Molecular Microbiology (1990) 4(2), 245–252

# Molecular characterization of the nodulation gene, *nodT*, from two biovars of *Rhizobium leguminosarum*

**B. P. Surin,[1†] J. M. Watson,[1] W. D. O. Hamilton,[2] A. Economou[3] and J. A. Downie[1,3*]**
[1]*C.S.I.R.O., Division of Plant Industry, Canberra, GPO Box 1600, ACT 2601, Australia.*
[2]*Agricultural Genetics Company, c/o Cambridge Laboratory, AFRC Institute of Plant Science Research, Maris Lane, Cambridge CB2 2JB, UK.*
[3]*John Innes Institute and AFRC Institute of Plant Science Research, Colney Lane, Norwich NR4 7UH, UK.*

## Summary

**DNA sequencing of the *nodIJ* region from *Rhizobium leguminosarum* biovar *trifolii* revealed the *nodT* gene immediately downstream of *nodJ*. DNA hybridizations using a *nodT*-specific probe showed that *nodT* is present in several *R. leguminosarum* strains. Interestingly, a flavonoid-inducible *nodT* gene homologue in *R. leguminosarum* bv. *viciae* is not in the *nodABCIJ* operon but is located downstream of *nodMN*. The sequence of the *nodT* gene from bv. *viciae* was determined and a comparison of the predicted amino-acid sequences of the two *nodT* genes shows them to be conserved; the predicted protein sequences appear to have a potential transit sequence typical of outer-membrane proteins. Mutations affecting *nodT* in either biovar had no observed effect on nodulation of the legumes tested.**

## Introduction

The nodulation of leguminous plants requires the co-ordinated expression of several *Rhizobium* nodulation (*nod*) genes arranged in several different operons. The nodulation process, *nod* gene arrangement and gene structure have been reviewed recently (Downie and Johnston, 1988; Rolfe and Gresshoff, 1988; Long, 1989), and a comparison of the *nod* gene regions from the symbiotic plasmids of *Rhizobium leguminosarum* biovars (bv.) *viciae* and *trifolii* is shown in Fig. 1. In essence, plant-made flavonoids (or isoflavonoids) induce bacterial *nod* gene expression via *nodD* gene products which are positively acting transcriptional regulatory proteins. Several *nod* operons are induced, some of which specify host-specific determinants in nodulation, while other *nod* genes appear to be functionally equivalent, in that *nod* genes from one *Rhizobium* species or biovar can complement *nod* mutations in a different species. The *nodABC* genes belong to this latter group of so-called 'common' nodulation genes and in bv. *viciae* are in an operon containing *nodABCIJ* (Evans and Downie, 1986; Downie *et al.*, 1987; Canter-Cremers *et al.*, 1988).

The DNA sequence of the bv. *viciae nodIJ* genes has been determined and DNA hybridization experiments have shown that strongly homologous genes exist in bv. *trifolii* immediately downstream of *nodC*. However, mutations in the *nodIJ* region of bv. *trifolii* (Djordjevic *et al.*, 1985; Huang *et al.*, 1988) cause a different phenotype from those in the *nodIJ* genes of bv. *viciae* (Downie *et al.*, 1985; Canter-Cremers *et al.*, 1988) in that nodulation of clover by *nodI* or *nodJ* mutants of bv. *trifolii* is much more severely affected than the nodulation of *Vicia hirsuta* and peas by *nodI* or *nodJ* mutants of bv. *viciae*. These markedly different phenotypes could be due either to different responses of the various host plants used in the nodulation tests, or to differences that polar mutations in the *nodIJ* region have on a hitherto unidentified downstream gene.

In one strain of biovar *viciae*, a novel nodulation gene (*nodX*) has already been identified in the *nodABCIJ* operon (Götz *et al.*, 1985; Davis *et al.*, 1988). The *nodX* gene extends the host range of bv. *viciae* enabling it to nodulate cv. Afghanistan peas, a cultivar not nodulated by most western strains of bv. *viciae*. Since genetic studies have implicated a region downstream of *nodIJ* in bv. *trifolii* as being able to stimulate nodulation of cv. Afghanistan when transferred to bv. *viciae* (Canter-Cremers *et al.*, 1988), the DNA sequence of the *nodIJ* region from bv. *trifolii* was determined, in order to identify gene(s) downstream of *nodJ*. The nodulation gene *nodT* was identified downstreaam of *nodJ* and was found to have a homologue in bv. *viciae* in which it was found to be located downstream of the *nodMN* genes. However, the subcloned *nodT* gene from bv. *trifolii* did not confer on a strain of bv. *viciae* the ability to nodulate cv. Afghanistan peas.

Received 28 August, 1989; revised 26 October, 1989. †Present address: Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia. *For correspondence. Tel. (0603) 52571; Fax (0603) 56844.

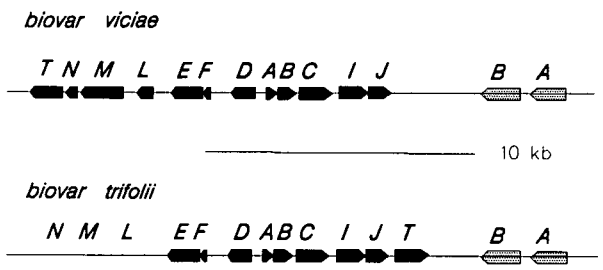


**Fig. 1.** Arrangement of *nod* and *nif* genes in two biovars of *R. leguminosarum*. The *nod* and *nif* gene regions from bv. *viciae* and bv. *trifolii* are aligned to show the sizes and orientations of coding regions for *nod* genes (black) and *nif* genes (shaded). The map of the bv. *viciae* genes is based on work by Rossen *et al.* (1984a,b), Shearman *et al.* (1986), Evans and Downie (1986), Surin and Downie (1988) and the results presented here. The map of the bv. *trifolii* genes is based on work by Schofield and Watson (1986), Djordjevic *et al.* (1986), Iismaa and Watson (1989) and the results presented here.

## Results

### *DNA sequence of the* nodIJT *region of bv.* trifolii

The DNA sequence of the *nodIJ* region of bv. *trifolii* is shown in Fig. 2. Analysis of the sequence revealed three open reading frames; the predicted amino acid sequence terminating at nucleotide 218 is very similar to the predicted carboxy-terminal sequence of NodI from bv. *viciae*. The following open reading frame (nucleotides 218–1003, Fig. 2) is 85% homologous to *nodJ* from bv. *viciae* (Evans and Downie, 1986); those amino acids not identical in biovars *viciae* and *trifolii* are marked with dots in Fig. 2. However, the considerable DNA homology between the sequences from bv. *trifolii* and bv. *viciae* breaks down immediately after the end of *nodJ* (see below).

A third long open reading frame was identified extending between nucleotides 1131 and 2531 (Fig. 2) encoding a predicted protein of molecular weight 50345. Analysis of this sequence (using the programs of Staden, 1984) confirmed that it has a very high probability of encoding a polypeptide. Because this gene is not preceded by a *nod*-box sequence and is immediately downstream of *nodJ* (Fig. 2), it is probably in the same operon as *nodABCIJ*. We propose to call this new nodulation gene *nodT*.

The *nodT* DNA sequence was found to have no homology with the ~2 kb region of DNA downstream of *nodJ* on pRL1JI (I. J. Evans and J. A. Downie, unpublished) or pRL5JI (Davis *et al.*, 1988), the symbiotic plasmids from two *viciae* biovars. It was shown previously that the *nodX* gene, which allows nodulation of cv. Afghanistan peas, is downstream of *nodJ* on pRL5JI, and has no homologue on pRL1JI (Davis *et al.*, 1988). Significantly, there is strong DNA-sequence homology throughout the *nodJ* genes of bv. *trifolii* and the two strains of bv. *viciae*. However, the DNA homology stops abruptly at the end of the *nodJ* translated region in all three strains (Fig. 3). On the basis of these limited comparisons, it appears that the region downstream of *nodJ* may be a site of genetic rearrangements even though the positions of the *nifA* and *nifB* genes further downstream of *nodJ* are conserved in bv. *viciae* and bv. *trifolii* (Fig. 1).

### nodT *from bv.* viciae *is in the* nodMN *operon*

Hybridization experiments using a *nodT*-specific probe were carried out with DNA isolated from several *Rhizobium* strains in order to see if *nodT* genes were present in other species. Strongly hybridizing bands were found in several strains of bv. *trifolii* and weak hybridization was found to a 2.85 kb *Eco*RI fragment of DNA from pRL1JI in bv. *viciae* strain 8401pRL11JI. Restriction mapping of the hybridizing region of pRL1JI showed that this *Eco*RI fragment was located on pIJ1089 (Downie *et al.*, 1983) immediately downstream of the previously described gene, *nodN* (Surin and Downie, 1988; Fig. 1). The DNA sequence of the 2.85 kb *Eco*RI fragment from bv. *viciae* was determined (Fig. 4); the nucleotide positions correspond to those described previously from the *nodMN* operon (Surin and Downie, 1988). Downstream of *nodN* (168 nucleotides from the translational stop) a large open reading frame was identified, with the most probable start codon being the TTG at position 7713. This codon, unlike the ATG codon further downstream, is preceded by a good potential ribosome-binding site. The choice of the TTG start was also influenced by protein sequence comparisons (see below). The open reading frame ends at position 9158, and the deduced gene product ($M_r$ 51471) is homologous to NodT of bv. *trifolii* (see below).

The sequenced region downstream of *nodT* is unlikely to encode protein. Three (overlapping) potential reading frames starting at ATG codons at positions 9262, 9390 and 9506 (Fig. 4), and encoding polypeptides of 108, 114 and 96 amino acids, respectively, did not show normal codon usage when studied using the ANALYSEQ programs (Staden, 1984).

### *The predicted protein sequence of NodT suggests that it is targeted to the outer membrane*

The predicted NodT protein sequences from the two biovars are aligned in Fig. 5 and the identical residues or conservative substitutions are marked; the two protein sequences are about 60% homologous, as calculated by the GAP program (Devereux *et al.*, 1984) using a pair score of 1.0.

Initially it appeared that the amino-terminal sequence of the bv. *viciae nodT* gene product was different from that of bv. *trifolii* because of the relative position of the first ATG codon in the open reading frame found in bv. *viciae* (Figs 4

and 5). However, the predicted amino acid sequence preceding this methionine codon in the bv. *viciae nodT* gene shows significant homology to the amino-terminal region of bv. *trifolii* NodT (Fig. 5).

The amino-terminal regions of the NodT proteins have some characteristics typical of transit peptides that enable proteins to be secreted through the bacterial inner membrane. The *N*-terminal regions are somewhat hydrophobic, and have two positively charged residues (positions 2 and 5, Fig. 5) preceding the hydrophobic domain. In

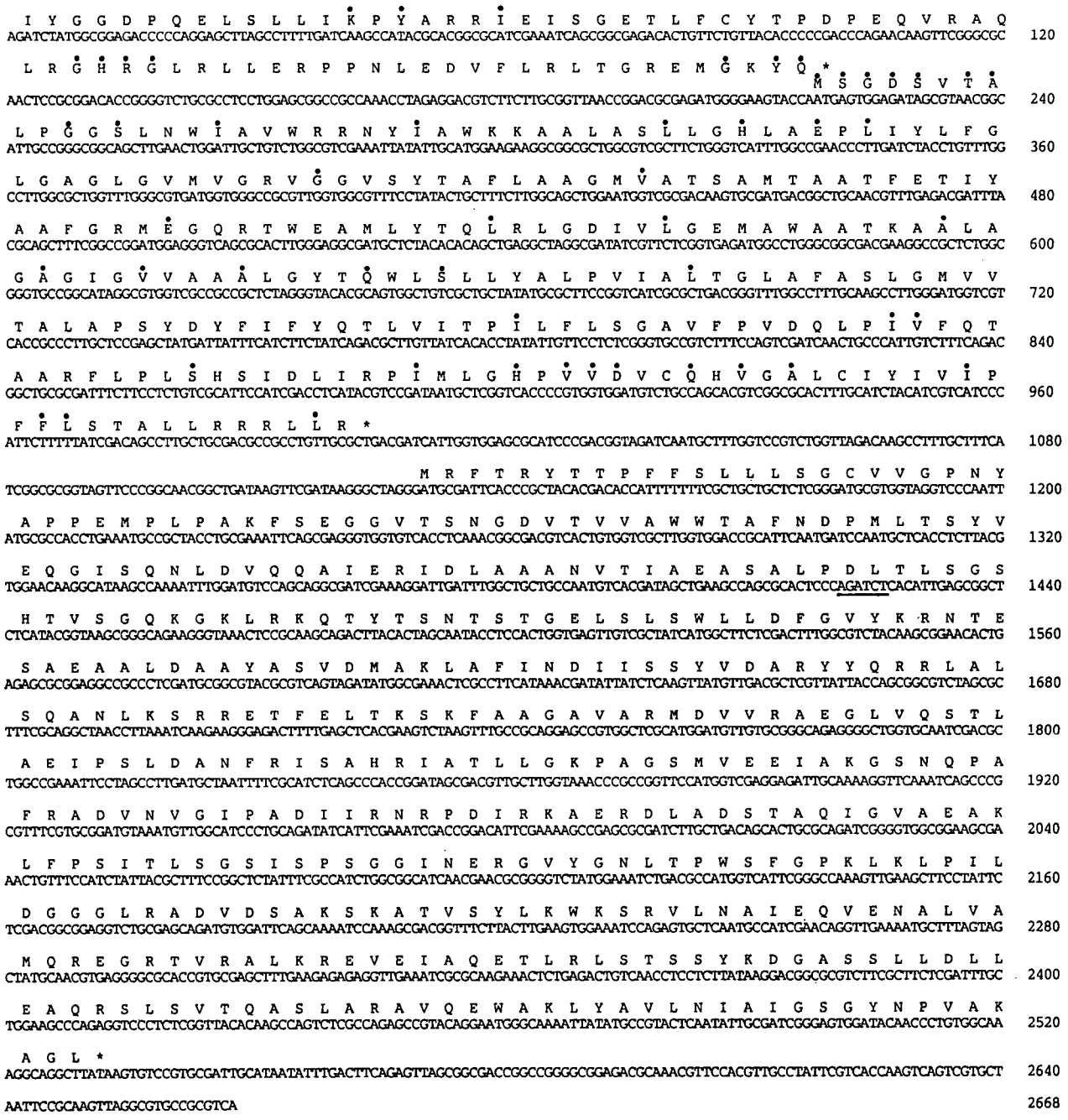

**Fig. 2.** DNA and predicted amino acid sequence of the *nodIJT* region from bv. *trifolii*. The sequence extends from a *Bgl*II site (nucleotides 1–6, Fig. 2) within the *nodI* gene and was determined by sequencing overlapping fragments derived from pRt732, a plasmid containing the *nodCIJ* region (region II described in Schofield and Watson, 1986). The nucleotide sequence of the open reading frame extending to nucleotide 218 is homologous to *nodI* from bv. *viciae* and the second coding region extending from nucleotide 217 to 1003 is homologous to *nodJ* from bv. *viciae* (Evans and Downie, 1986). Non-identical amino acid residues found in the comparisons of the predicted *nodIJ* gene products are indicated by dots above the residues. The *nodT* open reading frame extends from nucleotide 1131 to 2531. The *Bgl*II site used in the construction of the *nodT21*::kan allele is underlined. These sequence data will appear in the EMBL/GenBank/DDBJ Nucleotide Sequence Data Libraries under the accession number X51411.

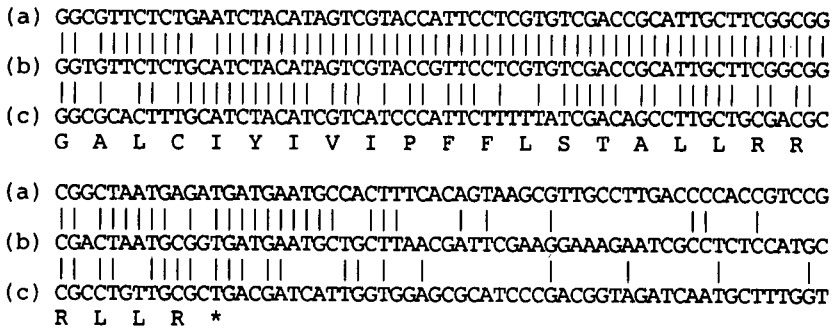


**Fig. 3.** Lack of DNA sequence homology at the ends of three *nodJ* genes. The sequences at the ends of three *nodJ* genes are aligned, showing that the sequence homology ends abruptly at the end of the coding region. The sequence (b) from bv. *viciae* pRL1JI (Evans and Downie, 1986) is compared with the sequences from (a) bv. *viciae* pRL5JI (Davis *et al.*, 1988) and (c) bv. *trifolii*. The amino acid sequence of the end of *nodJ* from bv. *trifolii* is shown below the DNA sequence.

```
                                                                  * L H S F R L A
GAATTCTCGATCTGCTGCGTCGGCAATTGCGCGCTGGCATCATCCTTTGGGCAGGCGAAACGGGTTTGAGTTGAATTTCGCTGTAATGGAGGCTGATTGATTGCATTCGTTCCGTCTGGC   7732

 A A V L P L L L S S C M L G P D H A P P E T P L P E K F S E G A K Q S A G D V A
CGCTGCGGTGTTGCCGCTACTTTTGTCATCTTGCATGCTCGGGCCCGACCATGCACCGCCGGAAACGCCTTTGCCGGAGAAGTTTTCCGAGGGTGCAAAACAGAGCGCCGGCGATGTTGC   7852

 V S A W W D S F S D R T L N Q Y V A S G L D E N L S V Q Q A L E R V N A A A A D
GGTGTCGGCGTGGTGGGATTCGTTTTCAGATCGTACGCTCAACCAGTACGTTGCCTCCGGCCTGGATGAGAACCTCAGTGTTCAGCAGGCGCTCGAAAGGGTTAATGCCGCCGCCGCGGA   7972

 V T I A G A G G L P K A S H T T S G E I G K G G D I T S T Q N I S S V Q L S L T
CGTCACCATCGCTGGCGCCGGGGGCCTTCCCAAGGCATCACACACGACAAGCGGCGAGATCGGCAAAGGGGGAGATATCACGAGCACACAGAATATCTCCAGCGTGCAACTATCGTTGAC.   8092

 W L L D V F G Q Y R R S T E S A L A S L D S A H A A V D A A K L A L I K D L V S
CTGGCTGCTGGACGTATTCGGTCAGTACCGGCGCAGCACTGAAAGCGCACTGGCTTCTCTTGATTCTGCGCATGCGGCTGTCGACGCCGCAAAGCTAGCTCTAATCAAAGATCTTGTTTC   8212

 S Y I D A R Y Y Q Q R V S I S R A N L K S R Q E T Y D F T N L Q V E A G A A S R
CTCCTACATCGACGCCCGGTACTACCAGCAGCGCGTCTCGATTTCCCGAGCGAACCTGAAGTCCCGCCAGGAGACCTACGATTTCACGAATCTTCAAGTCGAAGCGGGAGCGGCTTCCCG   8332

 Q D V L Q A E G L V R S T I A E I P R L E L N F R V S A H H I A A L L A L P S E
GCAGGATGTCTTGCAAGCCGAAGGACTTGTCCGATCGACGATTGCCGAAATACCAAGGCTAGAACTGAATTTTCGCGTGTCGGCGCATCACATTGCGGCCCTCCTCGCTTTGCCGTCGGA   8452

 T V I K Q L Q K S E G Q P V Y R G K I N A G I P A D L I R N R F D I R Q A E R D
AACAGTGATCAAGCAACTGCAGAAAAGTGAAGGGCAGCCGGTCTATCGCGGAAAGATCAATGCCGGCATTCCCGCCGATCTTATTCGCAATCGATTCGATATTCGCCAGGCTGAACGGGA   8572

 L A A A T A Q I G V A E A Q L Y P A I T L S G S I T P S Y I K Q R G R H G G I L
TCTCGCCGCCGCGACTGCACAGATCGGCGTGGCGGAAGCACAGCTTTACCCGCGATAACGCTCTCCGGCTCGATCACCCCTTCCTATATCAAGCAGCGCGGCCGTCACGGGGGTATCTT   8692

 K W S F G P S L D L P I L D G G R L R A N V E T S K S D A A A A Y I S W K L T V
GAAATGGTCCTTCGGGCCGAGCCTTGACCTCCCAATTCTGGACGGCGGCCGTTTGCGGGCAAACGTGGAGACTTCTAAATCGGACGCCGCGGCAGCCTACATAAGCTGGAAATTAACCGT   8812

 L T A V Q E V E D A L T A V R R D V H T E N S R R R Q V E T I E E A L K L S T A
CCTGACAGCTGTGCAGGAAGTCGAGGACGCTTTGACAGCCGTCCGACGCGACGTTCACACAGAGAACTCACGTCGCAGGCAAGTAGAAACGATCGAAGAAGCACTGAAGCTTTCGACGGC   8932

 S Y T D G A S S L L D V L E A Q R Q V S S A Q A S L A A A I Q Q L A K D H V R L
CTCGTACACGGATGGCGCCTCTTCTCTCCTCGATGTTCTCGAAGCGCAAAGGCAGGTGTCCAGCGCTCAGGCGAGCCTCGCTGCGGCGATCCAGCAACTGGCCAAGGATCACGTGCGGTT   9052

 N V A I R G G F A A P K V A S P R E A S T V A A A N A N I Q S A H S P *
GAATGTAGCAATCCGCGGTGGTTTTGCGGCTCCCAAAGTCGCGTCACCCAGGGAGGCCTCGACGGTTGCAGCCGCAAATGCAAACATCCAATCCGCACATTCGCCATAGCCGAGCGCTGG   9172

ATGCGGGATGGTTTAACCTGATCAAGTCATTAGCGCGCCAGCGCCTGAGTTGTATTGGCGCCGGCCGGGAGGTCGTCCAATCGCGGCACATGCCCTTCACCACCTCCCGCTTACGAGCAG   9292

GGCTTAGAGCTAATGGGATGGTCCGCCCTGTCTCCCGAGAGCATAACGGAGGTGGGCAAAGGAGGGCTGAAATTATGGATCGCAACCAACTGCGCCAATGCGCGTCTTCGGCATCGATAT   9412

CGGAAAGAACCTCTTCCACGTGGTCGGCCTGGATAGTTCCGGCGCGAGACGCTACTGCAGTTCTTCGAACGAGCGAACAGAACGCTTGTCGGAATGGAATCCCGTCCGGGTTCCCAATGG   9532

CGACCGACGATGCGCTTCGTTGGCATTAAGCAAGCCGATCAGGTCGATCTTCAGATGTTGCATTGGGTGCGTATTCCGATGAAACCGGCCACCGATTCCGATTTGAAGCCGGCTATTTTT   9772

CGGGCTGATCCTGGGTCTGTTTGATGTTTGGATCGGGTTTTGTGTCAGGTCAAGCTTTGCTGTGTTTCAAGCGCCATCGGTGAACGCTGTTTGCGCATACTTTCGCCAATGAGATCGATG   9892

CCTCGGCGTTGTGAACGAGGCGATCCAGTATCGCATCGGCGGCGTTGGGTTTGAAATAATTTCCCACCAGCGATCCAGGGGCACCTGGCTTGTGACGATCGTCGAGCGACGCTCGTAGCG   10012

GTCTTCGATGATCTCGAGAAGATCGCGGCGCTGATCATCGTTGAGTTTTTCCGGTCCCTAGTCATCGAGGATCAGCAGGTCGGTCTTGGCGATAGATTTGGGGATTCTGCCGTATCGCCC   10132

GTCACCCCTTGCAAGCGCGAGCGTGGCGAACAGCCGGGGAACGCGGTGGTATGCAACGGAGAAATCCTCGCGGCATGCTTTGACCGATGGATGCGGCCCAGCGCCGAAAGCGATCTGGTG   10252

TCCACTCAGCGTATCGCCGGTGGGAACTGGGCATGTGATCTGGGTCGGTTCCATGACGACAGCCGCCATAGCGGCGCTGATGGACGGCGATGCGCTTGCCGCGGTAAAAGATCTCGATGG   10372

TGCGCGCCGTTGCTCTAAGATCGACCTGCTGGCGAATGAGGGCATGAGGGACGGAATAGAAGAAGCTTTTGAACTCGACGTGATAATCGGTCGAGACACGGAGCAGACGCCATTCGGCGA.   10492

ATTC  10496
```

**Fig. 4.** DNA and predicted protein sequence of *nodT* from bv. *viciae*. The DNA sequence of the 2884 nucleotide *Eco*RI fragment from pRL1JI is shown, with the nucleotide positions numbered to follow those described previously (Surin and Downie, 1988). The sequence was determined by sequencing a series of ordered deletions of plasmids pIM243 and pIM244, which contain the 2.85 kb *Eco*RI fragment cloned in either orientation in the vector pUC118. The predicted protein sequence of NodT is shown initiating at a TTG codon (see text). The DNA sequence downstream of *nodT* did not have a high probability of encoding proteins when studied using ANALYSEQ programs (Staden, 1984) and no long open reading frames were evident. The *nodT*92::Tn*3lac* mutation was mapped to be around nucleotide position 8000. These sequence data will appear in the EMBL/GenBank/DDBJ Nucleotide Sequence Data Libraries under the accession number X17285.

**Fig. 5.** Alignment of NodT protein sequences. The predicted NodT protein sequences from bv. *trifolii* (upper sequence) and bv. *viciae* (lower sequence) were aligned using the GAP program (Devereux *et al.*, 1984) with a matching pair score of 1.0. The matching residues scoring 1.0 or above are indicated by vertical lines, and introduced gaps are indicated by dots. The potential leader sequence is indicated by the positively charged residues (+) and a hydrophobic region ending at a cysteine residue (asterisk) which could be a site of lipoylation after cleavage of the putative signal peptide (see text).

addition, in the bv. *trifolii* sequence there is the potential signal peptidase cleavage sequence, Ser-Gly/Cys (SG/C at positions 16–18, Fig. 5) typical of outer-membrane lipoprotein cleavage sites (Wu and Tokunaga, 1986). At the equivalent position in NodT from bv. *viciae* there is the sequence Ser-Ser-Cys; although this is slightly different from the typical cleavage site, it does retain the important cysteine residue that is normally cross-linked with a lipid after cleavage of the outer membrane transit sequence. We note that this is the only cysteine residue present in the two NodT protein sequences.

### *The* nodT *gene is induced by flavonoids*

Plasmid pIJ1086, which contains the *nodMNT* genes from bv. *viciae* on a ~30 kb region of cloned DNA (Downie *et al.*, 1983), was mutagenized with the transposon Tn*5lac* and the mutated derivatives were transferred to strain 8401-pRL1JI. About 200 pIJ1086::Tn*5lac* derivatives were screened on plates for the induction of β-galactosidase activity by the flavonoid hesperetin, to identify Tn*5lac* insertions in *nod* genes. Twelve flavonoid-inducible derivatives of pIJ1086 were found and the Tn*5lac* insertion sites were mapped using the restriction enzymes *Eco*RI, *Hin*dIII and *Bam*HI. One of the Tn*5lac* insertions was found to be located within *nodT* and orientated appropriately for the proposed direction of transcription of *nodT*. The approximate position of the Tn*5lac* insertion (called *nodT92*::Tn*5lac*) is indicated in the legend to Fig. 4.

Quantitative measurements of β-galactosidase expression were made as described previously (Rossen *et al.*, 1985) using strain 8401pRL1JI containing pIJ1747 (which carries *nodT92*::Tn*5lac* on pIJ1086). When the strain was grown in the absence of hesperetin, a background level of β-galactosidase activity (30 Units) was observed, whereas in the presence of 1 μM hesperetin the strain expressed 60 Units of β-galactosidase activity, confirming that *nodT* is flavonoid-inducible. This level of *nodT* expression is rather low, being about half that of a similar fusion in *nodN* and about 10% of that observed with Tn*5lac* inserted in other *nod* genes such as *nodO* (A. Economou, unpublished observations).

The region of DNA between the *nodN* and *nodT* genes (Fig. 4) did not contain the highly conserved *nod*-box sequence characteristic of *nod* gene promoters (Rostas *et al.*, 1986; Shearman *et al.*, 1986; Schofield and Watson, 1986). Since no such sequence is present (Fig. 4) it is concluded that *nodT* from bv. *viciae* is in the same operon as *nodMN* and is induced under the control of the *nod* promoter upstream of *nodM*.

### *Mutation of* nodT *had no observed effect on nodulation phenotype*

The *nodT92*::Tn*5lac* allele on pIJ1747 was recombined onto pRL1JI to form the mutant strain A66 which is a derivative of bv. *viciae* strain 8401pRL1JI. In parallel, a *nodT* derivative of bv. *trifolii* (strain BS21) was constructed (see the *Experimental procedures*). Strains A66 and BS21 were scored for their nodulation phenotypes on appropriate host legumes. No impairment of nodulation was observed with strain BS21 compared with its isogenic control inoculated on the *Trifolium* species *subterraneum*, *glomeratum*, *cherleri* or *repens* using 20 plants for each test. Similarly, no difference in nodulation was observed

between strains A66 and 8400pRL1JI inoculated on peas (20 plants per test) or *Vicia hirsuta* (40 plants per test).

It was concluded previously that a gene downstream of *nodJ* in bv. *trifolii* strain ANU843 conferred on a strain of bv. *viciae* the ability to nodulate peas of cultivar Afghanistan (Canter-Cremers *et al.*, 1988). The intact *nodT* gene from bv. *trifolii* was subcloned into a broad host-range vector behind a vector promoter to form pBS21, and this plasmid was transferred to bv. *viciae* strain 8401pRL1JI. The transconjugant was inoculated on to cv. Afghanistan peas but no significant nodulation was observed, indicating that, in contrast to the conclusions of Canter-Cremers *et al.* (1988), the gene (*nodT*) downstream of *nodJ* in bv. *viciae* does not confer nodulation of cv. Afghanistan peas to bv. *viciae*.

## Discussion

### *The* nodI *and* nodJ *genes*

The DNA sequence presented here confirms that the *nodI* and *nodJ* genes are present downstream of *nodC* in bv. *trifolii*. It was observed previously that the *nodI* gene product is strongly homologous to a group of bacterial proteins that appear to be ATP-binding proteins (Evans and Downie, 1986; Higgins *et al.*, 1986). Many of these proteins are known to be involved in transmembrane transport of substrates such as maltose or phosphate, although others are involved in processes such as protein secretion (of haemolysin) or cell division. The NodJ protein is hydrophobic and it was suggested previously that it may be transmembranous and may function in conjunction with NodI in membrane transport (Evans and Downie, 1986). Further evidence supporting the transmembranous nature of NodJ was obtained using the transposon Tn*phoA*, which can be used to identify secreted proteins and transmembrane proteins (Manoil and Beckwith, 1985). Two classes of Tn*phoA* insertions that expressed membrane alkaline phosphatase activity were identified within the *nodJ* gene of bv. *trifolii* after mutagenesis of pRt744, a plasmid which carries the *nodJ* gene from bv. *trifolii* (J. Watson and B. Surin, unpublished observations). However, the precise biochemical functions of the *nodIJ* gene products remain to be determined.

### *The* nodT *genes*

It is clear that the two *nodT* genes described here are closely related, despite the fact that in one biovar (*viciae*) *nodT* is in the *nodMN* operon whereas in bv. *trifolii*, *nodT* appears to be in the *nodABCIJ* operon. The amino-terminal region of the protein encoded by the bv. *trifolii nodT* gene has characteristics typical of signal peptides required for the translocation of proteins across the bacterial inner membrane. These features include a positively charged region followed by a hydrophobic domain and a signal peptide cleavage site (Wu and Tokunaga, 1986). A potential transit sequence in the bv. *viciae* NodT protein is found upstream of the first ATG codon and it appears possible that an alternative translational start codon is used to initiate the bv. *viciae nodT* gene. The amino acid sequence in Fig. 4 begins at a TTG codon (position 7709) that immediately follows a translational stop site, and is downstream of a potential ribosome-binding site (GGAGG). The TTG codon is a potential translational initiation codon in about 1% of known *Escherichia coli* genes (Stormo, 1986), and if used here the NodT protein would retain those sequences typical of a secreted protein. The amino acid sequence SGC in bv. *trifolii* NodT (position 18, Fig. 5) is typical of a cleavage site for outer-membrane proteins that are lipoylated on the cysteine residue that is *N*-terminal after cleavage of the transit sequence (Wu and Tokunaka, 1986). It is probable that the amino-acid sequence in the equivalent position in bv. *viciae* NodT (SSC) functions as a similar cleavage site, but clearly protein sequence analyses will be required to substantiate these proposals.

Mutations in the *nodT* gene do not significantly affect nodulation of either biovar under the conditions used. Similar observations have been made previously for mutations in other *nod* genes such as *nodM* and *nodN* (Surin and Downie, 1988). It is possible that the phenotype of mutations in *nodT* is masked by other *nod* genes which compensate for the lack of *nodT* expression, and to test this with the *nodT* gene would require the construction of a series of double mutants. Such experiments should lead to clarification of the symbiotic role of the NodT protein in its proposed outer-membrane location.

## Experimental procedures

### *Microbiological techniques*

Bacterial strains and plasmids are listed in Table 1 or referenced in the text. Media and general growth conditions were as described by Beringer (1974); matings and selection for homologous recombination were as described previously (Surin and Downie, 1988). Antibiotics were added to media as described previously (Surin and Downie, 1988). Plant tests were as described by Knight *et al.* (1986) or Iismaa *et al.* (1989) using peas (variety Wisconsin Perfection), and *V. hirsuta* or subterranean clover (*Trifolium subterraneum*) and white clover (*Trifolium repens*). Bacteria were isolated from surface-sterilized nodules as described previously (Knight *et al.*, 1985), and identified on the basis of their antibiotic resistance. Assays of β-galactosidase were done as described by Rossen *et al.* (1985) and the units of activity calculated as described by Miller (1972).

### *Manipulation of DNA*

Enzymes were from Boehringer Mannheim and DNA manipulations were carried out using standard procedures (Maniatis *et al.*,

**Table 1.** Bacterial strains and plasmids.

| Strain/Plasmids | Relevant properties | Source |
|---|---|---|
| *Rhizobium leguminosarum* | | |
| 8401pRL1JI | bv. *viciae* | Downie *et al.* (1983) |
| A66 | 8401 pRL1JI, *nodT*92::Tn*5lac* | This work |
| ANU843 | bv. *trifolii* | Rolfe *et al.* (1980) |
| BS21 | ANU843, *nodT*21::*kan* | This work |
| pRt732 | *nodCIJT* region from ANU843 cloned on a 6.0kb *Eco*RI fragment in pUC18 | This work |
| pRt744 | *nodIJ* region from ANU843 cloned on a 1.4kb *Bgl*II fragment in the *Bam*HI site of pUC18 | This work |
| pBS50 | pRt732 carrying *nodT*21::*kan* | This work |
| pBS51 | pRK290 carrying the *Eco*RI fragment from pBS50 | This work |
| pBS21 | *nodT* from ANU843 cloned on a 2.3kb *Sph*I fragment in pKT230 | This work |
| pIM243 | *nodT* from pRL1JI cloned on a 2.85kb *Eco*RI fragment in pUC118 | This work |
| pIM244 | As pIM243 but in the opposite orientation | This work |
| pIJ1086 | 30kb region of pRL1JI cloned in pLAFR1; carries *nodL*, *nodMNT* | Downie *et al.* (1983) |
| pIJ1747 | pIJ1086 carrying *nodT*92::Tn*5lac* | This work |

1982). DNA was sequenced essentially as described by Bankier and Barrell (1983) on both strands using overlapping sequence determinations from fragments subcloned in the M13 derivatives mp18 and mp19 (Norrander *et al.*, 1983). The *nodT* gene from bv. *trifolii* was sequenced using a series of ordered deletions generated from pRt744 by Bal31-exonuclease digestion, essentially as described by Iismaa and Watson (1989). The sequence of the bv. *viciae nodT* gene was determined using a similar strategy with fragments subcloned from pIM243 and pIM244. The DNA sequence was compiled and analysed using the programs of Staden (1984). Plasmid pBS21 carrying *nodT* from bv. *trifolii* was constructed by subcloning a 2.3kb *Sph*I fragment into pKT230 (Bagdasarian *et al.*, 1981). Plasmids pIM243 and pIM244 were made by subcloning a 2.85kb *Eco*RI fragment from pIJ1089 (Downie *et al.*, 1983) into the *Eco*RI site of pUC118 (Vieira and Messing, 1987).

### *Construction of mutants*

The mutated form of the bv. *trifolii* gene was made by partially digesting pRt732 with *Bgl*II and ligating gel-purified linear plasmids with a 1.9kb *Bam*HI fragment of DNA from pANU1, which carries the kanamycin resistance gene from Tn*5* cloned in pUC8 (J. M. Watson, unpublished). Transformants were selected using kanamycin and an appropriate recombinant carrying the kanamycin resistance gene within *nodT* (at position 1421; Fig. 2) was identified by restriction endonuclease mapping of recombinant plasmids. The *Eco*RI fragment of DNA from one such recombinant (carrying the *nodT21*::kan allele) was subcloned into the *Eco*RI site of pRK290 (Ditta *et al.*, 1980) to form pBS51. This plasmid was transferred to ANU843 by conjugation, selecting for neomycin (100 $\mu$g ml$^{-1}$) resistance, and the *nodT21*::kan allele was recombined into its homologous location.

To mutagenize pIJ1086, the transposon Tn*5lac* (Simon *et al.*, 1989) was first transferred from phage lambda to a chromosomal location in *E. coli* strain CC118 (Manoil and Beckwith, 1985) by selecting a kanamycin-resistant transductant in which the Tn*5lac* had transposed to an unmapped location to form strain A118. pIJ1086 was then transferred into A118 and Tn*5lac* derivatives of pIJ1086 were selected following conjugation to strain 8401-pRL1JI, essentially as described previously (Downie *et al.*, 1985); the mutants were selected on Y medium containing kanamycin (20 $\mu$g ml$^{-1}$). Mutations in *nod* genes were screened after replica plating to Y plates with or without 1 $\mu$M hesperetin and containing 40 $\mu$g ml$^{-1}$ of 5 bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal). The mutated plasmids containing inducible Tn*5lac* insertions were transferred to *E. coli* strain 803 by transformation and the sites of Tn*5lac* insertions mapped using the restriction endonucleases *Bam*HI, *Eco*RI and *Hin*dIII.

## Acknowledgements

We are indebted to S. Patton, J. C. Timans, D. Harrison and A. Davies for skilled technical assistance. We thank R. Simon for the generous gift of the transposon Tn*5lac*, and S. Iismaa and C. Hayes for assistance in sequencing the 3′ end of NodT from bv. *trifolii*. This work was supported in part by a post-doctoral fellowship to B.P.S. from the Wool Research Trust Fund on the recommendation of the Australian Wool Corporation. A.E. was supported by a sectoral training grant from the E.E.C. J.C.T. and D.H. are members of the Agricultural Genetics Company. We thank D. A. Hopwood, K. F. Chater and N. J. Brewin for critical comments on the manuscript.

## References

Bagdasarian, M., Lurz, R., Ruckert, B., Franklin, F.C.H., Bagdasarian, M.M., Frey, J., and Timmis, K.N. (1981) Specific-purpose plasmid cloning vectors. II. Broad host range, high copy number RSF1010-derived vectors and a host-vector system for cloning in *Pseudomonas*. *Gene* **16:** 237–247.

Bankier, A.T., and Barrell, B.G. (1983) Shotgun DNA sequencing. In *The Life Sciences*. Flavell, R.A. (ed.). Shannon: Elsevier, pp. 137–144.

Beringer, J.E. (1974) R factor transfer in *Rhizobium leguminosarum*. *J Gen Microbiol* **84:** 188–198.

Canter-Cremers, H.C.J., Wijffelman, C.A., Pees, E., Rolfe, B.G., Djordjevic, M.A. and Lugtenberg, B.J.J. (1988) Host specific

nodulation of plants for the pea cross-inoculation group is influenced by genes in fast growing *Rhizobium* downstream of *nodC*. *J Plant Physiol* **132:** 398–404.

Davis, E.O., Evans, I.J., and Johnston, A.W.B. (1988) Identification of *nodX*, a gene that allows *Rhizobium leguminosarum* biovar *viciae* strain TOM to nodulate Afghanistan peas. *Mol Gen Genet* **212:** 531–535.

Devereux, J., Haeberli, P., and Smithies, O. (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucl Acids Res* **12:** 387–395.

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

Djordjevic, M.A., Schofield, P.R., and Rolfe, B.G. (1985) Tn*5* mutagenesis of *Rhizobium trifolii* host-specific nodulation genes result in mutants with altered host-range ability. *Mol Gen Genet* **200:** 463–471.

Djordjevic, M.A., Innes, R.W., Wijffelman, C.A., Schofield, P.R., and Rolfe, B.G. (1986) Nodulation of legumes is controlled by several distinct loci in *Rhizobium trifolii*. *Plant Mol Biol* **6:** 389–401.

Downie, J.A., and Johnston, A.W.B. (1988) Nodulation of legumes by *Rhizobium*. *Plant Cell Environ* **11:** 403–412.

Downie, J.A., Knight, C.D., Johnston, A.W.B., and Rossen, L. (1985) Identification of genes and gene products involved in the nodulation of peas by *Rhizobium leguminosarum*. *Mol Gen Genet* **198:** 255–262.

Downie, J.A., Ma, Q.-S., Knight, C.D., Hombrecher, G., and Johnston, A.W.B. (1983) Cloning of the symbiotic region of *Rhizobium leguminosarum*: the nodulation genes are between the nitrogenase genes and a *nifA*-like gene. *EMBO J* **2:** 947–952.

Downie, J.A., Surin, B.P., Evans, I.J., Rossen, L., Firmin, J.L., Shearman, C.A., and Johnston, A.W.B. (1987) Nodulation genes of *Rhizobium leguminosarum*. In *Proceedings of the III International Symposium on Plant-Microbe Interactions*. Verma, D.P.S., and Brisson, N. (eds). Amsterdam: Martinus Nijhoff, pp. 225–228.

Evans, I.J., and Downie, J.A. (1986) The *nodI* gene product of *Rhizobium leguminosarum* is closely related to ATP-binding bacterial transport proteins; nucleotide sequence analysis of the *nodI* and *nodJ* genes. *Gene* **43:** 95–101.

Götz, R., Evans, I.J., Downie, J.A., and Johnston, A.W.B. (1985) Identification of the host-range DNA which allows *Rhizobium leguminosarum* strain TOM to nodulate cv. Afghanistan peas. *Mol Gen Genet* **201:** 296–300.

Higgins, C.F., Hiles, I.D., Salmond, G.P.C., Gill, D.R., Downie, J.A., Evans, I.J., Holland, I.B., Gray, L., Buckel, S.D., Bell, A.W., and Hermodson, M.A. (1986) A family of related ATP-binding subunits coupled to many distinct biological processes in bacteria. *Nature* **323:** 448–450.

Huang, S.Z., Djordjevic, M.A., and Rolfe, B.G. (1988) Characterization of aberrant infection events induced on *Trifolium subterraneum* by *Rhizobium trifolii* region II mutants. *J Plant Physiol* **13:** 16–24.

Iismaa, S.E., and Watson, J.M. (1989) The *nifA* gene product from *Rhizobium leguminosarum* biovar *trifolii* lacks the *N*-terminal domain found in other NifA proteins. *Mol Microbiol* **3:** 943–956.

Iismaa, S.E., Ealing, P.M., Scott, K.F., and Watson, J.M. (1989) Molecular linkage of the *nif/fix* and *nod* gene regions in *Rhizobium leguminosarum* biovar *trifolii*. *Mol Microbiol* **3:** 1753–1764.

Knight, C.D., Rossen, L., Robertson, J.G., Wells, B., and Downie, J.A. (1986) Nodulation inhibition by *Rhizobium leguminosarum* multicopy *nodABC* genes and analysis of early stages of plant infection. *J Bacteriol* **166:** 552–558.

Long, S.R. (1989) *Rhizobium*-legume nodulation: life together in the underground. *Cell* **56:** 203–214.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Manoil, C., and Beckwith, J. (1985) Tn*phoA*: a transposon probe for protein export signals. *Proc Natl Acad Sci USA* **82:** 8129–8133.

Miller, J.H. (1972) *Experiments in Experimental Genetics*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Norrander, K., Kempe, T., and Messing, K. (1983) Construction of improved M13 vectors using oligonucleotide-directed mutagenesis. *Gene* **26:** 101–106.

Rolfe, B.G., and Gresshoff, P.M. (1988) Genetic analysis of legume nodule initiation. *Annu Rev Plant Phys Plant Mol Biol* **39:** 297–319.

Rolfe, B.G., Gresshoff, P.M., and Shine, J. (1980) Rapid screening for symbiotic mutants of *Rhizobium* on white clover. *Plant Sci Lett* **19:** 277–284.

Rossen, L., Johnston, A.W.B., and Downie, J.A. (1984a) DNA sequence of the *Rhizobium leguminosarum* nodulation genes *nodAB* and *C* required for root hair curling. *Nucl Acids Res* **12:** 9497–9508.

Rossen, L., Ma, Q.-S., Mudd, E.A., Johnston, A.W.B., and Downie, J.A. (1984b) Identification and DNA sequence of *fixZ* a *nifB*-like gene from *Rhizobium leguminosarum*. *Nucl Acids Res* **12:** 7123–7134.

Rossen, L., Shearman, C.A., Johnston, A.W.B., and Downie, J.A. (1985) The *nodD* gene of *Rhizobium leguminosarum* is autoregulatory and in the presence of plant exudate induces the *nodABC* genes. *EMBO J* **4:** 3369–3373.

Rostas, K., Kondorosi, E., Horvath, B., Simoncsits, A., and Kondorosi, A. (1986) Conservation of extended promoter regions of nodulation genes in *Rhizobium*. *Proc Natl Acad Sci USA* **84:** 1757–1761.

Schofield, P.R., and Watson, J.M. (1986) DNA sequence of *Rhizobium trifolii* nodulation genes reveals a reiterated and potentially regulatory sequence preceding *nodABC* and *nodFE*. *Nucl Acids Res* **14:** 2891–2903.

Shearman, C.A., Rossen, L., Johnston, A.W.B., and Downie, J.A. (1986) The *Rhizobium leguminosarum* nodulation gene *nodF* encodes a polypeptide similar to acyl-carrier protein and is regulated by *nodD* plus a factor in pea root exudate. *EMBO J* **5:** 647–652.

Simon, R., Quandt, J., and Klipp, W. (1989) New derivatives of transposon Tn*5* suitable for mobilization of replicons, generation of operon fusions and induction of genes in Gram-negative bacteria. *Gene* **80:** 161–169.

Staden, R. (1984) Measurements of the effects that coding for a protein has on DNA sequence and their use for finding genes. *Nucl Acids Res* **12:** 9509–9524.

Stormo, G.D. (1986) Translation initiations. In *Maximizing Gene Expression*. Reznikoff, W., and Gold, L. (eds). Boston: Butterworth, pp. 195–224.

Surin, B.P., and Downie, J.A. (1988) Characterization of the *Rhizobium leguminosarum* genes *nodLMN* involved in efficient host specific nodulation. *Mol Microbiol* **2:** 173–183.

Vieira, J., and Messing, J. (1987) Production of single-stranded plasmid DNA. *Meth Enzymol* **153:** 3–11.

Wu, H.C., and Tokunaga, M. (1986) Biogenesis of lipoproteins in bacteria. *Curr Top Microbiol Immunol* **125:** 127–157.