JOURNAL OF BACTERIOLOGY, July 1990, p. 3695–3700
0021-9193/90/073695-06$02.00/0

Vol. 172, No. 7

DNA Sequence and Translational Product of a New Nodulation-Regulatory Locus: SyrM Has Sequence Similarity to NodD Proteins

MELANIE J. BARNETT AND SHARON R. LONG*

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

Received 8 December 1989/Accepted 25 April 1990

***Rhizobium meliloti* nodulation (*nod*) genes are expressed when activated by *trans*-acting proteins in the NodD family. The *nodD1* and *nodD2* gene products activate *nod* promoters when cells are exposed to plant-synthesized signal molecules. Alternatively, the same *nod* promoters are activated by the *nodD3* gene when *nodD3* is carried in *trans* along with a closely linked global regulatory locus, *syrM* (symbiotic regulator) (J. T. Mulligan and S. R. Long, Genetics 122:7–18, 1989). In this article we report the nucleotide sequence of a 2.6-kilobase *Sph*I fragment from *R. meliloti* SU47 containing *syrM*. Expression from this locus was confirmed by using in vitro transcription-translation assays. The open reading frame encoded a protein of either 33 or 36 kilodaltons whose sequence shows similarity to NodD regulatory proteins.**

The bacterium *Rhizobium meliloti* forms nitrogen-fixing nodules on certain members of the genera *Medicago*, *Trigonella*, and *Melilotus*. The bacteria invade the host plant through root hairs, stimulate cortical cells to divide, and occupy nodule cells, where the plants and bacteria establish a metabolic symbiosis (28). This process requires bacterial nodulation (*nod*) genes, including *nodA*, *nodB*, and *nodC*, which are required for early events in this sequence, and also *nodE*, *nodF*, and *nodH*, which influence efficiency and host specificity (4, 8, 9, 24, 25, 49).

Regulation of *nod* genes is complex, and each of the two symbiotic partners has a role. Induction of the *nod* genes in *R. meliloti* requires both plant-produced flavones such as luteolin (38) and a regulatory gene, *nodD1* (34, 35). The deduced amino acid sequence of *nodD* shows it to be in the newly defined *lysR* family of positive gene activator proteins (20). NodD1 stimulates *nod* gene expression at the level of transcription, probably through an interaction of NodD protein with a specific and conserved sequence called the *nod* box (11, 12, 42), which is located upstream of many of the *nod* genes, including *nodABC*, *nodFE*, and *nodH*. *R. meliloti* possesses additional, highly conserved copies of *nodD* (*nodD2* and *nodD3*), each of which is functional under certain conditions (15, 21, 22). Activation of gene expression by NodD2 occurs in the presence of plant signal molecules other than luteolin (16, 21). When *nodD3* is carried in *trans* on plasmid pRmJT5, it allows elevated expression of a *nodC-lacZ* fusion in the absence of plant inducer, from the same promoter as that activated by *nodD1* (35). A locus designated *syrM* (symbiotic regulator), which maps between *nodH* and *nodD3* on pRmJT5, is required for this *nodD3*-mediated expression: *syrM nodD1 nodD2* mutants fail to express *nodC-lacZ* constitutively at high levels, nor do they initiate nodule development (35). Bacteria which lack all three copies of *nodD* but not *syrM* are also Nod$^-$ (22).

The genetic constructs which revealed the *syrM-nodD3* regulatory circuit also revealed a second activity of *syrM*: bacteria carrying plasmid pRmJT5 form unusually mucoid colonies (35). Transposon Tn*5* insertions in *syrM* abolish the enhanced mucoid effect. Tn*5* insertions in another locus on pRmJT5, *syrA*, also suppress the mucoid phenotype but do not affect *nodD3*-mediated expression of the *nodC-lacZ* fusion (35). The *syrM-syrA* mucoid phenotype acts through the known *exo* genes *exoA*, *exoB*, and *exoE* and perhaps others (35). Therefore, *syrM* acts in *trans* to stimulate *nod* gene expression via *nodD3* and *exo* gene expression via *syrA* (35).

In this study we determined the DNA sequence of *syrM* and verified expression from the *syrM* locus by in vitro transcription-translation assays. We report that the SyrM amino acid sequence shares similarity to the sequence of NodD proteins and to other proteins in the LysR family.

MATERIALS AND METHODS

Plasmids and strains. pRmT17 is pBR322 with the 5.5-kilobase (kb) *Cla*I fragment of *R. meliloti* 1021 Sym plasmid. pMB1 and pMB2 consist of the 2.6-kb *Sph*I fragment of pRmT17 cloned in both orientations into pUC119. *Escherichia coli* strains DH5α (17) and JM101 (32) were grown as described before.

DNA sequencing. Both strands of the 2.6-kb insert of pMB1 and pMB2 were sequenced by dideoxy-chain termination (43) with a modified T7 polymerase (Sequenase; U.S. Biochemical Corp.). Exonuclease III (19) was used to generate a set of nested deletions. pMB1 and pMB2 were digested with *Kpn*I and *Bam*HI so that the 2.6-kb insert would be progressively deleted from the *Bam*HI site. Single-stranded DNA was prepared as described previously (13) with the aid of helper bacteriophage M13K07 (50). Overlapping deletions were sorted by using the DNA sequence analysis program SEQSORT (8). Analysis of the completed sequence was done with the University of Wisconsin Genetics Computer Group software (5) and the TULLA program (48). Tn*5* insertions were subcloned, and their locations were determined by sequencing from a primer specific for the Tn*5* end.

In vitro transcription-translation. Single-stranded DNA from selected exonuclease III-generated deletions was transformed into DH5α. Double-stranded DNA was isolated and purified on a CsCl gradient, and 1 μg of each of the resulting DNAs was incubated in an *R. meliloti* S-30 extract (13) at 30°C for 70 min. Radiolabeled proteins were electrophoresed on 12.5% polyacrylamide gels (27) and visualized by autoradiography.

Southern analysis. A 1-μg amount of bacterial DNA (a gift of Jean Swanson) from each strain was digested to comple-

* Corresponding author.

tion and resolved on a 1% agarose gel. DNA was transferred to a Zeta-Probe (Bio-Rad Laboratories) membrane by standard procedures (30). Hexamer-labeled probe was prepared by the method of Feinberg and Vogelstein (10), with gel-purified DNA fragments as the template. Hybridization was performed at 65°C for 16 h in 0.5 M NaH_2PO_4 (pH 7.2)–1 mM EDTA–7% sodium dodecyl sulfate (SDS)–1% bovine serum albumin (3). One wash in low-stringency buffer (40 mM NaH_2PO_4 [pH 7.2], 1 mM EDTA, 5% SDS, 0.5% bovine serum albumin) was conducted at room temperature and was followed by two washes in high-stringency buffer (40 mM NaH_2PO_4 [pH 7.2], 1 mM EDTA, 1% SDS). Autoradiographs were exposed at −80°C with an intensifying screen.

RESULTS AND DISCUSSION

A set of Tn*5* insertions in pRmJT5 (49) genetically defined the *syrM* locus (Fig. 1) (35). The location of the *syrM* gene, subcloned on a 2.6-kb *Sph*I fragment, is shown in Fig. 2B. Both strands of this *Sph*I insert were sequenced. A search for open reading frames (ORFs) in the 2.6-kb *Sph*I fragment revealed only one ORF of reasonable length, shown in Fig. 2C. The nucleotide sequence containing this ORF is presented in Fig. 1. The DNA sequence begins 850 base pairs (bp) from the rightmost *Sph*I site of Fig. 2B. The ORF, which terminated at bp 1119 (Fig. 1), had ATG codons at positions 142 and 217. If these ATG codons were used as translation initiation sites, the ORFs would encode polypeptides of 35,952 and 33,063 daltons (Da), respectively. A Shine-Dalgarno-like sequence was present upstream of the first ATG codon but not the second. However, this may not predict the translation start site, since most *Rhizobium nod* genes do not have a Shine-Dalgarno consensus and the corresponding rRNA sequence has not been reported. The direction of transcription of this ORF was opposite that of each of the known flanking loci, *nodH* and *nodD3* (Fig. 2B). The insertion points of Tn*5* mutations 701, 311, and 712, which correlate with the *SyrM*$^-$ phenotype (35), were determined by sequencing the Tn*5*-genome junctions (triangles in Fig. 1). Tn*5* insertion sites are consistent with the positions deduced from restriction mapping (49); all three lay within the putative ORF (Fig. 1).

That this putative ORF does indeed encode a protein product was confirmed by in vitro transcription-translation analysis with *R. meliloti* S-30 extracts. The autoradiograph in Fig. 2A shows that a protein product which migrated at approximately 35 kDa was obtained when pMB1 or pMB2 DNA was used but not when the pUC119 vector lacking insert was used. Similarly, reaction mixtures to which no DNA was added did not have products migrating at this position (data not shown).

Deletion analysis was used to define the region of the 2.6-kb *Sph*I fragment that allowed expression of the *syrM* product (SyrM). The deletions used were chosen from among the exonuclease III-generated derivatives of pMB1 and pMB2 used for dideoxy sequencing. The restriction map of the 2.6-kb fragment is shown in Fig. 2B along with the truncated clones represented by bold lines 1 through 8. Results obtained with these clones are shown in Fig. 2A, lanes 1 to 8. Expression of the 35-kDa product was seen when deletions 1, 7, and 8 were used as templates for coupled transcription-translation but not when deletions 2 through 6 were used. These results define a region, represented by a hatched line in Fig. 2C, containing the locus necessary for expression of SyrM. These results are consistent with the location of the ORF shown in Fig. 2C by a thin

```
GTACTTCTGATTAACGGAAAATCACGACGCTTCCCATTTGTCGGCGGGCCAGGAAAGTCT 60

TTGGCACGCACACCCATATTGTGACGACCTGGATCTTCGGGATGTGGCGCGGGAACATTA 120

CCTGCCGAGGAGAATAGATCCATGGATCAGCCCACTTGGAAGCGGCCGCATAGGGCCAAA 180
                    M  D  Q  P  T  W  K  R  P  H  R  A  K

TTTGCCGGTGTCAGTGACGCCGCACAACAGCGGCAAATGCCGAACCTTGCGTCCATCGAC 240
F  A  G  V  S  D  A  A  Q  Q  R  Q  M  P  N  L  A  S  I  D

CTAAACCTACTGGTGGATCTTGAGGCTCTGCTGCAGTATCGGCACATTACTCAAGCGGCC 300
L  N  L  L  V  D  L  E  A  L  L  Q  Y  R  H  I  T  Q  A  A

CAGCACGTCGGCCGTAGCCAACCGGCGATGAGCAGAGCCTTGTCAAGGCTGCGCGGCATG 360
Q  H  V  G  R  S  Q  P  A  M  S  R  A  L  S  R  L  R  G  M
          701
TTAAAGGACGATCTTCTGGTTGCCGGCTCTAGAGGCTTGGTCCTGACGCCTCTGGCTGAA 420
L  K  D  D  L  L  V  A  G  S  R  G  L  V  L  T  P  L  A  E
    311
TGCTTGACCCAGATGCTGCCTTCAGTACTAGACGCGATTCGCCAGATGATGAACCTCAGC 480
C  L  T  Q  M  L  P  S  V  L  D  A  I  R  Q  M  M  N  L  S

TTGGCTCCGGCGCAACGGCGATGGAAGGTGACAATGGCTATGCCCGATCACCAAGCGGTG 540
L  A  P  A  Q  R  R  W  K  V  T  M  A  M  P  D  H  Q  A  V

GTTCTGCTGCCACATCTGTTGCCGCGGTTACACGAGCGCGCCCCTCATCTCGACATTGTG 600
V  L  L  P  H  L  L  P  R  L  H  E  R  A  P  H  L  D  I  V

ACCGATCCGCTTTTGGGCGGCGCGCTGGGGTTACTTGAACAAGGTGAGATCGACGTGGTC 660
T  D  P  L  L  G  G  A  L  G  L  L  E  Q  G  E  I  D  V  V

GTCGGGCAGATGGGCGCCGCTCCGCTCGGCTACTTGCGGCGCAGACTCTACGCCGACAGC 720
V  G  Q  M  G  A  A  P  L  G  Y  L  R  R  R  L  Y  A  D  S

TTCACCTGCGTGCTACGCCACAATCACCCGGCCTTGGCGCAGGAGTGGACGATCGAGGCT 780
F  T  C  V  L  R  H  N  H  P  A  L  A  Q  E  W  T  I  E  A
    712                                                  BglII
TTCGCGGCCTTGCGCCACGTCGCCATTGCCTCGGAACCTGACGAGCTTTTCGGTCAGATC 840
F  A  A  L  R  H  V  A  I  A  S  E  P  D  E  L  F  G  Q  I

TATGACAGGTTAACCAAACTCGGACTGCAGCGTGGTGATCCGATGGTGGTTTCCACCGTA 900
Y  D  R  L  T  K  L  G  L  Q  R  G  D  P  M  V  V  S  T  V

CTGACCGCGGCGGTCTTGATCGCGGCGACTGACTCAGTGCTGGTTGTACCAAGCCGCGTG 960
L  T  A  A  V  L  I  A  A  T  D  S  V  L  V  V  P  S  R  V

GCGACCCGTGTCGCCGCCATGCTGTCGCTCGCGGTGATTCCCCCGCCTGTGGAACTGAGA 1020
A  T  R  V  A  A  M  L  S  L  A  V  I  P  P  P  V  E  L  R

CCATACGAAGTCGCGCTGATCTGGCACGAGCGGTGCCATCGAGACCCAGAGCATCGCTGG 1080
P  Y  E  V  A  L  I  W  H  E  R  C  H  R  D  P  E  H  R  W

CTGCGTGGCGAAATCGCCGCCGCAGCGTCCACGGCCGGTTAGACGAAGACTAGCGAATAG 1140
L  R  G  E  I  A  A  A  A  S  T  A  G
```

FIG. 1. DNA sequence of the *syrM* locus. The sequence begins upstream from the start of the putative *syrM* reading frame, 850 bp from the rightmost *Sph*I site as indicated in Fig. 2B. The two possible N-terminal methionine residues are boxed. Tn*5* insertion points that have been sequenced with a primer specific for the ends of Tn*5* are shown with triangles. By convention, the triangle represents the beginning of the 9-bp repeat. Only one endpoint of Tn*5* 712 was determined by DNA sequencing. The nucleotide sequence of the entire 2.6-kb *Sph*I fragment has been submitted to GenBank and has been assigned accession number M33495.

arrow. It is interesting that the amount of SyrM appeared to be reduced in reaction 7 compared with reaction 8. Deletion 8 had the first 171 bases of the *Sph*I fragment deleted, whereas deletion 7 had 786 bases deleted. This suggests that the sequence missing from deletion 7, upstream of the ORF translation start site, is important for efficient expression of SyrM.

The SyrM ORF possessed two translation start codons in proximity to each other. As stated earlier, depending on

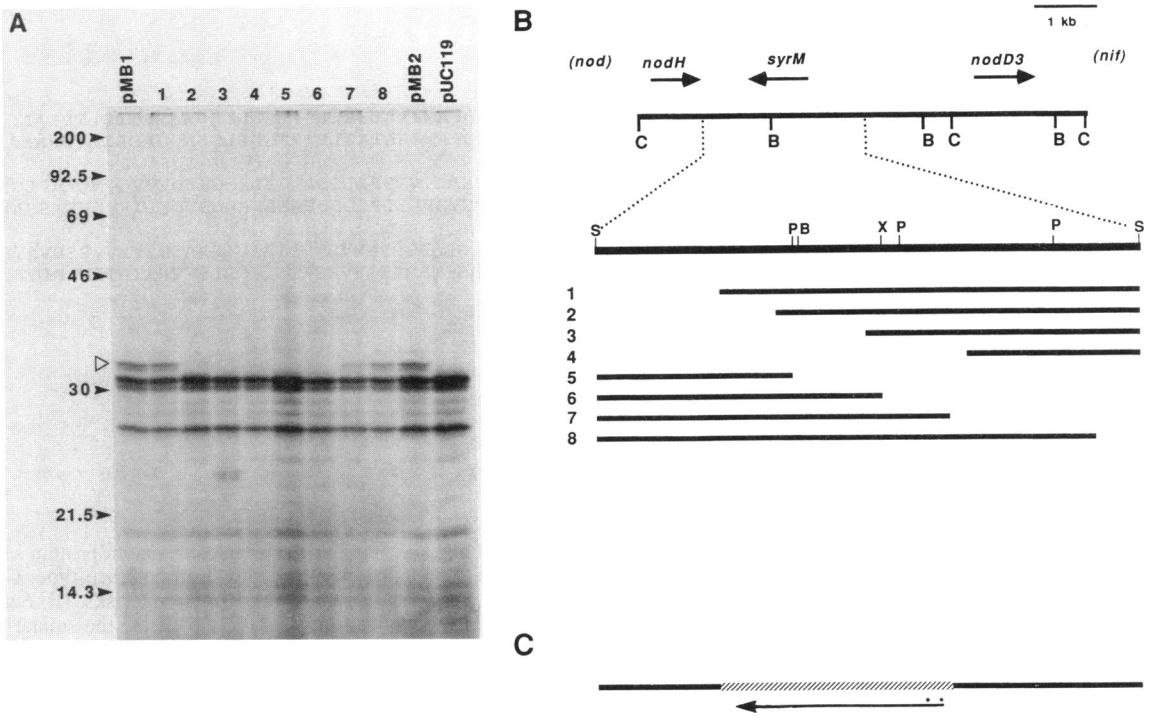

FIG. 2. In vitro expression of the *syrM* product. (A) SDS-polyacrylamide gel electrophoresis of coupled transcription-translation reactions. Assays were performed as stated in Materials and Methods. Solid arrowheads indicate the migration of Rainbow (Amersham) molecular weight markers (mass in kilodaltons). The putative *syrM* product is indicated by an open arrowhead and migrates at about 35 kDa. pMB1 and pMB2 contain the entire 2.6-kb *Sph*I fragment cloned in opposite orientations. Lanes numbered 1 to 8 show the products obtained when the corresponding deletions illustrated in panel B were used. pUC119 was used as a negative control. (B) Map of the 2.6-kb *Sph*I fragment showing the deletions used to determine the region encoding *syrM*. These exonuclease III deletions were sequenced as described in Materials and Methods. Lanes 1 to 8 indicate the amount of DNA remaining in each deletion used in the corresponding in vitro transcription-translation assays shown in panel A. (C) *Sph*I fragment as in panel B. The hatched segment represents the maximum physical limits of the coding region established by deletion analysis. The computer-determined ORF is designated with a thin arrow. The two possible start codons are represented by asterisks. Abbreviations: S, *Sph*I; P, *Pst*I; X, *Xba*I; B, *Bgl*II; C, *Cla*I.

which start site is actually used, SyrM may be either 35.9 or 33.1 kDa. It is difficult to determine unambiguously which ATG represents the true start codon from the results obtained by SDS-polyacrylamide gel electrophoresis. In fact, SyrM migrated at about 35 kDa, intermediate between the two predicted sizes.

The Protein Identification Resource (Release 19.0) was searched with Genetics Computer Group software (5) and FASTA-MAIL (37) for proteins with amino acid sequences similar to the SyrM sequence. It is interesting that SyrM showed significant similarity to NodD protein. NodD proteins have been characterized from a wide variety of *Rhizobium* species and show great similarity to each other. The similarity of SyrM to NodD1 from *R. meliloti* is shown in Fig. 3A. Identical and highly conserved amino acids are in boldface. Whereas NodD1, NodD2, and NodD3 from *R. meliloti* show an amino acid sequence conservation of greater than 80% (15, 21; B. Rushing and S. R. Long, unpublished data), SyrM and NodD1 are about 30% similar, including conservative substitutions. NodD1 and SyrM sequence similarity seemed to be clustered in the N-terminal regions of the polypeptides. NodD1 and NodD3 bind *nod* box sequences (42) as demonstrated by gel retardation assays (12) and DNA footprinting (11). The N-terminal region of NodD has been proposed to contain a DNA-binding domain, based on similarities to *cro*-like DNA-binding proteins (20).

The amino-terminal region of SyrM was analyzed with the same frequency matrix (6) used in the Henikoff et al. study (20). This analysis assigns SyrM a 60% probability of having a helix-turn-helix motif in the N-terminal region, a greater probability than was determined for NodD. The above results and additional sequence analyses place SyrM in the LysR (46) family of DNA-binding proteins (20). Other members of this family include NahR (44), MetR (39), NodD (8), OxyR (2), CysB (36), LeuO (18, 20), IlvY (51), AmpR (23), AntO (14, 29), TrpI (1), and TfdO (47). Results obtained from TULLA alignments of SyrM and LysR to several of these LysR family members are shown in Fig. 3B. TULLA is a program designed to perform multiple sequence alignments that best reproduce the three-dimensional structural alignments determined by X-ray crystallography (48). The matrix of relatedness used for this comparison was that of McLachlan (31).

Each number in Fig. 3B represents the standard deviation above the random mean, i.e., how the similarity score for each pair compares to a similarity score where the sequence of one member of the pair is randomized. These data indicate that SyrM has significant similarity to LysR, NahR, LeuO, IlvY, and AmpR. The homology between SyrM and NahR (44) is especially significant, the standard deviation being slightly higher than that of the NodD1-SyrM alignment. NahR is a positive regulatory protein that activates the genes necessary for utilization of naphthalene and salicylate by *Pseudomonas putida* (45). Other members of the LysR family, such as CysB and TrpI, are not significantly similar

A

```
                                                    **********
SyrM ---------------MPNLASIDLNLLVDLEALLQYRHITQAAQHVGRSQPAMSRALSRLRGMLKDDLLVAGSRGLVLTPLAECLTQMLP
NodD1 -------------MPHMRFRGLDLNLLVALDALMTERKLTAAARRINLSQPAMSAAIARLRTYFGDELFSMQGRELIPTPRAEALAPAVR

       SVLDAIRQMM-N-LSLAPAQRRWKVTMAMPDHQAVVLLPHLLPRLHERAPHLDIVTDPLLGGALGLLEQGEIDVVVGQMGAAPLGYLRRRL
       DALLHIQLSVIAWDPLNPAQSDRRFRIILSDFMILVFFARIVERVAREAPGVSFELLPLDDDPHELLRRGDVDFLIFPDVFMSSAHPKAKL

       YADSFTCVLRHNHPALAQEWTIEAFAALRHVAIASEPDELFGQIYDRLTKLGLQRGDPMVVSTVLTAAVLIAATDSVLVVPSRVATRVAAML
       FDEALVCVGCPTNKKLLGNISFETYMSMGHVAAQFGREMKPSVEQWLLLEHGFNRRIELVVPGFTLIPRLLSGTNRIATLPLRLVKYFEQTI

       SLAVIPPPVELRPYEVALIWHERCHRDPEHRWLRGEIAAAASTAG------
       PLRIVTSPLPPLFFTEAIQWPALHNTDPGNIWLREILLQEASRIDPQSDTC
```

B

	LysR	NodD1	NahR	LeuO	IlvY	TrpI	AmpR	CysB
SyrM	7.54	24.72	25.51	9.07	10.55	4.45	7.09	1.94
LysR	---	8.61	7.75	7.52	11.80	6.40	8.13	10.04

FIG. 3. Similarity of SyrM to members of the LysR-NodD family. (A) The TULLA program (48) was used to match protein sequences. SyrM is numbered starting at the second methionine. Identical and highly conserved amino acids are shown in boldface type. Conserved groupings are as follows: Ile, Val, Leu; Arg, Lys; Glu, Asp. A possible helix-turn-helix motif (6) is indicated by asterisks. (B) Significance of SyrM similarity to LysR family members. SyrM was compared with selected members of the LysR family by using the similarity matrix of McLachlan (31) as contained within the TULLA program (48). Pairings with LysR are shown for comparison. Each number represents a standard deviation (sigma) above the random mean. Sigma values must be at least 3 to 5 to be considered significant.

to SyrM. This is consistent with current theories that the LysR family may contain one or more subfamilies (20, 44) whose members share greater similarity. Biochemical experiments are now needed to confirm that SyrM is a positive regulatory protein and to characterize its DNA target(s).

The common *nod* genes *nodDABC* have been shown to be highly conserved among various *Rhizobium* species (40). Therefore, we wanted to determine whether homologs of *syrM* are present in different *Rhizobium* species. Genomic DNAs from several species were separately digested with *Bam*HI and *Cla*I, transferred to nylon membrane, hybridized with a 448-bp *Xba*I-*Bgl*II fragment wholly internal to the *syrM* coding sequence, and washed at low stringency, as described in Materials and Methods. Figure 4 shows that the probe hybridized to the expected 7.8-kb *Bam*HI (the size was determined from a lighter exposure of the same blot) and 5.5-kb *Cla*I fragments of *R. meliloti* 1021 DNA. In addition, a 3-kb *Bam*HI fragment and large *Cla*I fragment showed a lesser degree of hybridization. This may indicate the presence of a second *R. meliloti* locus having some sequence similarity to *syrM*. The *syrM* probe hybridized with low intensity to three *Bam*HI fragments and several *Cla*I fragments of DNA from ANU843, an *R. leguminosarum* biovar *trifolii* strain. Similarly, the *syrM* probe hybridized weakly to several *Cla*I fragments and more strongly to at least one *Bam*HI fragment of the two *R. leguminosarum* bv. *viciae* strains p8401(pRL1) and RL300 and to digests of ANU265, a Nod$^-$, pSym-cured derivative of a broad-host-range *Rhizobium* species. *NodDABC*, which are well conserved at the amino acid and DNA sequence levels among many *Rhizobium* species, cross-hybridize with *R. meliloti*-derived probes at somewhat higher stringencies (40).

We wondered whether the *syrM* probe cross-hybridized to *nodD* sequences because the hybridizations were done at such low stringency. This is unlikely for several reasons. First, the 448-bp probe does not include the N-terminal region which contains much of the *syrM-nodD* conserved sequence and has only a 36% similarity to *nodD1* at the DNA sequence level. Second, we observed that the *syrM* probe did not hybridize to fragments containing *nodD1* or *nodD3* in nanogram amounts (data not shown), even at the same low stringencies used for the genomic Southern blots. Finally, the size of the faint 3-kb *Bam*HI fragment seen in 1021 did not match that expected for *nodD1* or *nodD3* based on the known *R. meliloti* restriction map. Since cross-hybridization between the *syrM* probe and *nodD1* of *R. meliloti* did not occur, we conclude it is unlikely that cross-reaction occurs

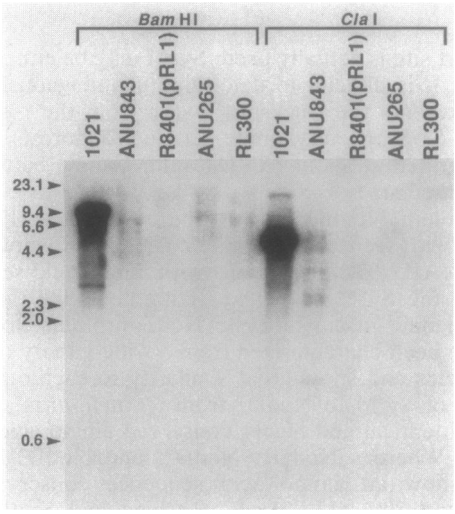

FIG. 4. Hybridization of *syrM* probe to *Rhizobium* genomic DNA. Genomic DNA from various *Rhizobium* strains was digested separately with *Bam*HI or *Cla*I, transferred to a membrane, and hybridized, as described in Materials and Methods. The *syrM*-specific probe is a 448-bp *Xba*I-*Bgl*II fragment (see map in Fig. 2). Strain 1021 is an Str^r derivative of *R. meliloti* SU47. ANU843 is an *R. leguminosarum* bv. *trifolii* strain (41). ANU265 is a broad-host-range *Rhizobium* strain that has been cured of its Sym plasmid (33). RL300 (26) and R8401(pRL1) (7) are *R. leguminosarum* bv. *viciae* strains. Molecular sizes (in kilobases) of lambda *Hin*dIII markers are indicated by arrows.

between the *syrM* probe and *nodD2* or with the *nodD* genes of other species. We also found by Southern analysis that two additional *R. meliloti* strains, Rm41 (provided by A. Kondorosi) and 102F51 (Nitragin Corp.), have sequences homologous to *syrM* (data not shown). To determine whether any of the detected homologies correspond to a *syrM*-like gene, it will be necessary to clone the homologous segments of DNA for use in functional tests.

We determined the nucleotide sequence of *syrM* and confirmed by in vitro transcription-translation that it encodes a protein of either 33 or 36 kDa. SyrM shows similarity to the amino acid sequence of *Rhizobium* NodD proteins and to several members of the LysR family of bacterial activators (20). The locations of the Tn*5* insertions previously analyzed by the genetic level (49) have been confirmed to lie within the *syrM* ORF. These insertions were shown in another study to abolish *nodD3*-mediated expression of the *nod* genes as well as *syrA*-mediated expression of the *exoA*, *exoB*, and *exoE* genes (35).

The function of *syrM* in nodulation is still largely a puzzle. The *syrM* locus is not absolutely required for nodulation of alfalfa in our standard nodulation assays (*syrM nodD3* mutant strains are Nod^+). Perhaps *syrM* is important in some manner not detected in our plant assays, for example, in competitiveness or host range specificity. Additionally, there may exist a protein with redundant function that is not highly conserved at the DNA sequence level and therefore not readily detectable by Southern analysis. There is some evidence showing how *syrM* may interact with *nodD3*. Western immunoblot experiments show that *syrM* regulates the amount of NodD3 protein and that extracts from strains containing *syrM* retard *nodD3* promoter fragments in a gel mobility shift assay, whereas extracts from strains lacking *syrM* do not (Yelton et al., unpublished data). Further experiments will be needed to determine how and when expression of *syrM* is mediated.

ACKNOWLEDGMENTS

We are indebted to R. Fisher, B. Rushing, and M. Yelton for critical reading of this manuscript and to A. Bloom for its skillful preparation. We thank M. Yelton, T. Egelhoff, and J. Schwedock for materials and advice on transcription-translation experiments and J. Schwedock for advice on DNA sequencing. We are grateful to J. Mulligan for DNA sequence of the Tn*5* insertion points and to J. Swanson for the *Rhizobium* genomic DNAs.

This research was supported by Public Health Service grant RO1-GM30962 from the National Institutes of Health to S.R.L. M.J.B. was supported by a Training Grant in Cell and Molecular Biology from the National Institutes of Health to Stanford University.

LITERATURE CITED

1. **Chang, M., A. Hadero, and I. P. Crawford.** 1989. Sequence of the *Pseudomonas aeruginosa trpI* activator gene and relatedness of *trpI* to other procaryotic regulatory genes. J. Bacteriol. **171**:172–183.
2. **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.
3. **Church, G. M., and W. Gilbert.** 1984. Genomic sequencing. Proc. Natl. Acad. Sci. USA **81**:1991–1995.
4. **Debellé, F., F. Maillet, J. Vasse, C. Rosenberg, F. De Billy, G. Truchet, J. Dénarié, and F. M. Ausubel.** 1988. Interference between *Rhizobium meliloti* and *Rhizobium trifolii* nodulation genes: genetic basis of *R. meliloti* dominance. J. Bacteriol. **170**:5718–5727.
5. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
6. **Dodd, I. B., and J. B. Egan.** 1987. Systematic method for the detection of potential *cro*-like DNA-binding regions in proteins. J. Mol. Biol. **194**:557–564.
7. **Downie, J. A., C. D. Knight, A. W. B. Johnston, and L. Rossen.** 1985. Identification of genes and gene products involved in the nodulation of peas by *Rhizobium leguminosarum*. Mol. Gen. Genet. **198**:255-262.
8. **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.
9. **Faucher, C., F. Maillet, J. Vasse, C. Rosenberg, A. A. N. Van Brussel, G. Truchet, and J. Dénarié.** 1988. *Rhizobium meliloti* host range *nodH* gene determines production of an alfalfa-specific extracellular signal. J. Bacteriol. **170**:5489–5499.
10. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
11. **Fisher, R., and S. R. Long.** 1989. DNA footprint analysis of the transcriptional activator proteins NodD1 and NodD3 on inducible *nod* gene promoters. J. Bacteriol. **171**:5492–5502.
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., 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.
14. **Goldberg, E. B., T. Arbel, J. Chen, R. Karpel, G. A. Mackie, S. Schuldiner, and E. Padan.** 1987. Characterization of a Na^+/H^+ antiporter gene of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **84**:2615–2619.
15. **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. f
16. **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.
17. **Hanahan, D.** 1985. Techniques for transformation of *E. coli*, p. 109–114. *In* D. M. Glover (ed.), DNA cloning, vol. I. A practical approach. IRL Press, Oxford.
18. **Haughn, G. W., S. R. Wessler, R. M. Gemmill, and J. M. Calvo.** 1986. High A+T content conserved in DNA sequences upstream of *leuABCD* in *Escherichia coli* and *Salmonella typhimurium*. J. Bacteriol. **166**:1113–1117.
19. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28**: 351–359.
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. A., M. Asomaning, and F. M. Ausubel.** 1990. *Rhizobium meliloti nodD* genes mediate host-specific activation of *nodABC*. J. Bacteriol. **172**:901–911.
22. **Honma, M. A., and F. M. Ausubel.** 1987. *Rhizobium meliloti* has three functional copies of the *nodD* symbiotic regulatory gene. Proc. Natl. Acad. Sci. USA **84**:8558–8562.
23. **Honoré, N., M. H. Nicolas, and S. T. Cole.** 1986. Inducible cephalosporinase production in clinical isolates of *Enterobacter cloacae* is controlled by a regulatory gene that has been deleted from *Escherichia coli*. EMBO J. **5**:3709–3714.
24. **Horvath, B., E. Kondorosi, M. John, J. Schmidt, I. Torok, L. Z. Györgypal, 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.
25. **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.
26. **Johnston, A. W. B., and J. E. Beringer.** 1975. Identification of the *Rhizobium* strains in pea root nodules using genetic markers. J. Gen. Microbiol. **87**:343–350.
27. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head 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. **Mackie, G. A.** 1986. Structure of the DNA distal to the gene for ribosomal protein S20 in *Escherichia coli* K12: presence of a strong terminator and an IS*1* element. Nucleic Acids Res. **14**:6965–6981.
30. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
31. **McLachlan, A. D.** 1971. Tests for comparing related amino acid sequences. Cytochrome *c* and cytochrome c_{551}. J. Mol. Biol. **61**:409–424.
32. **Messing, J., R. Crea, and P. H. Seeburg.** 1981. A system for shotgun DNA sequencing. Nucleic Acids Res. **9**:309–321.
33. **Morrison, N. A., C. Y. Hau, M. J. Trinick, J. Shine, and B. G. Rolfe.** 1983. Heat curing of a Sym plasmid in a fast-growing *Rhizobium* sp. that is able to nodulate legumes and the nonlegume *Parasponia* sp. J. Bacteriol. **153**:527–531.
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. **Ostrowski, J., G. Jagura-Burdzy, and N. M. Kredich.** 1987. DNA sequences of the *cysB* regions of *Salmonella typhimurium* and *Escherichia coli*. J. Biol. Chem. **262**:5999–6005.
37. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85**:2444–2448.
38. **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.
39. **Plamann, L. S., and G. V. Stauffer.** 1987. Nucleotide sequence of the *Salmonella typhimurium metR* gene and the *metR-metE* control region. J. Bacteriol. **169**:3932–3937.
40. **Rodriguez-Quinones, F., Z. Banfalvi, P. Murphy, and A. Kondorosi.** 1987. Interspecies homology of nodulation genes in *Rhizobium*. Plant Mol. Biol. **8**:61–75.
41. **Rolfe, B. G., P. M. Gresshoff, and J. Shine.** 1980. Rapid screening for symbiotic mutants of *Rhizobium* and white clover. Plant Sci. Lett. **19**:277–284.
42. **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.
43. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
44. **Schell, M., and M. Sukordhaman.** 1989. Evidence that the transcription activator encoded by the *Pseudomonas putida nahR* gene is evolutionarily related to the transcription activators encoded by the *Rhizobium nodD* genes. J. Bacteriol. **171**:1952–1959.
45. **Schell, M. A., and P. Wender.** 1986. Identification of the *nahR* gene product and nucleotide sequences required for its activation of the *sal* operon. J. Bacteriol. **166**:9–14.
46. **Stragier, P., and J.-C. Patte.** 1983. Regulation of diaminopimelate decarboxylase synthesis in *Escherichia coli*. III. Nucleotide sequence and regulation of the *lysR* gene. J. Mol. Biol. **168**: 333–350.
47. **Streber, W. R., K. N. Timmis, and M. H. Zenk.** 1987. Analysis, cloning, and high-level expression of 2,4-dichlorophenoxyacetate monooxygenase gene *tfdA* of *Alcaligenes eutrophus* JMP134. J. Bacteriol. **169**:2950–2955.
48. **Subbiah, S., and S. C. Harrison.** 1989. A method for multiple sequence alignment with gaps. J. Mol. Biol. **209**:539–548.
49. **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.
50. **Vieira, J., and J. Messing.** 1987. Production of single-stranded plasmid DNA. Methods Enzymol. **153**:3–11.
51. **Wek, R. C., and G. W. Hatfield.** 1986. Nucleotide sequence and *in vivo* expression of the *ilvY* and *ilvC* genes in *Escherichia coli* K12. J. Biol. Chem. **261**:2442–2450.