic construct in some cases influences or determines the host range of the Rhizobium strain (Spaink et al. 1987b; Horvath et al. 1987; Gyorgypal, Iyer and Kondorosi 1988). However, there are as yet no direct biochemical demonstrations of NodD-inducer interaction. The nodD gene product, NodD, appears to be a DNA-binding protein by the criterion that NodD-containing extracts (Hong et al. 1987) and substantially purified NodD (Fisher et al. 1988) affect the electrophoretic mobility of nod gene promoters. Competition studies show that this binding occurs at least partially at the nod box (Fisher et al. 1988), a conserved sequence found upstream of each inducible nod operon (Rostas et al. 1986; Schofield and Watson 1986; Spaink et al. 1987b). Transcription of the inducible nod genes initiates 26-28 bp downstream from the nod box (Fisher et al. 1988).

In this study, we examined the regulatory properties of the three R. meliloti nodD genes. We found that while each of the nodD genes is capable of affecting the expression of a nodC-lacZ fusion, the three genes are not equivalent.

nodD3 requires a second gene, syrM (symbiotic regulator), for its activating function, but is not dependent on flavones to activate expression of nodC-lacZ. The transcription start sites for both nodA and nodH are unchanged, whether their expression is mediated by nodD1 and luteolin, or by nodD3 and syrM. SyrM can affect both the expression of the nod genes and the expression or activity of the exo genes required for invasion of the plant roots, and thus may play a key role in a regulatory system which coordinates several symbiotic functions.

MATERIALS AND METHODS

Strains and plasmids: Strains and plasmids used in this study are listed in Table 1. The described insertions in all of the listed strains were confirmed by restriction digestion and Southern (1975) analysis.

Plasmid constructions: Plasmid pRmM111 is a derivative of pRmJT5 carrying an insertion of Tn5 which was produced as described in Swanson et al. (1987) and screened as described in Results. The construction of pRmM113 involved two steps; first, pRmSS507 was digested with ClaI and the fragment bearing the Tn5 insertion was cloned into the ClaI site of pBR322. The resulting plasmid was digested with SauI and XhoI, and the fragment carrying syrM and a segment of Tn5 which encodes neomycin resistance was ligated into pLAFR3 digested with BamHI, by the method of partial fill-in of each cohesive end (Zabrovskey and Alkimets 1986), to form pRmM113. pRmSS11 was digested with XhoI and relaxed at low concentration to form pRmM136, which lacks the sequences between the Tn5 and the left-most XhoI site of pRmJT5. pRmM157 (Fisher et al. 1988) was digested with XbaI, partially filled-in and ligated to a partially filled-in HindIII fragment which carries the Tn5 neomycin resistance gene to form pRmM141. pRmJT5 was digested with XbaI, partially filled-in and ligated to a partially filled-in HindIII fragment carrying the uidA gene and spectinomycin resistance to form pRmM142. A 2-kg BglII fragment from pRmSS11 was purified and cloned into the BamHI site of pUC119 to form pRmM151 and pRmM152. A 2.1-kb ClaI fragment from pRmSS11 was cloned into the AccI site of pUC119 to establish pRmM147.

Strain construction: Marker exchange was carried out as described by Jacobs, Egelhoff and Long (1985) except that pR751 was used to exclude the IncP plasmids. Transductions were carried out using phage N3 as described in Martin and Long (1984). JM139 was produced by marker exchange of the spectinomycin resistance insertion in pRmM139 into R. meliloti 1021. JM80 and JM86 were produced by marker exchange of the Tn5 insertions from pRmSS303 and pRmST01, respectively, into JM61. JM81 and JM83 were produced by marker exchange of the Tn5 insertions from pRmSB97 and pRmST01, respectively, into JM57. JM85 was produced by marker exchange of the neomycin resistance insertion of pRmM141 into JM57. JM88, JM204 and JM207 were produced by transduction of the Tn5 insertions from JT308, JT801 and JT709, respectively, into JM57. JM206 was produced by transduction of the Tn5 insertion from JT801 into JM61. JM90 was produced by cotransduction of the nodD3 Tn5 insertion and the nodD1-lacZ fusion of JM80 into JM139. JM96 was produced by cotransduction of the syrM Tn5 insertion and the nodD1-lacZ fusion of JM86 into JM139. JM201 was produced by cotransduction of the nodD3 Tn5 insertion and the nodD-lacZ fusion of JM88 into JM139. The described insertions in all of the strains were confirmed by Southern blot analysis.

Assays: β-galactosidase assays were carried out as described by Miller (1972) with the modifications previously described (Mulligan and Long 1985). Four replicate samples were assayed for each condition and each strain. The β-galactosidase units reported in the text and tables are the average of the results from the four assays. The results of the individual assays are all within 15% of the reported averages. The endogenous β-galactosidase activity was 2 units under all the conditions tested. Seed wash was prepared as described by Mulligan and Long (1985). Nodulation tests were carried out as described by Jacobs, Egelhoff and Long (1985). Western blots were performed as previously described by Egelhoff and Long (1985). For Southern blots, DNA was transferred to Biotrans nylon membrane (ICN) by the method of Reed and Mann (1985). DNA fragments used as probes were labeled by the hexamer labeling method (Feinberg and Vogelstein 1983), with the modification that reactions were carried out in Klenow buffer (50 mM Tris (pH 8.0) and 50 mM MgCl2). High
stringency washes were 0.1× SSC at 65°C, and low stringency washes were 1× SSC at 65°C.

Primer extensions: To determine the transcription start sites, an oligonucleotide primer complementary to a section of the coding sequence of each gene was synthesized. The sequence of the primer for nodA was 5'-TAG CTTCCACTGAC-3', and for nodH, 5'-GCAGCG-TGGAATGGG-3'. End-labeling was carried out in a volume of 10 μl with 2 picomoles of DNA primer and 8 picomoles (405 Ci/mmol) of [γ-32P]ATP for 4 hr at 37°C using T4 polynucleotide kinase in the buffer suggested by the supplier. RNA was prepared as described by FISHER et al. (1987a). RNase-free DNase was prepared as described by TULLIS and RUBIN (1980). DNase (Sigma) and proteinase K (Boehringer Mannheim), each at a concentration of 1 mg/ml, were incubated together in 20 mM Tris (pH 7.5) and 10 mM CaCl2 for 2 hr at 37°C, and then added fresh to the RNA preparation to a final concentration of 0.1 mg/ml DNase. Primer extensions were carried out by the method of WILLIAMS and MASON (1985), except that the primer was hybridized to 10 μg of RNA overnight at 42°C. Sequencing ladders were obtained by using the same primers in dideoxy sequencing reactions on single stranded clones of the upstream region for each gene. The sequencing ladder and RNA-complementary primer extension products were run in parallel on sequencing gels to establish the position of the transcription initiation site as previously described for nodF and nodH (FISHER et al. 1987b).

RESULTS

Identification and isolation of nodD2 and nodD3: In R. meliloti 1021, nodD1 is adjacent to nodABC (EGELHOFF et al. 1985).
Göttfert et al. (1986) and Honma and Ausubel (1987) demonstrated that two homologous genes, designated nodD2 and nodD3, are present, respectively, on a 6.8-kb and a 15.5-kb EcoRI fragment in R. meliloti.

Swanson et al. (1987) describe the isolation of a cosmide clone, pRMJT5, which carries nodH, nodF and nodE on a 15.5-kb EcoRI fragment. This same fragment hybridizes to an internal nodD1 probe (data not shown), indicating that nodD3 resides on this same 15.5-kb fragment. The remaining 6.8-kb nodD1 homologous fragment was not present on any of our previously identified clones (data not shown). As described by Fisher et al. (1988), we identified a clone containing this 6.8-kb fragment from a cosmide clone bank of R. meliloti 1021. The 6.8-kb fragment was subcloned into pLAFR3 to create pRM137.

Plasmid-borne copies of nodD2 and nodD3 stimulate expression of nodC: Our previous work (Mulligan and Long 1985) showed that a plasmid-borne copy of nodD1 enhanced the induction of the nodC-lacZ fusion by plant exudate (see also Table 2, lines 1 and 9). Peters, Frost and Long (1986) showed that the flavone luteolin is the major inducing molecule in alfalfa plant exudate when nodD1 is plasmid-borne. We therefore tested the effect of plasmid-borne copies of the other two nodD genes on the induction of the nodC-lacZ fusion by luteolin and by alfalfa seed wash (exudate). Seed wash is a complex mixture including luteolin and a wide variety of related compounds (S. Long and N. K. Peters, unpublished observations), some of which may potentially interact with the three nodD gene products.

The plasmid carrying nodD2, pRM137, enhanced the expression of the nodC-lacZ fusion in the absence of any inducer (Table 2, lines 1 and 2). Luteolin stimulated expression of the nodC-lacZ fusion when no plasmid was present (therefore reflecting the action of the normal genomic copies of nodD) but not when the nodD2 plasmid was present. In contrast, seed wash stimulated nodC expression four fold (line 2). The increase in nodC expression in the presence of the plasmid was dependent on nodD2: an insertion in the plasmid borne copy of the nodD2 coding region, reduced the expression of nodC (Table 2, line 3). This mutated plasmid, pRM141, had a consistent small (2-fold) effect on the uninduced level of nodC expression, but caused no enhancement of seed wash induction of nodC (Table 2, lines 1 and 3).

The cosmide clone pRMJT5, which carries nodD3, in addition to nodE, nodF and nodH, caused very high expression of the nodC-lacZ fusion with or without inducer (Table 2, line 4). In addition, strains which contain pRMJT5 displayed a mucoid colonial morphology (Figure 1). Ninety derivatives of pRMJT5 with different Tn5 insertions (Swanson et al. 1987; we were tested for their ability to induce constitutive expression of the chromosomal nodC-lacZ fusion and to cause a mucoid colony morphology. Insertions in three regions of the plasmid affected one or both of these phenotypes (Figure 2) and defined three loci: syrM, syrA and nodD3 (syr for symbiotic regulation). Each locus is delimited by flanking Tn5 insertions with no regulatory phenotype. Plasmids bearing insertions in syrM (pRM5701, pRM5712, pRM5715 and pRM5711) showed reduced nodC expression and failed to cause the mucoid colony morphology (Table 2, line 5, shows results representative of the three insertions in syrM). The syrA insertion (pRM111) had no effect on the expression of nodC, but failed to cause the mucoid colonial morphology (Table 2, line 6). Three Tn5 insertions in the 2.2-kb nodD3-containing BglII fragment of pRMJT5 (pRM5803, pRM5709 and pRM5801) reduced expression of nodC but still rendered the strain mucoid (Table 2, lines 7 and 8, show results of two of the three insertions in nodD3). None of the remaining 83 Tn5 insertions in pRMJT5 affected nodC expression or colony morphology (data not shown). Since several Tn5 insertions lying between syrM and nodD3 display no regulatory phenotype, the two genes are probably in separate transcriptional units. Genomic Tn5 insertions at each of these genes individually did not block or substantially delay nodulation (Swanson et al. 1987; Honma and Ausubel 1987). The lack of requirement for syrA and syrM may be due to redundant function, as in the case of the nodD group of genes; alternatively, syrM and syrA may affect regulation of functions in a way which is not critical for nodulation, at least in standard laboratory assays.

The same exo genes, such as exoA, exoB and exoE, which are required for bacterial invasion of the plant roots (Finan et al. 1985; Leigh, Signer and Walker 1985) are also required for the mucoid colonial phenotype of strains containing pRMJT5. We found that Tn5 insertion mutations in these genes suppress the mucoid colonial morphology caused by pRMJT5 (data not shown), suggesting that pRMJT5 causes the mucoid phenotype by affecting the activity or expression of these exo genes. This suggests that syrM can affect the regulation or activity of two classes of symbiotic genes, the nod genes and the exo genes.

Transcription initiates at the same site whether induced by nodD1 plus luteolin or by syrM plus nodD3: We examined the transcription start sites for nodABC and for nodH under two regulated conditions. Figure 3 shows that for both nodA and nodH, transcription initiates at the same nucleotide whether induction is mediated by nodD1 in the presence of luteolin, or by plasmid-borne copies of nodD3 and syrM. The reactions in panel A of Figure 3 use an oligonucleotide primer that is complementary to the
TABLE 2

Effect of cloned genes on inducible nod gene expression and colony morphology

<table>
<thead>
<tr>
<th>Line No.</th>
<th>Strain</th>
<th>Plasmid Genotypea</th>
<th>Minimal β-Galactosidase activity</th>
<th>Minimal + luteolin β-Galactosidase activity</th>
<th>Minimal + seed wash β-Galactosidase activity</th>
<th>Colonial morphology</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>JM57</td>
<td>None</td>
<td>3</td>
<td>15</td>
<td>15</td>
<td>Normal</td>
</tr>
<tr>
<td>2</td>
<td>JM57 pRmM137</td>
<td>D2</td>
<td>16</td>
<td>9</td>
<td>67</td>
<td>Normal</td>
</tr>
<tr>
<td>3</td>
<td>JM57 pRmM141</td>
<td>D2</td>
<td>6</td>
<td>10</td>
<td>13</td>
<td>Normal</td>
</tr>
<tr>
<td>4</td>
<td>JM57 pRmJT5</td>
<td>D3, M+, A+</td>
<td>260</td>
<td>265</td>
<td>240</td>
<td>Mucoid</td>
</tr>
<tr>
<td>5</td>
<td>JM57 pRmS701</td>
<td>D3, M+, A+</td>
<td>8</td>
<td>8</td>
<td>14</td>
<td>Normal</td>
</tr>
<tr>
<td>6</td>
<td>JM57 pRmM111</td>
<td>D3, M+, A−</td>
<td>220</td>
<td>NT</td>
<td>240</td>
<td>Normal</td>
</tr>
<tr>
<td>7</td>
<td>JM57 pRmS303</td>
<td>D3, M+, A−</td>
<td>45</td>
<td>NT</td>
<td>75</td>
<td>Mucoid</td>
</tr>
<tr>
<td>8</td>
<td>JM57 pRmS801</td>
<td>D3, M+, A−</td>
<td>70</td>
<td>NT</td>
<td>75</td>
<td>Mucoid</td>
</tr>
<tr>
<td>9</td>
<td>JM57 pRmJ30</td>
<td>D1</td>
<td>3</td>
<td>87</td>
<td>75</td>
<td>Normal</td>
</tr>
</tbody>
</table>

a Gene abbreviations: D1 = nodD1; D2 = nodD2; D3 = nodD3; M = syrM; A = syrA.

In this and the following tables the activities listed are the averages of the results from four assays. In each case the individual assays gave values within 15% of the average. The endogenous β-galactosidase activity in the absence of introduced fusions is 2 units. NT = not tested.

Figure 1.—Mucoid colonial morphology of an R. meliloti strain carrying plasmid pRmJT5. R. meliloti 1021 carrying pRmSL26 (left) or pRmJT5 (right) was streaked for single colonies on an LB plate.

initial portion of the nodA coding sequence. This primer was used in dideoxy sequencing reactions (SANGER, NICKLEN and COULSON 1977) on a clone of the putative promoter region to provide the size standards in lanes 1–4. The same primer was end-labeled and used in primer extension reactions on RNA from two strains. In lane 5 the RNA was purified from a strain carrying a nodD1 plasmid and induced with luteolin. In lane 6 the RNA was isolated from a strain with plasmid copies of nodD3 and syrM. The primer extension products from the two reactions are similar and indicate that transcription initiates 26 bp from the nodA nod box. Panel B shows analogous reactions carried out with an oligonucleotide primer that is complementary to the beginning of the nodH coding sequence. Again, the primer extension products synthesized using the two induced RNA samples are similar (lanes 11 and 12), and indicate that transcription initiates 28 bp from the nodH nod box. Thus inducible nod promoters are affected indistinguishably by the two regulatory systems: nodD1 in the presence of luteolin, and nodD3 with syrM.

Effect of genomic mutations on nodC expression:
The experiments described above used the overexpression of plasmid-borne copies of regulatory genes to demonstrate gene location and function. We also studied regulatory consequences of genomic inser-}

tional mutations in these genes. A chromosomal copy of the nodC-lacZ fusion in an otherwise wild-type background was induced 5-fold by purified luteolin and 5–12-fold by alfalfa seed wash (Table 3, line 1).
A strain containing a Tn5 insertion in nodD1 showed almost no induction of nodC by luteolin or seed wash (Table 3, line 2). Tn5 insertions in syrM have little effect on the regulation of nodC (line 3). An insertion in nodD2 enhances the induction of nodC by luteolin (line 4).

Three different Tn5 insertions in the nodD3 region enhanced the induction of nodC by luteolin (Table 3, lines 5, 6, 7 compared with line 1). The enhancement of nodC induction by nodD3 insertions is unlikely to be due to polar effects on another gene because our mapping indicates that insertion #1005 (SWANSON et al. 1987), which has no regulatory phenotype (data not shown), is fewer than 500 bp downstream of the nodD3 coding region.

Luteolin-mediated induction of the nodC-lacZ fusion by nodD1 was dramatically enhanced in a background in which nodD2 and nodD3 both were absent. nodC induction in JM201, a derivative of the nodC fusion strain with only nodD1 intact (nodD2::uidA insertion, nodD3::Tn5 #303), was similar to that allowed by a plasmid copy of nodD1 (Table 3, lines 8 and 9). The increased induction in JM201 (nodD2^− nodD3^−) relative to JM57, which has intact copies of all three nodDs, could be due either to enhanced expression of nodD1 or to an increase in the activity of a constant amount of NodD1. To test these possibilities, we analyzed the level of NodD1 in the two strains. Figure 4 shows a Western blot loaded with an equal amount of protein from the nodC-lacZ fusion strain JM57 (lane

**TABLE 3**

Regulatory behavior of mutations in nodD and syrM genes

<table>
<thead>
<tr>
<th>Line No.</th>
<th>Strain</th>
<th>D1</th>
<th>D2</th>
<th>D3</th>
<th>Minimal</th>
<th>Minimal + luteolin</th>
<th>Minimal + seed wash</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>JM57</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>3</td>
<td>15</td>
<td>15−35</td>
</tr>
<tr>
<td>2</td>
<td>JM81</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>3</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>JM83</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>3</td>
<td>22</td>
<td>18</td>
</tr>
<tr>
<td>4</td>
<td>JM85</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>3</td>
<td>75</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>JM88^a</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>3</td>
<td>30</td>
<td>18</td>
</tr>
<tr>
<td>6</td>
<td>JM207^b</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>2</td>
<td>33</td>
<td>NT</td>
</tr>
<tr>
<td>7</td>
<td>JM204^c</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>5</td>
<td>84</td>
<td>17</td>
</tr>
<tr>
<td>8</td>
<td>JM201</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>4</td>
<td>110</td>
<td>20^d</td>
</tr>
<tr>
<td>9</td>
<td>JM57</td>
<td>++</td>
<td>+</td>
<td>−</td>
<td>3</td>
<td>130</td>
<td>145</td>
</tr>
</tbody>
</table>

^a In all the experiments reported here the bacterial cultures were exposed to inducer for 4 hr. When JM201 (line 8) was induced for 8 hours with seed wash the final activity was 80 units.

^b Batches of seed wash are variable. Repetition of inductions with a single batch gave consistent results, but occasional batches gave enhanced induction. The remaining inductions were done with seed wash which induced JM57 to 15 units.

^c These strains each carry a different nodD3::Tn5 insertion. JM88 carries Tn5 insertion #305, JM207 carries Tn5 insertion #709, and JM204 carries Tn5 insertion #801 (SWANSON et al. 1987).

**FIGURE 3.**—Primer extension to determine transcription start sites for nodA (lanes 1–6) and nodH (lanes 7–12) when induced by nodD1 plus luteolin or nodD3 plus syrM. Oligonucleotide primers were used to direct DNA synthesis complementary to transcripts isolated from luteolin-induced cells carrying pRnJ826, which contains nodD1 (lanes 5 and 11), and luteolin uninduced cells carrying pRnJ15, which contains nodD3 and syrM (lanes 6 and 12). The nodA complementary oligomer was used in lanes 1–6 and the nodH complementary oligomer in lanes 7–12. For both nodA and nodH, transcription is initiated at the same site in cells induced by nodD1 plus luteolin, and in cells induced by nodD3 plus syrM. Corresponding sequencing ladders (lanes 1–4 and 7–10) permit identification of the transcription start sites.

**FIGURE 4.**—Use of a Western blot to estimate nodD1 protein levels. Protein extracts from R. meliloti strains were subjected to SDS-polyacrylamide gel electrophoresis, transferred to nitrocelulose and probed with a NodD-specific polyclonal antibody. Strains from which extracts were prepared are as follows: Lane 1, MJ57; lane 2, JM57 (pRnJ30); lane 3, JM201 (nodD2^− nodD3^−). The large arrow denotes the nodD1 protein; the small arrows indicate the expected positions of the nodD2 and nodD3 proteins.
TABLE 4

**nodD1** expression in mutant backgrounds

<table>
<thead>
<tr>
<th>Line No.</th>
<th>Strain</th>
<th>Chromosome genotype</th>
<th>Plasmid genotype</th>
<th>β-Galactosidase activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>JM61</td>
<td>lacZ + + +</td>
<td>None</td>
<td>54</td>
</tr>
<tr>
<td>2</td>
<td>JM80</td>
<td>lacZ + - +</td>
<td>None</td>
<td>52</td>
</tr>
<tr>
<td>3</td>
<td>JM90</td>
<td>lacZ - - +</td>
<td>None</td>
<td>58</td>
</tr>
<tr>
<td>4</td>
<td>JM96</td>
<td>lacZ - - +</td>
<td>None</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>JM216</td>
<td>lacZ + - +</td>
<td>None</td>
<td>61</td>
</tr>
<tr>
<td>6</td>
<td>JM61 pRmE43</td>
<td>lacZ + + + nodD1</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>JM61 pRmM157</td>
<td>lacZ + + + nodD2</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>JM61 pRmJ75</td>
<td>lacZ + + + nodD3</td>
<td>105</td>
<td></td>
</tr>
</tbody>
</table>

*a nodD3::Tn5 insertion #303.

* nodD3::Tn5 insertion #801.

The behavior of plasmid-borne **syrm** or **nodD3** was not affected by the presence or absence of other **nodD** genes (Table 5, panels 4 and 5).

**Definition of **nodD3** gene**: The phenotypes and regulatory effects of the different **nodD3** insertions were correlated with their likely position within **nodD3** by a variety of criteria. The insertions in JM88 and JM207, #303 and #709 respectively, are near the left border of a 1.6-kb BgII fragment (Figure 2). The insertion in JM204, #801, is near the middle of the same fragment. The **nodD1** homologous **ClaI** and **BgII** fragments were subcloned into pUC119 in *Escherichia coli* (Figure 5) and protein extracts were prepared from these strains. The extracts were separated by SDS polyacrylamide gel electrophoresis, electroblotted onto nitrocellulose, and hybridized with polyclonal anti-NodD antiserum (FISHER et al. 1988).

The Western blot shown in Figure 5A indicates that the **nodD3** coding sequence is located between the leftmost **ClaI** and rightmost **BgII** sites (Figure 5B). A product the size of authentic NodD3 (FISHER et al. 1988) was seen in extracts from strains in which the vector promoter directs transcription from the left of the **BgII** and **ClaI** fragments (Figure 5A, lanes 1 and 3). No such product was detected in the extract from the strain with a non-coding insert vector alone (lane 4) or in the extract from the strain with the vector promoter reading from right to left across the **BgII** fragment (lane 2). This indicates that the open reading frame for **nodD3** reads from left to right.

**nodD3** coding region is bounded by the **ClaI** site on one end and **Tn5** insertion #1005, 1.6 kb to the right. Based on the apparent molecular weight of the **nodD3** gene product (FISHER et al. 1988) the coding sequence is probably slightly longer than the 1-kb **nodD1** coding sequence (EgELHOFF and LONG 1985). Insertion #801, which is internal to this 1.6 kb region, is thus likely to interrupt the gene. The two insertions with the less pronounced regulatory effects, #303 and #709, are probably not in the **nodD3** coding sequence, although as they are upstream of **nodD3**, they may interrupt or interfere with its transcription. This is consistent with nodulation assays which suggest that insertion #303 allows residual expression of **nodD3**: JM90, a strain with insertion #303 and also insertions in **nodD1** and **nodD2**, nodulates *R. meliloti* at low frequency (Figure 6). HONMA and AUSUBEL (1987) have shown that strains with insertions in all three of the **nodD** genes are **Nod**

**The interaction of **syrm** and **nodD3**: Two plasmids were constructed to test the interdependence of **nodD3** and **syrm**: one carries only **syrm** (pRmM113), and one carries only **nodD3** (pRmM186). In a strain with an intact genomic copy of **nodD3**, the **syrm** plasmid causes high uninduced expression (50 or more units); the luteolin or seed wash induced expression is twofold higher, about 100–200 units (Table 5, panel 5, lines 1–3 and line 6). However, in the strains lacking an intact **nodD3** gene (Table 5, lines 4 and 5), the **syrm** plasmid had almost no effect on **nodC** expression (compare panel 5 with panel 1). Similarly, the **nodD3** plasmid caused high, constitutive expression of **nodC** when **syrm** was intact but had almost no effect in a **syrm**

**background (Table 5, panel 4, compare line 1 and line 6). The behavior of plasmid-borne **syrm** or **nodD3** was not affected by the presence or absence of other **nodD** genes (Table 5, panels 4 and 5).
Expression of *syrM* was also required for *nodD3*-dependent nodulation. *R. meliloti* strains which lack *nodD1*, *nodD2* and *nodD3* are *Nod*⁻ (HONMA and AUSUBEL 1987); we observed low nodulation for strain JM90, which we ascribe to the leakiness of *nodD3::Tn5* dependent nodulation.

### Table 5

<table>
<thead>
<tr>
<th>Line No.</th>
<th>Strain</th>
<th>D1</th>
<th>D2</th>
<th>D3</th>
<th>M</th>
<th><em>nodC-lacZ</em> activity</th>
<th>Plasmid, genotype, and <em>nodC-lacZ</em> fusion activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>JM57</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>3 15 15</td>
<td>None 1</td>
</tr>
<tr>
<td>2</td>
<td>JM81</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>3 5 4</td>
<td>nodD1</td>
</tr>
<tr>
<td>3</td>
<td>JM85</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>3 7 15 3 94 81 14 19 9 58</td>
<td>3 87 75 16 9 67</td>
<td>107 113 105 51 150 120</td>
</tr>
<tr>
<td>4</td>
<td>JM88</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>3 50 18 3 97 89 17 10 60</td>
<td>125 137 NT 7 14 11</td>
</tr>
<tr>
<td>5</td>
<td>JM204</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>5 84 17 NT NT NT NT NT NT</td>
<td>155 155 NT 8 57 16</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>JM85</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>3 22 18 3 62 45 16 10 61</td>
<td>9 18 11 79 140 121</td>
</tr>
<tr>
<td>7</td>
<td>JM201</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>4 110 20 5 111 138 17 12 25</td>
<td>NT NT NT NT NT</td>
</tr>
</tbody>
</table>

* Numbers represent β-galactosidase assay repeated at least twice. NT = not tested.

* D1, D2, D3, M as in Table 2.

* Lut = luteolin.

* sw = alfalfa seed wash.

DISCUSSION

Of the *Rhizobium* nodulation genes identified to date, only the *nodD* genes have a well characterized function: regulation of the expression of the inducible *nod* genes. Most of the species of *Rhizobium* tested have two or three copies of sequences which are highly homologous to *nodD* (RODRIGUEZ-QUINONES et al. 1987). Mutations in *nodD* are *Nod*⁻ in *R. leguminosarum* bv. *trifolii* and *R. leguminosarum* bv. *viciae* but are *Nod*⁺ in *R. meliloti* (DOWNIE et al. 1984; SCHOFIELD and WATSON 1985; JACOBS, EGGELHOFF and LONG 1985; EGGELHOFF et al. 1985). This difference is postulated to occur because the two *R. leguminosarum* strains each have a single copy of *nodD*, but *R. meliloti* harbors three *nodD* homologs (PUTNOKY and KONDOROSI 1986; HONMA and AUSUBEL 1987). Each of the three homologous *R. meliloti* *nodD* genes will allow some nodulation in the absence of the other two (HONMA and AUSUBEL 1987; GYORGYPAL, IYER and KONDOROSI 1988), although strains which have only *nodD2* cause delayed and poor nodulation on alfalfa and are *Nod*⁻ on *Melilotus alba* (GÖTTFERT et al. 1986; HONMA and AUSUBEL 1987). The copy of *nodD* adjacent to *nodABC* in *R. meliloti* is referred to as *nodD1* and the other two loci are referred to as *nodD2* and *nodD3*. Strains which lack all three *nodD* homologs are *Nod*⁻ (HONMA and AUSUBEL 1987).

We found that regulation of the *nod* genes may be coordinated with the regulation of another set of symbiotic genes, those required for exopolysaccharide synthesis. One gene, *syrm*, can increase the expression of the *nod* genes in conjunction with *nodD3*, and also can increase the expression or activity of the *exo* genes in conjunction with another locus, *syrA*. Insertions in *syrm* or *nodD3* have little effect on nodulation in a wild-type strain (SWANSON et al. 1987) but completely block nodulation in an otherwise wild-type strain which lacks *nodD1* and *nodD2* (Figure 6). In the presence of *syrm*, *nodD3* induces sufficient *nod* gene expression to allow normal nodulation. Together, *syrm* and *syrA* may induce the expression of *exo* genes. Plasmid pRm[JM5], carrying *syrm* and *syrA*, can suppress the *Exo*⁻ phenotype of Tn5 insertions in two different *exo* genes which are probably regulatory (J. MULLIGAN and S. LONG, unpublished observations; J. LEIGH and G. WALKER, personal communications). This suggests that, like the *nod* genes, expression of the *exo* gene functions can be controlled by more than one locus.

The data presented here indicate that all three copies of *nodD* in *R. meliloti* are functional as regulators, and at least *nodD1* and *nodD3* affect the same promoters (Figure 3). Each copy is capable of affecting the expression of the *nod* genes when it is overproduced (Table 2), and insertions in any copy will alter the expression of a *nodC-lacZ* fusion (Table 3).

The regulatory activities of the three *nodD* homologs are consistent with the nodulation phenotypes of their mutants. Enhanced expression of *nodD1* or *nodD3* dramatically affects *nodC* expression, and strains which lack both of these genes nodulate very poorly. Strains with insertions in *nodD2* display a
gene products may induce the nod genes in situ in response to conditions which are not imitated by our induction assays.

The physiological significance of this pattern of two or three regulatory genes with similar functions in *R. meliloti* is not yet clear, particularly in light of the fact that many *R. leguminosarum* biovars have only one copy. Gößfert et al. (1986) proposed that *nodD* binds nod boxes and that *R. meliloti* requires more copies of the gene because it has six nod boxes and *R. leguminosarum* has only three. Another possibility is that several copies of *nodD* may allow recognition of a broader range of plant compounds than a single copy. Several recent reports indicate that *nodD* copies from different species determine the inducer specificity (Spaink et al. 1987b; Gößfert et al. 1986; Burn and Johnston 1987). With three *nodD* homologs *R. meliloti* may be able to respond to a wider range of plant compounds than those species that have only a single copy of *nodD*. The multiple homologs of *nodD* may each react to a given spectrum of compounds encountered by *R. meliloti* during symbiosis with a variety of its host plants (Gyorgypal, Iyer and Kondorosi 1988).

Recent studies provide direct evidence that the *nodD1* gene product, NodD1, binds DNA. Hong, Burn and Johnston (1987) found that extracts of *R. leguminosarum* bound to nod gene promoter DNA, as assayed by shift of DNA electrophoretic mobility; the ability of these extracts to interact with promoters correlated with the *nodD* genotype of the cells. Fisher et al. (1988) observed this behavior both with *R. meliloti* cell extracts and with substantially purified NodD1 and NodD3 proteins. They found that both NodD1- and NodD3-promoter interactions could be competed by a nod-box duplex oligomer. These results strongly point to a direct interaction of NodD1 and NodD3 with the nod box and possibly other DNA.
However, neither the degree of binding of the two proteins to nod promoters nor their footprints on promoter DNA is affected by the presence of inducer (FISHER et al. 1988; R. FISHER and S. LONG, manuscript submitted for publication).

As accompaniment to a discussion of possible nodD mechanisms, it is useful to restate several generalizations from our observations of nodD behavior. First, the three nodD genes have distinguishable properties with respect to activating inducible nod gene operons; second, a nodD gene on a plasmid has a dominant effect; and third, the products of the three nodD genes can, at least in some constructs, interfere with each other’s actions. A general type of model which can account for these observations incorporates the idea of competition of nodD protein protomers for a site. Formally, this could be envisioned as competition of nodD protomers for a binding site on DNA, for multimerization sites on other nodD subunits, or both.

The ability of the chromosomal copies of nodD2 and nodD3 to reduce nodABC induction by NodD1 and luteolin could be explained if all three nodD proteins bind to the nodABC promoters in the presence or absence of their inducers. In this model, the inducer would affect the regulatory protein's ability to stimulate transcription, but would not alter its binding affinity for the promoter. Inactive nodD proteins (those which have not bound an inducer) would be competing with the active proteins (those with bound inducer) for operator binding sites, perhaps at the nod box. The E. coli araC protein, which is similar to sequence to nodD, is both an activator and a repressor, and binds to the same specific sites in the presence and absence of its effector (MARTIN, HUO and SCHLEIFF 1986). As our model would predict, the NodD1 and NodD3 footprints are similar and are unaffected by the presence of inducer (R. F. FISHER and S. R. LONG, unpublished data). Strains which overexpress nodD1 and the strain with mutations in both nodD2 and nodD3 are induced to a similar extent by luteolin; in both cases NodD1 would be the major or only occupant of the activator binding site. In each case, the transcription rate would be proportional to the percentage of NodD1 which has been activated by luteolin.

Overexpression of nodD1 and nodD2 enhances the induction of nodC expression by seed wash (Table 5, panels 2 and 3), which is consistent with the model that the activity of these proteins is affected by plant compounds. The fact that NodD1 can induce nodABC expression in response to both seed wash and purified luteolin (a component of seed wash), while NodD2 only enhances induction by seed wash suggests that NodD1 and NodD2 differ in their specificity for plant compounds. The fact that seed wash does not enhance nodABC expression when nodD1 is mutated (Table 3, line 2 and Table 5, line 2, panel 3) suggests that NodD1 and NodD2 interact.

Overexpression of nodD3 from a plasmid copy of the gene results in very high constitutive expression of nodABC which is unresponsive to plant compounds. syrM is required for nodD3-dependent nodulation (Figure 6) and for the stimulation of nodABC expression but plasmid copies of nodD3 (Table 5, panel 4, lines 1 and 6), but is not required for the apparent competition of NodD3 with NodD1 (Table 3, lines 1, 3 and 7). These effects could be accounted for by postulating that SyrM induces the expression of nodD3 from a low basal level which is SyrM independent. In this addition to the model, when the concentration of NodD3 is high it activates nodABC expression. At low basal concentrations NodD3 would compete with NodD1 for access to the nodABC operator but would be unable to activate expression of the gene. SyrM may normally control the abundance of NodD3 in response to an environmental or developmental cue which we have yet to detect. syrM and nodD3 would together duplicate the function of nodD1, with SyrM responding to the presence of a signal by affecting the transcription of nodD3, and NodD3 concentration controlling the expression of the nod box genes.

Overexpression of NodD2 also enhances the basal level of nodABC expression, although to a much lesser extent than does NodD3. Concentration-dependent activity of the nodD proteins could result if nodABC induction requires a protein-protein interaction which could be driven by increased protein concentration, such as NodD dimerization or cooperative binding at two sites on the promoter.

Our observation that the various nodD genes display different activities with respect to the presence or absence of luteolin or crude seed wash is consistent with the observations and proposals that nodD genes determine the spectrum of compounds active as inducers (HONMA and AUSUBEL 1987; SPAINK et al. 1987b; GYORGYPAL, IVER and KONDOROSI 1988). However, the biochemical mechanism responsible for the complex interactions of the nodD alleles with each other and with inducers remains unknown. Detailed knowledge of nodD protein structure and action is essential to gain an understanding of these regulatory behaviors.

We are grateful to THOMAS EGELHOFF for purification of NodD1 protein and for obtaining anti-NodD serum; to JEAN SWANSON for strain construction and mapping; to ROBERT FISHER, PHILIPPA WEBSTER and STEVEN ELLEDGE for a critical reading of the manuscript and to A. BLOOM for its skillful preparation. We also thank M. HONMA, F. AUSUBEL and G. WALKER for useful discussions and for providing data prior to publication. This work was supported by the National Institute of Health research grant GM30962-06 and Department of Energy contract (AS03-82-ER12084) to S.R.L.

J.T.M. was supported by a NIH Cell and Molecular Biology Training Grant to Stanford University.
LITERATURE CITED

APPELBAUM, E. R., D. V. THOMPSON, K. IDLER and N. CHARTRAIN, 1988 *Rhizobium japonicum* USDA 191 has two nodD genes that differ in primary structure and function.


J. T. Mulligan and S. R. Long


Communicating editor: G. MOSIG