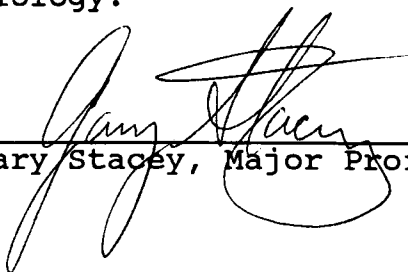


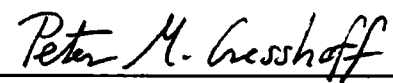
To the Graduate Council:

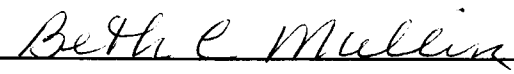
I am submitting herewith a dissertation written by Sohail Saad Luka entitled "Genetic and Functional Studies on Novel Nodulation Genes in *Bradyrhizobium japonicum*". I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfilment of the requirements for the degree of Doctor of Philosophy with a major in Microbiology.

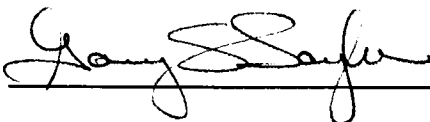


Gary Stacey, Major Professor

We have read this Dissertation
and recommend its acceptance:







Accepted for the Council



Associate Vice Chancellor and
Dean of the Graduate School

GENETIC AND FUNCTIONAL STUDIES ON NOVEL NODULATION GENES
IN *BRADYRHIZOBIUM JAPONICUM*

A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Sohail Saad Luka

August 1993

This dissertation is dedicated to
my Mother, my Father and my Sister Reem

ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Gary Stacey for his support throughout the course of this study. I also wish to thank the rest of my committee, Dr. Peter Gresshoff, Dr. Beth Mullin and Dr. Gary Sayler for their guidance and constructive criticism. I would also like to acknowledge the help of Dr. Juan Sanjuan in the purification of Nod Factors, as well as Dr. Russell Carlson for the chemical analysis of Nod factors. I wish to thank my colleagues Jong-Yoon Chun and Dr. Tom Dockendorff for constantly providing help throughout our graduate school experience. Special thanks go to Maria Schell for her outstanding support and for her valuable friendship throughout the years. I am indebted to my good friends in Egypt, USA and other parts of the world for their genuine concern and for standing by me during stressful times that accompanied this study. I wish to thank my Judo instructor Michael Carter who has allowed me to discover the elegance of the gentle way.

Finally, I am grateful to my parents, my sister and the rest of my family for their unlimited love and support and for all of their sacrifices throughout the years.

ABSTRACT

The DNA region downstream of the *nodYABCSUIJ* operon in *Bradyrhizobium japonicum* strain USDA110 was analysed. The DNA sequence of four novel genes (i.e., *noIMNO* and *ORFA*) and a previously identified genetic locus (i.e., *nodZ*) was determined and the function and transcriptional control of *noIN*, *ORFA* and *nodZ* were studied. *noIMNO* were shown to be co-transcribed with the common nodulation genes, *nodABC*, thus, establishing the *nodYABCSUIJnoIMNO* operon. *noI* showed similarity to *nodU* from various rhizobia on the amino acid level. The lipo-oligosaccharide nodulation signals produced by strains SL67 (*noI*⁻) and SL65 (*noIN*⁻) were purified and their chemical structures were determined (chemical analysis was done by Russell Carlson, The University of Georgia). In addition to the wild-type signal molecule, both mutants produced modified compounds that are not produced by the parent strain USDA110. The most prevalent difference observed was the absence of the 2-O-methylfucosyl residue from the mutant structures. In addition, metabolites were found in which the N-acetylglucosamine residue at the reducing end was glycosidically linked to glycerol. These alterations in the profiles of nodulation signals produced by strains SL67 and SL65 were accompanied by reduced nodulation efficiency on all hosts tested.

The expression of *nodZ* and ORFA was studied using *lacZ* gene fusions. Both genes appear to be constitutively expressed under laboratory conditions. The transcriptional start site of ORFA was determined revealing a 128 nucleotide overlap between the transcripts of the divergently transcribed *nodZ* and ORFA. Mutational inactivation of ORFA had no effect on nodulation. DNA Southern hybridization studies suggested that *nolO*, ORFA and *nodZ* may be linked in other rhizobial strains.

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION AND LITERATURE REVIEW	
<i>Rhizobium</i> Symbiosis and Taxonomy	1
Stages of Nodule Development	4
Bacterial Genes Required for Efficient Nodulation	10
i. <i>nod/nol</i> Genes	13
ii. <i>nif</i> and <i>fix</i> Genes	16
iii. Other Genes Involved in <i>Rhizobium</i> - Legume Symbiosis	19
iv. Summary and Perspective	20
Biochemical Function of <i>nod</i> Gene Products	21
i. Lipo-oligosaccharide Signal Molecules Produced by Rhizobia	21
ii. Role of <i>nod</i> Genes in Nod Factor Biosynthesis	25
Regulation of Nodulation Genes	29
Rationale for Present Work	31
II. MATERIALS AND METHODS	
Bacterial Strains, Plasmids and Bacteriophages	33
Growth Conditions	39
DNA Sequencing and Sequence Analysis	40
DNA Southern Hybridization	41
RNA Dot Blot Hybridization	42
B-Galactosidase Assays	43
Primer Extension	43
Plant Tests	44
Purification and Chemical Analysis of Nod Metabolites	45
III. RESULTS	
A) Discovery AND CHARACTERIZATION OF <i>nolMNO</i>	
DNA Sequencing of an Isoflavone-Inducible Locus Downstream to <i>nodJ</i>	48
Interstrain and Interspecies Hybridization Using a <i>nolO</i> Probe	58
Transcriptional Control of <i>nolMNO</i>	61
Nodulation Phenotype of <i>nolO</i> and <i>nolNO</i> Mutants	64
Lipo-oligosaccharide Nodulation Signals Produced by <i>nolO</i> and <i>nolNO</i> Mutants	69

B) INVESTIGATION OF TWO GENETIC LOCI LINKED
TO *noIMNO*

Expression and DNA Sequence of ORFA	75
Transcriptional Control of ORFA	81
Determining the Transcriptional Start of ORFA	81
Nodulation Phenotype of ORFA Mutant Strain SL68	83
DNA Sequencing and Sequence Analysis of <i>nodZ</i>	85
Expression of <i>nodZ</i>	91
IV. DISCUSSION	94
REFERENCES	107
APPENDIX 1	122
VITA	127

LIST OF TABLES

TABLE	PAGE
1. Predicted biochemical function of Nod proteins based on sequence homology	26
2. Strains, plasmids and phages used	34
3. β -Galactosidase activity of <i>nolO-lacZ</i> fusions	63
4. β -Galactosidase activity of <i>ORFA-lacZ</i> fusions	80
5. Interstrain hybridization to <i>ORFA</i> , <i>nodZ</i> or <i>nolO</i> probes	82
6. β -Galactosidase activity of <i>nodZ-lacZ</i> fusions	92

LIST OF FIGURES

FIGURE	PAGE
1. Major symbiotic clusters from <i>R.meliloti</i> AK631, <i>R.leguminosarum</i> bv. <i>viciae</i> 248, <i>R.leguminosarum</i> bv. <i>trifolii</i> ANU843 and <i>B.japonicum</i> USDA110	11
2. Structures of Nod factors from <i>B.japonicum</i> , <i>R.meliloti</i> and <i>R.leguminosarum</i> bv. <i>viciae</i> . . .	22
3. Genetic and restriction map of the major symbiotic cluster of <i>B.japonicum</i> chromosome . .	37
4. DNA sequence of <i>nolMNO</i>	49
5. Analysis of <i>nolMNO</i> sequence by TestCode	53
6. Multiple sequence allignments of the predicted amino acid sequences of the following genes: <i>nolO</i> and <i>nodU</i> from <i>B.japonicum</i> USDA110, <i>A.caulinodans</i> ORS571 and <i>R.fredii</i> USDA257	56
7. Autoradiogram of a nylon filter carrying total bacterial DNA cut with <i>EcoRI</i> and probed with ³² P-labeled 0.7 kb fragment internal to <i>nolO</i>	59
8. RNA dot blot probed with ³² P-labeled probes: a) <i>nolO</i> , b) 23S rRNA	62
9. Percentage of nodulated plants after inoculation with <i>B.japonicum</i> strains USDA110 or SL65	65
10. Number of nodules/plant 20 days after inoculation with <i>B.japonicum</i> strains USDA110 and SL65	67
11. Growth curve of <i>B.japonicum</i> strains USDA110, SL65, SL67 and SL68	68
12. Structure of the Nod signal produced by the wild-type <i>B.japonicum</i> strain USDA110	70
13. Nod metabolites produced by strain SL65 (<i>nolNO</i> ⁻)	72
14. FAB-MS and MS-MS spectra of SL65 (Peak 3) . . .	73

15.	Genetic and restriction map showing ORFA and <i>nodZ</i>	76
16.	DNA sequence of ORFA	77
17.	Analysis of ORFA sequence by TestCode	78
18.	Primer extension analysis of the ORFA transcript	84
19.	Percentage of nodulated <i>G.max</i> cv. Essex plants after inoculation with <i>B.japonicum</i> strains USDA110 or SL68	86
20.	DNA sequence of <i>nodZ</i>	88
21.	Model for the biosynthetic pathway of the <i>B.japonicum</i> USDA110 Nod factor	100

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

Rhizobium Symbiosis and Taxonomy

Bacteria from the family *Rhizobiaceae* are Gram negative, aerobic rods that establish a symbiotic relationship with plant hosts. This relationship is of significant benefit to both the plant and the bacterium. The microsymbionts reside inside outgrowths developed from the host root where they convert atmospheric nitrogen into ammonia which is subsequently utilized by the plant. Likewise, inside the host, bacteria receive plant-derived carbon for their active metabolism and to support the energy demanding process of nitrogen fixation. During this symbiosis, rhizobia induce the formation of root nodules which are the specialized plant organs that harbor bacteria and where nitrogen fixation takes place.

The *Rhizobium* nitrogen-fixing symbiosis is highly specific, a feature which is manifested on multiple levels. With one known exception, the host is always a member of the family *Leguminosae*. The tropical woody tree *Parasponia* (family *Ulmaceae*) is the only known non-legume nodulated by rhizobia (Akkermans et al., 1978). Each rhizobial species

has a defined host range. Furthermore, within the same bacterial species, different strains may have different cultivar specificities. For example, *Bradyrhizobium japonicum* falls into two distinct taxonomic groups (i.e., groups I and II) that differ in their ability to nodulate the restrictive Rj4 cultivars of soybeans (see below).

The first attempts to classify rhizobia relied solely on host range as the criterion to define species (Fred et al., 1889). This system for classifying rhizobia has proven to be inadequate; for example, it was realized that rhizobia with a similar host range can be too diverse to fall into the same genus (Elkan, 1992; Young, 1992). Furthermore, only a small fraction of legumes has been tested for nodulation by rhizobia. Therefore, adopting host range as the criterion for classification can only result in an everchanging rhizobial taxonomy. Indeed, studies using DNA:DNA and RNA:DNA hybridization, EPS composition, antibiotic resistance, 16S rRNA sequence and other methods have revealed differences among members of the same species (and genera) that indicate that the current taxonomy and nomenclature will have to change (e.g., Hollis et al., 1981; Young et al., 1991). Currently, the accepted genera in the family *Rhizobiaceae* are *Rhizobium* (fast-growers), *Bradyrhizobium* (slow-growers) and *Azorhizobium* (stem-nodulating rhizobia) (Young, 1992). Other genera that have

been proposed include *Sinorhizobium* (fast-growing soybean rhizobia) (Chen et al., 1988) and *Photorhizobium* (photosynthetic, stem-nodulating rhizobia; Eaglesham et al., 1990). These latter designations, however, have been challenged (Young, 1992) and have not been widely accepted.

Slow-growing bacterial isolates that nodulate soybeans are grouped in a single species and referred to as *Bradyrhizobium japonicum*. DNA:DNA hybridization studies revealed that *B. japonicum* strains fall into at least 2 groups (i.e., groups I and II) that are taxonomically distinct (Hollis et al., 1981). This finding was subsequently confirmed by numerous studies emphasizing differences between the two groups in aspects such as EPS structure (Huber et al., 1984), resistance to different antibiotics (Kuykendall et al., 1988) and the production of lipo-oligosaccharide signal molecules (Carlson et al., 1993). Groups I and II of *B. japonicum* also display differences in their host range in that most of the members of group I strains are able to nodulate the *Rj4* restrictive cultivars of soybeans, while group II strains cannot (Vest and Caldwell, 1972; Devive et al., 1990). The extensive differences between groups I and II have led Kuykendall et al. (1992) to propose a new species, *B. elkanii*, to replace group II strains. However, this suggestion has yet to be

widely accepted. Strain USDA110 used in this study is a *B.japonicum* group I strain.

Stages of Nodule Development

In addition to their symbiosis, rhizobia exist in a free-living form and can be recovered from most soils in the absence of compatible hosts. Chemotaxis towards host root exudates or towards chemicals that induce the ability of bacteria to elicit nodulation (i.e., flavonoids) has been demonstrated (Caetano-Anollés et al., 1988; Gaworzewska and Carlile, 1982). This chemotaxis, however, does not appear to be specific (Gaworzewska and Carlile, 1982) and the motility of *R.leguminosarum* bv. *viciae* was shown to be dispensable for nodulation under laboratory conditions (Smit et al., 1987). Once bacterial cells are in the vicinity of the plant root, their nodulation genes (described below) are induced by the plant-produced flavonoids (Smit et al., 1992; Banfalvi et al., 1988; Spaink et al., 1987; Zaat et al., 1987). However, the induction of nodulation genes does not appear to be the first signaling event in the communication between *Rhizobium* and its host. Apparently, the presence of a compatible *Rhizobium* in the rhizosphere can affect flavonoid production by the host. For example, root exudates from *Vicia sativa* ssp. *nigra* pre-treated with its homologous microsymbiont, *R.leguminosarum* bv. *viciae*, were

more capable of inducing nodulation genes when compared to exudates from uninoculated roots (van Brussel et al., 1990).

Induction of the early events of nodulation does not require direct contact between the bacterium and plant root (Banfalvi and Kondorosi, 1989; Faucher et al., 1988; Schmidt et al., 1988; Debelle et al., 1986; van Brussel et al., 1986). For example, cell-free culture extracts of flavonoid-induced *R.meliloti* specifically induced root-hair curling on its homologous host alfalfa, but not on the non-host vetch (Faucher et al., 1988; Debelle et al., 1986). Furthermore, van Brussel et al. (1986) demonstrated the induction of thick and short roots (TSR) on vetch using cell-free extracts of flavonoid-induced *R.l. bv. viciae* or *trifolii*. Although TSR is not a feature associated with nodule formation, it is significant since this hormone-like effect on the plant root was *nod*-gene dependent (van Brussel et al., 1986). The notion that direct contact is not a prerequisite for the early events of the *Rhizobium*-legume symbiosis was further tested using purified fractions of rhizobial extracts that were able to simulate the early events of nodulation in the absence of bacterial cells (Truchet et al., 1991).

Root-hair curling (or deformation) and cortical cell division are the first visible events in the nodulation

process. The function of root-hair curling may be to entrap the bacteria inside the curl; thus, increasing their density and subsequently facilitating the penetration of the host root (Nutman, 1959). However, there are several reported examples where root hairs are not involved in the infection process (Kijne, 1992; Tsien et al., 1983; Finan et al., 1985). Nevertheless, root-hair curling is *nod*-gene dependent (Faucher et al., 1988; Develle et al., 1986; Finan et al., 1985) and appears to be host specific (Spaink, 1992; Faucher et al., 1988; Debelle et al., 1986). The initiation of cortical-cell division, another early event in nodulation that can be induced by cell-free extracts of induced rhizobia, probably precedes infection (Dudley et al., 1987; Libbenga and Harkes, 1973). The origin of cortical-cell division is determined by the host plant (see below).

Following the initial plant-bacterium cross-talk described above, bacterial cells attach to the root surface in a process thought to be mediated by a bacterial calcium-binding adhesin (i.e., rhicadhesin; Smit et al., 1989). Attachment is followed by degradation of the plant cell-wall and a new layer of host synthesized wall material is deposited, initiating the infection thread (IT; Callaham and Torrey, 1981). Recently, van Brussel et al. (1992) reported that during the infection of vetch or pea plants, or following the treatment with signal molecules produced by

R. l. bv. viciae, the nuclei and the cytoplasm of outer cortical cells migrate to the cell center forming radial cytoplasmic bridges. The formation of these structures precedes the formation of the IT and, therefore, they were termed the pre-infection threads (PIT; van Brussel et al., 1992). The bacteria penetrate into the plant cortex via the IT (Kijne, 1992; Libbenga and Harkes, 1973), which grows and ramifies into dividing cortical cells where bacteria are released. Upon release of the bacteria into the plant cell, they are surrounded by a membrane of plant origin forming a vacuolar structure termed a symbiosome (Roth et al., 1988). The symbiosome is the compartment where nitrogen fixation takes place.

Nodule formation and structure are genetically determined by the host plant. This claim is best supported by studies where genuine nodules were formed on plant roots in the absence of rhizobia or their culture extracts (Truchet et al., 1989). This implies that the legume host is pre-programmed to form nodules and that the timing of the event is dictated by compatible rhizobia. Legume nodules fall into two groups (i.e., determinate and indeterminate nodules) based on their development, structure and persistence (Nap and Bisseling, 1990; Newcomb et al., 1979). The fundamental difference between indeterminate and determinate nodules is that the former has an apical

meristem that continuously provides newly formed cortical cells for further infection; while the latter has a non-persistent meristem that ceases to divide after a defined period of time after which cortical cells grow in size, but not in number. Structural differences among the two kinds of nodules are imposed by the different meristems and by the origins of cortical-cell division (for recent reviews, see Kijne, 1992; Nap and Bissling, 1990). The same bacterial strain can induce determinate nodules on one host but indeterminate nodules on another (e.g., Dart, 1977). An important feature that distinguishes both kind of nodules from lateral roots is the occurrence of peripheral vascular bundles in contrast to the central vascular bundle present in lateral roots (Newcomb, 1981). Nodules formed by the non-legume *Parasponia* are an exception to this generality (Price et al., 1984).

The nodule environment is best described as microaerobic due to an oxygen barrier formed by the root tissues (Layzell et al., 1990; Tjepkema and Jocum, 1974). While this feature of the nodule is consistent with the extreme sensitivity of the nitrogenase enzyme to free oxygen, active oxidative phosphorylation is required by the bacteria to generate enough ATP for nitrogen fixation (The reduction of 1 mole of N_2 into NH_3 utilizes at least 16 moles of ATP; Thorneley and Lowe, 1985). This apparent

paradox is resolved partly by the presence of leghemoglobin, which binds oxygen reversibly and is abundant in infected plant cells (Wittenberg et al., 1974). Oxygen released from leghemoglobin is proportional to that consumed during bacterial respiration; thus, leaving little free oxygen to inactivate nitrogenase (Appleby, 1984; Bergersen, 1982). Furthermore, an alternative bacterial respiratory chain appears to be induced in the nodule which has a high affinity terminal oxidase that functions at nanomolar concentrations of free oxygen (Bergersen and Turner, 1990; O'Brian and Maier, 1989). Interestingly, the globin portion of leghemoglobin is plant produced, while the bacterial symbiont produces and releases the heme prosthetic group (Nadler and Avissar, 1977). However, a *B. japonicum* strain mutated in the gene encoding δ -aminolevulinic acid synthetase, an essential enzyme for heme synthesis, was able to induce nitrogen fixing nodules that were not heme deficient (Guerinot and Chelm, 1986). Sangwan and O'Brian (1991) provided evidence that this mutant can use the plant-produced δ -aminolevulinic acid as a precursor for heme biosynthesis.

In summary, nodules are very complex and specialized structures designed to accommodate the requirements of symbiotic nitrogen fixation. This complexity is reflected in the number of genetic determinants (both plant and

bacterial) involved in the formation and maintenance of a nitrogen fixing nodules as well as the tight transcriptional control of these genes.

Bacterial Genes Required for Efficient Nodulation

The biological events leading to the formation of a nitrogen fixing nodule require a very large number of bacterial and plant genes (Caetano-Anollés and Gresshoff, 1991). The bacterial genes are required for inducing the morphological changes associated with nodule formation, as well as adapting the bacterial cell to a new environment, the nodule, and a new physiological task, symbiotic nitrogen fixation. A compilation of all known genes in *Rhizobium* and *Bradyrhizobium* has recently been published (Sanjuan et al., 1993). Fig 1 shows the organization of symbiotic genes from representative rhizobial species (Fig 1). These genes were identified using insertional mutagenesis, reporter gene fusions, interspecies DNA hybridization to probes from previously identified genes, genetic complementation studies or simply by sequencing DNA regions flanking known symbiotic genes. These efforts were aided by the tendency of such genes to occur in clusters on the chromosome (as in *Bradyrhizobium* and *Azorhizobium*) or on plasmids (as in *Rhizobium*). The following sections will focus on the organization and symbiotic function of genes from *Rhizobium*

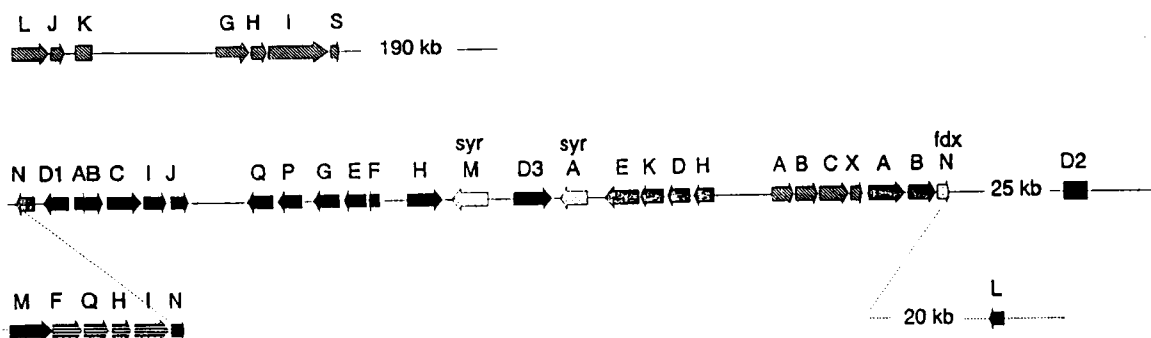
Figure 1

Major symbiotic clusters from *R.meliloti* AK631,
R.leguminosarum bv. *viciae* 248, *R.leguminosarum* bv. *trifolii*
ANU843 and *B.japonicum* USDA110.

Rhizobium meliloti RCR2011

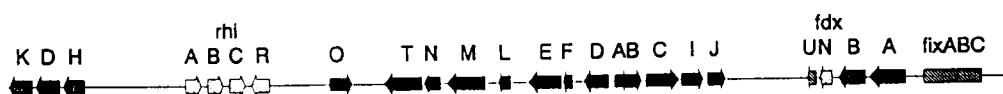
plasmid pSym

[strain AK631]



R. leguminosarum biovar *viciae* 248

plasmid pRL1Jl

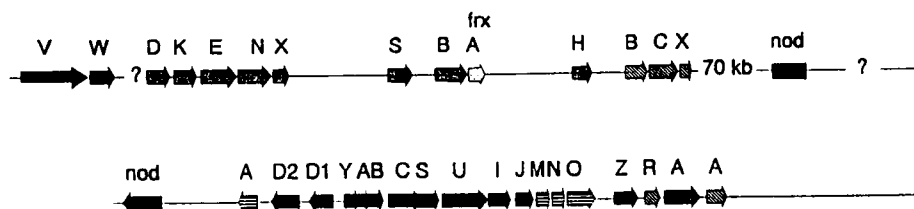


R. leguminosarum biovar *trifolii* ANU843

plasmid pRt843a



Bradyrhizobium japonicum USDA110



nif other nod fix nol

2 kb

and *Bradyrhizobium* species since more is known about these organisms than the genus *Azorhizobium*.

i. nod/nol Genes

Genes required for the morphological changes associated with the early events of nodule formation are referred to as nodulation (*nod/nol*) genes. Mutational inactivation of one or more of these genes can result in a variety of nodulation phenotypes, the most drastic of which is the complete inability to form nodules on any plant host (Nod^- phenotype). Other mutants may exhibit a delay in nodule formation compared to the wild-type (Nod^d phenotype). Inactivating some of the *nod/nol* genes can alter the host range of the strain (Hsn phenotype) or may not result in any detectable alteration in nodulation ability (Nod^+ phenotype). The identification of the latter genes as *nod/nol* often follows from the fact that they are co-transcribed with other known nodulation genes. Such cryptic *nod/nol* genes may be required for nodulation of unidentified hosts or play a role in nodulation in a field situation.

The *nodABC* gene products are indispensable for the ability of any rhizobial strain to nodulate its hosts. They are called common *nod* genes since they are conserved among rhizobia regardless of host range and because they can complement *nodABC* mutants of different species with varying

host ranges (Scott, 1986; Fisher et al., 1985; Marvel et al., 1985; Kondorosi et al., 1984). Two additional genes, *nodIJ*, seem to be conserved among all rhizobia as part of the *nodABC* operon (Göttfert et al., 1990; Nieuwkoop et al., 1987; Spaink et al., 1987; Evans and Downie, 1986). Although *nodIJ* mutations in *R.l. bv. viciae* cause a delayed nodulation phenotype on vetch (Canter Cremers et al., 1990; 1988), there is no absolute requirement for *nodIJ* in the nodulation process.

The *nodD* gene, which codes for an activator of *nod* gene transcription, is commonly found in all rhizobia, yet plays a role in host specificity (Gyorgypal et al., 1988; Horvath et al., 1987; Spaink et al., 1987). The NodD of a given *Rhizobium* is adapted to recognize the flavonoids produced by specific host species; thus, inducing *nod* operons in response to compatible hosts (Schlaman et al., 1992; Banfalvi et al., 1988; Horvath et al., 1987; Spaink et al., 1987). However, allelic forms of *nodD* are present in many rhizobia (Göttfert et al., 1989; Appelbaum et al., 1988; Honma and Ausubel, 1987). For example, *R.meliloti* possesses three *nodD* copies, the products of which have different flavonoid specificities (Gyorgypal et al., 1988; Honma and Ausubel, 1987).

In addition to the nodulation genes conserved in all rhizobia, some genes are limited to particular species and are expected to be determinants of host specificity. For example, *nodH*, found in *R.meliloti*, is required for nodulation of alfalfa but not sweet clover (Swanson et al., 1987; Horvath et al., 1986). Furthermore, mutations in the *nodFE* region of *R.l. bv. trifolii* impair nodulation of clover yet extend the host range to include pea, normally not a host plant (Djordjevic et al., 1985). In *R.leguminosarum bv. viciae*, *nodX* enables strain TOM to nodulate afghanistan pea, thus demonstrating specificity on a cultivar level (Davis et al., 1988). Two other examples of such genotype specific nodulation (GSN) genes are the *B.japonicum nola* and *nodZ* genes. *nola* enables *B.japonicum* strain USDA123 to nodulate soybean varieties that are normally restrictive to this strain (Sadowsky et al., 1990), while *nodZ*, in addition to its requirement for siratro nodulation, appears to be essential for the efficient nodulation of soybean cultivars Peking and Forest, but not Essex (Stacey et al., 1993).

Several common features are observed regarding the organization of *nod/nol* genes in *Rhizobium* and *Bradyrhizobium*. For example, whether on a symbiotic plasmid or on the chromosome, the nodulation genes tend to occur in clusters (Fig 1; Sanjuan et al., 1993). Furthermore, *nodABC*

usually occur in a single operon followed by *nodIJ* and preceded by a divergently transcribed *nodD* (Fig 1; Sanjuan et al., 1992). Nevertheless, interesting organizational differences are observed between *Rhizobium* and *Bradyrhizobium*. For example, unlike the case in *Rhizobium* species, *nodA* is preceded by an additional gene in *B.japonicum* (i.e., *nodY*; Nieuwkoop et al., 1987) and in *Bradyrhizobium* sp. *Parasponia* (i.e., *nodK*; Scott, 1986). Furthermore, the *nodIJ* genes are located immediately downstream of *nodABC* in *Rhizobium* species but separated from them by *nodSU* in *B.japonicum* (Göttfert et al., 1990). Additional unique features of the *B.japonicum* *nodABC* operon are presented in this work.

ii. nif and fix genes

The block in some symbiotic mutants can occur in stages following nodule formation, resulting in nodules (or nodule-like structures) that do not fix nitrogen and are thus referred to as ineffective nodules. Mutations in *nif* or *fix* genes can result in the formation of ineffective nodules. While *nif* genes are required for nitrogen fixation *per se*, *fix* genes are required for symbiotic nitrogen fixation. This distinction, however, is not very well defined since, for example, *fixABC* genes of *B.japonicum* are required for symbiotic as well as artificially induced nitrogen fixation (Gubler and Hennecke, 1986). One

criterion used to distinguish between *nif* and *fix* genes is that each *nif* gene should have a structural counterpart in a free-living diazotroph such as *Klebsiella pneumoniae*. In general, *nif* genes are directly involved in nitrogenase structure and function (recently reviewed in Dean and Jacobson, 1992). The Fix phenotype, on the other hand, is rather ill-defined since a Fix⁻ bacterial mutant can be defective in diverse physiological aspects ranging from electron transport for nitrogenase function (*fixABC*, Gubler and Hennecke, 1986) to respiration (*fbcFH*, Thöny-Meyer et al., 1989) to C₄ dicarboxylate transport (Bolton et al., 1986). Like *nod/nol* genes, *nif* and *fix* genes tend to occur in clusters (Fig 1), a feature that is more pronounced in *Rhizobium* than in *Bradyrhizobium*. For example, while *nifHDK* in *Rhizobium* (as well as in *Klebsiella*) form a single operon, *nifH* and *nifDK* of *Bradyrhizobium* are separated by several kilobases (Fig 1; Fischer and Hennecke, 1984). Furthermore, *fixA* and *fixBC* occur in two different genetic clusters in *B. japonicum* while *fixABC* represent a single operon in *Rhizobium* (Fig 1; Earl et al., 1987; Gubler and Hennecke, 1986).

Most of the *nif* and *fix* genes in rhizobia are transcriptionally activated by conditions of low oxygen similar to those prevailing inside the nodule where nitrogen fixation takes place (Batut et al., 1992; Fischer et al.,

1991). Oxygen regulation is mediated by NifA, a transcriptional activator of RpoN-dependent promoters. The mechanism by which changes in oxygen concentration are sensed by *Rhizobium* differs from that adopted by *Bradyrhizobium* (Anthamatten and Hennecke, 1991). In *R.meliloti*, NifA function is controlled by regulating both the level of *nifA* transcription and the activity of NifA protein. The transcriptional activation of *R.meliloti nifA* is mediated by FixL and FixJ proteins which are members of the two-component family of prokaryotic transcriptional regulators (David et al., 1988): FixL, a hemoprotein that senses oxygen, is a kinase that phosphorylates FixJ which in turn activates the transcription of *nifA* (Gilles et al., 1991). The post-translational regulation of the *R.meliloti* NifA protein is attributed to its oxygen sensitivity (Fischer et al., 1988). Although the regulation of *B.japonicum* NifA is also manifested at both the transcriptional and post-translational levels, it displays interesting differences when compared to that of *R.meliloti* (Fischer and Hennecke, 1987; Thöny et al., 1987). First, the *B.japonicum nifA* is expressed aerobically, yet it is slightly induced (i.e., 5-fold) under anaerobic conditions (Thöny et al., 1989). Second, the anaerobic induction of *B.japonicum nifA* is not mediated by FixLJ but instead by NifA itself (Thöny et al., 1989). As in *R.meliloti*,

B.japonicum NifA is rapidly and irreversibly inactivated by oxygen (Kullik et al., 1989).

iii. Other Genes Involved in Rhizobium-Legume Symbiosis

In addition to *nod*, *nif* and *fix* genes, other genetic determinants are required for the production and maintenance of a nitrogen fixing nodule. These genes can either be chromosomal or plasmid born and are not tightly linked to the known symbiotic clusters. Genes involved in the synthesis of exopolysaccharides (EPS) and lipopolysaccharides (LPS) have been implicated in rhizobial symbioses in numerous studies (e.g., Long et al., 1988; Puvanesarajah et al., 1987). Cell surface mutants can exhibit a wide variety of symbiotic phenotypes including the induction of deformed nodules and the failure of mutant strains to be released from infection threads. Interestingly, a transcriptional activator, SyrM, was found to be involved in the regulation of both *exo* and *nod* genes; thus, providing evidence for the functional relatedness of these two classes of genes (Mulligan and Long, 1989). Similarly, in *R.fredii*, multiple copies of *nodD₂* inhibited EPS synthesis and resulted in the formation of Muc⁻ (i.e., non-mucoid) colonies (Appelbaum, 1988).

Other bacterial genes are required for accessory functions associated with symbiotic nitrogen fixation. For

example, mutations in the *dct* genes required for the transport of C4 dicarboxylic acids (the most likely carbon sources for the bacteroid) result in a *Fix*⁻ phenotype (e.g., Finan et al., 1983). Furthermore, the respiratory demands imposed by nitrogen fixation in a microaerobic environment appear to be met by the induction of an alternative, nodule-specific bacterial respiratory chain with a high-affinity terminal oxidase that functions at nanomolar concentrations of free oxygen (Bergersen and Turner, 1990; O'Brian and Maier, 1989; Thöny-Meyer et al., 1989). Accordingly, mutations in respiratory genes can result in a *Fix*⁻ phenotype (e.g., Thöny-Meyer et al., 1989).

iv. Summary and Perspective

Although this section has only presented examples representative of genes involved in the *Rhizobium*-legume symbiosis, the genetic complexity of the nodulation process should be clear. The identification of *nod*, *nif* and *fix* genes has been facilitated by the fact that they occur in large clusters either on symbiotic plasmids (as in *Rhizobium*) or on the chromosome (as in *Bradyrhizobium* and *Azorhizobium*). Several interesting differences between *Rhizobium* and *Bradyrhizobium* were observed in the operon structures of *nod*, *nif* and *fix* genes. Features unique to the *nodABC* operon of *B.japonicum* are addressed in this work.

Biochemical Function of *nod* Gene Products

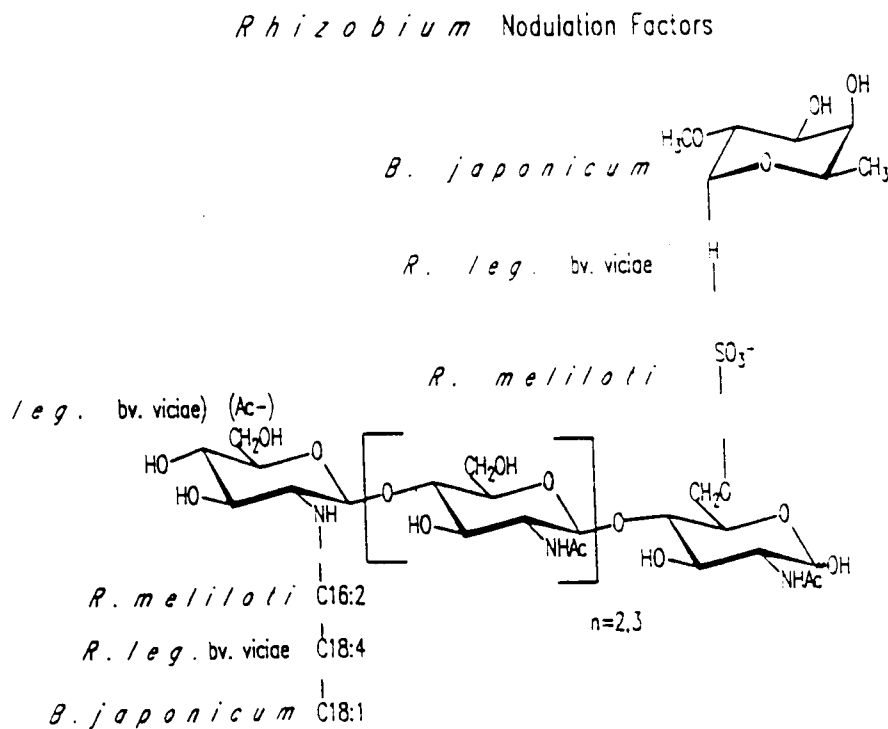
It is widely believed that most, if not all, of the known *nod/nol* genes are involved in the production of lipo-oligosaccharide signal molecules with host-specific biological activity (Spaink, 1992). To present the proposed role for *nod/nol* genes in the synthesis of these molecules, their chemical structure is reviewed in the following section.

i. Lipo-oligosaccharide Signal Molecules Produced by Rhizobia

Establishment of a *Rhizobium*-legume symbiosis requires the mutual exchange of signals between the partners involved (i.e., the bacterium and its homologous host). The signal received by rhizobia in the rhizosphere is the plant produced flavonoids that specifically induce the *nod* genes. Several reports have indicated that rhizobia produce a *nod*-gene dependent signal capable of eliciting root-hair deformation (Had) and curling (Hac) and cortical cell division (Coi) on homologous hosts (Faucher et al., 1988; van Brussel et al., 1986; Debelle et al., 1986). Had, Hac and Coi are among the earliest observable events of nodulation. Recently, the nodulation signal molecule was purified from culture supernatants of flavonoid-induced *R.meliloti* strain 2011. Chemical analysis of the purified factor showed it to be an oligomer of chitin (a β ,1-4 linked

Figure 2

Structures of the Nod factors from *B.japonicum*, *R.meliloti* and *R.leguminosarum* bv. *viciae*. (From Stacey et al., 1993).



polymer of *N*-acetyl glucosamine), *N*-acylated and *O*-acetylated on the non-reducing end and sulfated at the reducing end (Fig 2; Schultze et al., 1992; Roche et al., 1991; Lerouge et al., 1990). The subsequent elucidation of the chemical structures of lipo-oligosaccharide signal molecules (also referred to as Nod factors or Nod metabolites) from other rhizobia has proceeded rapidly since this first report (Fig 2; reviewed in Spaink, 1992).

Rhizobia with differing host ranges produce different Nod factors (Carlson et al., 1993; Mergaert et al., 1993; Price et al., 1992; Sanjuan et al., 1992; Schultze et al., 1992; Spaink et al., 1991; Roche et al., 1991; Lerouge et al., 1990). All the reported Nod factor structures have two common features; namely, the chitose backbone (as a tri-, tetra- or pentamer) and the *N*-acylation of the terminal non-reducing sugar with a fatty acyl residue whose length and unsaturation pattern vary among different strains. Figure 2 shows the different chemical structures of Nod factors from representative rhizobial strains. The purified factors exhibit a wide range of biological activity at nanomolar concentrations. For example, purified Nod factors have been shown to induce root-hair curling and cortical cell division (Sanjuan et al., 1992; Spaink et al., 1991; Truchet et al., 1991). However, the formation of nodules that are anatomically similar to genuine bacteria-

induced structures has only been demonstrated in the case of alfalfa treated with the Nod metabolite from *R.meliloti* (Truchet et al., 1991). This is likely a special case since alfalfa can form nodules spontaneously in the absence of bacteria or external signals (Truchet et al., 1989).

Each rhizobial strain produces at least one factor with a unique structure. The structural differences among the Nod factors appear crucial for host specific biological activity. For example, the sulfation of the reducing end of the *R.meliloti* Nod factors is required for the nodulation of *Medicago* but not *Melilotus* (Roche et al., 1991; Faucher et al., 1988). Furthermore, a highly unsaturated fatty acyl chain (C18:4) is characteristic of the *R.l. bv. viciae* Nod factor and is lacking in the Nod factors from *R.l. bv. trifolii* which possesses a different host range (Spaink et al., 1991). Some rhizobial species have a sugar moiety attached to position 6 of the terminal reducing sugar. Fucose, 2-O-methylfucose, sulfate or acetyl-substituted 2-O-methyl fucose or D-arabinose are known substituents on the terminal reducing sugars of Nod factors produced by *B.japonicum* USDA61, *B.japonicum* USDA110, *Rhizobium* sp. NGR234 and *A.caulinodans* ORS571, respectively (Carlson et al., 1993; Mergaert et al., 1993; Price et al., 1992; Sanjuan et al., 1992). Furthermore, a unique feature was found for one of the Nod metabolites produced by *B.japonicum*

USDA61; that is, a glycerol moiety was found glycosidically linked to the reducing hydroxyl group of the molecule (Carlson et al., 1993). The structure of a molecule with the same glycerol substituent is reported in this work.

ii. Role of nod Genes in Nod Factor Biosynthesis

nod/nol genes have been implicated in the synthesis and modification of various Nod factors (reviewed in Spaink, 1992). Interesting parallels can be drawn between *nod* genes and Nod factors regarding their function and distribution among different strains. The common *nod* genes are present in all rhizobia while the occurrence of the *hsn* genes varies among strains with different host ranges. Similarly, all Nod factors purified thus far have a common backbone (the acylated chitose molecule) with various substitutions characteristic of each strain. Furthermore, strains carrying mutations in the *nodABC* genes are unable to nodulate any of their original hosts and they also fail to produce any of the Nod metabolite derivatives (Sanjuan et al., 1992; Spaink et al., 1991; Truchet et al., 1991; Lerouge et al., 1990). On the other hand, strains mutated in *hsn* genes usually have an altered host range. These *hsn* mutations result in the production of Nod factors having the wild-type backbone structure yet with different substitution patterns that result in alterations in host-specific biological activity (Schultze et al., 1992; Roche et al.,

**Table 1: Predicted Biochemical Function of Nod Proteins
Based on Sequence Homology**

Nod Protein	Predicted Function
NodB	EnvA (a possible deacetylase) ²
NodC	<i>N</i> -acetylglucosaminyltransferase
NodE	β -ketoacylsynthase
NodF	acyl carrier protein
NodG	NADH-dependent dehydrogenase NADPH-dependent β -ketoacylreductase
NodH	Sulfotransferase
NodL	Acetyltransferase
NodM	D-glucosamine synthase
NodP	ATP sulfurylase subunit
NodQ	ATP sulfurylase subunit
NodS	SAM-utilizing enzymes ³

1, Adapted from Spaink, 1992.

2, Reference: Spaink et al., 1993.

3, Reference: Geelen et al., 1993.

1991; Spaink et al., 1991; Lerouge et al., 1990). Thus it is currently believed that the common *nod* genes are involved in the synthesis of the chitose backbone and that the *hsn* genes are required for the addition of specific "decorations" peculiar to each strain (reviewed in Spaink, 1992).

Comparing the predicted amino acid sequences of nodulation genes to those of previously identified sequences from prokaryotic or eukaryotic origins has shed some light on their possible biochemical functions. Nod proteins that show sequence similarity to other proteins with known functions are listed in Table 1. Some of these speculations are backed by experimental evidence. For example, the role of NodH and NodPQ in the sulfation of *R.meliloti* Nod factors is the best described. Table 1 shows that the amino acid sequence of NodH is similar to mammalian sulfotransferases consistent with the fact that *nodH* mutants produce Nod factors that lack the sulfate substitution on their reducing end (Roche et al., 1991). NodP and NodQ are similar to *E.coli* ATP sulfurylase subunits encoded by *cysD* and *cysN*, respectively (Table 1). A DNA fragment carrying *nodPQ* was found to complement the auxotrophy of *cysD* and *cysN* *E.coli* mutants (Schwedock and Long, 1990). Furthermore, the ATP sulfurylase activity of NodP and NodQ has been demonstrated both *in vivo* and *in vitro* (Schwedock and Long, 1990). The

Nod factors produced by *nodP* and *nodQ* mutants are a mixture of sulfated and non-sulfated compounds (Roche et al., 1991). Such a leaky phenotype could be justified by the fact that an additional copy of *nodPQ* is found on the *R.meliloti* chromosome (Schwedock et al., 1989). Another example is the NodS protein which shows similarity to a consensus sequence conserved in *S*-adenosylmethionine (SAM)-utilizing enzymes (Geelen et al., 1993). It is proposed to be a methyl transferase that mediates the *N*-methylation of *A.caulinodans* Nod factors (Geelen et al., 1993; Mergaert et al., 1993). Indeed, the *in vivo* methylation of Nod metabolites from *A.caulinodans* was *nodS* dependent (Geelen et al., 1993). Mutants lacking *NodL*, which shows sequence similarity to *E.coli* acetyl transferases, produced Nod factors lacking the *O*-acetylation on the terminal non-reducing sugar. Furthermore, *NodF* and *NodE* are similar in sequence to *E.coli* acyl carrier protein and β -ketoacylsynthase, respectively. Mutations in *nodFE* result in the production of Nod factors with alterations in the *N*-acylation of the terminal non-reducing sugar (Demont et al., 1993; Spaink et al., 1991).

In *B.japonicum*, mutations in *nodZ* result in the production of Nod metabolites lacking the 2-*O*-methylfucosyl substitution on the terminal reducing sugar (Carlson et al., 1993). To investigate the relationships of *NodZ* to other known proteins, the *nodZ* DNA sequence was determined and is

presented in this work. Furthermore, a novel *nol* gene linked to *nodZ* and involved in the 2-O-methylfucosylation of Nod factors is reported.

Regulation of Nodulation Genes

With few exceptions, *nod* genes are inactive until the bacterium is in the vicinity of the host root where they are transcriptionally induced by plant produced flavonoids (Smit et al., 1992; Banfalvi et al., 1998; Spaink et al., 1987; Zaat et al., 1987). In addition to the chemical inducer, *nod* gene induction requires a functional NodD (Banfalvi et al., 1988; Spaink et al., 1987; Rossen et al., 1985; Mulligan and Long, 1985) and a conserved *cis*-acting promoter element, the *nod*-box, which is the NodD binding site (Goethals et al., 1992; Kondorosi et al., 1989; Fisher et al., 1988; Hong et al., 1987; Rostas et al., 1986). NodD is a member of the LysR family of transcriptional regulators the members of which share several common properties including a conserved N terminal DNA-binding domain (Schlaman et al., 1992; Henikoff et al., 1988). The *nod*-box was shown to bind NodD *in vitro* and while this binding was found to be independent of the chemical inducer in some systems (Fisher et al., 1988; Hong et al., 1987), the presence of the inducer enhanced such binding in other systems (Goethals et al., 1992; Kondorosi et al., 1989). It is assumed that the chemical inducer directly interacts with

NodD facilitating *nod* gene transcription. Although direct NodD-flavonoid interaction is not supported by biochemical evidence, strong genetic data support such an interaction. For example, mutations in *nodD* alter the flavonoid specificity of *nod* gene induction or extend the host range of mutant strains (Burn et al., 1989; McIver et al., 1989; Burn et al., 1987; Spaink et al., 1987). Since NodD is capable of binding *nod*-box sequences in the absence of chemical inducers (Fisher et al., 1988; Hong et al., 1987), the role of the flavonoids may be to induce a conformational change in NodD that allows *nod* gene transcription. This hypothesis is supported by studies where *nodD* mutant alleles produced proteins that can activate *nod* genes in the absence of flavonoids (Spaink et al., 1989; Burn et al., 1987).

Although the role NodD plays in *nod* gene induction is well established, the regulatory circuitry possesses additional levels of complexity. For example, *nodD* itself is subject to regulation in several systems. NodD autoregulates its own expression, positively in *B.japonicum* (Smit et al., 1992; Banfalvi et al., 1988) and *R.l. bv. phaseoli* (Davis and Johnston, 1990) and negatively in *R.l. bv. viciae* and *trifolii* (Rossen et al., 1985; Spaink et al., 1987). A *nod* gene repressor has been reported in some *R.meliloti* strains and was found to bind the *nodD*₁ and *nodD*₂ promoters (Kondorosi et al., 1989). Furthermore, combined

nitrogen (such as ammonia) repressed the expression of *nodD*₁ in *B.japonicum* (Wang and Stacey, 1990) and *nodD*₃ in *R.meliloti* (Dusha et al., 1989). This is consistent with the previously reported inhibition of nodulation in the presence of combined nitrogen (Truchet and Dazzo, 1982).

In *B.japonicum*, *nodVW* are required for host-specific nodulation (Göttfert et al., 1990). Inspection of the predicted amino acid sequence of *nodVW* revealed that they are members of the two-component family of transcriptional regulators (Göttfert et al., 1990; Stock et al., 1989) and suggested that they may represent additional *nod* gene activators. Indeed, *nodW* was required for the induction of *nodD-lacZ* and *nodY-lacZ* fusions by isoflavones (Sanjuan et al., 1993). However, the mechanism by which NodVW activate *nod* genes is currently unknown.

Rationale for Present Work

The production of root nodules by various rhizobia on their corresponding hosts has been subject to extensive study for decades. Recently, the nature of this process has been analysed on the molecular level, an approach that provided considerable knowledge on the communication that occurs between the bacteria and its plant host resulting in the production of root nodules in a host-specific manner. Identification of the bacterial *nod* genes has been crucial

for the discovery of the lipo-oligosaccharide signal molecules produced by rhizobia after being induced by the host flavonoids. Since the *nodABC* genes are required for the synthesis of the *B.japonicum* USDA110 Nod factor, the adjacent DNA region was analysed.

The flavonoid inducible nature of *nod* genes and their tendency to occur in clusters were utilized in this study as criteria for identifying novel symbiotic genes. The sequence of five novel genes was determined and analyzed for features that may suggest their biochemical functions. Mutants were constructed and tested for their ability to nodulate their corresponding hosts and for Nod factor production. Transcriptional control of the studied genes was investigated using *lacZ* gene fusions and RNA hybridization.

CHAPTER II

MATERIALS AND METHODS

Bacterial Strains, Plasmids and Bacteriophages

Bacterial strains, plasmids and bacteriophages used in this study are described in Table 2. Plasmid pDG69 is pVK102 carrying a Tn3-*lacZ* fusion inserted in a 0.7 Kb *Pst*I fragment downstream to *nodJ* (Fig 3; Gerhold, 1989). Plasmids pSSL74 and pSSL75 carry a *Sal*I fragment from pDG69 that contains the Tn3-*lacZ* fusion and 2-3 Kb of flanking DNA but excluding the *nodY* *nod*-box from the construct (Fig 3). pSSL74 and pSSL75 have the Tn3-*lacZ* positioned in the opposite or the same orientation, respectively, relative to the *tet* gene of pVK102. Transfection of *E.coli* LE392 cells with pVK102-based constructs was done using the Packagene^R Lambda packaging system (Promega). pSSL74 and pSSL75 were mated into *B.japonicum* USDA110 to yield strains SSL74 and SSL75, respectively. The latter strains were confirmed by DNA Southern blot hybridization (see below). Plasmids pMGS102 (*ORFA-lacZ*) and pMGS104 (*nodZ-lacZ*) were constructed using Tn3-HoHo1 random mutagenesis of pMS95 which is the broad-host-range vector pPP375 carrying a 9.5 kb *Eco*RI fragment that includes *ORFA* and *nodZ* (Fig 3; Table 2). Mapping the position and orientation of the Tn3-*lacZ* insertions was done by *Eco*RI digestion followed by Southern

Table 2: Strains, plasmids and phages used

Designation	Relevant Characteristics	Source/Reference
Bacteria		
<i>E.coli</i>		
S17-1	Chromosomally integrated RP4 derivative	Simon et al., 1983
JM101	<i>lac proAB supE thi/F' traD36 lacI^q ZM15</i>	Messing et al., 1981
LE392	<i>F' lacY1 supE44 supF58</i>	Murray et al., 1977
<i>B.japonicum</i>		
USDA110	Wild-Type	Kuykendall, 1976
AN314	USDA110, <i>nodD::Tn5</i> , Km Sm	Nieuwkoop et al., 1987
SL66	USDA110, <i>nodS del/ins</i> Sp/Sm	This study
SL67	USDA110, <i>nodO::Sp/Sm</i>	This study
SL65	USDA110, <i>nodNO::del/ins</i> Sp/Sm	This study
SL68	USDA110, <i>ORFA::Sp/Sm</i>	This Study
USDA6	Wild-Type	D. Kuykendall
USDA46	"	"
USDA76	"	"
USDA94	"	"
USDA122	"	"
USDA123	"	"
AN6A	"	"

Table 2 (continued)

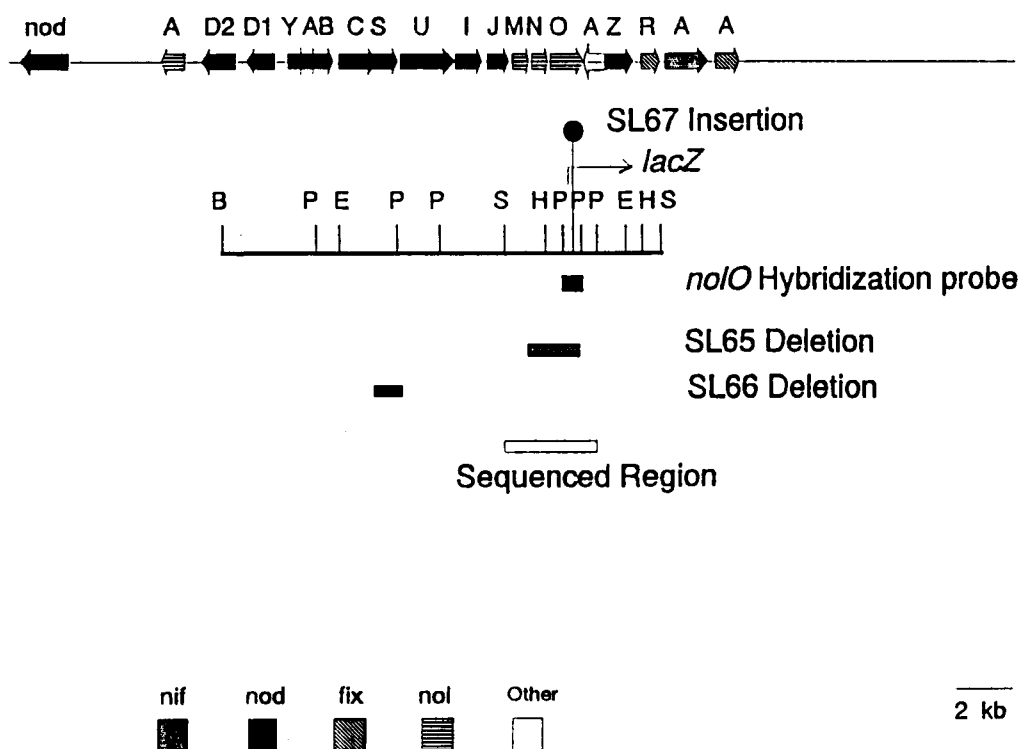
61A101C	Wild-Type	Nitragin
Bj2101	USDA110, <i>fixRnifA</i>	Chelm/Adams
Bj27147	USDA110, <i>ntrC</i>	Martin et al., 1988
Bj3263	USDA110, <i>ntrBC</i>	Martin et al., 1988
Bj-N50	USDA110, <i>rpoN</i> ₁	Kullik et al., 1991
Bj-N63	USDA110, <i>rpoN</i> ₂	Kullik et al., 1991
BjN50-97	USDA110, <i>rpoN</i> ₁ <i>N</i> ₂	Kullik et al., 1991
<i>Bradyrhizobium</i> sp.		
IRC2133	Wild-Type (peanut)	G. Elkan
32H1	,, (cowpea)	P. van Berkum
ANU289	,, (<i>Parasponia</i>)	B. Rolfe
<i>Rhizobium</i>		
R.sp. NGR234	Wild-Type	B. Rolfe
R.l. 897	,,	N. Brewin
<i>R.fredii</i>		
193	,,	H. Keyser
HH103	,,	,,
<i>R.l.bv.</i>		
<i>trifolii</i>	,,	B. Rolfe
<i>R.meliloti</i>		
AK2121	,,	A. Kondorosi
AK631	,,	,,
Plasmids/Phages		
pSUP202	pBR325 derivative containing the Mob site from RP4	Simon et al., 1983

Table 2 (continued)

pUC19	High copy-number cloning vector α -complementation	Vieria and Messing, 1982
PSL22	pUC19+a 6Kb <i>EcoRI</i> - <i>SmaI</i> fragment including <i>nodMNO</i>	This study
PHP450	Sp/Sm resistance cassette	Fellay et al., 1987
PHF360	pUN121+23S rRNA Probe from <i>P.aeruginosa</i>	Festl et al., 1986
pVK102	RP4 based, 1 <i>cos</i> site	Knauf and Nester, 1982
pDG69	<i>nolO</i> -Tn3HoHo1	Gerhold, 1989
pSSL74	<i>nolO</i> -Tn3HoHo1, Δ <i>nod</i> -box opposite to pVK102 <i>tet</i>	This study
pSSL75	<i>nolO</i> -Tn3HoHo1, Δ <i>nod</i> -box same as pVK102 <i>tet</i>	This study
pMGS102	ORFA-Tn3HoHo1	This study
pMGS104	<i>nodZ</i> -Tn3HoHo1	This study
M13mp18, 19	Sequencing cloning vectors	Messing et al., 1981
SL1, 2	0.2 kb <i>PstI</i> fragment in mp18 and 19, respectively	This study
SL3, 4	0.7 kb <i>PstI</i> fragment in mp18 and 19, respectively	This study
SL5, 6	0.5 kb <i>HindIII</i> - <i>PstI</i> fragment in mp18 and 19, respectively	This study
WS18, 19	2.0 kb <i>HindIII</i> - <i>PstI</i> fragment in mp18 and 19, respectively	This study

Figure 3

Genetic and restriction map of the major symbiotic cluster of *B. japonicum* chromosome. The location of a host-inducible *lacZ* fusion is shown. B, *Bam*HI; E, *Eco*RI; H, *H*IndIII; P, *Pst*I; S, *Sal*I. The locations of the insertion and deletions in strains SL67, 65 and 66, respectively, are indicated. The *nolO* hybridization probe is indicated by a solid bar.



hybridization using a 1 kb *Bgl*II fragment (Fig 16) as a probe. pMGS102 and pMGS104 were mated into *B.japonicum* strain USDA110 and also into a set of mutant strains listed in Table 1. Strain SL65 was constructed by replacing 3 *Sac*I fragments from pSL22 (Table 1) by a spectinomycin/streptomycin (sp/sm) resistance cassette from pHP450 (Table 1) resulting in the deletion of 50% of the *nolN* gene and most of the *nolO* gene (Fig 1). The resulting mutated fragment was cloned into the *Pst*I site of pSUP202 to yield pSL35. *E.coli* S17-1 (Table 1) was used to mobilize this plasmid into *B.japonicum* USDA110 and the deletion mutant strain SL65 was selected by adding spectinomycin and streptomycin to the medium (see growth conditions). Tetracycline sensitive strains were chosen since the absence of the suicide plasmid pSUP202 *tet* gene from *B.japonicum* chromosome indicates a double recombination event. Using the same strategy, strains SL67, 66 and 68 were constructed. Strain SL67 has a sp/sm cassette inserted in a unique *Bst*EII site within *nolO*. In strain SL66, 2 *Hinc*II fragments were replaced by the sp/sm marker resulting in the deletion of the last 7 amino acid residues of *nodC*, the entire *nodS* and the first 41 amino acid residues of *nodU*. Strain SL68 has the sp/sm marker inserted in a unique *Bal*I site within ORFA. All mutants constructed in this study were confirmed by Southern blot hybridization (see below).

Growth Conditions

B.japonicum strains were aerobically grown at 30° in RDY or HM medium as previously described (Nieuwkoop et al., 1987; So et al., 1987) with the appropriate antibiotics (Km, 150 µg/ml and 100 µg/ml of Sp, Sm and/or Tc). *E.coli* cells were grown at 37° in LB medium (Maniatis et al., 1982) using the appropriate antibiotics (Ap, 100 µg/ml, Cm, 15 µg/ml, Km, 25 µg/ml, Nal, 20 µg/ml, Sp, 100 µg/ml, Sm, 30 µg/ml and/or Tc, 20 µg/ml). *B.japonicum* strains were anaerobically grown in 12 ml Falcon tubes that were filled to the top with nitrate medium and sealed with parafilm. The latter medium has the following composition :

Yeast extract	0.25 g/l
K ₂ HPO ₄	0.12 g/l
KH ₂ PO ₄	0.12 g/l
MgSO ₄	0.10 g/l
KNO ₃	1.00 g/l
Na Gluconate	5.00 g/l
1000 X RDY Trace elements	1.00 ml/l
Adjust the pH to 7.0	

A *nifD-lacZ* fusion was used as a control for anaerobic conditions. The latter fusion was chosen since *nifD* is anaerobically induced (Kaluza and Hennecke, 1984).

E.coli JM101 cells harboring M13 phage derivatives were grown in TB (Tartof and Hobbs, 1987) broth for DNA sequencing template preparation.

DNA Sequencing and Sequence Analysis

DNA sequencing was done using the dideoxynucleotide chain termination method (Sanger et al., 1977) using M13 phage single-stranded templates (Table 2). A 3948 base pair (bp) fragment (Fig 3) was sequenced in both orientations using the "Sequenase" kit (United States Biochemicals) according to the manufacturer's protocol and using ^{32}P or ^{35}S as the radioactive isotope. The sequence was confirmed using the "Taquence" kit (United States Biochemicals), which utilizes the heat stable *Taq* polymerase. Sequence analysis was performed using Microgenie program (Beckman) and the GCG program (Genetics Computer Group - University of Wisconsin).

M13 phage derivatives SL1, 2, 3, 4, 5 and 6 and WS18 and 19 (Table 2) were constructed by inserting different *B.japonicum* USDA110 restriction fragments (listed in Table 2) into the multiple cloning site (MCS) of M13 mp18 or mp19 and (Table 2) transforming *E.coli* JM101 (Messing et al., 1981). The insertions in the MCS of M13 phages abolished their ability to undergo α -complementation of *E.coli* JM101. Therefore, the phages of interest formed colorless plaques

when allowed to infect *E.coli* JM101 plated on LB medium supplemented with the β -galactosidase enzyme chromogenic substrate 5-bromo-4-chloro-3-indolyl- β -D-galactoside (X-gal). The universal M13 primer was used for the sequencing of phages SL1, 2, 3, 4, 5 and 6 (Table 2), while the sequencing of phages WS18 and 19 (Table 2) was done using 12 additional synthetic primers (see appendix 1).

DNA Southern Hybridization

A Nylon filter carrying *Eco*RI digests of total DNA from various rhizobial strains was kindly provided by Maria Schell (University of Tennessee). The strains used were: *Bradyrhizobium* strains: 61A101C, IRC2133, USDA 6, 46, 76, 94, 110, 122 and 123, AN6A, 32H1 and ANU289; *Rhizobium* sp. NGR234, *R. leguminosarum* 897, *R. fredii* USDA 193, *R. fredii* HH103, *R. trifolii*, *R. meliloti* AK2121, *R. meliloti* 631 and *Azotobacter vinelandii* UV (Table 2). The blot was used in two hybridization experiments using two different probes corresponding to *nolO* and ORFA, respectively (Fig 3, 16). Hybridization was performed using low stringency (i.e., 37°C hybridization temperature and a final wash with 2 X SSC at 37°C) followed by high stringency (i.e., 65°C hybridization temperature and a final wash with 0.1 X SSC 65°C) conditions as previously described (Maniatis et al., 1982). Radioactive labeling of the probe was done using α -

³²P-labeled dATP in nick translation as previously described (Rigby et al., 1977).

All *B.japonicum* mutants constructed in this study in addition to strains SSL74 and 75 (Table 1) were confirmed by Southern hybridization.

RNA Dot Blot Hybridization

For the transcriptional studies of *nolO*, RNA was isolated from uninduced and SSE-induced cells of the following *B.japonicum* strains: USDA110, AN314, SL66 and SL65 (Table 2). Induction was done using SSE as described above and RNA was isolated using the hot phenol extraction method as previously described (Wang et al., 1991; Kullik et al., 1989). Aliquots containing 1-10 µg of RNA were blotted onto nitrocellulose or nylon filters using a 96 well blotting manifold. Blotting was performed using standard procedures (Maniatis et al., 1982). The RNA blot was hybridized to a ³²P-labeled 0.7 Kb *Pst*I fragment internal to *nolO* (Fig 3). Hybridization was allowed to proceed for 48 hours at 42°C in the presence of 10% dextran sulfate and 50% formamide. The blot was subsequently stripped by washing in a 0.005M Tris-HCl, pH8.0, 0.002 Na₂EDTA, 0.1 X Denhardt's solution for 2 hours at 65°C. The blot was subsequently reprobed with a DNA fragment coding for 23S ribosomal RNA from *Pseudomonas aeruginosa* (Festl et al., 1986).

Using the same extraction procedure, RNA was isolated from aerobically grown *B.japonicum* USDA110 cells to be used for the determination of the transcriptional start of ORFA.

β -Galactosidase Assays

Soybean seed extract (SSE) used for inducing *lacZ* gene fusions was prepared as follows: Seeds were shaken overnight at 30°C in 50% ethanol using 1 ml/seed. Extracts were cleared by filtration through a 5 μ m pore size filter. The cleared extracts were concentrated 10-times by rotary evaporation then filter sterilized using 0.25 μ m pore size filters. Induction of *nod* genes was done as described by Banfalvi et al. (1988) using 2 μ M genistein or 10 μ l/ml SSE. β -galactosidase activity was assayed as previously described (Banfalvi et al., 1988), using chlorophenol red- β -D-galactopyranoside (CPRG) as the chromogenic substrate. *B.japonicum* strains carrying *nolO-lacZ* fusions were grown in RDY medium and were induced for *nod* gene expression as described above, while strains carrying ORFA-*lacZ* or *nodZ-lacZ* fusions were grown in either RDY or Nitrate medium.

Primer Extension

The transcriptional start of ORFA was determined using primer extension as described by Kassavetis et al. (1982). The following primer, P19-2 (Appendix 1), was used:

5'-ATGGGAGCTTCGGGGCGTTC-3'. The primer was radiolabeled using γ -³²P ATP and polynucleotide kinase (USB). The radiolabeled primer was co-precipitated with 50 μ g of RNA isolated from aerobically grown *B.japonicum* USDA110 as described above then resuspended in the following hybridization buffer: 20 mM Tris-pH=8.0, 200 mM NaCl, 0.1 mM EDTA. The hybridization mixture was denatured at 100°C for 3 min then incubated at 63°C for 1.5 h. The mixture was adjusted to a final concentration of 6 mM MgCl₂, 6 mM DTT, 0.5 mM dNTP then mixed with 30 units of AMV reverse transcriptase (USB) and incubated for 1 h at 37°C. After the completion of the reaction, RNA was digested using RNase A and the newly synthesized DNA was precipitated and run on a 6% polyacrylamide sequencing gel next to a ladder from a sequencing reaction using the same primer (i.e., P19-2; appendix 1) and phage WS19 (Table 2) as template.

Plant Tests

Plants used for nodulation tests were soybeans (*Glycine max* cv. Essex and *G.soja* PI468397), siratro (*Macroptilium atropurpureum*), cowpeas (*Vigna unguiculata* cv. Caloona) and mungbeans (*Vigna radiata* cv. King). Seeds were surface sterilized, germinated and grown in clear plastic pouches as previously described (Halverson and Stacey, 1985). Seedlings were inoculated with 1×10^4 cells/ml of either *B.japonicum* USDA110, SL65, SL67 or SL68, using 1

ml/seedling. The bacterial concentration of each inoculum used was confirmed by direct colony count on RDY plates. Nodules/plant were counted every day from days 7 to 20 after inoculation.

Purification and Chemical Analysis of Nod Metabolites

Purification and chemical analysis of Nod metabolites produced by strains SL65 (*nolNO⁻*) and SL67 (*nolO⁻*) were performed as previously described (Sanjuan et al., 1992) with minor modifications introduced to the procedure. Modified Bergensen's minimal medium (Bergersen, 1961), used to grow cells for preparation, had the following composition:

K ₂ HPO ₄	0.23 g/l
MgSO ₄ .7H ₂ O	0.1 g/l
Na Glutamate	1.1 g/l
Glycerol	4.0 ml/l
1000 x RDY Trace Elements	1.0 ml/l
pH to 6.8-7.0	

Cells were grown in RDY to an O.D.₆₀₀=0.8-1.0 and diluted into 500 ml of minimal medium where cells were subsequently grown to an O.D.₆₀₀=0.8. This culture was then diluted into 5 liters (5x1 liter in 4 liter flasks) of minimal medium to a final O.D.₆₀₀=0.05-0.08. After each inoculation, samples were plated on both RDY and LB to check for contamination. After inoculation, the 5x1 liter cultures were grown for 4-6

hours before induction. Cells were induced using 10 ml/liter SSE prepared as described above and equivalent to 100 seeds/liter of culture. Cells were shaken at 30°C for 40 hours then extracted with 300 ml/liter of distilled *n*-butanol by shaking at 30°C for 3 hours. The butanol extract was concentrated to dryness by rotary evaporation, then resuspended in 50 ml of 50% acetonitrile. The sample concentration was brought up to 82% acetonitrile then loaded onto a silica gel 60 column (Pharmacia LKB, 1.6x100 cm). The sample was washed with 82% acetonitrile then eluted with 60% acetonitrile. Fractions used for HPLC analysis were chosen according to the approximate elution time of the wild-type Nod signal which has previously been determined (Sanjuan et al., 1992).

Nod metabolites were purified for chemical analysis by reverse-phase HPLC using 2 schemes; First to separate the putative Nod factors from the rest of the fraction components and second to purify each peak individually. In the first procedure (i.e., NodFact 5), a uniform gradient from 20% to 82% acetonitrile is used. The second procedure (i.e., NodFact 6) starts with 20% acetonitrile, then shifts to 46% acetonitrile.

The chemical structures of the purified compounds were determined by Russell Carlson (Complex Carbohydrates

Research Center, University of Georgia, Athens, GA) using GC-MS, FAB-MS and MS-MS studies. Glycosyl composition was performed by GC-MS analysis of alditol acetates prepared as described by York et al. (1985). FAB-MS analysis was carried out using a VG ZAB-SE instrument at an accelerating voltage of 8 KV in the positive mode. Tandem MS-MS analysis was performed using a JEOL HX110/HX110 mass spectrometer operated at 10 KV accelerating potential.

CHAPTER III

RESULTS

A) DISCOVERY AND CHARACTERIZATION OF *noIMNO*

DNA Sequencing of an Isoflavone-Inducible Locus Downstream to *nodJ*

A Tn3-*lacZ* fusion inducible by the isoflavone genistein or SSE was previously mapped to a 0.7 kb *PstI* fragment downstream to *nodJ* (Gerhold, 1989; Fig 3). The inducible nature of this locus suggested its involvement in the nodulation process, and raised the question of whether it is regulated in a manner similar to known nodulation genes (i.e., via NodD-*nod* box interaction). To investigate these two possibilities, a 1948 bp fragment directly downstream to *nodJ* was sequenced in both orientations (Fig 4). The sequence revealed 3 open-reading-frames, *noIMNO*, having the same orientation as the *nodYABCSUIJ* operon with *noIM* starting 32 bp downstream to the stop codon of *nodJ*. The coordinates of *noIMNO* in Figure 4 are 37-207, 260-640 and 664-1737, respectively. The sizes of *noIMNO* are 171, 381 and 1074 bp, respectively (Fig 4).

Figure 4

DNA sequence of *nolMNO*. The putative ribosomal binding sites are underlined. The stop codon of *nodJ* is boxed. An inverted repeat downstream to *nolO* is indicated by 2 arrows in opposite orientation with the only mismatch marked (*). The coordinates of *nolM*, *nolN* and *nolO* are 37-207, 260-640 and 664-1737, respectively.

Figure 4

1 TUGACGTGAGTTCAGCGAAACGGAGAAGTCAAAGCGATGCAGTGTGCATGATGCGTTCCGC 60
 * M Q C H D A F R

61 GGTAGCTCGCGAGTGACCTGGTCTGTCTCAGAGCAGCTTGTGAAACGGCCTTGAGAAGCCGC 120
 G S S R V T W S V R A A C E T A L R S R

121 GATCGAGCCAGCTCTACGGCCGCTGTTCGGCCGATCGGACCTTGCAAAGTCGTGTTGTG 180
 D R A S S T A A C S A D R T L Q S R V V

181 CTGCCTGGCAGCGCACCGGCTGCGCGTGAAGTGTACAGATTAGATCGGTGGGACGGCATG 240
 L P G T R T G C A *

241 GCCTTCCATGGAGAACGGGATGTTGAGCATAGGACTGAGTGGCGGCCTTGACAGAATCTA 300
 M L S I G L S G G L D R I Y

301 TGAAAGCTCTCCGGAGCTGCCGAATACATTTCTTCACGATGGCGCTGCGGTACTTGTGCA 360
 E S S P E L P N T F L H D G A A V L V Q

361 GGACGGCCGCGTAATTGCTGCTGTGCGAAGAGGAGCGCCTCAACCGCGTCAAGCAGTCCAA 420
 D G R V I A A V E E E R L N R V K Q S N

421 CAAGTTCCCCAGCAATTCGATCCGATACTGCCTTTCTACCGCAGGAGTCGAGCTCGGCGA 480
 K F P S N S I R Y C L S T A G V E L G D

481 CATCGACCGTATCGCGTTCTACGCGACTGAAGCTTACTGCAAAGCCATGCTCGAGCGCCT 540
 I D R I A F Y A T E A Y C K A M L E R L

541 TTCTGTTTCTCAGCCCGTTCCGTTGGACCCCAAAGTGTGCTGCGGCAGCTGTTGGCAGG 600
 S V S Q P V P L D P K L L L R Q L L A R

601 GGAATTAGGGGCCGAGATCGATCCGTCGGGTTTCTTCGTGAATCACCACGAAGTACAT 660
 E L G A E I D P S G F P S *

661 GCCGTGAGCGCGTTTGCCATGTCCGGGTTTCGAGCAAAGTCTGGTGTTCGCGATCGACGGC 720
 V S A F A M S G F E Q S L V F A I D G

721 GGGGGCGATTTTCTCTCAGGCTTAATGGCGGTAGGGTCCGGGACCGTAATGCGCAGCTT 780
 G G D F L S G L M A V G S G T V I A Q L

781 GTAAGTTTCCCGGAACACAACCTCTCTGGGCTGTTCTATCTCGAGACGATCCGATATCTC 840
 V S F P E H N S L G L F Y L E T I R Y L

841 GGCTGCGGGATGTTTCGACGAATACAAAGTGATGGGGCTTGCAACCGTATGGTGATCCCGAT 900
 G C G M F D E Y K V M G L A P Y G D P D

901 CGCTATCGCGAGCTCTTCGCGCAGTTCTACGAGTTGTTGGACAGCGGTGGCTATCGCGTT 960
 R Y R E L F A Q F Y E L L D S G G Y R V

961 CACCTGGATCGAATCGGACCGCCCTTTTTCGCAACATCCAGGTTTCGGCGAAAGGGAATG 1020
 H L D R I G P A L L R N I Q V R R K G M

1021 CCATTACCCCAACAGCATCGAGATGTCAGCGCGTTCGCTGCAGGAAGCGCTGGAGCGGATC 1080
 P F T Q Q H R D V S A S L Q E A L E R I

1081 GTGTTTTCACACCCTGCGGCACCATCGTAAAGCGACGGGCATGACACCGCTGTGCCTAGCT 1140
 V F H T L R H H R K A T G M T R L C L A

Figure 4 (Cont.)

1141	GGCGGGGTGGCCCATAATTGCACGCTGAACGGTAAGCTGCTTTATTCCGGGCTATTTGAC	1200
	G G V A H N C T L N G K L L Y S G L F D	
1201	GACATCTTTGTTCAACCCGCGGCGCACGACGCCGGTTGCGCACTCGGCGCTGCATTAATG	1260
	D I F V Q P A A H D A G C A L G A A L M	
1261	GTATCAAGCGAGCTAGGCAGGCCAGCGCCGCGGAGCGACTGCCGGACGTTTATTGGGGG	1320
	V S S E L G R P A P R E R L P D V Y W G	
1321	CCCGATCTTGGGAACGAGCAAGCCACAGAGCTTGAGCTCAACGCATGGTCAGGCCACCTC	1380
	P D L G N E Q A T E L E L N A W S G H L	
1381	GACATTCAACGAAGTGATGACATAGCATCCAGTGCGGCGGACTGGATGGCAAATGGCGCT	1440
	D I Q R S D D I A S S A A D W M A N G A	
1441	GTAATCGGCTGGGTTGAGGGGCGCTCTGAGTTGCGGCCACGTGCGCTTGGTAACCGCAGC	1500
	V I G W V Q G R S E F G P R A L G N R S	
1501	ATTCTAGCGGACCCGCGGCCTGCGGAAAACAAGGATCGCATCAACGCGGTGGTCAAGAAG	1560
	I L A D P R P A E N K D R I N A V V K K	
1561	CGCGAGACGTATCGCCCCTTTGCGCCATCTGCATTAGAGGAGGATGCGAGCGAGTTCTTT	1620
	R E T Y R P F A P S A L E E D A S E F F	
1621	GAGCTCCCCGACGGCACGCGGCGAGTCCCCTTCATGAACCTTCGTAGTTGCGGTGCGTGAA	1680
	E L P D G T R Q L P F M N F V V R V R E	
1681	GCCAAGGGTAACGTGCTCGGTGCGATCACGCACGTTGATGGGACGGCTCGGCTGCAGTGA	1740
	A K G N V L G A I T H V D G T A R L Q *	
1741	TGCTGGCCGCTTCCTGACGAGGAACGGCCCAACAACAAGTCCGTCCAGATCAGTCGTCAA	1800
1801	AAACGTGGTGATCGCATCTGACACGGACTGCACGATCGGCTCGGCATTGTTGTTAAATGA	1860
1861	GGTATTTAGAAGAATCGGAAGGCTCGTTCGCTTCTTGAACGCATTAATGACGTCCCAGTA	1920
1921	GCCGGGGTTTGTCTTGCGGATACGGTCTGCAG	1953

Sequence Analysis of *nolMNO*

Figure 4 shows that each of the three genes, *nolMNO*, is preceded by a putative ribosomal binding site and that *nolO* is followed by a GC rich inverted repeat that could function as a *rho*-independent transcriptional terminator (Fig 4). The DNA regions upstream to *nolM*, *nolN* and *nolO* and within the *nodJ* sequence (Göttfert et al., 1990) did not show sequence similarity to any of the known promoters including the *nod*-box motif conserved in flavonoid inducible promoters. Based on the non-randomness of every third base in the sequence, the "TestCode" computer program (GCG package, Fickett, 1982) predicted *nolNO* to be coding sequences to a 95% level of confidence (Fig 4a). The size of *nolM* is too small for such analysis. This program was chosen as it does not require a table of codon frequencies since the codon usage in *B.japonicum* was shown to vary considerably among different classes of genes (Ramseier and Göttfert, 1991). In Figure 5a, lines above the window of uncertainty indicate coding sequences which were predicted for both *nolN* and *nolO*. As a negative control, the sequence was randomized then subjected to the same test and, as expected, was predicted to be non-coding (Fig 5b). *nolMNO* were predicted to code for proteins of approximately 6 kd, 14 kd and 39 kd, respectively. Using the "MicroGenie" package, computer analysis of the predicted peptides

Figure 5

Analysis of *noIMNO* sequence by TestCode. Lines above the window of uncertainty predict for the corresponding sequence to be coding to a 95% level of confidence, while the lines below the window indicate non-coding sequence to the same level of confidence. **a.** *noIMNO* sequence. **b.** Randomized *noIMNO* sequence.

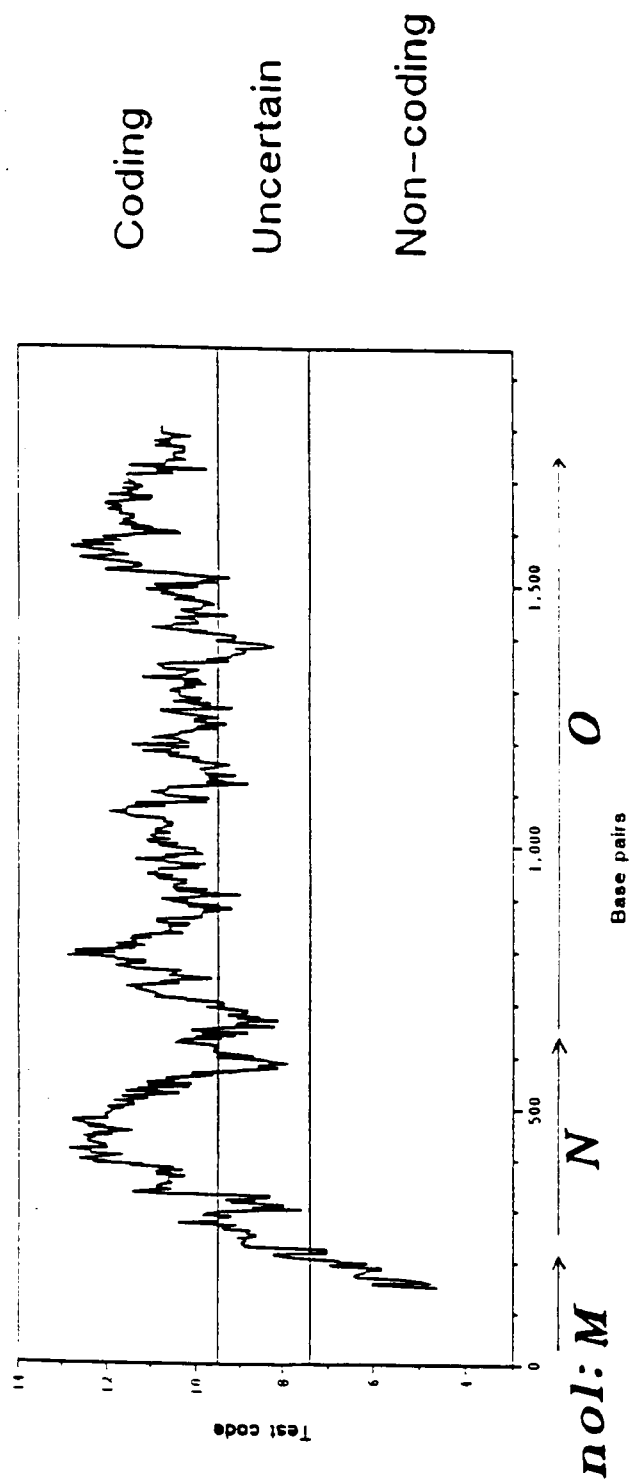


Figure 5a

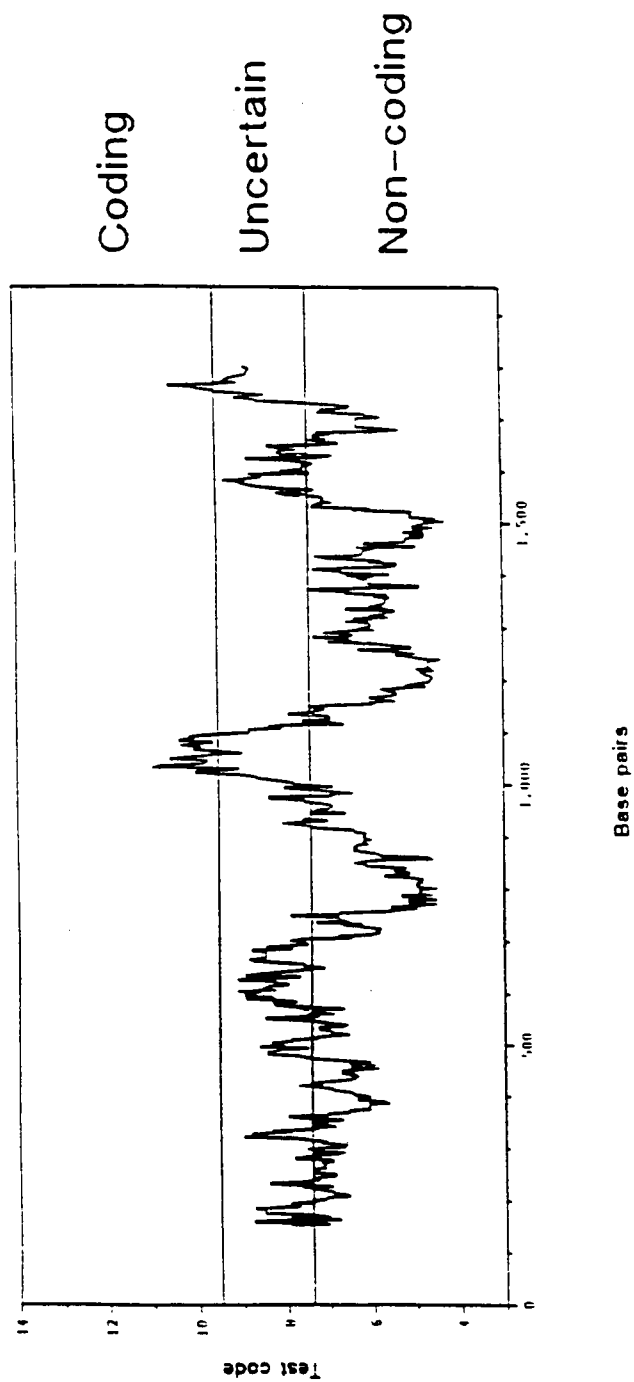
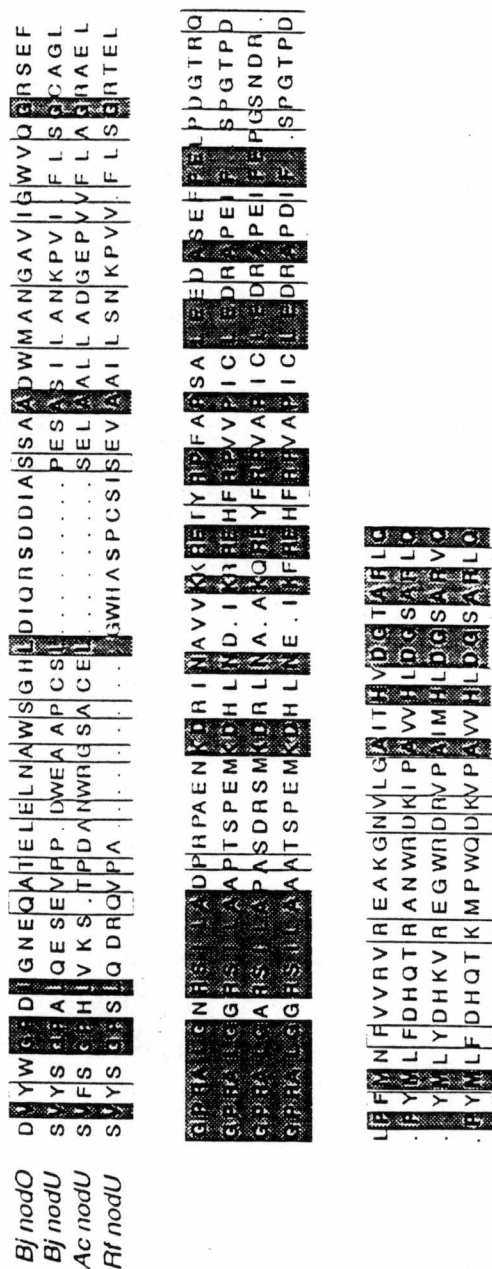


Figure 5b

Multiple sequence allignments of the predicted amino acid sequences of the following genes: *nolO* and *nodU* from *B.japonicum* USDA110, *A.caulinodans* ORS571 and *R.fredii* USDA257. Identical residues are in shaded boxes and similar residues are in open boxes.



corresponding to *nolMNO* did not reveal any membrane spanning domains or DNA binding motifs.

The comparison of the *nolMN* DNA or predicted amino acid sequence to sequences within the Genbank and Swissprot databases did not reveal any significant similarities. The predicted amino acid sequence of *nolO*, however, showed significant similarity (i.e., 60.1% similarity and 35.5% identity within 401 bp) to that of the *nodU* gene of *B.japonicum* USDA110 which is located about 4 kb upstream of *nolO* (Fig 3). In order to investigate the significance of the similarity between the *nolO* and *nodU* predicted gene products, the sequence of *nolO* was randomized then compared again to that of *nodU*. A marked decline in the percentage of identity and similarity (i.e., 44.5% similarity and 17% identity within 401 bp) between the two sequences was observed. Similar regions of sequence similarity were found between the *nolO* amino acid sequence and those of *nodU* genes of *R.fredii* USDA257 and *A.caulinodans* ORS571. Multiple sequence alignment of all 4 genes revealed an 144 amino acid stretch that is particularly conserved (Fig 6). On the DNA level, however, no corresponding similarity was found between *nolO* and *nodU* sequences consistent with the fact that a *nolO* probe did not hybridize to *nodU* from *B.japonicum* USDA110, either in Southern or northern hybridization studies (e.g., Fig 8a).

Interstrain and Interspecies Hybridization Using a *nolO* Probe

To investigate the occurrence of *nolO* in rhizobial strains other than *B. japonicum* USDA110 and in order to determine if it exists in more than one copy on the *B. japonicum* USDA110 chromosome, DNA from 12 *Bradyrhizobium* strains, 7 *Rhizobium* strains and the free-living diazotroph *Azotobacter* UW (Table 2) was digested with *EcoRI*, run onto a 0.7% agarose gel, then blotted onto a nylon filter (kindly provided by M. Schell, University of Tennessee). The filter was subsequently hybridized to a ³²P-labeled 0.7 kb *PstI* fragment internal to *nolO* (Fig 3) under conditions of low and high stringency, respectively. Figure 7 shows that *nolO* is conserved among *Bradyrhizobium* strains and that no detectable hybridization signal was observed for any of the *Rhizobium* strains (Fig 7). After a longer exposure time, *Bradyrhizobium* strains 61A101C, IRC2133A and ANU289 showed positive hybridization signals each to a band of an approximate size of 4 kb. It is noteworthy that the *Bradyrhizobium* strains used, all of which hybridized to *nolO*, include members of both Groups I and II of *B. japonicum* which differ in many biological aspects including their ability to nodulate restrictive *Rj4* cultivars of soybeans (Hollis et al., 1981; Vest and Caldwell, 1972). For each of the strains that showed homology to *nolO*, only one hybridizing band was observed.

Figure 7

Autoradiogram of a nylon filter carrying total bacterial DNA cut with *EcoRI* (kindly provided by Maria Schell) and probed with ³²P-labeled 0.7 kb fragment internal to *nolO*.

Bradyrhizobium strains: 1: 61A101C, 2: IRC2133A, 3, 4, 5, 6, 7, 9 and 10: USDA46, 123, 122, 94, 76, 6 and 110, respectively, 8: AN6A, 11: 32H1, 13: ANU289. Other strains: 12: *Azotobacter* UW, 14: *Rhizobium* sp. NGR234, 15: *R.l.* bv. *viciae* 897, 16: *R.l.* bv. *trifolii*, 17: *R.meliloti* AK2121, 18: *R.meliloti* 631, *R.fredii* USDA193, 20: *R.fredii* HH103.

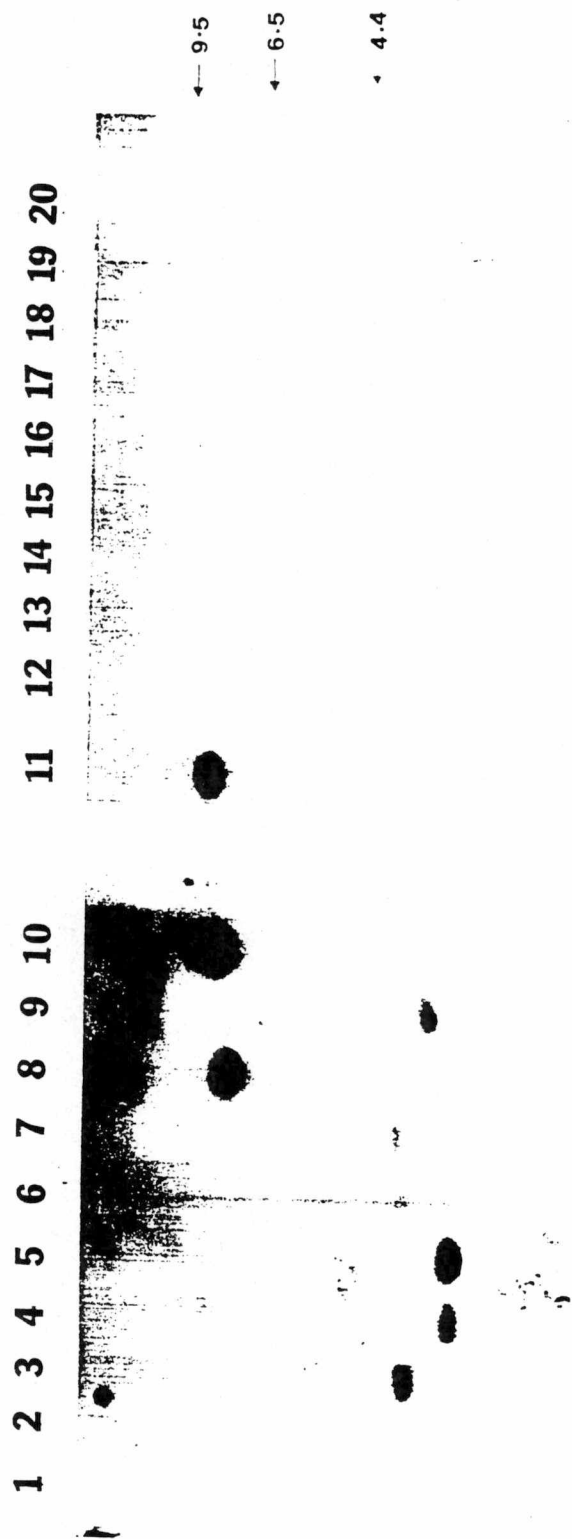


Figure 7

Transcriptional Control of *noLMNO*

The flavonoid inducibility of the *nolO-lacZ* fusion suggested that it is transcriptionally regulated in a manner similar to that described for previously identified nodulation genes (Long, 1989; Schlaman et al., 1992). This mechanism of regulation requires the transcriptional activator NodD. The nearest putative NodD binding site (i.e., *nod*-box) is upstream of the *nodY* gene suggesting that *nolO* may be part of the *nodYABCSUIJ* operon. Such an assumption implies that *nolO* transcription should be *nodD*₁-dependent and that a polar genetic insertion anywhere within the *nodYABCSUIJ* operon should block *nolO* expression. Indeed, RNA extracted from SSE-induced cultures of strains AN314 (*nodD*₁⁻) and SL66 (Δ *nodS*, *sp/sm* Ω) showed no detectable signal when hybridized to a *nolO*-specific probe (Fig 8a). Strain SL65 contains a *noINO* deletion and was included as a negative control. When the blot was stripped and rehybridized to a 23S rRNA probe (Table 2), RNA from all strains used gave comparable hybridization signals (Fig 8b).

Another approach used to investigate the *nod*-box dependence of *nolO* expression utilized subclones derived from pDG69 (*nolO-lacZ*) (Table 2). A fragment carrying the *nolO-lacZ* fusion and 2-3 kb of flanking DNA (Fig 3) was subcloned from pDG69 into pVK102 to create plasmid pSSL74 lacking the *nodY*-proximal *nod*-box (Table 2). β -

Figure 8

RNA dot blot probed with ^{32}P -labeled probes: **a)** *nolO*, **b)** 23S rRNA. USDA110, WT; AN314, *nodD₁⁻*; SL66, *nodS⁻*; SL65, *nolNO⁻*. **U**: RNA from uninduced cultures; **I**: RNA from cultures induced with soybean seed extract.

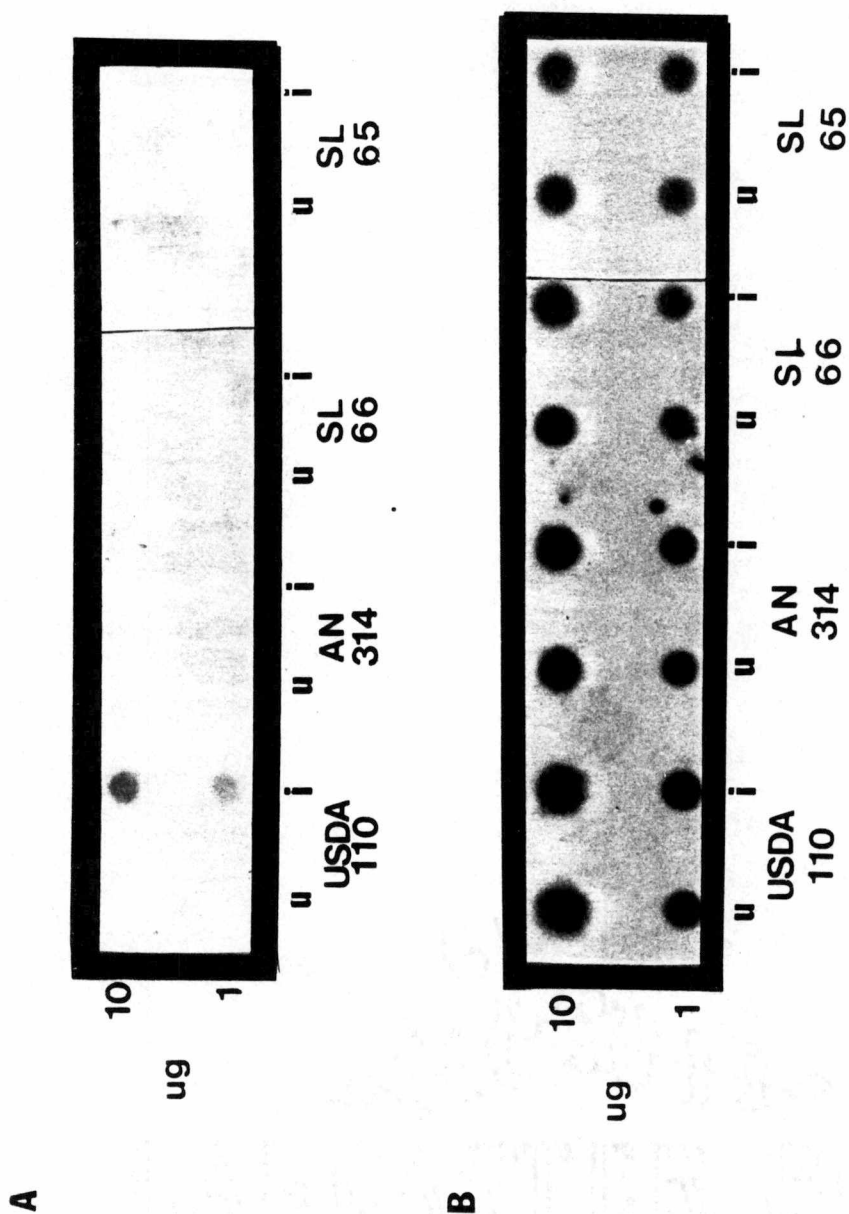


Table 3: β -Galactosidase Activity of *nolO-lacZ* Fusions

β -Galactosidase Units			
STRAIN	GENOTYPE	UNINDUCED	2 μ M GENISTEIN
USDA110	Wild-Type	<1	<1
SSL71	USDA110 (pDG69)	5 $\pm$ 1	40 $\pm$ 3
SSL74	USDA110 (pSSL74)	39 $\pm$ 3	41 $\pm$ 4
SSL75	USDA110 (pSSL75)	34 $\pm$ 1	37 $\pm$ 2

Each entry is the average of 4 different experiments.

galactosidase activity expressed from this plasmid in *B.japonicum* USDA110 showed that it no longer exhibited inducible expression (Table 3). The uninduced levels, however, were higher than those of the parent plasmid pDG69 (Table 3). To investigate whether the high uninduced levels of pSSL74 were due to a promoter from within the cloning vector pVK102, the same fragment cloned in pSSL74 was recloned in the same restriction site of pVK102 but in the opposite orientation to yield plasmid pSSL75. Table 3 shows that both plasmids exhibited the same levels of β -galactosidase activity (Table 3).

Nodulation Phenotype of *nolO* and *nolNO* Mutants

Mutant strains SL67 (*nolO*⁻) and SL65 (*nolNO*⁻) were constructed using a sp/sm⁰ resistance cassette (Fig 3, Table 2). The ability of the mutants to nodulate a number of host plants [i.e., soybean (*G.max* and *G.soja*), cowpea, mungbean and siratro] was compared to that of the wild-type strain *B.japonicum* USDA110 using a low inoculum concentration of 10⁴ cells/ml/seedling. Both strains SL65 and SL67 showed the same nodulation phenotype and therefore, only data for SL65 is shown. The mutant strains showed a 3 day delay in nodulation (Fig 9) and a reduction in the maximum number of nodules per plant 20 days after inoculation (the difference in nodule number between strains USDA110 and SL65 is significant to a 90% level of confidence; Fig 10). This

Figure 9

Percentage of nodulated plants after inoculation with *B. japonicum* strains USDA110 (open squares) or SL65 (solid squares). Each point represents the average of two experiments each using 20 plants/strain.

Figure 9

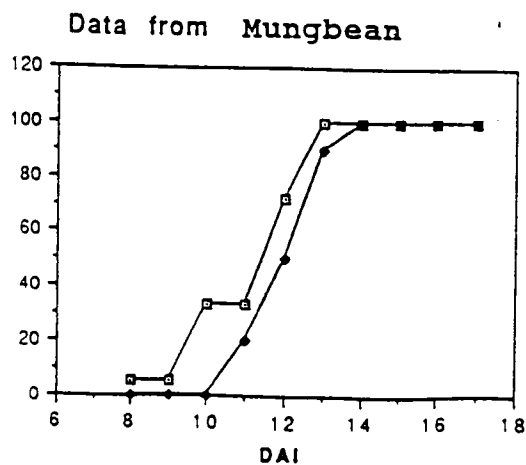
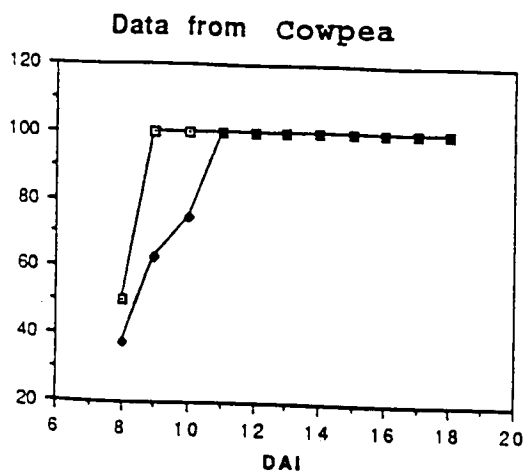
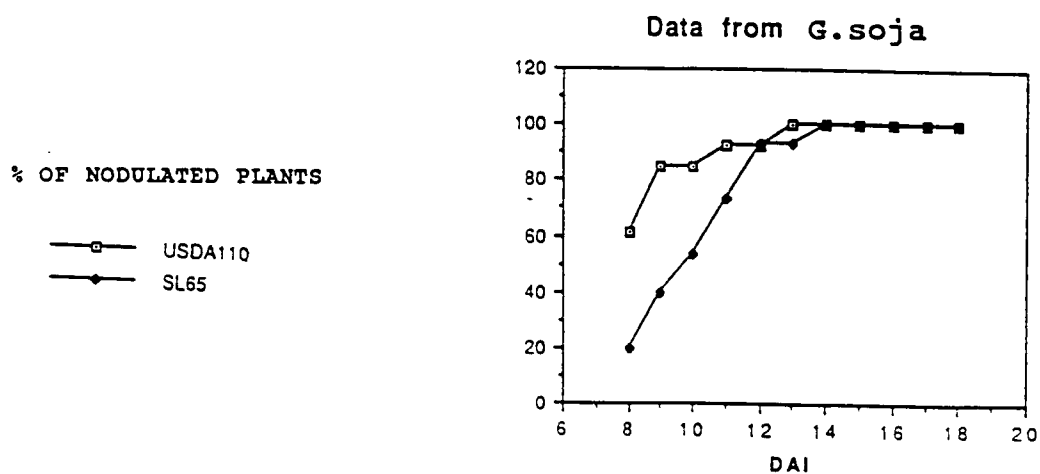
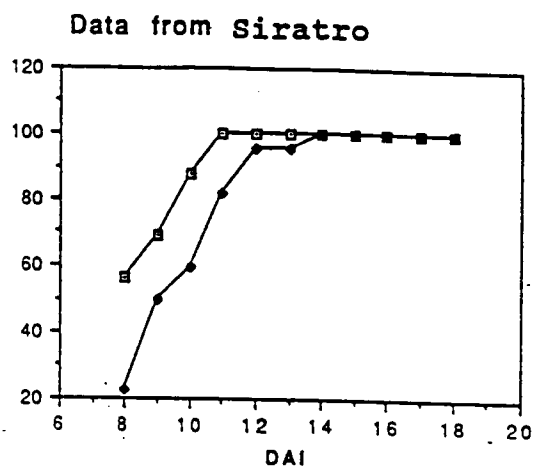
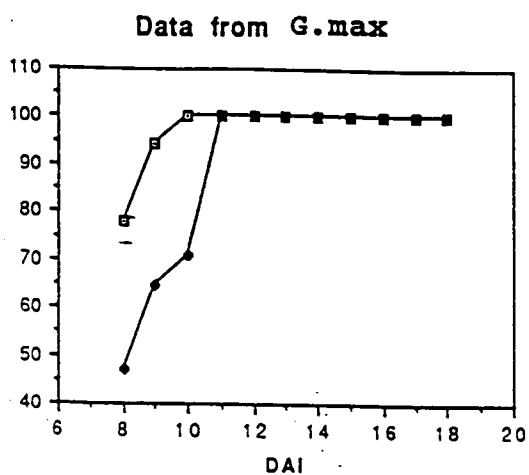


Figure 10

Number of nodules/plant 20 days after inoculation with *B. japonicum* strains USDA110 or SL65. 20 plants were used for each strain. The significance of the differences between wild-type and mutant strain was tested using the t test; $\alpha = 0.1$. For all plant species, the differences in nodule number between strains USDA110 and SL65 were significant to a 90% level of confidence.

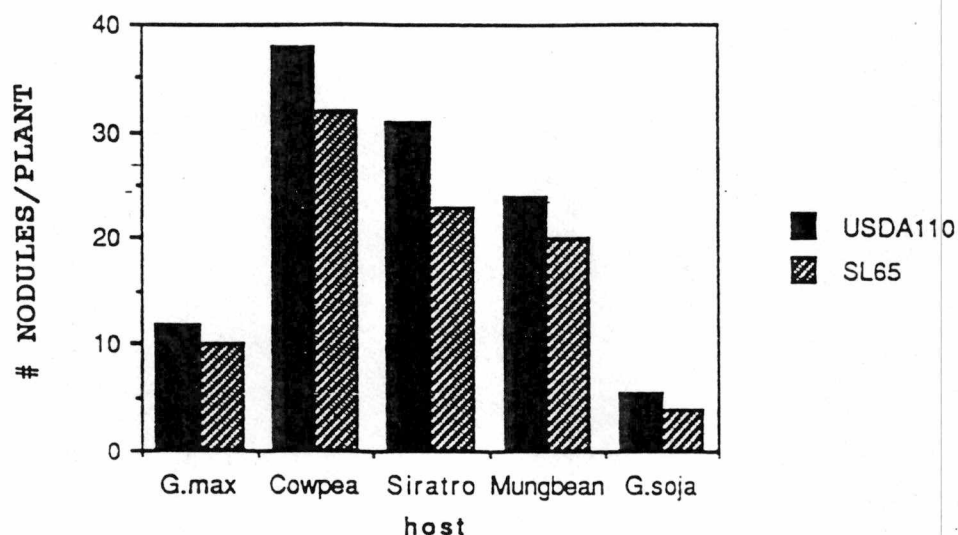
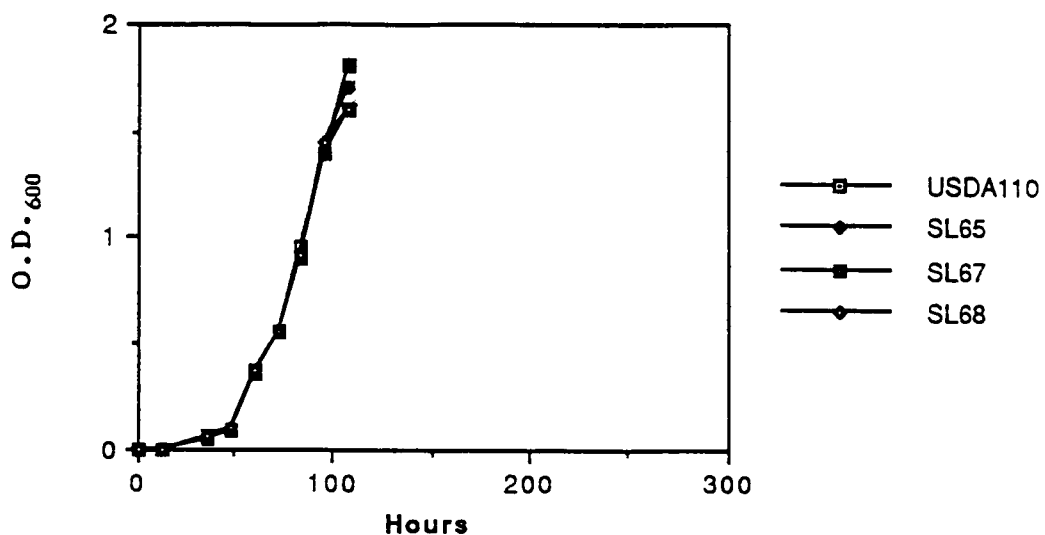


Figure 11

Growth curve of *B. japonicum* strains USDA110, SL65, SL67 and SL68. Growth was estimated by measuring O.D.₆₀₀ of RDY broth cultures.



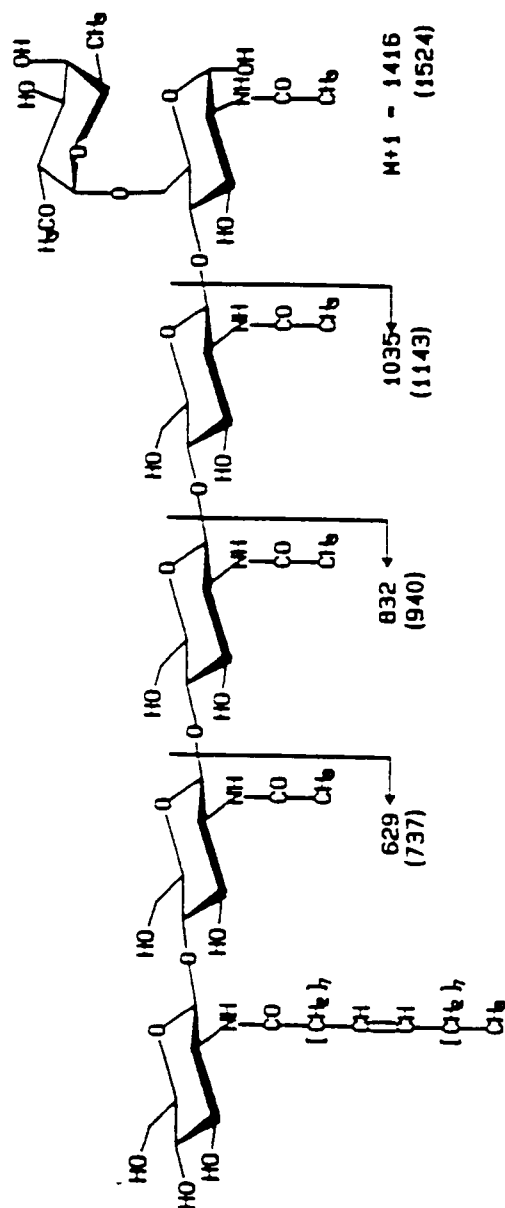
reduction in the nodulation efficiency was not observed when an inoculum of a higher concentration (i.e., 10^6 cells/ml/seedling) was used. No detectable differences among strains SL67, SL65 and USDA110 were found with respect to colony morphology or growth rate (Fig 11).

Lipo-oligosaccharide Nodulation Signals Produced by *nolO* and *nolNO* Mutants

The nodulation ability of *B.japonicum* was previously shown to be dependent on the production of a lipo-oligosaccharide signal molecule (Sanjuan et al., 1992; Carlson et al., 1993). In *B.japonicum* USDA110, this molecule is a pentasaccharide of β 1-4 linked, N-acetylglucosamine with a $C_{18:1}$ fatty acyl group attached to the terminal non-reducing sugar and a 2-O-methylfucose substituent on the terminal reducing sugar (Fig 12; Sanjuan et al., 1992). Other *B.japonicum* strains produced additional derivatives of the basic USDA110 structure (Carlson et al., 1993). The ability of *B.japonicum* USDA110 to produce such a nodulation factor was dependent on the common nodulation genes *nodABC* (Sanjuan et al., 1992). The fact that *nolMNO* are co-transcribed with *nodABC* (Fig 8) in addition to the reduced nodulation ability exhibited by *nolO* and *nolNO* mutants (Fig 9, 10) suggested the involvement of these genes in the production of the *B.japonicum* USDA110 lipo-oligosaccharide signal molecule.

Figure 12

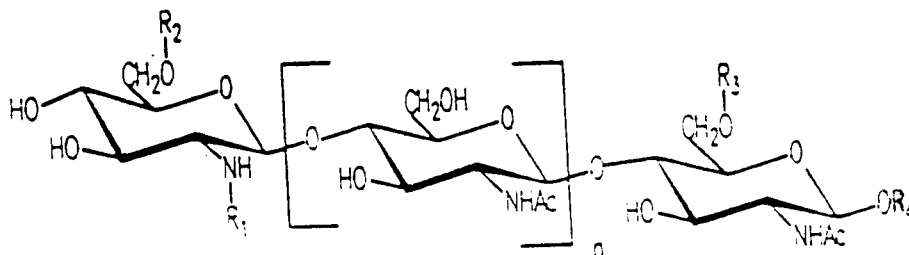
Structure of the Nod signal produced by the wild-type *B. japonicum* strain USDA110. (From Sanjuan et al., 1992).



The nodulation signals produced by the mutant strains SL67 (*nolO*⁻) and SL65 (*nolNO*⁻) were purified from butanol extracts of SSE-induced cultures by silica gel chromatography followed by HPLC purification as described in materials and methods. The chemical structure of all purified compounds was determined by Russell Carlson (The University of Georgia) using FAB-MS, MS-MS and NMR analyses (see materials and methods). Both mutants produced the wild-type nodulation factor (Fig 13, Peak 1). However, in addition, several other molecules were identified that were not produced by the parent strain USDA110 (Fig 13). The ratio between the areas of the 5 peaks calculated from the HPLC chromatograms was 1:0.26:0.16:0.16:0.05 for peaks 1-5, respectively. FAB-MS analysis of peak 1 (spectrum not shown) revealed an (M+H)⁺*m/z*=1416 with fragment ions of *m/z*=1035, 832, 629 and 426. This fragmentation pattern is identical to that previously published for the *B.japonicum* USDA110 Nod factor (i.e., NodBj-V(C_{18:1},MeFuc; Sanjuan et al., 1992). The major compound that was produced only in the mutant strains corresponded to peak 2 (Fig 13). FAB-MS analysis [(M+H)⁺*m/z*=1256 with fragment ions of *m/z*=1035, 832, 629 and 426] and glycosyl composition analysis (data not shown) showed that peak 2 differs from the wild-type nodulation factor in its lack of the 2-O-methylfucose residue [i.e., NodBj-V(C_{18:1}); Fig 13, Peak 2]. Other compounds produced by the mutants are shown in figure 13

Figure 13

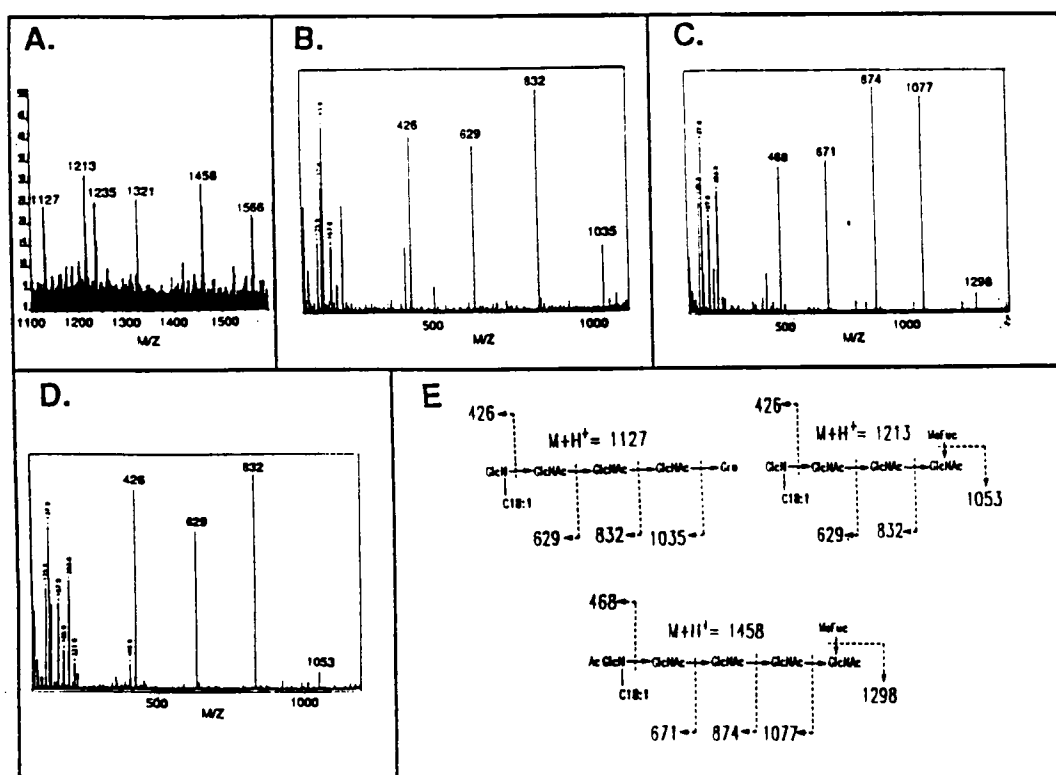
Nod metabolites produced by strain SL65 (*noI*^{NO}). **Ac**, Acetyl; **2-O-MeFuc**, 2-O-Methylfucose; **Gro**, Glycerol. Chemical analysis of Nod factors was done by Russell Carlson, Univ. of Georgia. (Figure adapted from Stacey et al., 1993b).



<u>Peak</u>	<u>R1</u>	<u>R2</u>	<u>R3</u>	<u>R4</u>	<u>n</u>
1	C _{18:1}	H	2-O-Me-Fuc	H	3
2	C _{18:1}	H	H	H	3
3 (a)	C _{18:1}	Ac	2-O-Me-Fuc	H	3
(b)	C _{18:1}	H	2-O-Me-Fuc	H	2
(c)	C _{18:1}	H	H	Gro	2
4	C _{16:0}	H	2-O-Me-Fuc	H	3
5	C _{16:0}	H	H	H	3

Figure 14

FAB-MS and MS-MS spectra of SL65 (Peak 3). Panel A shows a FAB-MS spectrum of SL65 peak 3. Panels B-D show tandem MS-MS analyses of the $m/z = 1127$, 1458 and 1213, respectively. Panel E shows the structures and their expected fragmentation patterns for the above molecules. (Chemical analysis of Nod factors was done by Russell Carlson, Univ. of Georgia).



(i.e., Peaks 3, 4 and 5). FAB-MS analysis of peak 3 (Fig 14A) revealed that it consists of 3 different components of $(M+H)^+m/z=1127$, 1213 and 1458. MS-MS spectra of the $m/z=1127$, 1458 and 1213 ions are shown in figures 14B, 14C and 14D, respectively. MS-MS analysis of the $m/z=1127$, 1213 and 1485 revealed structures corresponding to NodBj-IV($C_{18:1}$,Gro), NodBj-IV($C_{18:1}$,MeFuc) and NodBj-V(Ac, $C_{18:1}$,MeFuc), respectively (Fig 14B,C,D,E). These three molecules are not produced by the parent strain USDA110 (Sanjuan et al., 1992). The presence of a glycerol residue on the reducing end of the lipo-oligosaccharide molecule has previously been reported in *B.japonicum* strain USDA61, a group II strain which is taxonomically different from strain USDA110, and was suggested to be an intermediate in the synthesis of nodulation signals (Carlson et al., 1993). FAB-MS analysis of peak 4 revealed a fragmentation pattern [$(M+H)^+m/z=1390$ with fragment ions of $m/z=1009$, 806, 603 and 400] identical to that previously published for one of the Nod metabolites produced by *B.japonicum* strain USDA135, namely, NodBj-V($C_{16:0}$,MeFuc) (Carlson et al., 1993). This molecule has previously been reported in the *B.japonicum* group I strain USDA135 (Carlson et al., 1993) and is likely to be present in the parent strain USDA110 in trace amounts. FAB-MS analysis of peak 5 indicated an $(M+H)^+m/z=1230$ which is 160 a.m.u. less than that corresponding to peak 4. This

difference corresponds to the loss of a 2-*O*-methylfucosyl residue and indicates that peak 5 is NodBj-V(C_{16:0}).

In summary, strains SL65(*noI*NO⁻) and SL67(*noI*O⁻) produced the previously reported wild-type *B. japonicum* USDA110 Nod factor (i.e., NodBj-V(C_{18:1},MeFuc), but in addition, they produced the following molecules (Fig 13): NodBj-V(C_{18:1}), NodBj-V(C_{16:0},MeFuc), NodBj-V(C_{16:0}), NodBj-V(Ac,C_{18:1},MeFuc), NodBj-IV(C_{18:1},Gro) and NodBj-IV(C_{18:1},MeFuc).

B) INVESTIGATION OF TWO GENETIC LOCI LINKED TO *noLMNO*

Expression and DNA Sequence of ORFA

An active Tn3-*lacZ* fusion was mapped downstream to *noI*O that has an orientation opposite to that of *noI*O (Fig 15, Table 4). The location of this locus within the major symbiotic cluster on the *B. japonicum* chromosome suggested its involvement in the symbiosis. Sequencing the DNA region where the fusion was mapped revealed a 350 bp open-reading-frame that was given the name ORFA (Fig 16, bases 140-490). The most probable start codon is a GTG which is preceded by a possible ribosomal binding site (Fig 16). The ORFA start codon is located 13 bp upstream to that of the divergently transcribed *nodZ* (see below) while the stop codon of ORFA is located 348 bp downstream to that of *noI*O (Fig 16).

"TestCode" program (GCG computer package, see materials and

Figure 15

Genetic and restriction map showing ORFA and *nodZ*. The location of an active *lacZ* fusion is indicated by an arrow. B, *Bgl*II; E, *Eco*RI; H, *Hind*III; P, *Pst*I.

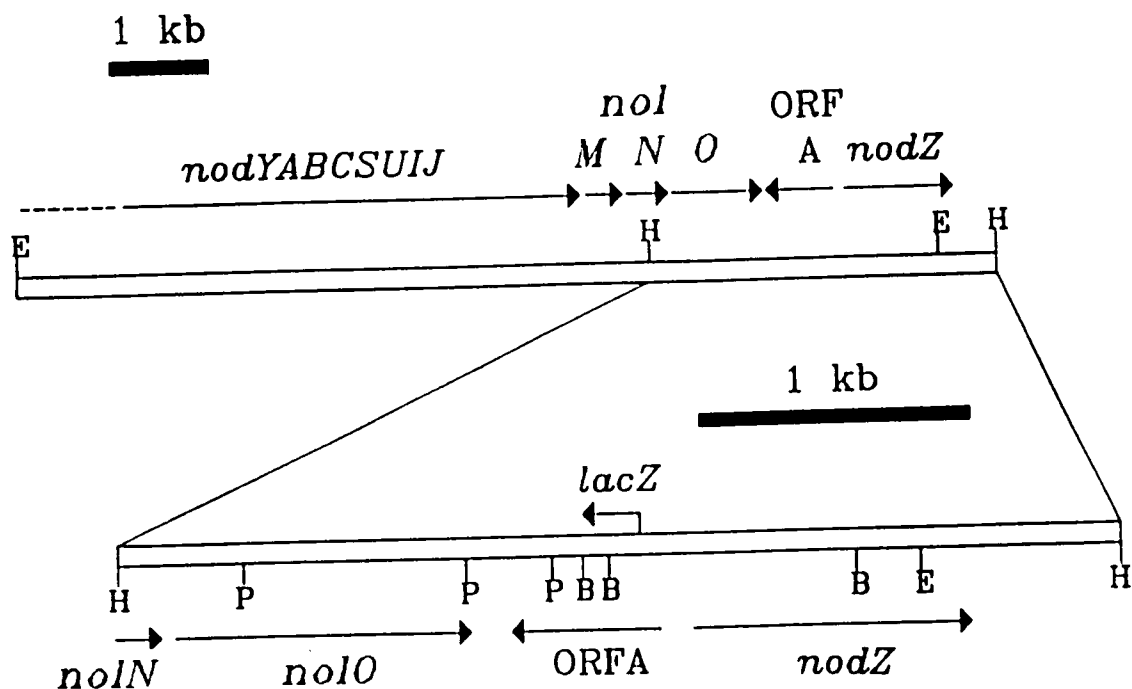


Figure 16

DNA sequence of ORFA. The putative ribosomal binding site is underlined. The translational start of *nodZ* is 13 bases 3' to that of ORFA and is indicated by an arrow. The end of *nodZ* is indicated by a blocked arrow. The transcriptional start sites of *nodZ* and ORFA are indicated by arrow heads.

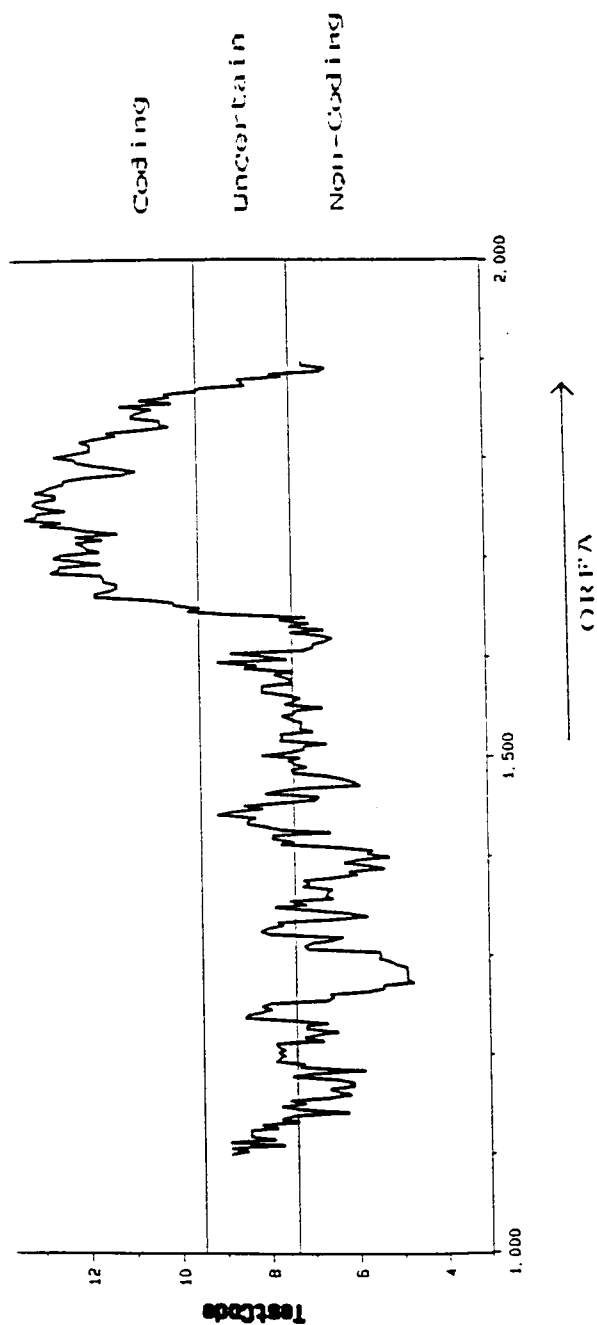
```

1  CACAGACAGT.CGCCAAACCCCGTACGTCGCCGGGACACTACGAACCGATCATTGCTCGAA 60
61  CCGCTAACCATCTGCCTTGCCTTTTCGCGCGCTCCTGGTTGGACAAGAGACGTCAGTACA 120
121 GACGTCTCGCGTATCTTCTCGCCCGGCGTCACGGCCTGAAATCGCGGTGCCGGCGAGCTG 180
181 CATCGGTAGAACTTCATAGCAGGGCTATCCGTGACGGGCCGCGTTCTCTCAATACGATCG 240
      nodZ
      V T G R V L S I R S
nodZ
241 TCCTTTCTTCAACGGTGTTCATCGACCTGCCCGAAGCCCGATGCCCGATGCCCGCTCT 300
      S F L Q R C C I D L P E A R C P M P R S
301 GCATCGGATGGCCTGGTCCCTCACATGCACGCATGGGCGTGGCCATGCGCCATGAGTCA 360
      A S D G L V P S H A R M G V A M R H E S
361 TGGCTATGTTCCGCTTTGTACATTATCGTGACGGCAAGCCGCGTGCAGCGGATGCAGGGG 420
      W L C S A L Y I I V T A S R V Q R M Q A
421 AACGCCCCGAAGCTCCCATACGAGCCGCAATTCCTTCACAAGATCTTCACGCTTGGCCTG 480
      N A P K L P Y E P Q F L H K I F T L G L
481 GTCCAGCAGGGCTGTGTGCGAAAAGCGACCCGAGCGACGCCCTCGCCCTCGATTTCGCATCAA 540
      V Q Q G C V E K R P E R R L A L D S H Q
541 GATCTCGAAAAGCTCAGGTGAGATACGCATACTAGAGTGGGCGGAATGTCCCATGCGGAT 600
      D L E K L R *
601 CTCGCAGACCGTTTCCTGCCGATCTGGCGCTATGTGGGAGCGAACACCGATGAAGGGATGC 660
661 GTACGGCGGCAATGACGCAGCCAGCGCACTCCAATCCTGCAGACCGTATCGCGCAAGACA 720
721 AACCCCGGCTACTGGGACGTCATTAATGCGTTCAAGAAGCGAACGAGCCTTCCGATTCTT 780
781 CTAAATACCTCATTTAACAACAATGCCGAGCCGATCGTGCAGTCCGTGTCAGATGCGATC 840
841 ACCACGTTTTTTGACGACTGATCTGGACGGACTTGTGTTGGGCCGTTCCCTCGTCAGGAAG 900
901 CGGCCAGCATCACTGCAGCCGAGCCGTCCTCATCAACGTGCGTGATCGCACC 951
      ←
      nodZ

```

Figure 17

Analysis of ORFA sequence by TestCode. Lines above the window of uncertainty predict for the corresponding sequence to be coding to a 95% level of confidence.



methods) strongly suggested that ORFA represents a coding sequence (Fig 17). ORFA was predicted to code for a protein of approximately 12 kd. The predicted ORFA protein was analysed using the "MicroGenie" package (see Materials and Methods). This analysis did not suggest any membrane spanning domains or DNA binding motifs.

Comparing ORFA DNA or amino acid sequence to sequences in the Genbank and Swissprot databases did not reveal any significant similarities. Using a ^{32}P -labeled 0.1 kb *BglIII* fragment internal to ORFA (Fig 15) as a hybridization probe, bands that gave a positive signal were found in some *Bradyrhizobium* strains but not in any of the *Rhizobium* strains tested (Table 5). Among the *B.japonicum* strains that gave a positive hybridization signal were members of both groups I and II which are taxonomically different. For each of the strains that gave a positive signal, only one hybridizing band was observed indicating that ORFA is likely to be present in a single copy in these strains. Interestingly, in all strains, the restriction fragments that hybridized to ORFA had the same sizes as those that hybridized to the *nolO* probe (Table 5) suggesting a possible linkage between the two genes in other rhizobia.

Table 4: β -Galactosidase Activity of ORFA-lacZ Fusions

STRAIN	GENOTYPE	β -Galactosidase Units	
		AEROBIC	ANAEROBIC
BSL102	USDA110 (pMGS102)	123	120
BSL119	Bj2101-NifA ⁻ (pMGS102)	88	110
BSL121	Bj27147-NtrC ⁻ (pMGS102)	137 [*]	nd
BSL120	Bj3263-NtrBC ⁻ (pMGS102)	151 [*]	nd
SL105	Bj-N50-RpoN ₁ ⁻ (pMGS102)	82	102
SL106	Bj-N63-RpoN ₂ ⁻ (pMGS102)	87	93
SL107	BjN50-97-RpoN ₁ N ₂ ⁻ (pMGS102)	73	na
GS300	USDA110 (<i>nifD-lacZ</i>)	148	3723 ± 154

- Strains USDA110, Bj2101, Bj27147, Bj3263, Bj-N50, Bj-N63 and BjN50-97 had less than 1 unit of activity.

- Cultures were grown in nitrate medium (materials and methods).

- Each entry is the average of, at least, 2 experiments.

- Standard deviation for all strains except GS300 (anaerobic) was less than 10 units.

- **nd**: not determined, **na**: not applicable.

* Grown in RDY.

Transcriptional Control of ORFA

The majority of the genes known to be involved in the *Rhizobium*-legume symbiosis are under tight transcriptional control (Long, 1992; Merrick, 1992). For example, *nod/nol* genes are inducible by host-produced flavonoids while *nif/fix* genes are induced by anaerobic conditions such as those prevailing inside root nodules. Among the proteins mediating the transcription of the latter genes are NodD, NifA, NtrC/NtrB and the alternative sigma factor RpoN. Although ORFA was not found to be inducible by the isoflavone genistein or by anaerobic conditions, the requirement for the latter proteins for ORFA expression was investigated. The ORFA-*lacZ* fusion plasmid pMGS102 (Table 2) was mated into *B. japonicum* strains lacking NodD, NifA, NtrC, NtrB, NtrB/C, RpoN₁, RpoN₂ and RpoN₁/N₂, respectively, and the β -galactosidase activity of each strain was assayed either under aerobic or anaerobic conditions. Table 4 shows that none of the proteins studied was required for ORFA expression (Table 4). This is consistent with the fact that none of the *cis*-acting elements associated with these proteins was found within the DNA region upstream of ORFA (Fig 16).

Determining the Transcriptional Start of ORFA

Analysis of the DNA region upstream to the most probable translational start codon of ORFA did not reveal

Table 5: Interstrain Hybridization to *ORFA*, *nodZ* or *nolO* Probes

Strain	Probe		
	<i>ORFA</i>	<i>nodZ</i>	<i>nolO</i>
USDA 110	9.5	9.5	9.5
USDA 6	4	4	4
USDA 46	4.5	4.5	4.5
USDA 76	-	4.5	4.5
USDA 94	-	4.5	4.5
USDA 122	4	4	4
USDA 123	4	4	4
61A101C	-	4	4
IRC:2133A	-	4	4
AN6A	9.5	9.5	9.5
32 H1	9.5	9.5	9.5
ANU 289	-	4	4

- Hybridization was performed at low stringency followed by high stringency. Both conditions gave similar results.
- Entries represent in kb the sizes of *EcoRI* fragments which hybridized to each probe.
- *nodZ* hybridization data was kindly provided by Maria Schell.

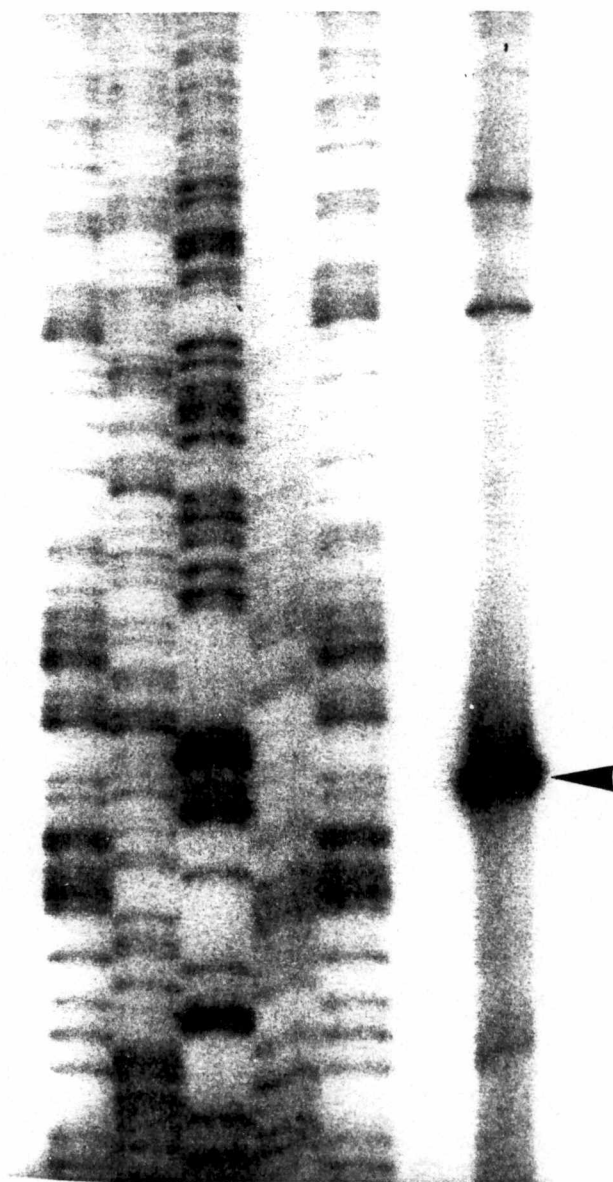
any of the sequences conserved in different prokaryotic promoters (Fig 16). Furthermore, all the *B.japonicum* DNA binding proteins tested for their involvement in ORFA expression were found not to be required for the activity of the ORFA-*lacZ* fusion plasmid pMGS102. Therefore, it was not possible to predict the size of ORFA transcript from sequence data. Determining the transcriptional start of ORFA was especially important because of the possible overlap between the ORFA transcript and that of the divergently transcribed *nodZ* (Fig 15; Stacey et al., 1993). Using a synthetic primer complementary to 20 nucleotides within the coding region of ORFA transcript (P19-2, Appendix 1), the transcriptional start of ORFA was determined by primer extension (as described in materials and methods) and found to be located 95 bp upstream to the putative GTG translational start of ORFA (Fig 16, 18). The transcriptional start of *nodZ* has previously been determined (Stacey et al., 1993; Fig 16) and would accordingly indicate an 128 nucleotide overlap between the divergently transcribed genes ORFA and *nodZ*; thus, raising the possibility of post-transcriptional down-regulation of both genes by anti-sense mRNA.

Nodulation Phenotype of ORFA Mutant Strain SL68

The location of ORFA within the major symbiotic cluster on the *B.japonicum* USDA110 chromosome (Fig 15) and

Figure 18

Primer extension analysis of the ORFA transcript. P19-2 (appendix 1) was used as a primer to USDA110 RNA and the M13 phage WS19 as template for a DNA sequencing ladder.



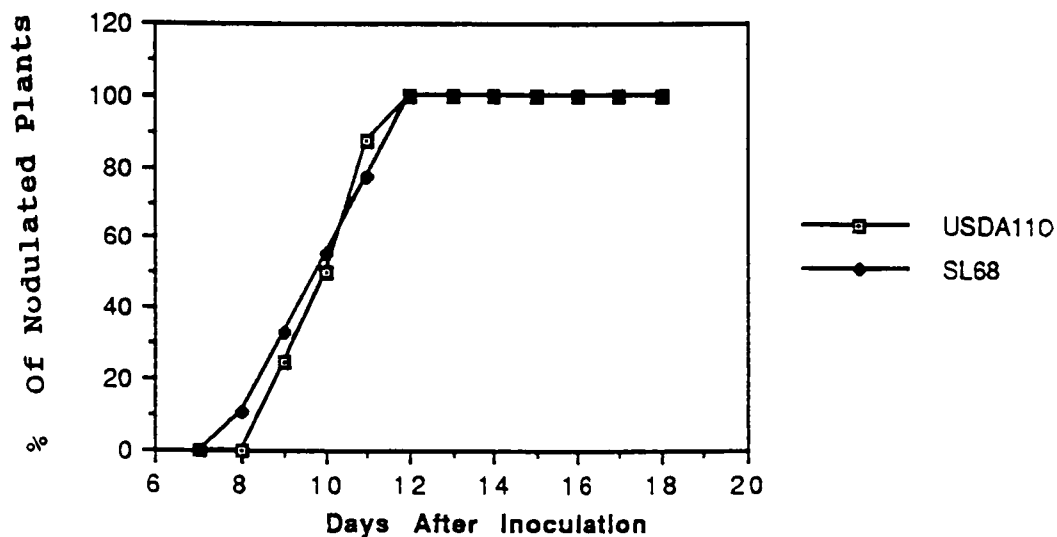
its linkage to *nolO* and *nodZ* in other *Bradyrhizobium* strains (Table 5) suggested its involvement in the nodulation process. To test this hypothesis, a *sp/sm* resistance cassette (Table 2) was inserted in a *BalI* site internal to ORFA resulting in strain SL68 (ORFA⁻). Nodulation assays were done in plastic pouches using a low concentration inoculum (i.e., 10⁴ cells/ml) as described above and the hosts tested were soybean (*G.max* and *G.soja*), siratro, mungbean and cowpea. No detectable differences in the nodulation rate or the maximum number of nodules per plant were observed between strains SL68 and USDA110 (e.g., *G.max*, Fig 19). Plants inoculated with strain SL68 looked as healthy as those inoculated with strains USDA110 and formed pink nodules indicating the presence of leghemoglobin.

DNA Sequencing and Sequence Analysis of *nodZ*

The Tn5 insertion in *B.japonicum* strain NAD138 was previously mapped to a 1.0 kb *BglII* fragment upstream to ORFA (Fig 15; Nieuwkoop et al., 1987; Stacey et al., 1993). When the nodulation ability of strain NAD138 was tested on siratro and a number of cultivars of soybean, the mutant exhibited a host-specific phenotype, i.e., the nodulation efficiency of NAD138 was similar to that of the wild-type strain USDA110 on soybean cv. Essex, reduced on soybean cv. Peking and Forest and severely reduced on siratro (Nieuwkoop et al., 1987; Stacey et al., 1993). Furthermore, the

Figure 19

Percentage of nodulated *G.max* cv. Essex plants after inoculation with *B.japonicum* strains USDA110 or SL68. Each point represents the average of two experiments, each using 20 plants/strain.



nodulation signal produced by strain NAD138 was found to be different from that of the parent strain USDA110 in its lack of the 2-*O*-methylfucose residue on the terminal reducing sugar (Stacey et al., 1993).

A 2.0 kb *Pst*I-*Hind*III fragment that spans the region where the Tn5 insertion in strain NAD138 was mapped was sequenced in both orientations using the M13 clones WS18 and WS19 (Table 2, Fig 20) and a set of synthetic primers (Appendix 1). The sequence revealed a 972 bp open reading frame corresponding to *nodZ* and starting 13 bp upstream of the start codon of the divergently transcribed ORFA (Fig 15, 20). The sequence predicts the NodZ protein to be 36.6 kd, which is approximately the same size previously estimated from expression of NodZ in *E.coli* maxicells (Stacey et al., 1993). The predicted start codon is preceded by a putative ribosomal binding site and the previously determined transcriptional start site of *nodZ* (Stacey et al., 1993) is located 48 bp upstream to the translational start (Fig 20). An RpoN -12.-24 consensus sequence is present in the promoter region and a NifA upstream activator sequence is found approximately 120 bp upstream to the transcriptional start site (Fig 20). According to the transcript size of ORFA, which was determined above, such a transcript could pair with the Shine-Dalgarno sequence and start codon of the *nodZ* transcript and may interfere with the efficiency of its

Figure 20

DNA sequence of *nodZ*. The putative ribosomal binding site is underlined. The transcriptional start is indicated by an arrow head. The -12, -24 RpoN consensus and a NifA putative upstream activator sequence are indicated. The ORF translational start is marked by an arrow.

Figure 20

1 CTGCAGGATTGGAGTGCCTGGCTGCGTCATTGCCGCCGTACGCATCCCTTCATCGTGT 60
61 CGCTCCACATAGCGCCAGATCGGCAGGAAACGGTCTGCGAGATCCGCATGGGACATTCC 120
121 GCCCCTCTAGTATGCGTATCTCACCTGAGCTTTTCGAGATCTTGATGCGAATCGAGGGC 180
181 GAGGCGTCGCTCGGGTCGCTTTTCGACACAGCCCTGCTGGACCAGGCCAAGCGTGAAGAT 240
241 CTTGTGAAGGAATTGCGGCTCGTATGGGAGCTTCGGGGCGTTCCGCTGCATCCGCTGCAC 300
301 GCGGCTTGCCGTACGATAATGTACAAAGCGGAACATAGCCATGACTCATGGCGCATGGC 360
361 CACGCCCCATGCGTGCATGTGAGGGGACCAGGCCATCCGATGCAGAGCGGGGCATCGGGCA 420
421 TCGGGCTTCGCGCAGGTTCGATCAACACCGTTGAAGAAAGGACGATCGTATTGAGAGAAC 480
481 GCGGCCCGTCACGGATAGCCCTGCTATGAAGTTCTACCGATGCAGCTCGCCGGCACCGCG 540
M K F Y R C S S P A P R
541 ATTCAGGCCGTGACGCCGGGCGAGAAGATACGCGAGACGTCTGTACTGACGTCTCTGT 600
F Q A V T P G E K I R E T S V L T S L V
601 CCAACCAGGAGCGCGGAAAGGCAAGGCAGATGGTTAGCGGTTGAGCAATGATCGGTT 660
Q P G A R E K A R Q M V S G S S N D R F
661 CGTAGTGTCCCGGCGACGTACGGGGTTTGGCGACTGTCTGTGGTCACTCGCGGCAGCCTG 720
V V S R R R T G F G D C L W S L A A A W
721 GCGATTTGCAAGCAGACAGGACGGACGCTAGCCATTGATTGGCGGGGATCTTGTTATCT 780
R F A K Q T G R T L A I D W R G S C Y L
781 TGATGAACCATTACCAATGCCTTTCCAGTCTTCTCGAACCAGGTTGAAGACATTGGCGG 840
D E P F T N A F P V F F E P V E D I G G
841 AGTGCAGTAATTGCGATGACGACATCAATACCGGTTTCGTTCCAGGACCATTTCTTCC 900
V R V I C D D D I N T R S F P G P F F P
901 AACATGGTGGAAACAAGCCATCTTTTCGATTGCATCTATCGACCGGACGAGCAGATTTTCAG 960
T W W N K P S F D C I Y R P D E Q I F R
961 GGAACGCGATCAACTCGATCAACTGTTCCAAAGTCAGAGAGATAGTGACGCTAACACGGT 1020
E R D Q L D Q L F Q S Q R D S D A N T V
1021 TGTGTGTGATGCCTGCCTGATGTGGCGCTGCGATCAGGAAGCGGAACGCGAAATCTTCCG 1080
V C D A C L M W R C D Q E A E R E I F R
1081 GAGTATTAAGCCGCGCCTGAAATTCAGGCCCGGATTGATGCTATCTACCGCGAGCACTT 1140
S I K P R P E I Q A R I D A I Y R E H F
1141 TGAGCCGTACAGCGTAATCGGCATTGATGTCGGCATGGCAATGGCGAAGATATTATGGC 1200
E P Y S V I G I H V R H G N G E D I M A
1201 ATGCCCTTACTGGGCGGACACGGAGCCGCTTGGCGCAGATCTACAATGCCATTGATGA 1260
C P L L G G H G A A L R Q I Y N A I D E
1261 GGCCAGGAGTTTATCCCACGCAAAACCTGTGAGGGCGTTCTGTGCACGGACAGTGCGCT 1320

Figure 20 (cont.)

	A R S L S H A K P V R A F L C T D S A L	
1321	CGTCCTCGAGCAGGTTTCGGTTAAATTTCCCGATGTCTTCGCGATTCCAAAGCAGTTCCA V L E Q V S V K F P D V F A I P K Q F Q	1380
1381	GGCGCCTCAAGCGGGGCCATTGCAACCATCCGCCCTCGGAGCTGAAGGCGGCTTTTCTGCC A P Q A G P L H H P P S E L K A A F L P	1440
1441	CTCACCGAGATGTATCTCCTTGCCCGCTGCGACACGGTGATCCGTTTTCCGCCAACGAGC S P R C I S L P A A T R *	1500
1501	GCTTTTACGCGCTATGCGCGTCTTTTCGCCCCACGCGTTATAGAATTCGATTTGAACGAT	1560
1561	CCGGGCCCGGTTGATCCTGATTGAAGACAACCTCGCAAGCGCTTATGGCTTCGTGAAAGTCG	1620
1621	CAAACCCGCGTGATCGATCTCACCGCCGCTCTTTACGTCCGATAGCCGTGGAAGTTCTTC	1680
1681	TCCGGGCGACGATTTTCTGGCCGCGGGCTGTCAATTGCCGCAACGTTCCAGGGGCGAAGAG	1740
1741	AACGAAGCCGCGGCCGAATGCGTTTCGGCTAAGGCGCGGCTTCGACAACAGAGCCCTAAAG	1800
1801	TGCCGCTGCGTTTTACAGAAGGAGCACGCCTTCGTCTTCGCTTCGGGCGGGCCGCAGACG	1860
1861	GAAGCTAAACAACGCTCAATCACGGATGAACATTTTCAGCTCGCCAGCTCATGCTTGGAAAC	1920
1921	TAGCGCCGCGAAATCCAGTGCCGCTCAGAGCATGGCGCGATAGATCAAAAGTGGCCGCCG	1980
1981	ACGAATCTAACGTGCAAGCTT 2001	

translation. The reverse situation is also possible, thus interfering with the efficient translation of ORFA.

Computer analysis of the *nodZ* predicted protein structure and the comparisons of DNA or amino acid sequences to the Genbank and Swissprot data bases did not reveal any significant similarities. Interestingly, however, compiling interstrain DNA hybridization data for *nodZ* (M. Schell, unpublished data), ORFA and *nolO* (this study) revealed that in every strain that hybridized to the different probes for all 3 genes, hybridization occurred to the same size band (Table 5) suggesting linkage between the three genes.

Expression of *nodZ*

Inspection of the DNA region upstream to the *nodZ* transcriptional start revealed a conserved -12, -24 region characteristic of RpoN-dependent promoters. In addition, a putative NifA upstream activator sequence (UAS) (Fig 20; Merrick, 1992) was observed 120 bp 3' of the transcriptional start. This suggested that *nodZ* may be regulated by NifA under anaerobic conditions. However, Table 6 shows that the β -galactosidase activity of the *nodZ-lacZ* fusion plasmid pMGS104 (Table 6) is not altered under anaerobiosis and that NifA is not required for its activity under the conditions used (Table 6). This is consistent with the fact that the *nodZ-lacZ* fusion was equally active in *B. japonicum* strain BjN50-97 lacking both copies of the sigma factor RpoN

Table 6: β -Galactosidase Activity of *nodZ-lacZ* Fusions

STRAIN	GENOTYPE	β -Galactosidase Units	
		AEROBIC	ANAEROBIC
BSL104	USDA110 (pMGS104)	45	39
BSL123	Bj2101-NifA ⁻ (pMGS104)	75	91
BSL125	Bj27147-NtrC ⁻ (pMGS104)	104	158
BSL124	Bj3263-NtrBC ⁻ (pMGS104)	69	62
SL112	Bj-N50-RpoN ₁ ⁻ (pMGS104)	29	80
SL113	Bj-N63-RpoN ₂ ⁻ (pMGS104)	40	81
SL114	BjN50-97-RpoN ₁ N ₂ ⁻ (pMGS104)	29	na
GS300	USDA110 (<i>nifD-lacZ</i>)	50	3569

- Strains USDA110, Bj2101, Bj27147, Bj3263, Bj-N50, Bj-N63 and BjN50-97 had less than 1 unit of activity.

- Each entry is the average of, at least, 2 experiments.

- Variation in the pMGS104 activity in different backgrounds was found not to be significant: *t*-test with corresponding $\alpha=0.1$.

- Cultures were grown in either RDY or Nitrate medium.

- **na**: not applicable.

(Tables 2, 6). Nevertheless, the possibility that there is a third copy of the *rpoN* gene in *B.japonicum* cannot be ruled out. Another activator protein that can mediate RpoN-dependent transcription is NtrC after being phosphorylated by the protein kinase NtrB (Merrick, 1992). Table 6 shows that *nodZ* expression is also not dependent on NtrC/NtrB (Table 6).

CHAPTER IV

DISCUSSION

Previously, Tn3-*lacZ* gene fusions and Tn5 mutagenesis indicated the presence of additional symbiotic loci between *nodJ* and *fixR* within symbiotic cluster II of the *B.japonicum* USDA110 chromosome (Fig 3). Sequencing of this region (4 kb) revealed five novel genes namely, *nolM*, *nolN*, *nolO*, ORFA and *nodZ*. The function and transcriptional control of these genes were studied in light of their location within a genetic cluster involved in symbiotic nitrogen fixation.

The structure of the *nodABC* operon in *B.japonicum* is different from that of *Rhizobium* strains in that *nodA* is separated from its promoter by *nodY* (Nieuwkoop et al., 1987) and that *nodSU* are found between *nodC* and *nodI*. In this work, *nolMNO* were shown to be part of this operon; thus establishing it as *nodYABCSUIJnolMNO* (approximately 9 kb long). The predicted *nolO* amino acid sequence was found to show sequence similarity to that of the upstream *nodU*. *nolO* appears to be the last gene in the operon since it is followed by a GC-rich inverted repeat that precedes an A-rich region, a structure characteristic to *rho*-independent

transcriptional terminators. The close proximity of the divergently transcribed ORFA also argues for *nolO* being the terminal gene in the operon. Another reported case where the *nodABC* operon extends beyond *nodJ* is *nodX* in *R.l. bv. viciae* strain TOM. The *nodX* gene is required for genotype-specific nodulation of Afghanistan pea (Davis et al., 1988).

The *nodABC* genes are absolutely required for the production of lipo-oligosaccharide signals produced by all rhizobia and also for nodulating compatible hosts (Long, 1992; Spaink, 1992). Since *nolMNO* are transcriptionally coupled to *nodABC*, they were predicted to have related functions. Indeed, strains SL65(*nolNO*⁻) and SL67(*nolO*⁻) produced several Nod metabolites that are not produced by the parent strain. Furthermore, these mutants were affected in their ability to nodulate all tested hosts. The fact that both mutants displayed the same phenotype suggests that the role of *NolN* in the biosynthesis of Nod factors may be in a step following that in which *NolO* is involved. However, the possibility that *NolN* may not be involved in Nod factor production cannot be ruled out. Although mutations in *nolNO* or *nolO* have an altered pattern of Nod factor production, considerable amounts of the wild-type BjNod-V(C_{18:1};MeFuc) nodulation signals were still produced. Accordingly, the reduction in nodulation efficiency by strains SL65 and SL67 can be explained by one of two

possibilities. First, the Nod metabolites produced by the mutants may be biologically inactive but compete with the wild-type Nod signal. This possibility is supported by the fact that unlike the wild-type Nod factor, the major Nod metabolite characteristic to the mutant strains [i.e., NodBj-V(C_{18:1})] was unable to induce cortical cell division when applied to *G. soja* roots (Stacey et al., 1993). Alternatively, the quantity of wild-type Nod factor produced by the mutant strains may be insufficient for efficient nodulation. The latter possibility appears more likely since the delay in nodulation by strains SL65 and SL67 can be overcome by increasing the inoculum concentration (data not shown).

In addition to the wild-type Nod factor, six different Nod metabolites produced by mutant strains SL65 and SL67 were identified and their structure determined. Two of these structures [i.e., NodBj-V(Ac,C_{18:1},MeFuc) and NodBj-V(C_{16:0},MeFuc)] are produced by the group I strain *B. japonicum* USDA135 and may be produced by strain USDA110 in minor quantities (Carlson et al., 1993). Two other compounds [i.e., NodBj-IV(C_{18:1},MeFuc) and NodBj-IV(C_{18:1},Gro)] were found to be produced only in the group II strain USDA61 (Carlson et al., 1993). Indeed, it was suggested that the production of these latter Nod signals may be a differentiating trait between wild-type group I and II

strains (Carlson et al., 1993). However, clearly the group I USDA110 mutant strains SL65 and SL67 make these compounds. This fact adds greater credibility to the suggestion of Carlson et al. (1993) that the glycerol-containing derivative may be an intermediate in Nod factor biosynthesis. This is consistent with the fact that all the glycerol-linked Nod metabolites reported in *B.japonicum* strains SL65, SL67 and USDA61 possess tetrameric (and not pentameric) chitose backbones (this work and Carlson et al., 1993). Since mutations in *nolO* of the group I strain USDA110 resulted in the production of Nod metabolites characteristic to the group II strain USDA61, it would be tempting to speculate that *B.japonicum* group II strains should lack *nolO*. However, in this study, a 0.7 *PstI* fragment internal to *nolO* hybridized to chromosomal digests from both groups I and II. Therefore, it seems more likely that the *nolO* mutation disrupts the biosynthetic pathway so intermediates accumulate not normally seen in the wild-type. Strain USDA61 may produce more Nod signal and its biosynthetic pathway may result in a detectable quantity of biosynthetic intermediates being produced.

It is reasonable to speculate that genes transcriptionally coupled to the *nodABC* may be involved in the production of Nod factors. In addition to the role of *NodABC* (presumably the synthesis of the chitose backbone),

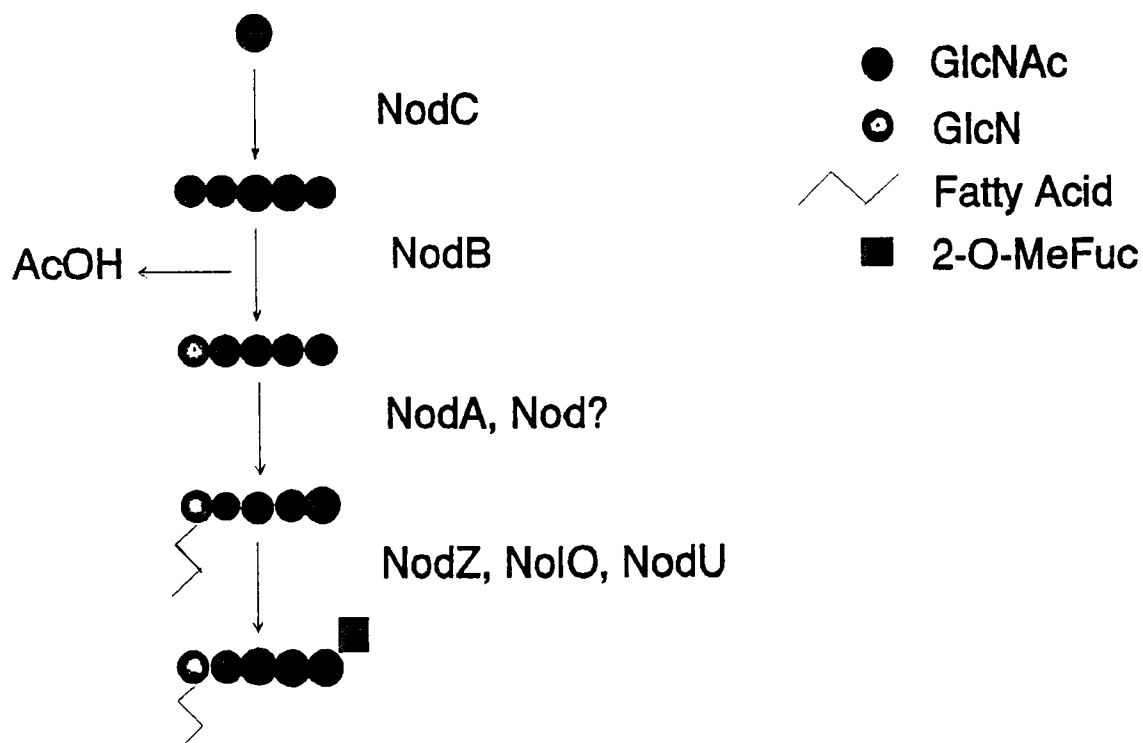
NodS from *A. caulinodans* is proposed to be an N-methyltransferase based on sequence homologies and biochemical data (Geelen et al., 1993). Although the N-methylation of the terminal non-reducing sugar of Nod metabolites has been reported in *B. japonicum* strain USDA61 (Carlson et al., 1993), such modification is not present in strain USDA110 which possesses a copy of *nodS* (Sanjuan et al., 1992; Göttfert et al., 1990). Accordingly, Geelen et al. (1993) suggested that *nodS* in *B. japonicum* is not translated due to the overlap between its putative ribosomal binding site and the 3' end of *nodC* (Geelen et al., 1993; Göttfert et al., 1990). The biochemical functions of NodYUIJ are not known. However, the work presented here implicates NodU in the Nod factor synthesis due to similarity to the protein sequence of Nolo. However, Nolo and NodU may not have the same biochemical function. For example, the region of similarity between the Nolo and NodU proteins may only represent a domain important for interaction with the chitose backbone. The major Nod factor produced by *nolo* mutants but not by the wild-type strain is lacking the 2-O-methylfucose substitution at the C6 position of the terminal reducing sugar. Therefore, it is possible that Nolo, in conjunction with NodZ, may be involved in fucosylation of the Nod factor. Nolo could be involved in production of the fucose donor, methylation of the fucose or some other step. The lack of a strong nodulation phenotype

for *NolO*⁻ mutants could be explained by phenotypic complementation by *NodU*. Of course, such a case would presuppose similar biochemical functions for *NodU* and *NolO*. That is, one would predict that *NodU* is also involved in the substitution of the terminal reducing sugar on the C6 position by a sugar (e.g., fucose). This possibility is supported by the fact that all rhizobia known to possess a copy of *nodU* produce Nod factors that are substituted with a sugar moiety at the C6 position of their terminal reducing sugar. These sugars are 2-*O*-methylfucose, substituted (Ac, or -SO₄) 2-*O*-methylfucose, D-arabinose or fucose found in the Nod factors from *B.japonicum* USDA110, *Rhizobium* sp. NGR234, *A.caulinodans* ORS571 and *R.fredii*, respectively (Mergaert et al., 1993; Price et al., 1993; Prome et al., 1992; Sanjuan et al., 1992). According to this model, *NodU* would be required for the addition of a sugar moiety *per se* and another genetic determinant(s) would determine the nature of the added sugar. It would be interesting to determine the effect of *nodU* mutations on the structure of Nod factors produced by mutant strains. It would also be useful to identify mutations that result in the addition of the "wrong" sugar to the C6 position of the terminal reducing sugar.

An interesting parallel can be drawn between *nodZ/nolO* in *B.japonicum* and *nodH/nodPQ* in *R.meliloti*.

Figure 21

Model for the biosynthetic pathway of the *B. japonicum* USDA110
Nod factor.



NodZ⁻ mutants produce Nod metabolites that lack the 2-O-methylfucose substitution of the terminal reducing sugar (Stacey et al., 1993), while NodO⁻ mutants produce a mixture of molecules with or without the 2-O-methylfucose residue (this work). Similarly, NodH⁻ strains produce Nod metabolites that lack the sulfate group substitution on their reducing end (Roche et al., 1991), while NodPQ⁻ strains produce a mixture of sulfated and non-sulfated molecules (Roche et al., 1991). Several lines of evidence indicate that NodH is a sulfotransferase and that NodP and NodQ are ATP sulfurylase subunits (Spaink, 1992; Roche et al., 1991; Schwedock and Long, 1990). ATP sulfurylase produces phosphoadenosine phosphosulfate, the presumed sulfate donor used by NodH. The leaky phenotype of the NodPQ mutants is attributed to the presence of, at least, one more copy of *nodP* and *nodQ* on the *R.meliloti* chromosome (Schwedock et al., 1989). Likewise, it is possible that NodZ mediates the 2-O-methylfucosylation of the *B.japonicum* Nod metabolites utilizing a NodO-dependent substrate. The leakiness of the phenotype of strains SL65 and SL67 may indicate the presence of a functional homolog to *nolO* (perhaps *nodU*) in *B.japonicum*. A model for Nod factor biosynthesis in *B.japonicum* USDA110 is presented in figure 21.

Downstream to *nolO* and in an opposite orientation, ORFA was identified after an actively transcribed Tn3-*lacZ* fusion was mapped to that region (BSL102, Table 4). ORFA was studied because of its location within a cluster of nodulation genes and since a DNA probe internal to ORFA hybridized to other *Bradyrhizobium* strains including members of the distantly related group II strains. Accordingly, ORFA was predicted to play a role in the nodulation process. However, strain SL68(ORFA⁻) displayed a wild-type nodulation phenotype on all hosts tested. Nevertheless, it is possible that ORFA may be required for nodulation under certain, unknown conditions (e.g., under field conditions) or for nodulation of a yet unidentified plant host(s).

The expression of an ORFA-*lacZ* fusion was not increased by the addition of isoflavones known to induce *nod* transcription. Therefore, the possible involvement of a variety of regulators (i.e., NodD₁, NifA, NtrC, NtrB, RpoN₁ and RpoN₂) in mediating ORFA expression was investigated. However, the ORFA-*lacZ* fusion gave similar activity in all the tested genetic backgrounds. Therefore, ORFA appears to be constitutively expressed and not subject to significant regulation. This is consistent with the fact that none of the known *cis*-acting promoter elements associated with the regulatory proteins tested were found upstream of the transcriptional start. However, it is possible that ORFA is

regulated by conditions (such as those prevailing inside the nodule or within the rhizosphere) that were not provided in this study.

nodZ was previously shown to be required for host-specific nodulation and for the 2-*O*-methylfucosylation of the *B. japonicum* Nod factor (Stacey et al., 1993). With the exception of some *nodD* alleles, *nod* genes are known to be induced by host-produced flavonoids (Long, 1992; Schlaman, 1992). Surprisingly, *nodZ* expression was not induced by genistein or SSE and was found to be NodD-independent (Stacey et al., 1993). The sequence of *nodZ* and its promoter region is presented here and, as expected, no *nod*-box sequence was found upstream of the *nodZ* transcriptional start. The promoter region of *nodZ* contains the -12, -24 consensus sequence characteristic to RpoN-dependent promoters and a NifA upstream activator sequence (UAS) is present approximately 120 bp upstream to the transcriptional start site. However, a *nodZ-lacZ* fusion tested in strains mutated in regulatory proteins associated with RpoN-dependent promoters (i.e., NifA, NtrC, NtrB, RpoN₁ and RpoN₂) gave similar activity in different genetic background. Nevertheless, RpoN/NifA may control *nodZ* expression under certain unknown conditions. Similar to ORFA, *nodZ* appears to be constitutively expressed under free-living conditions. However, *nodZ* expression is

elevated in bacteroids (Stacey et al., 1993). Elevated expression of ORFA in bacteroids was not tested and remains a possibility.

The transcriptional start of *nodZ* has previously been determined (Stacey et al., 1993). The start codons of the divergently transcribed ORFA and *nodZ* are separated by only 13 bp making it very likely that the transcripts of the two genes have a region of overlap. Indeed, determining the location of the transcriptional start site reveals a 128 nucleotide overlap between the two transcripts. In the light of the lack of a known regulation mechanism for ORFA and *nodZ*, this transcriptional overlap may have regulatory significance. Indeed, when a *nodZ-lacZ* fusion was mated into strain SL68(ORFA⁻) and assayed for β -galactosidase activity, a 3-fold decrease in activity was observed when compared to the wild-type strain containing the same fusion (data not shown). If ORFA expression modulates *nodZ* expression, then strain SL68 may show an altered pattern of Nod metabolites. This possibility was not examined in the current study.

In addition to their adjacent location on the *B. japonicum* USDA110 chromosome, *nolO*, ORFA and *nodZ* appear to be also linked in other *Bradyrhizobium* strains (see Table 5). The evidence for this comes from the fact that probes

from all genes consistently hybridized to the same size restriction fragments within individual strains. This apparent genetic linkage, the fact that *NolO*⁻ and *NodZ*⁻ mutants are defective in the 2-O-methylfucosylation of Nod metabolites, and the overlap observed between the *ORFA* and *nodZ* transcripts suggest that the three genes may be functionally related as well.

In summary, the sequences of five genes linked to the common nodulation genes were determined, and the possible function and transcriptional control of *nolNO*, *ORFA* and *nodZ* were studied. *nolMNO* were shown to be co-transcribed with the common nodulation genes; thus, adding a novel feature to the *nodABC* operon structure in *B.japonicum* USDA110. *nolO* was found to be involved in the synthesis of the Nod factor produced by strain USDA110 and to be required for efficient nodulation. These findings support the dogma implicating most, if not all, *nod/nol* genes in the production of Nod metabolites and correlating the structure of Nod factors produced by a rhizobial strain with its ability to nodulate specific hosts. Furthermore, the sequence similarity observed between *NolO* and *NodU* suggests that *NodU* could also be involved in Nod factor synthesis and allows for speculation on its function. The uniqueness of *nodZ* regulation was addressed and the lack of a *nod*-box sequence in its promoter region was demonstrated. The

regulation of ORFA was studied as well and the possibility of a coordinated regulatory mechanism between ORFA and *nodZ* was discussed.

REFERENCES

- Akkermans, A., Abdulkadir, S. and Trinick, M. (1978) Nitrogen fixing root nodules in Ulmaceae. *Nature* **274**, 190
- Anthamatten, D. and Hennecke, H. (1991). The regulatory status of the *fixL*- and *fixJ*-like genes in *Bradyrhizobium japonicum* may be different from that in *Rhizobium meliloti*. *Mol. Gen. Genet.* **225**, 38-48.
- Appelbaum, E.R., Thompson, D.V., Idler, K., and Chartrain, N. (1988). *Rhizobium japonicum* USDA 191 has two *nodD* genes that differ in primary structure and function. *J. Bacteriol.* **170**, 12-20.
- Appleby, C. (1984) Leghemoglobin and *Rhizobium* respiration. *Ann. Rev. Plant Physiol.* **35**, 433-478.
- Banfalvi, Z., Nieuwkoop, A.J., Schell, M.G., Besl, L., and Stacey, G. (1988). Regulation of *nod* gene expression in *Bradyrhizobium japonicum*. *Mol. Gen. Genet.* **214**, 420-424.
- Banfalvi, Z. and Kondorosi, A. (1989). Production of root hair deformation factors by *Rhizobium meliloti* nodulation genes in *Escherichia coli*: HsND (NodH) is involved in the plant host-specific modification of the NodABC factor. *Plant Mol. Biol.* **13**, 1-12.
- Batut, J., de Philip, P., Reytrat, J., Waelkens, F. and Boistard, P. (1992) Oxygen regulation of nitrogen fixation gene expression in *Rhizobium meliloti*. In *Advances in Molecular Genetics of Plant-Microbe Interactions*, Nester, E. and Verma, D. P. (eds), Kluwer Academic Publishers, Dordrecht-Boston-London, pp 183-192.
- Bergersen, F. (1982) *Root nodules of legumes: Structure and functions*. Chechester: Research Studies Press, UK.
- Bergersen, F.J. and Turner, G.L. (1990). Bacteroids from soybean root nodules - Respiration and N₂- fixation in flow-chamber reactions with oxyleghaemoglobin. *P. Roy. Soc. B.* **238**, 295-320.
- Bolton, E., Higgisson, B., Harrington, A. and O'Gara, F. (1986) Dicarboxylic acid transport in *Rhizobium meliloti*: Isolation of mutants and cloning of dicarboxylic acid transport genes. *Arch. Microbiol.* **144**, 142-146.
- Burn, J.E., Hamilton, W.D., Wootton, J.C., and Johnston, A.W.B. (1989). Single and multiple mutations affecting properties of the regulatory gene *nodD* of *Rhizobium*. *Mol. Microbiol.* **3**, 1567-1577.
- Caetano-Anollés, G., Crist-Estes, D.K., and Bauer, W.D. (1988). Chemotaxis of *Rhizobium meliloti* to the plant flavone luteolin

requires functional nodulation genes. *J. Bacteriol.* **170**, 3164-3169.

Caetano-Anollés, G. and Gresshoff, P. (1991) *Annu. Rev. Microbiol.* **45**, 345-382.

Callaham, D.A. and Torrey, J.G. (1981). The structural basis for infection of root hairs of *Trifolium repens* by *Rhizobium*. *Can. J. Bot.* **59**, 1647-1664.

Canter Cremers, H., Wijffelman, C., Pees, E., Rolfe, B., Djordjevic, M. and Lugtenberg, B. (1988) Host specific nodulation of plants of the pea cross-inoculation group is influenced by genes in fast growing *Rhizobium* downstream *nodC*. *J. Plant Physiol.* **132**, 398-404.

Canter Cremers, H., Spaink, H., Wijffjes, A., Pees, E., Wijffelman, C., Okker, R. and Lugtenberg, B. (1990) Additional nodulation genes on the Sym plasmid of *R. leguminosarum* biovar *viciae*. *Plant Mol. Biol.* **13**, 163-174.

Carlson, R.W., Sanjuan, J., Bhat, U.R., Glushka, J., Spaink, H.P., Wijffjes, A.H.M., van Brussel, A.A.N., Stokkermans, T.J.W., Peters, N.K., and Stacey, G. (1993). The structures and biological activities of the lipo-oligosaccharide nodulation signals produced by Type I and Type II strains of *Bradyrhizobium japonicum*. *J. Biol. Chem.* (submitted).

Chen, W.X., Yan, G.H., and Li, J.L. (1988). Numerical taxonomic study of fast-growing soybean rhizobia and a proposal that *Rhizobium fredii* be assigned to *Sinorhizobium* gen. nov. *Int. J. Sys. Bacteriol.* **38**, 392-397.

Dart, P. (1977). Infection and development of leguminous nodules. In A treatise on dinitrogen fixation. Section III. Biology, R.W.F. Hardy and W.S. Silver, eds. (New York: John Wiley and Sons),

David, M., Daveran, M.L., Batut, J., Dedieu, A., Domergue, O., Ghai, J., Hertig, C., Boistard, P., and Kahn, D. (1988). Cascade regulation of *nif* gene expression in *Rhizobium meliloti*. *Cell* **541**, 671-6830.

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.

Davis, E.O. and Johnston, A.W.B. (1990). Regulatory functions of the three *nodD* genes of *Rhizobium leguminosarum* biovar *phaseoli*. *Mol. Microbiol.* **4**, 933-941.

Dean, D. and Jacobson, M. (1992) Biochemical genetics of nitrogenase. In *Biological nitrogen fixation*, Stacey, G., Burris, R. and Evans, H. (eds), Chapman and Hall, New York-London, pp 763-834.

Debellé, F., Rosenberg, C., Vasse, J., Maillet, F., Martinez, E., Denarié, J., and Truchet, G. (1986). Assignment of symbiotic developmental phenotypes to common and specific nodulation (*nod*) genetic loci of *Rhizobium meliloti*. *J. Bacteriol.* **168**, 1075-1086.

Demont, N., Roche, P., Aurelle, H., Talmont, F., Promé, D., Promé, J.C., Price, N., Relic, B., Broughton, W., Debellé, F., Ardourel, M., Maillet, F., Rosenberg, C., Truchet, G. and Denarié, J. (1993) *Rhizobium* nodulation signals: Variations on a theme. In *Advances in Molecular Genetics of Plant-Microbe Interactions*, Nester, E. and Verma, D.P. (eds), Kluwer Academic Publishers, Dordrecht, pp133-142.

Devine, T.E., Kuykendall, L.D., and O'Neill, J.J. (1990). The Rj4 allele in soybean represses nodulation by chlorosis-inducing Bradyrhizobia classified as DNA homology group-II by antibiotic resistance profiles. *Theor. Appl. Genet.* **80**, 33-37.

Djordjevic, M.A., Schofield, P.R., and Rolfe, B.G. (1985). Tn5 mutagenesis of *Rhizobium trifolii* host-specific genes result in mutants with altered host range ability. *Mol. Gen. Genet.* **200**, 463-471.

Dudley, M., Jacobs, T. and Long, S. (1987) Microscopic studies of cell divisions induced in alfalfa roots by *Rhizobium meliloti*. *Planta* **171**, 289-301.

Eaglesham, A.R.J., Ellis, J.M., Evans, W.R., Fleischman, D.E., Hungria, M. and Hardy, R.W.J. (1990). The first photosynthetic N₂-fixing *Rhizobium*: Characteristics. In *Nitrogen fixation: achievements and Objectives*. Gresshoff, P., Roth, E., Stacey, G. and Newton, W. (eds), Chapman and Hall, New York. pp. 805-811.

Earl, C.D., Ronson, C.W., and Ausubel, F.M. (1987). Genetic and structural analysis of the *Rhizobium meliloti* *fixA*, *fixB*, *fixC*, and *fixX* genes. *J. Bacteriol.* **169**, 1127-1136.

Elkan, G.H. (1992). Taxonomy of the rhizobia. *Can. J. Microbiol.* **38**, 446-450.

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.

Faucher, C., Maillet, F., Vasse, J., Rosenberg, C., van Brussel, A.A.N., Truchet, G., and Denarié, J. (1988). *Rhizobium meliloti* host range *nodH* gene determines production of an alfalfa-specific extracellular signal. *J. Bacteriol.* **170**, 5489-5499.

Fellay, R., Frey, J., and Krisch, H. (1987). Interposon mutagenesis of soil and water bacteria: a family of DNA fragments designed for in vitro insertional mutagenesis of Gram-negative bacteria. *Gene* **52**, 147-154.

Festl, H., Ludwig, W. and Schleifer, K.H. (1986) DNA hybridization probe for the *Pseudomonas fluorescens* group. *Appl. Environ. Microbiol.* **52**, 1190-1194.

Finan, J.M., Hirsch, A.M., Leigh, J.A., Johansen, E., Kuldau, G.A., Deegar, S., Walker, G.C., and Signer, E.R. (1985). Symbiotic mutants of *Rhizobium meliloti* that uncouple plant from bacterial differentiation. *Cell* **40**, 869-877.

Finan, T.M., Hirsch, A.M., Leigh, J.A., Johansen, E., Kuldau, A., Wood, J.M., and Jordan, D.C. (1983). Symbiotic properties of C4-dicarboxylic acid transport mutants of *Rhizobium leguminosarum*. *J. Bacteriol.* **154**, 1403-1413.

Fischer, H.-M. and Hennecke, H. (1984). Linkage map of the *Rhizobium japonicum nifH* and *nifDK* operons encoding the polypeptides of the nitrogenase enzyme complex. *Mol. Gen. Genet.* **196**, 537-5400.

Fischer, H.-M. and Hennecke, H. (1987). Direct response of *Bradyrhizobium japonicum nifA*-mediated gene regulation to cellular oxygen status. *Mol. Gen. Genet.* **209**, 621-626.

Fischer, H.-M., Bruderer, T., and Hennecke, H. (1988). Essential and non-essential domains in the *Bradyrhizobium japonicum* NifA protein: identification of indispensable cysteine residues potentially involved in redox reactivity and/or metal binding. *Nucl. Acids Res.* **16**, 2207-2224.

Fischer, H.M., Anthamatten, D., Kullik, I., Morett, E., Acuna, G. and Hennecke, H. (1991) Complex regulatory network for *nif* and *fix* gene expression in *Bradyrhizobium japonicum*. In *Advances in Molecular Genetics of Plant-Microbe Interactions* Vol. I, Hennecke, H. and Verma, D.P. (eds), Kluwer Academic Publishers, Dordrecht, pp 203-210.

Fisher, R.F., Tu, J.K., and Long, S.R. (1985). Conserved nodulation genes in *R. meliloti* and *R. trifolii*. *Appl. Env. Microbiol.* **490**, 14320-14350.

Fisher, R.F., Egelhoff, T.T., Mulligan, J.T., and Long, S.R. (1988). Specific binding of proteins from *Rhizobium meliloti*

cell-free extracts containing NodD to DNA sequences upstream of inducible nodulation genes. *Genes Devel.* **2**, 282-293.

Fred, E.B., Baldwin, I.L. and McCoy, E. (1932). Root nodule bacteria and leguminous plants. University of Wisconsin Press, Madison, Wis.

Gaworzewska, E.T. and Carlile, M.J. (1982). Positive chemotaxis of *Rhizobium leguminosarum* and other bacteria towards roots exudates from legumes and other plants. *J. Gen. Microbiol.* **128**, 1179-1188.

Geelen, D., Mergaert, P., Geremia, R., Goormachtig, S., van Montagu, M., and Holsters, M. (1992). Identification of *nodSUIJ* genes in Nod locus 1 of *Azorhizobium caulinodans*: evidence that *nodS* encodes a methyltransferase involved in Nod factor modification. *Mol. Microbiol.* (in press).

Gerhold, D. (1989) Host regulation of nodulation genes in *Rhizobium meliloti* and *Bradyrhizobium japonicum*. Ph.D. Dissertation, Univ. of Tennessee, Knoxville.

Gilles-Gonzalez, M., Ditta, G. and Helinski, D. (1991) A hemoprotein with kinase activity encoded by the oxygen sensor of *Rhizobium meliloti*. *Nature* **350**, 170-172.

Goethals, K., Mergaert, P., Gao, M., Geelen, D., van Montagu, M., and Holsters, M. (1992). Identification of a new inducible nodulation gene in *Azorhizobium caulinodans*. *Mol. Plant-Microbe Int.* **5**, 405-411.

Göttfert, M., Grob, P., and Hennecke, H. (1990a). Proposed regulatory pathway encoded by the *nodV* and *nodW* genes, determinants of host specificity in *Bradyrhizobium japonicum*. *Proc. Natl. Acad. Sci. (USA)* **87**, 2680-2684.

Göttfert, M., Hitz, S., and Hennecke, H. (1990b). Identification of *nodS* and *nodU*, two inducible genes inserted between the *Bradyrhizobium japonicum* *nodYABC* and *nodIJ* genes. *Mol. Plant-Microbe Int.* **3**, 308-316.

Gubler, M. and Hennecke, H. (1986). *fixA*, *B*, and *C* genes are essential for symbiotic and free-living microaerobic nitrogen fixation. *FEBS Lett.* **200**, 186-1920.

Guerinot, M. and Chelm, B. (1986) Bacterial δ -aminolevulinic acid synthase activity is not essential for leghemoglobin formation in the soybean/*Bradyrhizobium japonicum* symbiosis. *Proc. Natl. Acad. Sci. USA* **83**, 1837-

Gyorgypal, Z., Iyer, N., and Kondorosi, A. (1988). Three regulatory *nodD* alleles of divergent flavonoid-specificity are

- involved in host-dependent nodulation by *Rhizobium meliloti*. *Mol. Gen. Genet.* **212**, 85-92.
- Halverson, L.J. and Stacey, G. (1985). Host recognition in the *Rhizobium*-soybean symbiosis. Evidence for the involvement of lectin in nodulation. *Plant Physiol.* **77**, 621-625.
- Henikoff, S., Haughn, G.W., Calvo, J.M., and Wallace, J.C. (1988). A large family of bacterial activator proteins. *Proc. Natl. Acad. Sci. (USA)* **85**, 6602-6606.
- Hollis, A.B., Kloos, W.E., and Elkan, G.H. (1981). DNA:DNA hybridization studies of *Rhizobium japonicum* and related Rhizobiaceae. *J. Gen. Microbiol.* **123**, 215-222.
- Hong, G.-F., Burn, J.E., and Johnston, A.W.B. (1987). Evidence that DNA involved in the expression of nodulation (*nod*) genes in *Rhizobium* binds to the product of the regulatory gene *nodD*. *Nucl. Acids Res.* **15**, 9677-9689.
- Honma, M.A. and Ausubel, F.M. (1987). *Rhizobium meliloti* has three functional copies of the *nodD* symbiotic regulatory gene. *Proc. Natl. Acad. Sci. (USA)* **84**, 8558-8562.
- Horvath, B., Kondorosi, E., John, M., Schmidt, J., Torok, I., Gyorgypal, Z., Barabas, I., Wieneke, U., Schell, J., and Kondorosi, A. (1986). Organization, structure, and symbiotic function of *Rhizobium meliloti* nodulation genes determining host specificity for alfalfa. *Cell* **46**, 335-343.
- Horvath, B., Bachem, C.W.B., Schell, J., and Kondorosi, A. (1987). Host-specific regulation of nodulation genes in *Rhizobium* is mediated by plant-signal, interacting with the *nodD* gene product. *EMBO J.* **6**, 841-848.
- Huber, T., Agarwal, A and Keister, D. (1984) Extracellular polysaccharide composition, ex planta nitrogenase activity and DNA homology in *Rhizobium japonicum*. *J. Bacteriol.* **158**, 1168-1171.
- Kaluza, K. and Hennecke, H. (1984) Fine structure analysis of the *nifHDK* operon encoding the α and β subunits of dinitrogenase from *Rhizobium japonicum*. *Mol. Gen. Genet.* **196**, 35-42.
- Kassavetis, G.A. and Geiduschek, E.P. (1982). Bacteriophage T4 late promoters: Mapping 5' ends of T4 gene 23 mRNAs. *EMBO J.* **1**, 107-114.
- Kijne, J. (1992). The *Rhizobium* infection process. In *Biological Nitrogen Fixation*, G. Stacey, R.H. Burris, and H.J. Evans, eds. (New York: Chapman and Hall), pp. 349-398.

Knauf, V. and Nester, E. (1982) Wide host range cloning vectors: a cosmid clone bank of an *Agrobacterium* Ti plasmid. *Plasmid* **8**, 45-

Kondorosi, A., Kondorosi, E., Pankhurst, C.E., Broughton, W.J., and Banfalvi, Z. (1984). Physical and genetic analysis of a symbiotic region of *Rhizobium meliloti*: Identification of nodulation genes. *Mol. Gen. Genet.* **193**, 445-452.

Kullik, I., Hennecke, H., and Fischer, H.-M. (1989). Inhibition of *Bradyrhizobium japonicum* *nifA*-dependent *nif* gene activation by oxygen occurs at the NifA protein level and is irreversible. *Arch. Microbiol.* **151**, 191-197.

Kullik, I., Fritsche, S., Knobel, H., Sanjuan, J., Hennecke, H., and Fischer, H.-M. (1991). *Bradyrhizobium japonicum* has two differentially regulated, functional homologs of the γ^{54} gene (*rpoN*). *J. Bacteriol.* **173**, 1125-1138.

Kuykendall, L., Roy, M., O'Neill, J. and Devine, T. (1988) Fatty acids, antibiotic resistance and deoxyribonucleic acid homology groups of *Bradyrhizobium japonicum*. *Int. Syst. Bacteriol.* **8**, 358-361.

Kuykendall, L.D., Saxena, B., Devine, T.E., and Udell, S.E. (1992). Genetic Diversity in *Bradyrhizobium japonicum* Jordan 1982 and a Proposal for *Bradyrhizobium elkanii* sp. nov. *Can. J. Microbiol.* **38**, 501-505.

Kuykendall, L.D. and Elkan, G.H. (1976). *Rhizobium japonicum* derivatives differing in nitrogen-fixing efficiency and carbohydrate utilization. *Appl. Env. Microbiol.* **32**, 511-519.

Layzel, D., Hunt, S. and Palmer, G. (1990) Mechanism of nitrogenase inhibition in soybean nodules. Pulse-modulated spectroscopy indicates that nitrogenase activity is limited by O_2 . *Plant Physiol.* **92**, 1101-1107.

Lerouge, P., Roche, P., Faucher, C., Maillet, F., Truchet, G., Promé, J.C., and Denarié, J. (1990). Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulphated and acylated glucosamine oligosaccharide signal. *Nature* **344**, 781-784.

Libbenga, K. and Harkes, P. (1973) Initial proliferation in the formation of root nodules in *Pisum sativum*. *Planta* **114**, 17-28.

Long, S., Reed, J.W., Himawan, J., and Walker, G.C. (1988). Genetic analysis of a cluster of genes required for synthesis of the calcofluor-binding exopolysaccharide of *Rhizobium meliloti*. *J. Bacteriol.* **170**, 4239-4248.

- Maniatis, T., Fritsch, E. and Sambrook, J. (1982) *Molecular Cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, New York.
- Martin, G.B., Chapman, K.A., and Chelm, B.K. (1988). Role of the *Bradyrhizobium japonicum ntrC* gene product in differential regulation of the glutamine synthetase II gene (*glnII*). *J. Bacteriol.* **170**, 5452-5459.
- Marvel, D.J., Torrey, J.G., and Ausubel, F.M. (1987). *Rhizobium* symbiotic genes required for nodulation of legume and nonlegume hosts. *Proc. Natl. Acad. Sci. (USA)* **84**, 1319-1323.
- McIver, J., Djordjevic, M.A., Weinman, J.J., Bender, G.L., and Rolfe, B.G. (1989). Extension of host range of *Rhizobium leguminosarum* biovar *trifolii* caused by point mutations in *nodD* which result in alterations in regulatory function and recognition of inducer molecules. *Mol. Plant-Microbe Int.* **2**, 97-106.
- Mergaert, P., van Montagu, M., Promé, J.-C., and Holsters, M. (1992). Three unusual modifications, a D-arabinosyl, a N-methyl, and a carbamoyl group, are present on the Nod factors of *Azorhizobium caulinodans* strain ORS571. *Proc. Natl. Acad. Sci. (USA)* (in press).
- Merrick, M. (1992) Regulation of nitrogen fixation genes in free-living and symbiotic bacteria. In *Biological Nitrogen Fixation*, Stacey, G., Burris, R. and Evans, H. (eds), Chapman and Hall, New York-London, pp 835-876.
- Messing, J., Crea, R., and Seeburg, R.H. (1981). A system for shotgun DNA sequencing. *Nucl. Acids Res.* **9**, 309-321.
- Mulligan, J.T. and Long, S.R. (1985). Induction of *Rhizobium meliloti nodC* expression by plant exudate requires *nodD*. *Proc. Natl. Acad. Sci. (USA)* **82**, 6609-6613.
- Mulligan, J.T. and Long, S.R. (1989). A family of activator genes regulates expression of *Rhizobium meliloti* nodulation genes. *Genetics* **122**, 7-18.
- Murray, N., Brammar, W. and Murray, K. (1977) Lambdoid phages that simplify the recovery of in vitro recombinants. *Mol. Gen. Genet.* **150**, 53
- Nadler, K. and Avissar, Y. (1977) Heme synthesis in soybean root nodules. II. On the role of bacteroid δ -aminolevulinic acid synthase and δ -aminolevulinic acid dehydrogenase in the synthesis of the heme of leghemoglobin. *Plant Physiol.* **60**, 433-

Nap, J.P. and Bisseling, T. (1990). Nodulin function and nodulin gene regulation in root nodule development. In *Molecular biology of symbiotic nitrogen fixation*, P. Gresshoff (ed), CRC press Inc., Boca Raton, Fl, pp 181-229.

Newcomb, W., Sipell, D. and Peterson, R. (1979) The early morphogenesis of *Glycine max* and *Pisum sativum* root nodules. *Can. J. Bot.* **57**, 2603-2616.

Newcomb, W. (1981) Nodule morphogenesis. In *International Review of Cytology*, Suppl. 13, Bourne, G. and Danielli, J. (eds) New York: Academic Press, pp 246-298.

Nieuwkoop, A.J., Banfalvi, Z., Deshmane, N., Gerhold, D., Schell, M.G., Sirotkin, K.M., and Stacey, G. (1987). A locus encoding host range is linked to the common nodulation genes of *Bradyrhizobium japonicum*. *J. Bacteriol.* **169**, 2631-2638.

Nutman, P. (1959) Some observations on root-hair infection by nodule bacteria. *J. Exp. Bot.* **10**, 250-263.

O'Brian, M.R., Kirschbom, P.M., and Maier, R.J. (1987). Bacterial heme synthesis is required for expression of the leghemoglobin holoprotein but not the apoprotein in soybean root nodules. *Proc. Natl. Acad. Sci. (USA)* **84**, 8390-8393.

Price, G. Mohapatra, S and Gresshoff, P. (1984) Structure of nodules by *Rhizobium* strain ANU289 in the nonlegume *Parasponia* and the legume siratro (*Macroptilium atropurpureum*). *Bot. Gaz.* **145**, 444-451.

Price, N.P.J., Relic, B., Talmont, F., Lewin, A., Promé, d., Pueppke, S.G., Maillet, F., Denarié, J., Promé, J.-C., and Broughton, W.J. (1992). Broad-host-range *Rhizobium* species strain NGR234 secretes a family of carbamoylated and fucosylated nodulation signals that are O-acetylated or sulphated. *Mol. Microbiol.* **6**, 3575-3584.

Promé, J., Bec-Ferte, Broughton, W., Denarié, J., Lewin, A., Martinez, E, Poupot, R., Price, N., M., Promé, D., Puepke, S., Relic, B., Savagnac, A. and Talmont, F. (1992) Structural studies of Nod factors from *Rhizobium* NGR234, *Sinorhizobium fredii*, *Rhizobium tropici* and *Rhizobium leguminosarum* bv. *phaseoli*. Abstract book, 9th International Congress On Nitrogen Fixation, Cancun, Mexico.

Puvanesarajah, V., Schell, F.M., Gerhold, D., and Stacey, G. (1987). Cell surface polysaccharides from *Bradyrhizobium japonicum* and a nonnodulating mutant. *J. Bacteriol.* **169**, 137-141.

Ramseier, T.M. and Göttfert, M. (1991). Codon Usage and G + C Content in *Bradyrhizobium japonicum* Genes Are Not Uniform. *Arch. Microbiol.* **156**, 270-276.

Rigby, P., Dieckmann, M., Rhodes, C. and Berg, P. (1977) Labeling deoxynucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. *J. Mol. Biol.* **113**, 237

Roche, P., Debellé, F., Maillet, F., Lerouge, P., Faucher, C., Truchet, G., Denarié, J., and Promé, J.-C. (1991). Molecular basis of symbiotic host specificity in *Rhizobium meliloti*: *nodH* and *nodPQ* genes encode the sulfation of lipo-oligosaccharide signals. *Cell* **67**, 1131-1143.

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 rhizobia. *Proc. Natl. Acad. Sci. (USA)* **83**, 1757-1761.

Roth, L.E., Jeon, K., and Stacey, G. (1988). Homology in endosymbiotic systems: The term "symbiosome". In *Molecular Genetics of Plant-Microbe Interactions*, R. Palacios and D.P.S. Verma, eds. (St. Paul, MN: The American Phytopathological Society), pp. 220-225.

Sadowsky, M.J., Cregan, P.B., Göttfert, M., Sharma, A., Gerhold, D., Rodriguez-Quinones, F., Keyser, H.H., Hennecke, H., and Stacey, G. (1991). The *Bradyrhizobium japonicum nola* gene and its involvement in the genotype-specific nodulation of soybeans. *Proc. Natl. Acad. Sci. (USA)* **88**, 637-641.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. (USA)* **74**, 5463-5467.

Sangwan, I. and O'Brian, M.R. (1991). Evidence for an inter-organismic heme biosynthetic pathway in symbiotic soybean root nodules. *Science* **251**, 1220-1222.

Sanjuan, J., Carlson, R.W., Spaink, H.P., Bhat, U.R., Barbour, W.M., Glushka, J., and Stacey, G. (1992). A 2-O-methylfucose moiety is present in the lipo-oligosaccharide nodulation signal of *Bradyrhizobium japonicum*. *Proc. Natl. Acad. Sci. (USA)* **89**, 8789-8793.

Sanjuan, J., Luka, S. and Stacey, G. (1993a) Genetic maps of *Rhizobium* and *Bradyrhizobium* species. In *Genetic maps*, 6th

edition, O'Brien, S. (ed) Cold Spring Harbour Laboratory Press, New York. (in press).

Sanjuan, J., Grob, P., Göttfert, M., Dockendorff, T., Hennecke, H. and Stacey, G. (1993b) Regulation of the *Bradyrhizobium japonicum* common nodulation genes by a two-component regulatory system. *Proc. Natl. Acad. Sci. USA* (submitted).

Schlaman, H.R.M., Okker, R.J.H., and Lugtenberg, B.J.J. (1992). Regulation of nodulation gene expression by NodD in rhizobia. *J. Bacteriol.* **174**, 5177-5182.

Schultze, M., Quicletsire, B., Kondorosi, E., Virelizier, H., Glushka, J.N., Endre, G., Gero, S.D., and Kondorosi, A. (1992). *Rhizobium meliloti* produces a family of sulfated lipo-oligosaccharides exhibiting different degrees of plant host specificity. *Proc. Natl. Acad. Sci. USA* **89**, 192-196.

Schwedock, J. and Long, S.R. (1989). Nucleotide sequence and protein products of two new nodulation genes of *Rhizobium meliloti*, *nodP* and *nodQ*. *Mol. Plant-Microbe Int.* **2**, 181-194.

Schwedock, J. and Long, S.R. (1990). ATP sulphurylase activity of the *nodP* and *nodQ* gene products of *Rhizobium meliloti*. *Nature* **348**, 644-646.

Scott, K.F. (1986). Conserved nodulation genes from the non-legume symbiont *Bradyrhizobium japonicum* sp. (*Parasponia*). *Nucl. Acids Res.* **14**, 2905-2910.

Simon, R., Priefer, U., and Pühler, A. (1983). A broad host range mobilization system for in vitro genetic engineering: transposon mutagenesis in gram negative bacteria. *Biotechnol.* **1**, 784-791.

Smit, G., Kijne, J and Lugtenberg, B. (1987) Both cellulose fibrils and a Ca^{2+} -dependent adhesin are involved in the attachment of *Rhizobium leguminosarum* to pea root hair tips. *J. Bacteriol.* **169**, 4294-4301.

Smit, G., Logman, T.J.J., Boerrigter, M.E.T.I., Kijne, J.W., and Lugtenberg, B.J.J. (1989). Purification and partial characterization of the Ca-dependent adhesin from *Rhizobium leguminosarum* biovar *viciae*, which mediates the first step in attachment of Rhizobiaceae cells to plant root hair tips. *J. Bacteriol.* **171**, 4054-4062.

Smit, G., Puvanesarajah, V., Carlson, R.W., Barbour, W.M., and Stacey, G. (1992). *Bradyrhizobium japonicum nodD1* can be specifically induced by soybean flavonoids that do not induce the *nodYABCSUIJ* operon. *J. Biol. Chem.* **267**, 310-318.

So, J.-S., Hodgson, A.L.M., Haugland, R., Leavitt, M., Banfalvi, Z., Nieuwkoop, A.J., and Stacey, G. (1987). Transposon-induced symbiotic mutants of *Bradyrhizobium japonicum*: Isolation of two gene regions essential for nodulation. *Mol. Gen. Genet.* **207**, 15-23.

Spaink, H.P., Okker, R.J.H., Wijffelman, C.A., Pees, E., and Lugtenberg, B.J.J. (1987a). Promoters in the nodulation region of the *Rhizobium leguminosarum* Sym plasmid pRL1JI. *Plant Mol. Biol.* **9**, 27-39.

Spaink, H.P., Wijffelman, C.A., Pees, W., Okker, R.J.H., and Lugtenberg, B.J.J. (1987b). *Rhizobium* nodulation gene *nodD* as a determinant of host specificity. *Nature* **328**, 337-340.

Spaink, H., Sheeley, D., van Brussel, A., Glushka, J., York, W., Tak, T., Geiger, O., Kennedy, E., Reinhold, V. and Lugtenberg, B. (1991) A novel highly unsaturated fatty acid moiety of lipo-oligosaccharide signals determine host specificity of *Rhizobium*. *Nature* **354**, 125-130.

Spaink, H., Wijffjes, A., Geiger, O., Bloemberg, G., Ritsema, B. and Lugtenberg, B. (1993) The function of the rhizobial *nodABC* and *nodFEL*. In *New Horizons in Nitrogen Fixation*, Palacios, R., Mora, J. and Newton, W. E. (eds), Kluwer Academic Publishers, Dordrecht-Boston-London, pp 165-170.

Stacey, G., Sanjuan, J., Spaink, H., van Brussel, T., Lugtenberg, B., Glushka, J. and Carlson, R. (1992) *Rhizobium* lipo-oligosaccharides: Novel plant growth regulators.

Stacey, G., Luka, S., Sanjuan, J., Carlson, R., Banfalvi, Z. and Nieuwkoop, A. (1993) *nodZ*, a unique host-specific nodulation gene, is involved in the fucosylation of the lipo-oligosaccharide nodulation signal of *Bradyrhizobium japonicum*. *Cell* (Submitted).

Stock, J.B., Ninfa, A.J., and Stock, A.M. (1989). Protein phosphorylation and regulation of adaptive responses in bacteria. *Microbiol. Rev.* **53**, 450-490.

Swanson, J.A., Tu, J.K., Ogawa, J., Sanga, R., Fisher, R.F., and Long, S.R. (1987). Extended region of nodulation genes in *Rhizobium meliloti* 1021. I. Phenotypes of Tn5 insertion mutants. *Genetics* **117**, 181-189.

Tartof, K. and Hobbs, C. (1987) Improved media for growing plasmid and cosmid clones. Bethesda Res. Lab. *Focus* **9**, 12

Thöny, B., Fischer, H.-M., Anthamatten, D., Bruderer, T., and Hennecke, H. (1987). The symbiotic nitrogen fixation regulatory operon (*fixRnifA*) of *Bradyrhizobium japonicum* is expressed

aerobically and is subject to a novel, *nifA*-independent type of activation. *Nucl. Acids Res.* **15**, 8479-8499.

Thöny, B., Anthamatten, D., and Hennecke, H. (1989). Dual control of the *Bradyrhizobium japonicum* symbiotic nitrogen fixation regulatory operon *fixRnifA*: analysis of *cis*- and *trans*- acting elements. *J. Bacteriol.* **171**, 4162-4169.

Thöny-Meyer, L., Stax, D., and Hennecke, H. (1989). An unusual gene cluster for the cytochrome *bcl* complex in *Bradyrhizobium japonicum* and its requirement for effective root nodule symbiosis. *Cell* **57**, 683-697.

Thorneley, R. and Lowe, D. (1985) Kinetics and mechanisms of the nitrogenase enzyme system. In *Molybdenum Enzymes*, Ch. 5, Spiro, T. (ed), Wiley, New York, pp 221-284.

Tjepkema, J. and Yocum, C. (1974) Measurement of oxygen partial pressure within soybean nodules by oxygen microelectrodes. *Planta* **119**, 351-360.

Truchet, G. and Dazzo, F. (1982) Morphogenesis of lucerne root nodules incited by *Rhizobium meliloti* in the presence of combined nitrogen. *Planta* **154**, 352-360.

Truchet, G., Barker, D.G., Camut, S., Debilly, F., Vasse, J., and Huguet, T. (1989). Alfalfa Nodulation in the Absence of *Rhizobium*. *Mol. G. Genet.* **219**, 65-68.

Truchet, G., Roche, P., Lerouge, P., Vasse, J., Camut, S., de Billy, F., Promé, J.-C., and Denarié, J. (1991). Sulphated lipo-oligosaccharide signals of *Rhizobium meliloti* elicit root nodule organogenesis in alfalfa. *Nature* **351**, 670-673.

Tsien, H., Dreyfus, B. and Schmidt, E. (1983) Initial stages in the morphogenesis of nitrogen fixing stem nodules of *Sesbania rostrata*. *J. Bacteriol.* **156**, 888-897.

van Brussel, A.A.N., Zaat, S.A.J., Canter-Cremers, H.C.J., Wijffelman, C.A., Pees, E., Tak, T., and Lugtenberg, B.J.J. (1986). Role of plant root exudate and Sym plasmid-localized nodulation genes in the synthesis by *Rhizobium leguminosarum* of Tsr factor, which causes thick and short roots on common vetch. *J. Bacteriol.* **165**, 517-522.

van Brussel, A.A.N., Recourt, K., Pees, E., Spaik, H.P., Tak, T., Wijffelman, C.A., Kijne, J.W., and Lugtenberg, B.J.J. (1990). A biovar-specific signal of *Rhizobium leguminosarum* bv. *viciae* induces increased nodulation gene-inducing activity in root exudate of *Vicia sativa* subsp. *nigra*. *J. Bacteriol.* **172**, 5394-5401.

- van Brussel, A.A.N., Bakhuizen, R., Van Spronsen, P.C., Spaink, H.P., Tak, T., Lugtenberg, B.J.J., and Kijne, J.W. (1992). Induction of pre-infection thread structures in the leguminous host plant by mitogenic lipo-oligosaccharides of *Rhizobium*. *Science* **257**, 70-72.
- Vest, G. and Caldwell, B.E. (1972). *Rj4-A* gene conditioning ineffective nodulation in soybean. *Crop Sci.* **12**, 692-694.
- Vieira, J. and Messing, J. (1982) The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**, 259-
- Wang, S.-P. and Stacey, G. (1990). Ammonia regulation of *nod* genes in *Bradyrhizobium japonicum*. *Mol. Gen. Genet.* **223**, 329-331.
- Wang, S.-P. and Stacey, G. (1991). Studies of the *Bradyrhizobium japonicum nodD1* promoter: A repeated structure for the *nod* box. *J. Bacteriol.* **173**, 3356-3365.
- Wittenberg, J., Bergersen, F., Appleby, C. and Turner, G. (1974) Facilitated oxygen diffusion. The role of leghemoglobin in nitrogen fixation by bacteroids isolated from soybean root nodules. *J. Biol. Chem.* **249**, 4057-4066.
- York, W.S., Darvill, A.G., McNeil, M., Stevenson, T.T., and Albersheim, P. (1985). Isolation and characterization of plant cell walls and cell-wall components. *Meth. Enzymol.* **118**, 3-40.
- Young, J.P.W. (1992). Phylogenetic classification of nitrogen-fixing organisms. In *Biological Nitrogen Fixation*, G. Stacey, R.H. Burris, and H.J. Evans, eds. (New York: Chapman and Hall), pp. 43-86.
- Young, J.P.W., Downer, H.L., and Eardley, B.D. (1991). Phylogeny of the phototrophic *Rhizobium* strain BTAi1 by polymerase chain reaction-based sequencing of a 16S rRNA gene segment. *J. Bacteriol.* **173**, 2271-2277.
- Zaat, S.A.J., Wijffelman, C.A., Spaink, H.P., van Brussel, A.A.N., Okker, R.J.H., and Lugtenberg, B.J.J. (1987). Induction of the *nodA* promoter of *Rhizobium leguminosarum* Sym plasmid pRL1JI by plant flavanones and flavones. *J. Bacteriol.* **169**, 198-204.

Appendix 1: Sequence of the DNA region between *nodJ* and *fixR* of *B.japonicum* USDA110. Primers used for sequencing this 3948 bp fragment are indicated.

Appendix 1:

1	TTGACGTGAGTTTCAGCGAAACGGAGAAGTCAAAGCGATGCAGTGTATGATGCGTTCCGC AACTGCACTCAAGTCGCTTTGCCTCTTCAGTTTCGCTACGTCACAGTACTACGCAAGGCG	60
61	GGTAGCTCGCGAGTGACCTGGTCTGTCTCAGAGCAGCTTGTGAAACGGCCCTTGAGAAGCCGC CCATCGAGCGCTCACTGGACCAGACAGTCTCGTCGAACACTTTGCCGGAACCTCTTCGGCG	120
121	GATCGAGCCAGCTCTACGGCCGCCTGTTTCGGCCGATCGGACCTTGCAAAGTCGTGTTGTG CTAGCTCGGTTCGAGATGCCGGCGGACAAGCCGGCTAGCCTGGAACGTTTCAGCACAAACAC	180
181	CTGCCTGGCAGCGCACCGGCTGCGCGTGACTGTACAGATTAGATCGGTGGGACGGCATG GACGGACCGTGCCTGGCCGACGCGCACTGACATGTCTAATCTAGCCACCCTGCCGTAC	240
241	GCCTTCCATGGAGAACGGGATGTTGAGCATAGGACTGAGTGGCGGCCCTTGACAGAATCTA CGGAAGGTACCTCTTGCCTACAACCTCGTATCCTGACTCACCGCCGGAACCTGTCTTAGAT	300
301	TGAAAGCTCTCCGGAGCTGCCGAATACATTTCTTACGATGGCGCTGCCGTACTTGTGCA ACTTTCGAGAGGCCTCGACGGCTTATGTAAAGAAGTGCTACCGCGACGCCATGAACACGT	360
361	GGACGGCCGCGTAATTGCTGCTGTCTGAAGAGGAGCGCCTCAACCGCGTCAAGCAGTCCAA CCTGCCGGCGCATTAACGACGACAGCTTCTCCTCGCGGAGTTGGCGCAGTTCGTACAGTT	420
421	CAAGTTCCCAGCAATTCGATCCGATACTGCCTTTCTACCGCAGGAGTCGAGCTCGGCGA GTTCAAGGGGTCTTAAGCTAGGCTATGACGGAAAGATGGCGTCTCAGCTCGAGCCGCT	480
481	CATCGACCGTATCGCGTTCTACGCGACTGAAGCTTACTGCAAAGCCATGCTCGAGCGCCT GTAGCTGGCATAGCGCAAGATGCGCTGACTTCGAATGACGTTTCGGTACGAGCTCGCGGA	540
541	TTCTGTTTCTCAGCCCCGTTCCGTTGGACCCCAAACCTGTTGCTGCGGCAGCTGTTGGCAG AAGACAAAGAGTCGGGCAAGGCAACCTGGGGTTTGACAACGACGCCGTCGACAACCGTGC	600
601	GGAATTAGGGGCGGAGATCGATCCGTCCGGGTTTCCTTCGTGAATCACCACGAAGTACAT CCTTAATCCCCGGCTCTAGCTAGGCAAGGCCAAAGGAAGCACTTAGTGGTGCTTCATGTA	660
661	GCCGTGAGCGCGTTTGCCATGTCCGGGTTTCGAGCAAAGTCTGGTGTGTTGCGATCGACGGC CGGCACTCGCGCAAACGGTACAGGCCCAAGCTCGTTTCAGACCACAAACGCTAGCTGCCG	720
721	GGGGGCGATTTTCTCTCAGGCTTAATGGCGGTAGGGTCCGGGACCGTAATTGCGCAGCTT CCCCCGCTAAAAGAGAGTCCGAATTACCGCCATCCAGGCCCTGGCATTAAACGCGTCGAA	780
781	GTAAGTTTCCCGGAACACAACCTCTCTTGGGCTGTTCTATCTCGAGACGATCCGATATCTC CATTCAAAGGGCCTTGTGTTGAGAGAACCCGACAAGATAGAGCTCTGCTAGGCTATAGAG	840
841	GGCTGCGGGATGTTTCGACGAATACAAAGTGATGGGGCTTGACCCGTATGGTGATCCCGAT CCGACGCCCTACAAGCTGCTTATGTTTCACTACCCGAACGTGGCATACCACTAGGGCTA	900
901	CGCTATCGCGAGCTCTTCGCGCAGTTCTACGAGTTGTTGGACAGCGGTGGCTATCGCGTT GCGATAGCGCTCGAGAAGCGCGTCAAGATGCTCAACAACCTGTCGCCACCGATAGCGCAA	960
961	CACCTGGATCGAATCGGACCGAGCCCTTTTTCGCGAACATCCAGGTTTCGGCGAAAGGGAATG GTGGACCTAGCTTAGCCTGGTCGGGAAAACGCGTTGTAGGTCCAAGCCGCTTTCCCTTAC	1020
1021	CCATTACCCAAACAGCATCGAGATGTCTCAGCGCTCGCTGCAGGAGCGCTGGAGCGGATC GGTAAGTGGGTTGTCGTAGCTCTACAGTCGCGCAGCGACGTCCTTCGCGACCTCGCCTAG	1080
1081	GTGTTTCACACCCTGCGGCACCATCGTAAAGCGACGGGCATGACACGCCTGTGCCTAGCT CACAAAGTGTGGACGCCGTGGTAGCATTTTCGTGCCCCGTACTGTGCGGACACGGATCGA	1140

Appendix 1 (cont.)

1141	GGCGGGGTGGCCATAATTGCACGCTGAACGGTAAGCTGCTTTATTCCGGGGCTATTTGAC CCGCCCCACCGGTATTAAACGTGCGACTTGCCATTGACGAAATAAGGCCCCGATAAACTG	1200
1201	GACATCTTTGTTCAACCCGCGGCGCACGACGCCGGTTGCGCACTCGGCGCTGCATTAATG CTGTAGAAACAAGTTGGGCGCCGCTGCTGCGGCCAACGCGTGAGCCGCGACGTAATTAC	1260
1261	GTATCAAGCGAGCTAGGCAGGCCAGCGCCGCGGAGCGACTGCCGGACGTTTATTGGGGG CATAGTTCGCTCGATCCGTCCGGTCCGCGCGCTCGCTGACGGCCTGCAAATAACCCCC	1320
1321	CCCGATCTTGGGAACGAGCAAGCCACAGAGCTTGAGCTCAACGCATGGTCAAGGCCACCTC GGGCTAGAACCCTTGCTCGTTCCGGTGTCTCGAACTCGAGTTGCGTACCAGTCCGGTGGAG	1380
1381	GACATTCAACGAAGTGATGACATAGCATCCAGTGCGGCGGACTGGATGGCAAATGGCGCT CTGTAAGTTGCTTCACTACTGTATCGTAGGTCACGCCGCTGACCTACCGTTTACCGCGA	1440
1441	GTAATCGGCTGGGTTCAAGGGCGCTCTGAGTTCGGCCCAAGTGCCTGGTAACCGCAGC CATTAGCCGACCCAAGTCCCCGCGAGACTCAAGCCGGGTGCACGCGAACCATTTGGCGTCG	1500
1501	ATTCTAGCGGACCCGCGGCTGCGGAAAAACAAGGATCGCATCAACGCGGTGGTCAAGAAG TAAGATCGCCTGGGCGCCGACGCCTTTTGTTCCTAGCGTAGTTGCGCCACCAGTTCTTC	1560
1561	CGCGAGACGTATCGCCCCCTTTGCGCCATCTGCATTAGAGGAGGATGCGAGCGAGTTCTTT GCGCTCTGCATAGCGGGGAAACGCGGTAGACGTAATCTCCTCCTACGCTCGCTCAAGAAA	1620
1621	GAGCTCCCCGACGGCACGCGGCGAGCTCCCCTTCATGAACTTCGTAGTTGCGGTGCGTGAA CTCGAGGGGCTGCCGTGCGCGCTCGAGGGGAAGTACTTGAAGCATCAAGCGCACGCACTT	1680
1681	GCCAAGGGTAACGTGCTCGGTGCGATCACGCACGTTGATGGGACGGCTCGGTGCGAGTGA CGGTTCCCATTTGCACGAGCCACGCTAGTGCGTGCAACTACCCTGCCGAGCCGACGCTACT	1740
1741	TGCTGGCCGCTTCTGACGAGGAACGGCCCCAACAAAGTCCGTCCAGATCAGTCGTCAA ACGACCGGCGAAGGACTGCTCCTTGCCGGGTTGTTGTTTCAAGCAGGTCTAGTCAGCAGTT	1800
1801	AAACGTGGTGATCGCATCTGACACGGACTGCACGATCGGCTCGGCATTGTTGTTAAATGA TTTGCACTAGCGTAGACTGTGCCTGACGTGCTAGCCGAGCCGTAACAACAATTTACT	1860
1861	GGTATTTAGAAGAATCGGAAGGCTCGTTGCTTCTTGAACGCATTAATGACGTCCCAGTA CCATAAATCTTCTTAGCCTTCCGAGCAAGCGAAGAACTTGCGTAATTAATGACGGGTCAT	1920
1921	GCCGGGGTTTGTCTTGCGCGATACGGTCTGCAAGATTGGAGTGCGCTGGCTGCGTCATTG CGGCCCCAAACAGAACGCGCTATGCCAGACGTCTTAACCTCACGCGACCGACGAGTAAC	1980
1981	CCGCCGTACGCATCCCTTCATCGTGTTCGCTCCCAATAGCGCCAGATCGGCAGGAAACG GGCGGCATGCGTAGGGAAGTAGCACAAAGCGAGGGTGTATCGCGGTCTAGCCGTCTTTGTC	2040
2041	GTCTGCGAGATCCGCATGGGACATTCCGCCCACTCTAGTATGCGTATCTCACCTGAGCTT CAGACGCTCTAGGCGTACCCTGTAAGGCGGGTGAGATCATACGCATAGAGTGGACTCGAA	2100
2101	TTGAGATCTTGATGCGAATCGAGGGCGAGGCGTCCGCTCGGGTCGCTTTTCGACACAGCC AAGCTCTAGAACTACGCTTAGCTCCCGCTCCGACGAGCCAGCGAAAAGCTGTGTCCG	2160
2161	CTGCTGGACCAGGCCAAGCGTGAAGATCTTGTGAAGGAATTGCGGCTCGTATGGGAGCTT GACGACCTGGTCCGGTTCGCACTTCTAGAACACTTCTTAACGCCAGCAGCATACCCTCGAA	2220
2221	CGGGGCGTTGCGCTGCATCCGCTGCACGCGGCTTGCGGTACAGATAATGTACAAAGCGGA GCCCCGCAAGCGGACGTAGGCGACGTGCGCCGAACGGCAGTGCTATTACATGTTTTCGCT	2280
2281	ACATAGCCATGACTCATGGCGCATGGCCACGCCCATGCGTGCATGTGAGGGGACCAGGCC TGTATCGGTACTGAGTACCGCGTACCGGTGCGGGTACGCACGTACACTCCCCTGGTCCGG	2340

Appendix 1 (cont.)

2341	ATCCGATGCAGAGCGGGGCATCGGGCATCGGGCTTCGGGCAGGTCGATGCAACACCGTTG TAGGCTACGTCTCGCCCCGTAGCCCGTAGCCCGAAGCCCGTCCAGCTACGTTGTGGCAAC	2400
2401	AAGAAAGGACGATCGTATTGAGAGAACGCGGCCCCGTACGGATAGCCCTGCTATGAAGTT TTCTTTCTGCTAGCATAACTCTCTTGCGCGGGCAGTGCTATCGGGACGATACTTCAA	2460
2461	CTACCGATGCAGCTCGCCGGCACCGCGATTTTCAGGCCGTGACGCCGGGCGAGAAGATACG GATGGCTACGTGAGCGGCCGTGGCGCTAAAGTCCGGCACTGCGGCCGCTCTTCTATGC	2520
2521	CGAGACGTCTGTACTGACGTCTCTTGTCACACCAGGAGCGCGGAAAAGGCAAGGCAGAT GCTCTGCAGACATGACTGCAGAGAACAGGTTGGTCCTCGCGCGCTTTTCCGTTCCGTCTA	2580
2581	GGTTAGCGGTTTCGAGCAATGATCGGTTTCGTAGTGTCCCGCGACGTACGGGGTTTGGCGA CCAATCGCCAAGCTCGTTACTAGCCAAGCATCACAGGGCCGCTGCATGCCCCAAACCGCT	2640
2641	CTGTCTGTGGTCACTCGCGGCAGCCTGGCGATTTGCAAAGCAGACAGGACGGACGCTAGC GACAGACACCAGTGAGCGCCGTGCGACCGCTAAACGTTTCGTCTGTCTGCTGCCGTGGATCG	2700
2701	CATTGATTGGCGGGGATCTTGTTATCTTGATGAACCATTACCAATGCCTTTCCAGTCTT GTAATAACCGCCCCCTAGAACAATAAGAACTACTTGTTAAGTGGTTACGGAAGGTGAGAA	2760
2761	CTTCGAACCGGTTGAAGACATTGGCGGAGTGCGAGTAATTTGCGATGACGACATCAATAC GAAGCTTGGCCAACTTCTGTAACCGCCTCACGCTCATTAACGCTACTGCTGTAGTTATG	2820
2821	GCGTTTCGTTTCCAGGACCATTCTTTCCAACATGGTGGAAACAAGCCATCTTTCGATTGCAT CGCAAGCAAAGTCTGGTAAGAAAGTTGTACCACCTTGTTTCGGTAGAAAAGCTAACGTA	2880
2881	CTATCGACCGGACGAGCAGATTTTCAGGGAACGCGATCAACTCGATCAACTGTTCCAAAG GATAGCTGGCCTGCTCGTCTAAAAGTCCCTTGCGCTAGTTGAGCTAGTTGACAAGGTTTC	2940
2941	TCAGAGAGATAGTGACGCTAACACGGTTGTGTGTGATGCCTGCCTGATGTGGCGCTGCGA AGTCTCTCTATCACTGCGATTGTGCCAACACACTACGGACGGACTACACCGCGACGCT	3000
3001	TCAGGAAGCGGAACCGGAAATCTTCCGGAGTATTAAGCCGCGGCCTGAAATTCAGGCCCG AGTCTTTCGCTTTCGCTTTAGAAAGCCCTCATAATTCGGCGCCGACTTTAAGTCCGGGC	3060
3061	GATTGATGCTATCTACCGCGAGCACTTTGAGCCGTACAGCGTAATCGGCATTTCATGTCCG CTAACTACGATAGATGGCGCTCGTGAAACTCGGCATGTGCGATTAGCCGTAAGTACAGGC	3120
3121	GCATGGCAATGGCGAAGATATTATGGCATGCCCCCTTACTGGGCGGACACGGAGCCGCGTT CGTACCGTTACCGCTTCTATAATACCGTACGGGGAATGACCCGCCTGTGCCTCGGCGCAA	3180
3181	GCGGCAGATCTACAATGCCATTGATGAGGCCAGGAGTTTATCCACGCAAAACCTGTGAG CGCCGTCTAGATGTTACGGTAATACTCCGGTCCTCAAATAGGGTGCCTTTTGGACACTC	3240
3241	GGCGTTCTGTGCACGGACAGTGCGCTCGTCTCGAGCAGGTTTCGGTTAAATTTCCCGA CCGCAAGGACACGTGCCTGTACGCGAGCAGGAGCTCGTCAAAGCCAATTTAAAGGGCT	3300
3301	TGTCTTCGCGATTCCAAAGCAGTTCCAGGCGCCTCAAGCGGGCCCATTCACCATCCGCC ACAGAAGCGCTAAGGTTTCGTCAAGGTCCGCGGAGTTTCGCCCCGGTAACGTGGTAGGCGG	3360
3361	CTCGGAGCTGAAGGCGGCTTTTCTGCCCTCACCGAGATGTATCTCCTTGCCCCGTGCGAC GAGCCTCGACTTCCGCCGAAAAGACGGGAGTGGCTCTACATAGAGGAACGGGCGACGCTG	3420
3421	ACGGTGATCCGTTTTCGCCAACGAGCGCTTTTACGCGCTATGCGCGTCTTTTCGCCCCA TGCCACTAGGCAAAAGGCGGTTGCTCGCGAAAATGCGCGATACGCGCAGAAAAGCGGGGT	3480
3481	CGCGTTATAGAATTGATTTGAACGATCCGGGCGGTTGATCCTGATTGAAGACAACCTCG GCGCAATATCTTAAGCTAACTTGCTAGGCCCGGCCAACTAGGACTAAGTTCTGTTGAGC	3540

Appendix 1 (cont.)

3541	CAAGCGCTTATGGCTTCGTGAAAGTCGCAAACCCGCGTGATCGATCTCACCGCCGCTCTT GTTTCGCGAATACCGAAGCACTTTCAGCGTTTGGGCGCACTAGCTAGAGTGGCGCGGAGAA	3600
3601	TACGTCCGATAGCCGTGGAAGTTCTTCTCCGGGCGACGATTTCTGGCCGCGGGCTGTCA ATGCAGGCTATCGGCACCTTCAAGAAGAGGCCCGCTGCTAAAAGACCGCGCCCCGACAGT	3660
3661	TTGCCGCAACGTTCCAGGGGCGAAGAGAACGAAGCCGCGGCCGAATGCGTTCGGCTAAGG AACGGCGTTGCAAGGTCCCCGCTTCTCTTGCTTCGGCGCCGGCTTACGCAAGCCGATTCC	3720
3721	CGCGGCTTCGACAACAGAGCCCTAAAGTGCCGCTGCGTTTTACAGAAGGAGCACGCCTTC GCGCCGAAGCTGTTGTCTCGGGATTTCACGGCGACGCAAAATGTCTTCCTCGTGCGGAAG	3780
3781	GTCTTCGCTTCGGGCGGGCCGCAGACGGAAGCTAAACAACGCTCAATCACGGATGAACAT CAGAAGCGAAGCCCCCGCGCTCTGCCTTCGATTTGTTGCGAGTTAGTGCCTACTTGTA	3840
3841	TTCAGCTCGCCAGCTCATGCTTGGAAGTAGCGCCGCGAAATCCAGTGCCGCTCAGAGCAT AAGTCGAGCGGTTCGAGTACGAACCTTGATCGCGGCGCTTTAGGTCACGGCGAGTCTCGTA	3900
3901	GGCGCGATAGATCAAAAGTGGCCGCCGACGAATCTAACGTGCAAGCTT CCGCGCTATCTAGTTTTACCGGCGGCTGCTTAGATTGCACGTTTCGAA	3948

VITA

Sohail Saad Luka was born in Cairo, Egypt on August 7, 1965. He developed an interest in microbiology during his last years of secondary school and subsequently joined the Botany Department at Ain-Shams University, Cairo, and earned his B. Sc. in microbiology and chemistry in the spring of 1986. He joined the Egyptian National Research Center for a year as part of his civil service, then left Egypt to USA in his pursuit of higher education. In the fall of 1987, he joined the Department of Microbiology at The University of Tennessee, Knoxville, where he joined the laboratory of Dr. Gary Stacey. There, he developed his interest in the process of legume nodulation by bacteria which has been the subject of his studies in the past five years.