Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

•

A thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy in molecular genetics at Massey University

Julie Marie Collins-Emerson

ABSTRACT.

The aim of the project was to characterise the *Rhízobium loti* Nod<sup>-</sup> Tn5 mutant strain, PN233. The Tn5 insertion had been previously localised to a 7.1 kb *Eco* RI chromosomal fragment. This fragment was sub-cloned and a *Bam* HI/*Sal* I endonuclease restriction map for the region was determined. *Hind* III digests were utilised to identify the approximate location of the Tn5 233 insertion and those of four other Tn5 insertions (4016, 4019, 4047 and 4053) in the 7.1 kb region. The 233 mutation was found to map to a 1.45 kb *Sal* I fragment and that of an overlapping 2.8 kb *Bam* HI fragment.

The 7.1 kb Eco RI fragment and a larger 22.7 kb fragment that encompassed this region, had been cloned into pLAFR1. The construct carrying the 22.7 kb fragment (pPN305) was crossed into four *R.1.* bv. trifolii strains, each mutant in one of the four common *nod* genes, A,B,C, and D. The construct was able to complement the *nod*C mutation indicating the presence of a *nod*C gene somewhere on the 22.7 kb region.

The mutations 4047 and 4053 had been found to map to either side of the 233 Tn5 insertion. Both insertions affected nodule formation and were thus included in further plant complementation tests. These experiments involved crossing both the pPN305 and a construct bearing the smaller 7.1 kb Eco RI fragment (pPN25) into the R. loti and R.1. bv. trifolii Tn5 mutants. What was unusual about the results was that, while the 7.1 kb fragment was able to complement the mutations, the larger 22.7 kb fragment which encompasses that region could complement PN4047 and PN4053 but was unable to complement the PN233 mutant.

The 2.8 kb Bam HI and 1.45 kb Sal I fragments, to which the 233 insertion was mapped, and that of an adjacent 1.2 kb Sal I fragment, were sub-cloned and then Bal 31 digested in both orientations to create a series of overlapping fragments. These fragments were then sequenced. The data revealed that the 233 Tn5 had inserted into the R. loti nodC gene. It was determined that the 4047 Tn5 was also located in this gene, slightly upstream of 233, while 4053 had inserted into

j

the 5'-region of nodI which is downstream of nodC. NodA was identified upstream of nodC indicating an arrangement of common nod genes different from the conventional nodABCIJ found in other rhizobia. The promoter for these nod genes, the nod box, was located upstream of the nodA gene.

A particularly puzzling aspect of the results is that, while PN4047 is complemented by both pPN305 and pPN25, PN233, which has an insertion in the same gene, could only be complemented by the smaller fragment carried by the pPN25 construct. To explain this result, it is proposed that PN233 is producing a mutant NodC protein and that this, in combination with doubled copies of a gene or genes present elsewhere on the 22.7 kb fragment, is responsible for interfering with complementation in this mutant. Alternatively, it may be that the imbalance of doubled copies of downstream, co-transcribed genes in the presence of one copy of a functional *nod*C gene causes complementation failure.

ii

#### ACKNOWLEDGEMENTS.

Sincere thanks goes to my chief supervisor Assoc. Prof. Eric Terzaghi for his guidance, patience, enthusiasm and friendship and to Prof. Barry Scott for his advice and assistance and for providing the research topic.

I wish to express my gratitude to the Department of Microbiology and Genetics for their support. Also, many staff members have been generous in their friendship and given of their time in either a technical or advisory capacity, in particular Robert C., Carolyn, Ron, Dawn, Shirley and Trish, while Laura Medhurst and Steven Thomas worked on the computer diagrams. This has been much appreciated.

The encouragement and friendship shown by Mike, Trish, Carolyn, David, Wil, Robert H., Grant, Mark and Chris and many members of the MGU past and present is most valued.

My parents are thanked for their encouragement and support throughout my education (and Dad for his proof reading).

Finally, sincere thanks go to my husband Stephen for his friendship, support and love - a vital ingredient in the success and completion of the project.

TABLE OF CONTENTS.

# page no.

ABSTRACT.		i
ACKNOWLED	GEMENTS.	iii
1.0	Introduction.	1
1.1	Di-nitrogen-fixing plant - microsymbiont associations.	1
1.1.1	Non-leguminous associations	1
1.1.2	Leguminoseae and Rhizobium	2
1.2	The biology of rhizobial nodulation	5
1.2.1	Nodule ontogeny	5
1.2.1.1	Rhizobial colonisation of root hairs and nodule	
	initiation	S
1.2.1.2	Further nodule development,	13
1.3	The genetics of nodulation	21
1.3.1	Nodulation genes	21
1.3.1.1	Common <i>nod</i> genes	21
1.3.1.2	Host-specificity genes	23
1.3.2	The nod box	25
1.3.3	The role of extracellular polysaccharides in	
	nodulation	28
1.3.4	Nitrogen fixation genes	28
1.3.5	Nodulins	29
1.4	Nitrogen fixation - the physical and biochemical	
	environment.	35
1.4.1	The physical environment of nitrogen fixation	35
1.4.2	Leghaemoglobin	35
1.4.3	Nitrogenase.	36
1.5	The project.	37
1.5.1	Background.	37
1.5.2	The project aim	40

2.0	Materials and methods	41
2.1	Media	41
2.1.1	Luria-Bertani broth. (LB)	41
2.1.2	Tryptone yeast extract broth. (TY)	41
2.1.3	YT medium	41
2.1.4	Yeast extract mannitol broth. (YEM)	42
2.1.5	S10-defined medium	42
2.1.6	M9 medium	43
2.1.7	Fahraeus medium. (FM)	44
2.2	Antibiotic stock solutions	45
2.3	Bacterial strains.	45
2.3.1	Maintenance of bacterial cultures	45
2.3.1.1	Short-term storage on plates	45
2.3.1.2	Medium-term storage on agar slopes	45
2.3.1.3	Long-term storage in glycerol	46
2.4	Triparental crosses	46
2.5	Trifolii repens (white clover) and Lotus pedunculatus	
	culture procedures	46
2.5.1	Sterilisation of clover and Lotus seeds	46
2.5.2	Inoculation of seedlings	52
2.6	Microscopy of nodule sections	52
2.6.1	Embedding procedure	52
2.6.2	Sectioning and microscopy	53
2.7	Buffers and solutions used in DNA preparation	54
2.7.1	Hae III buffer - a universal buffer	54
2.7.2	Tris borate EDTA buffer. (TBE)	54
2.7.3	Sequencing Tris borate EDTA buffer	54
2.7.4	TE buffer	54
2.7.5	10% Sodium dodecyl sulphate. (SDS)	55
2.7.6	Loading buffer	55
2.7.7	Standard saline citrate. (SSC)	55
2.8	Preparation of dialysis tubing for DNA purification	55
2.9	DNA preparation procedures	56
2.9.1	DNA extractions	56

2.9.1.1	Phenol and chloroform preparation for DNA extractions.	56
2.9.1.2	Phenol/chloroform extractions of DNA	56
2.9.2	DNA precipitation	57
2.10	Plasmid preparation procedures	57
2.10.1	Large scale plasmid preparation: Alkaline lysis method.	57
2.10.2	Medium scale plasmid preparation	60
2.10.3	Small scale plasmid preparation for E. coli. (Rapid	
	boil method)	61
2.11	DNA restriction endonuclease digests	62
2.12	Preparation of lambda standards	62
2.13	DNA agarose gel electrophoresis	63
2.14	Recovery of DNA from SeaPlaque (low-melting point)	
	agarose gels	65
2.14.1	Electrolution.	65
2.14.2	Phenol/freeze technique.	66
2.15	Ligations	66
2.15.1	CAP-treating vectors for ligation	67
2.15.2	DNA ligations	67
2.15.2.1	Ligation: Method 1	67
2.15.2.2	Ligation: Method 2. (in-gel ligations)	67
2.16	Bal 31 DNA digestion	68
2.17	Preparation of competent cells for transformations	70
2.18	Transformation of E. coli	71
2.18.1	Transformation of E. coli with plasmid DNA	71
2.18.2	Transformation of E. coli with M13 replicative form	
	(RF) DNA	71
2.19	Procedure for making $[^{32}P]$ -labelled probes	71
2.20	Screening of recombinant clones by hybridisation of	
	[ <sup>32</sup> P]-labelled probes to single M13 plaques	74
2.21	Preparation of template DNA and replicative form (RF)	
	DNA	75
2.22	$^{ m 35}$ S-labelling and sequencing of single-stranded (SS)	
	DNA	76
2.23	Polyacrylamide gel electrophoresis	78

. ...

vi

2.23.1	Sequencing reagents	79
2.23.2	Gel preparation	79
2.23.3	Running sequencing gels	80
Results		
3.0	Restriction mapping	81
3.1	Mapping strategy	81
3.2	Tn5 insertion site mapping	86
3.3	Discussion	90
4.0	Plant complementation experiments	93
4.1	Common nod gene complementation tests (Experiment 1).	93
4.2	Complementation of Tn5 mutations (Experiment 2)	99
4.3	Discussion	99
5.0	DNA company of the D lati achilation conce	107
5.0	DNA sequencing of the k. 10th hoditation genes	107
5.1	Fragment sub-cloning.	107
5.2	Del 21 di contre Sal I and Bam HI inserts	107
5.3	Bal 31 digestion of the Sal 1 and Bam HI fragments	113
5.4	Determination of the Tn5 insertion sites	113
5.5	Sequence analysis	115
5.5.1	A region upstream of the nod box demonstrates	
	similarity to the start of <i>nod</i> D	116
5.5.2	Nod box analysis	116
5.5.3	Comparison of the <i>nod</i> A sequences	125
5.5.4	NodC comparisons	132
5.5.5	Comparisons of the 5'-regions of nodI	132
5.5.6	Location of the Tn5 insertion sites	152
5.6	Discussion	155
6.0	Discussion	162
6.1	Organisation of the nod genes	162
6.2	Analysis of the Tn5 insertion sites and their	
	phenotypes	162
6.3	Plant complementation test results	164

vii

6.4	Nod gene products	172
6.5	Reflections on the presence of a putative PN233 nodC	
	product.	178
6.6	The nod box	180
6.7	Further investigation possibilities	187
REFEREN	ICES	190

•

PUBLISHED PAPERS.

LIST OF FIGURES.

1.1	The chemical structure of flavonoids and some	
	phytoalexins	6
1.2	The structure of determinate and indeterminate nodules	14
1.3	The arrangement of nod genes in R. meliloti and R.l. bv.	
	viciae	26
1.4	A map of the nif region of K. pneumoniae	30
1.5	The arrangement of fix and nif genes in R. meliloti and	
	Bradyrhizobium	30
1.6	The regulation of nif and fix genes in Bradyrhizobium and	
	R. meliloti	33
1.7	Nitrogenase turnover during conversion of $N_2$ to $NH_3$	38
3.1	An Eco RI (E) and Hind III (H) restriction map of the nod	
	region of R. loti NZP2037	82
3.2	An agarose gel of Sal I (S), Bam HI (B) and Eco RI (E)	
	single and double digests of pPN354 and pPN301	82
3.3	Sal I and Bam HI restriction maps for the 7.1 kb Eco RI	
	fragment	87
3.4	Tn5 insertion sites in pPN306 as determined from Hind III	
	restriction digests.	91
4.1	Plasmids pPN305 and pPN25	94
4.2	Complementation of a R.1. bv. trifolii nodC mutant by	
	pPN305	96

4.3	Photographs and micrographs of effective and ineffective	
	complementations	101
5.1	The vector pGEM-2.	108
5.2	Plasmids pJ14011 and pJ14012 that bear the 1.45 kb Sal I	
	fragment in opposite orientations	108
5.3	Plasmids pJ28003 and pJ28006 bearing the 2.8 kb Bam HI	
	fragments in opposite orientations	111
5.4	Determination of the orientations of the 1.45 kb Sal I	
	fragments in relation to the 2.8 kb Bam HI fragments	
	using Sph I sites	111
5.5	Plasmids pJ12009 and pJ12010 bearing the 1.2 kb Sal I	
	fragments in opposite orientations	111
5.6	Sequencing strategy for the nod box and nodA 5'-region of	
	R. loti.	117
5.7	Sequencing strategy for the region of DNA carrying the	
	3'-end of the <i>nod</i> A gene, the <i>nod</i> C gene and the 5'-end	
	of the <i>nod</i> I gene	119
5.8	DNA sequence of the R. loti nod box and the 5'-end of	
	nodA	121
5.9	DNA sequence of the R. loti nodA 3'-terminal, nodC and	
	nodI 5'-terminal.	121
5.10	A comparison between R. melíloti $nodD_1$ nucleotide sequence	
	and the region upstream of the nod box in NZP2037	126
5.11	Alignment of Rhizobium and Bradyrhizobium nod box	
	sequences and an associated conserved sequence	126
5.12	A comparison of the <i>nod</i> box sequences using the GAP	
	program	126
5.13	An alignment of R. meliloti and R.l. bv. viciae published	
	nodA sequences using the GAP program	128
5.14	A comparison of the published R. meliloti and R.I. bv.	
	viciae nodA genes and surrounding DNA sequences using	
	the GAP program	130
5.15	An alignment of the beginning of a number of <i>nod</i> A	
	nucleotide sequences	133

.

ix

5.16	A comparison of the start of the <i>nod</i> A sequences using the	
	GAP program	135
5.17	An alignment of the predicted amino acid sequences for the	
	start of the <i>nod</i> A genes using the GAP program	135
5,18	A comparison of the N-terminal of the predicted nodA amino	
	acid sequences using the GAP program	135
5.19	An alignment of the 3'-end of <i>nod</i> A nucleotide sequences	137
5.20	Comparative similarities for the 3'-end of <i>nod</i> A nucleotide	
	sequences as determined by the GAP program	137
5.21	An alignment of the predicted amino acid sequences for the	
	C-terminals of the nodA proteins	139
5.22	A comparison of the predicted amino acid sequences for the	
	C-terminals of the <i>nod</i> A proteins	139
5.23	An alignment of the published <i>nod</i> C nucleotide sequences	
	using the GAP program	141
5.24	A comparison of the <i>nod</i> C nucleotide sequences using the	
	GAP program	146
5.25	An alignment of the predicted amino acid sequences for the	
	nodC proteins	148
5.26	A comparison of the predicted <i>nod</i> C amino acid sequences	
	using the GAP program	150
5.27	An alignment of the nodI nucleotide sequences using the	
	GAP program	150
5.28	Comparative similarities of the 5'-terminal of the nodI	
	nucleotide sequences using the GAP program	153
5.29	An alignment of the predicted <i>nod</i> I N-terminal amino acid	
	sequences.	153
5.30	A comparison of the predicted nodI N-terminal amino acid	
_	sequences using the GAP program	153
5.31	An alignment between R. meliloti nocD <sub>l</sub> and sequences	
	upstream of the R. loti NZP2037 nodA nod box and the	_
	R. 10ti NZP2213 nodB nod box.	158
5.32	A comparison of R. meliloti nodD <sub>1</sub> and nodD-like sequences	
	in R. 10ti NZP2037 and NZP2213	160

A TTG codon in the 233 Tn5 insertion is in-frame with the 6.1 nodC coding region. .... 169 6.2 The proposed structure for the NodC protein. ..... 173 6.3 The alfalfa-specific nodulation signal, NodRm-1. ...... 176 6.4 Some transcription initiation sites for genes downstream from the nod box. ..... 182 6,5 Areas of the R. meliloti nod box protected from DNAse I digestion by NodD binding. ..... 182 6.6 Co-operative binding of the lambda repressor to the operator allows gene expression to be rapidly switched "off" and "on". ..... 182 6.7 A model for nod gene expression in R. loti. ..... 185

## LIST OF TABLES.

1.1	Microsymbionts and their respective hosts	4
1.2	The functions of K. pneumoniae nif genes	32
2.1	Bacterial strains and bacteriophage	47
2.2	Plasmids	50
2.3	Restriction endonucleases	64
3.1	Restriction fragments from Bam HI and Sal I cut pPN354	
	and pPN301 (kb)	84
3.2	Hind III digests of the Tn5 insertions in pPN306 (kb)	89
4.1	Complementation of R.l. bv. trifolii common nod genes	98
4.2	The phenotypes of four other R. loti Tn5 insertions (4016,	
	4019, 4047 and 4053) in the 7.1 kb Eco RI region	98
4.3	Complementation results following the introduction of	
	plasmids into the Tn5 mutants	100
6.1	Complementation of R. loti and R.l. bv. trifolii nod	
	mutations (Experiment 2) including the number and the	
	type of nod genes expected to be expressed	165
6.2	Complementation results of R. loti and R.I. bv. trifolii	
	nod mutations (Experiment 2) assuming PN233 is producing	
	a mutant NodC ( $\check{c}$ ).	171

xi

### 1.0 INTRODUCTION.

Microbes play an important role in the biological cycles of carbon, oxygen, nitrogen and sulphur which are fundamental to life on this planet. While nitrogen is abundant on Earth, most of it is not readily utilisable by plants. Consequently, one of the factors that most limits global agricultural productivity is the availability of fixed nitrogen. The dinitrogen bond is very strong. Consequently, industrial processes directed towards fixing nitrogen require specialised conditions and are energy-intensive. However, many microbes in either a free-living state and/or in a symbiotic relationship with a plant are capable of fixing atmospheric nitrogen by converting it to ammonia via the following chemical conversions.  $Mq^{2+}$ 

# $N_2 + 6e^- + 6H^+ + 12ATP ----->2NH_3 + 12ADP + 12Pi$ NITROGENASE

At present, a vast amount of money and effort is expended on applying industrial nitrogenous fertiliser, which is mainly a product of the Haber process, to agricultural crops. The impact of industrial fertilisers on the environment and the economic reality that nitrogen fertiliser is beyond the reach of many countries, is of worldwide concern. Research into biological nitrogen fixation is therefore of considerable interest, not only academically, but because of the potential applications it may have for crop improvement and productivity.

### 1.1 DINITROGEN-FIXING PLANT - MICROSYMBIONT ASSOCIATIONS.

### 1.1.1 NON-LEGUMINOUS ASSOCIATIONS.

A number of dicotyledonous, non-leguminous plants from phylogenetically unrelated families and genera, form nitrogen-fixing root nodules in symbiotic relationships with endophytes that mainly belong to the order Actinomycetales and to the genus *Rhizobium*. The actinomycete endophytes of non-legumes can be placed in one family Frankiaceae, with a single genus *Frankia* (Becking, 1975). Examples of

some plant genera that have members which bear non-leguminous nodules formed by actinomycete-host symbioses are *Casuarina*, *Myrica*, *Alnus*, *Cerocarpus*, *Coriaria*, *Comptonia*, and *Colletia* (Becking, 1975; Bowes et al., 1977; Callaham et al., 1979). Some non-leguminous, nodule-bearing dicotyledonous hosts capable of symbioses with *Rhizobium* or supposed *Rhizobium* species are *Trema*, *Parasponia*, *Zygophyllum*, *Fagonia*, *Viola* and *Opuntia* (Becking, 1975).

Two morphological types of actinomycetous nodules are known: 1) Alnus-type and 2) Myrica-Casuarina-type (Torrey and Callaham, 1978). Alnus-type nodules are found in representatives of Belutaceae, Elaegnaceae, Rhamnaceae, Coriariaceae and Rosaceae. The nodules comprise modified, often dicotomously-branched roots of arrested growth which usually have a coralloid appearance (Becking, 1975). Myrica/Casuarina-type nodule lobes give rise to a normal root which is negatively geotropic (Torrey and Callaham, 1978).

Actinomycetes enter the host plant via root hair infection. These hairs curl on invasion by the actinomycete (Torrey and Callaham, 1978; Callaham et al., 1979; Lalonde, 1980). The hyphae perforate the root cortical cells by local degradation of the cell walls and penetrate the host cell cytoplasm where it is then surrounded by host plasma membrane and a thick polysaccharide material termed the capsule. The hyphae branch extensively in specific layers of the cortex, penetrating most of the host cytoplasm (Newcomb et al., 1978; Lalonde, 1980). These nodules are highly modified lateral branches, both in their origin and in their development (Becking, 1975; Torrey and Callaham, 1978).

### 1.1.2 LEGUMINOSEAE AND RHIZOBIUM.

Much research effort is focused on nitrogen fixation in leguminous plants, many of which are of immense agricultural significance both as pasture and as food crops, examples being clover, lucerne (alfalfa), peas, beans, soya beans and peanuts. This symbiotic relationship enables them to grow in nitrogen-deficient soils and hence they are also of ecological importance, in that they can be used

to reclaim poor and nitrogen-deficient land. There are about 18,000 species in the family Leguminoseae and over 90 percent of plants in the sub-families Mimosodeae and Papilionoideae bear highly specialised root nodules which provide the appropriate microaerobic conditions necessary for the nitrogen-fixing bacterial symbiont, *Rhizobium* (Vincent, 1982).

Members of the genus *Rhizobium* characteristically invade the roots of leguminous plants and produce root nodules. Taxonomic classification of *Rhizobium* tends to be based on the plant affinity (cross-inoculation group) concept. Until recently, fast (generation time of under 6 hours) and slow (generation time greater than 6 hours) -growing rhizobia were grouped in the same genus *Rhizobium*. However studies of numerical taxonomy, RNA cistron similarities, DNA base ratio determination, nucleic acid hybridisation, immunology, composition of extracellular polysaccharides, carbohydrate utilisation and metabolism, bacteriophage and antibiotic susceptibilities, protein composition and types of intracellular inclusion bodies in bacteroids (Vincent, 1977; Elkan, 1981; Jordan, 1982; Trinick, 1980) have supported a major division of the genus into the fast-growing *Rhizobium* and the slow-growing *Bradyrhizobium* sp. (Buchanan-Wollaston *et al.*, 1980; Jordan, 1982).

Fast-growing R. japonicum isolated from Asian-type soya beans have physiological characteristics similar to other fast-growing rhizobia, but their symbiotic properties are similar to the cowpea miscellany (Stowers and Eaglesham, 1984). These types of rhizobia have been grouped into a new species, R. fredii (Scholla and Elkan, 1984; Sadowsky et al., 1987). Likewise, the fast-growing strains that nodulate Lotus sp. show low DNA:DNA homology (10-15%) both with other Rhizobium sp. and with slow-growing strains able to form symbiotic relationships with Lotus (Crow et al., 1981). These fast-growers have been grouped into a new Rhizobium species, R. loti (Jarvis et al., 1982). R. leguminosarum, R. trifolii and R. phaseoli are now categorised as different R. leguminosarum biovars, i.e. R.1. bv. viciae, R.1. bv. trifolii and R.1. bv. phaseoli. Table 1.1 lists host plants and their corresponding microsymbionts.

Bacterium	host plant
Rhizobium meliloti	lucerne (alfalfa)
R. leguminosarum bv. viciae	pea, vetch
bv. <i>trifolii</i>	clover
bv. phaseoli	bean
Rhizobium fredií	soya bean
Bradøyrhizobium japonicum	soya bean
Rhizobium loti	Lotus
Azorhizobium caulinodans	Sesbania (stem nodules)
Rhizobium NG234	Parasponia (a non-legume)
Bradyrhizobium sp. Parasponia	Parasponia (a non-legume)

Table 1.1: Microsymbionts and their respective host plants .