

Molecular and Functional Studies of Rhizopine Biosynthetic Genes

Chikezie I. Owuama

B.Sc (Hons), M.Sc (NIG); M.Sc (L' Pool)

A thesis submitted for the degree of Doctor of Philosophy in the Faculty of
Agriculture and Natural Resource Sciences at The University of Adelaide

Department of Applied and Molecular Ecology
Waite Agricultural Research Institute
The University of Adelaide

March 2000

TABLE OF CONTENTS

Summary	v
Declaration.....	viii
Acknowledgment	ix
Dedication	x
Abbreviations.....	xi
CHAPTER 1	1
Literature Review	1
1.1. General Introduction	1
1.2. Legumes	2
1.3. Systematics of <i>Rhizobiaceae</i>	3
1.4. Rhizobial Plasmids	7
1.5. Legume-Rhizobia Symbiosis	8
1.5.1 Rhizobial attraction to plants	9
1.5.2 Rhizobial attachment to root hairs	10
1.5.3 Infection thread development.....	11
1.5.4 Nodulation	11
1.5.5 Nodule specific proteins (nodulins) and leghaemoglobins.....	15
1.6. Regulation of <i>nod</i> genes	16
1.7. Nitrogen Fixation.....	21
1.8. Competition Among Rhizobia For Nodulation	27
1.9 Rhizopines in Rhizobia-Plant Interaction	28
1.9.1 Rhizopines	28
1.9.2 Rhizopine synthesis genes	29
1.9.3 Rhizopine catabolic genes	31
1.10 Project Aims and Objectives.....	32
CHAPTER 2	34
Materials and Methods	34
2.1 Bacterial strains, Plasmids and Media	34
2.2 General DNA Techniques	35
2.2.1 Isolation of plasmid DNA	35
2.2.1.1 Alkaline lysis method:	35
2.2.1.2 Plasmid DNA adsorption to matrix method	36
2.2.1.3 Small scale Sym plasmid preparation	36
2.2.2 Genomic DNA preparation	37
2.2.3 Transformation	37

2.2.3.1 Preparation of electro-competent <i>Escherichia coli</i> cells and electroporation	37
2.2.3.2 Electroporation and plating of transformants.....	38
2.2.4 Restriction digests and electrophoresis of DNA	38
2.2.5 Extraction of DNA from agarose gel	38
2.2.6 Cloning of DNA fragments into plasmid Φ	39
2.3 Eckhardt Gel Electrophoresis	39
2.4 High Voltage Paper Electrophoresis (HVPE)	41
2.5 Purification of Bacteroids from Nodules	41
2.6 Preparation of Peribacteroid Membrane	42
2.7 Preparation of Bacteroid Total Envelope Fraction and Cytoplasmic Contents	42
2.8 Southern Hybridisation.....	43
2.8.1 Southern blotting	43
2.8.2 Dot blotting.....	43
2.8.3 Colony hybridisation	44
2.8.4 Digoxigenin (DIG) labeling of DNA.....	44
2.8.5 Prehybridisation, hybridisation and washing conditions	44
2.8.6 Detection of Southern hybridisation using the DIG system	45
2.8.7 Colorimetric detection of DIG labeled DNA.....	45
2.8.8 Chemiluminescent detection of DIG labeled DNA	45
2.8.9 Autoradiography	45
2.8.10 Stripping and reprobing of membranes.....	46
2.9 Sequencing of DNA.....	46
2.9.1 Sequencing Reactions.....	46
2.9.2 Analysis of DNA sequences.....	46
2.10 Protein Analysis.....	47
2.10.1 Antibody preparation	47
2.10.2 Absorption of antibodies.....	47
2.10.3 Polyacrylamide gel electrophoresis (PAGE)	48
2.10.4 Western blotting.....	48
2.11 Electron Microscopy.....	49
2.12 Rhizopine Extraction	50
2.13 Rhizopine Catabolism Tests	50
2.14 Seed Sterilization	51
2.15 Root Extract.....	51
2.16. Mating of bacteria	51
CHAPTER 3	52
Analysis of <i>mosB</i> and <i>mosC</i> homologue gene sequences from.....	52
<i>R. leguminosarum</i> bv. <i>viciae</i> strain 1a.....	52
3.1 Introduction	52
3.2 Results	54
3.2.1 DNA purification and plasmid mapping.....	54
3.2.2 Southern blot with DIG labeling <i>mosB</i> and sub-cloning.....	54
3.2.3 Southern hybridisation of <i>BamH</i> I digested <i>R. leguminosarum</i> bv. <i>viciae</i> strain 1a pSym with DIG labeled <i>mosC</i>	55

3.2.4 Sequencing	56
3.3 Discussion	69
CHAPTER 4	77
Analysis of <i>R. leguminosarum</i> bv. <i>viciae</i> strains for <i>mos</i> gene homologues and complementation of DapA⁻ <i>E. coli</i> with MosA.....	77
4.1 Introduction	77
4.2 Results	81
4.2.1 Screening for a <i>mosA</i> homologue in <i>R. leguminosarum</i> bv. <i>viciae</i> strains.	84
4.2.2 Screening of <i>R. leguminosarum</i> bv. <i>viciae</i> strains for <i>mosB</i> and <i>mosC</i> homologues.....	87
4.2.3 Complementation of DapA ⁻ <i>E. coli</i> with MosA.	90
4.2.4 Sequencing regions upstream and downstream of the <i>mosA</i> homologue.....	91
4.2.5 Comparison of aligned sequences of <i>R. leguminosarum</i> bv. <i>viciae</i> strain 1a <i>mosA</i> homologue and <i>S. meliloti</i> strain L5-30 <i>mosA</i>	95
4.2.6 Screening the R primes (R'')	96
4.2.7 Mutagenesis strategy.....	96
4.2.8 Mutagenisation of the <i>mosA</i> homologue and selection of mutant.....	97
4.2.9 Screening of the <i>mosA</i> homologue mutant.....	97
4.3 Discussion	99
4.3.1 Presence and location of <i>mos</i> homologues in <i>R. leguminosarum</i>	99
4.3.2 MosA/DapA homology	100
4.3.3 Comparison of <i>mosA</i> and <i>mosA</i> homologue sequences	100
4.3.4 DNA sequence adjacent to the <i>mosA</i> homologue.....	103
CHAPTER 5	105
Immunological studies with	105
<i>Sinorhizobium meliloti</i> anti-MosC serum	105
5.1 Introduction	105
5.2 Results	106
5.2.1 Growing of plants, bacterial inoculation and harvesting of nodules	106
5.2.2 Preparation of antibodies	106
5.2.3 Western blot analysis of bacteroid fractions	107
5.2.4 Immunogold labeling and electron microscopy	109
5.3 Discussion	112
CHAPTER 6	115
Physiological and biochemical studies of rhizopine.....	115
6.1 Introduction	115
6.1.1 Effect of exogenously applied rhizopine on root nodulation.....	115
6.1.2 Antibiotic activity of rhizopine catabolic products.....	115
6.1.3 Microaerophilic induction of rhizopine	116
6.2 Results	117

6.2.1 Effect of addition of 3- <i>O</i> -MSI to lucerne roots upon nodulation by <i>S. meliloti</i> strains L5-30, Rm1021 and Rm2129	117
6.2.1.1 Results of the first experiment	119
6.2.1.2 Results of the second experiment.....	121
6.2.2 Antibiotic activity of rhizopine catabolic products	123
6.2.3 Microaerophilic induction of rhizopine	124
6.3 Discussion	125
6.3.1 Effect of exogenous rhizopine on root nodulation	125
6.3.2 Antibiotic activity of catabolised rhizopine	128
6.3.3 Microaerophilic induction of rhizopine synthesis.....	128
CHAPTER 7	129
General Discussion	129
7.1 Introduction	129
7.2 The <i>mosBC</i> Gene Homologues in <i>R. leguminosarum</i> strain 1a	129
7.3 MosA-DapA Hypothesis.....	130
7.4 Rhizopine and Root Nodulation	131
7.5 Rhizopine - A Nitrogen Sink?	133
CHAPTER 8	135
Bibliography	135
APPENDICES.....	171
Appendix 1 Media and Antibiotics.....	172
Appendix 2 Restriction maps of vectors	176
Appendix 2.1 Restriction map of pGEM®-7zf(+)	176
Appendix 2.2 Restriction map of pSUP202.....	177
Appendix 3 - BLAST analysis of <i>R. leguminosarum mosB</i> and <i>mosC</i>.....	178
Appendix 3.1 - BLAST analysis of <i>R. leguminosarum mosB</i>	178
Appendix 3.2 - BLAST analysis of <i>R. leguminosarum mosC</i>	179
Appendix 4 - Tables on rhizopine effect on nodulation.....	180
Appendix 5 Conference presentations and poster*	183

Summary

Rhizopine genes, which confer a competitive advantage on their rhizobial host in the nodulation of legume roots, are present on the symbiotic plasmid in rhizobia. Approximately, 10% of *Rhizobium leguminosarum* bv. *viciae* and *Sinorhizobium meliloti* strains induce the production of 3-*O*-methyl *scyllo*-inosamine in nodules. Strains capable of synthesising rhizopines also catabolise them. In *S. meliloti*, the rhizopine synthesis genes, *mosABC* form an operon structure regulated by the symbiotic nitrogen fixation gene, *nifA*.

In this study, the *mosB* and *mosC* homologues from *R. leguminosarum* bv. *viciae* strain 1a were sequenced and along with the *mosA* homologue shown to be present on the symbiotic plasmid. Sequence analyses of *mosB* and *mosC* homologues in *R. leguminosarum* bv. *viciae* strain 1a revealed 67% and 49% identity with *S. meliloti* strain L5-30 *mosB* and *mosC*, respectively. The hydrophobicity plot of the deduced amino acids of the MosC homologue revealed 12 membrane spanning regions, thus suggesting a possible membrane protein. MosB may be an aminotransferase as it shares a consensus sequence with a super family of secondary metabolic aminotransferases. The *mosB* and *mosC* homologues form an operon structure. This operon appears to be regulated by the NifA and NtrA regulatory system. Similar to the situation with the *mosABC* operon in *S. meliloti* strain L5-30, there is a duplicated promoter region upstream of the *mosBC* operon in *R. leguminosarum* bv. *viciae* strain 1a. Previously, it was shown that the *mosA* homologue is located at least 34 kb away from *mosBC* homologues in strain 1a.

Bahar (1997) showed that the *mosA*, *mosB* and *mosC* homologues were not sufficient to induce rhizopine production in nodules induced by *R. leguminosarum* bv. *viciae*. Nevertheless, in this study, Southern blot analyses of total DNA from *R. leguminosarum* species probed with *mos* homologues from *R. leguminosarum* bv. *viciae* strain 1a revealed that the *mosA*, *mosB* and *mosC* homologues only hybridise to rhizopine strains of *R. leguminosarum* bv. *viciae*.

The *S. meliloti mosA* gene shares 87% DNA homology with the *R. leguminosarum mosA* homologue. Comparison of the *S. meliloti mosA* and *R. leguminosarum mosA* homologue DNA sequence indicates that a point deletion in the *mosA* homologue sequence has resulted in the shortening of its predicted amino acid sequence by 44 amino acids when compared with *S. meliloti* MosA.

The MosA homologue has 37-47% homology with DapA from other bacterial species. The *dapA* gene encodes dihydrodipicolinate synthase which is involved in diaminopimelic acid synthesis. DNA sequence upstream and downstream of the *mosA* homologue indicate that it is co-located with genes associated with the synthesis of diaminopimelic acid, an unusual amino acid important in cell wall synthesis. However, attempts to complement a DapA⁻ mutant of *E. coli* with plasmid borne *mosA* from *S. meliloti* strain L5-30 and the *mosA* homologue from *R. leguminosarum* bv. *viciae* strain 1a were unsuccessful. This may be due to the lack of expression of these genes in *E. coli* or alternatively MosA does not function as DapA. Attempts to mutagenise the *mosA* homologue were incomplete.

Antibodies prepared to *S. meliloti* strain L5-30 MosC fusion protein hybridised to peribacteroid and bacteroid membrane fragments from nodules induced by the Mos⁺ strain L5-30 and Mos⁻ strain Rm1021 indicating that the antibodies were not specific for MosC. The antibodies did not however bind to free living *S. meliloti* strain L5-30 and *R. leguminosarum* strain 1a. Immunogold labeling studies support the location of MosC from L5-30 on the membrane but there was little gold labeling in nodule sections of Rm1021.

Exogenous application of the rhizopine, 3-*O*-methyl *scyllo*-inosamine (3-*O*-MSI) to lucerne rootlets inoculated with approximately 4×10^8 cells of *S. meliloti* showed an initial delay in nodulation followed by an enhanced level of nodulation. The effect was not related to the catabolism of rhizopine.

Catabolic by-products of rhizopine showed no inhibitory activity on *E. coli* and *S. meliloti* strain Rm1021 indicating that the intraspecies competitive nodulation

advantage of rhizopine catabolic strains is not due to antibiotic-like suppression of non-rhizopine strains. Attempts to induce rhizopine production in free-living *Rhizobium* by microaerophilic induction were not successful.

Declaration

This work contains no material which has been accepted for the award of any other degree or diploma in any University or other tertiary institutions, and to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

Signed

Date

15/03/00

Acknowledgment

I wish to acknowledge with gratitude the patience and objectivity of my supervisor, Dr Peter Murphy in our discussions during this study and in reading my thesis. I also acknowledge the contribution of my other supervisor, Prof. Otto Schmidt. I am grateful to John and Keith for their understanding and making the laboratory a wonderful environment for research. I remain grateful to Salubsri for her support during a very difficult period of my study. I also wish to acknowledge the camaraderie in the laboratory by other members of the Plant-Microbe Interaction laboratory, viz., Oka, Vira, Erna, Yuli, and Reza.

I am appreciative of the assistance from Anke, Terry and Gary towards a memorable period in the department. My appreciation is also to Nicole, Sandra, Ben and Anna for their support.

I cannot thank enough my wife, Patience and my children, Chinemere, Chioma, Ihuoma and Chidindu who had to endure my absence and remain in Nigeria for over 4½ years while I toiled in Australia for this program. I am particularly appreciative of my wife's encouragement as well as her effort in shielding me from all the bad news from home throughout the duration of this study.

I thank the Australian government for the Overseas Postgraduate Research Scholarship (OPRS) and the University of Adelaide for the University of Adelaide Scholarship (UAS) which enabled me study for this Ph.D.

Finally, I am immensely grateful to the Almighty God for His grace and infinite mercy in granting me this scholarship, rescuing me from my ordeal in the department of Microbiology and Immunology, and thereafter making it possible for me to complete this programme in a record time.

Dedication

To my late father, Mr. Christian Ibemere Owuama
and to my mother, Mrs Benice A. Owuama.

Abbreviations

(Note not in alphabetical order)

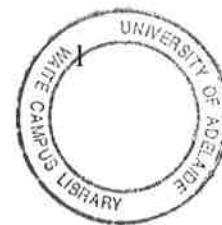
<i>SI</i>	<i>Scyllo</i> -inosamine
3-O-MSI	L-3- <i>O</i> -methyl <i>scyllo</i> -inosamine
sp. (spp.)	Species
bv.	Biovar
cv.	Cultivar
Amp	Ampicillin (sodium salt)
Str	Streptomycin sulphate
Tc	Tetracycline hydrochloride
Rif	Rifampicin
Sp	Spectinomycin
Km	Kanamycin monosulphate
DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
RNase	Ribonuclease
EDTA	Ethylenediamine tetraacetic acid (disodium salt)
MW	Molecular weight
M	Molar
g	Gram
mg	Milligram
µg	Microgram
bp	Base pair
kb	Kilobase pair
L	Litre
µl	Microlitre
w/v	Weight per volume
SDS	Sodium dodecyl sulphate
Tris	Tris (hydroxymethyl) aminomethane
ORF	Open reading frame
UV	Ultraviolet

Chapter 1

Literature review



THE UNIVERSITY OF ADELAIDE



CHAPTER 1

Literature Review

1.1. General Introduction

Nitrogen which makes up 78% of the earth's atmosphere is important in the synthesis of nucleic acid and protein by living organisms. However, plants are unable to utilise molecular nitrogen directly unless in the form of nitrate and ammonium ions or ~~in~~ inorganic compounds such as urea. Electrical discharge from lightning oxidises atmospheric nitrogen to nitrate which is washed down by rain into the soil. Lightning oxidation is responsible for about 20% of soil and water nitrogen (Werner, 1992). Approximately 30% of nitrogen is available to plants from fertilizer. Inorganic fertilizer production is expensive and its application to soil often causes pollution of rivers and streams due to fertilizer run-off. Fortunately, nitrogen-fixing microorganisms, the diazotrophs, are environmentally friendly because they use the nitrogenase enzyme to convert atmospheric nitrogen to ammonia, which can be utilised by plants.

The diazotrophs are found only in Eubacteria and Archaeobacteria (Young, 1992). Free-living diazotrophs fix approximately 40% of biological nitrogen (Gibson, 1990) and commonly associate with plants and in some instances with animals. These free-living diazotrophs include aerobes such as *Azotobacter* and photosynthetic *Rhodospirillum*, and anaerobes like *Clostridium*. Another group of diazotrophs, the cyanobacteria associate with plants. For example, *Anabaena azolla* is found within the heterocyst of the water fern, *Azolla*. However, the relationship between nodule forming bacteria and plants account for the rest of the fixed nitrogen. Among this group is the rhizobia-plant symbiotic association that has been shown to be the most efficient in nitrogen fixation with significant ecological and agricultural importance. In the rhizobia-legume symbiosis, the plant provides a low oxygen environment and ^{C₄-dicarboxylic acids} ~~carbohydrate~~ for the bacterial metabolism while the *Rhizobium* converts the gaseous nitrogen to compounds which can be utilised by the plants.

The importance of legumes was recognised as early as the 18th century, when it was used in crop rotation or cultivated between cereal crops. In the early 19th century, Beijerinck showed that nodules observed on legume roots contain bacteria responsible for the assimilation of atmospheric nitrogen and purified *Rhizobium leguminosarum* from nodules of pea root (Pueppke 1996). Later, Fred *et al.* (1932) demonstrated that legumes increased nitrogen at the expense of atmospheric nitrogen.

1.2. Legumes

The *Leguminosae*, one of the largest families of plants, has a broad geographical distribution. The soybean developed in Asia (Keyse *et al.*, 1982); *Phaseolus* species flourished in both Mesoamerica and the Andean region of South America (Gepts, 1990); *Trifolium*, *Medicago* and *Pisum* species originated from the Middle East (Lie, 1978; Lie *et al.*, 1987); and *Galega officinalis* and *G. orientalis* probably from the Caucasus (Lindström *et al.*, 1995). Bambara groundnut (*Vigna subterranea*), cowpea (*V. unguiculata*) and others were from Africa (Dakora, 1995; Martínez and Caballero-Mellado, 1996). However, older legumes, such as *Acacia*, with about 1500 species are distributed world wide (Martínez and Caballero-Mellado, 1996).

Legumes are now known to establish symbiotic interactions with nitrogen fixing bacteria from the genera *Rhizobium*, *Sinorhizobium*, *Mesorhizobium*, *Bradyrhizobium* and *Azorhizobium*, which form nodules on the root or stems of the plants (van Rhijn and Vanderleyden, 1995). Recently, model legumes, *Medicago truncatula* and *Lotus japonicus* were recommended for genetic studies on nitrogen fixation because they possess attributes that include having a diploid autogamous nature, short generation times and genome sizes which are about three or four times that of *Arabidopsis*. Whereas other legumes such as alfalfa, pea and soybean, are either tetraploid, have large genomes with abundant repetitive DNA, or are difficult to transform and regenerate (Barker *et al.*, 1990; Handberg and Stougaard, 1992). Both *Medicago truncatula* and *Lotus japonicus* can readily be transformed using *Agrobacterium tumefaciens* and then regenerated to yield fertile transgenic plants (Cook *et al.*,

1997). *M. truncatula* exhibits drought and salinity tolerance and grows in a wide range of soil and environmental conditions (Prosperi, 1993). *Medicago truncatula* has a haploid chromosome number of eight and a genome size of approximately 500 mbp/1C and can be efficiently nodulated by ~~Rhizobium~~^{Sinorhizobium} *meliloti* (Bonnin *et al.*, 1996). *Lotus japonicus* has a haploid chromosome number of 6 and a genome size of 450 mbp/1C. ~~Rhizobium~~^{Mesorhizobium} *loti* is the bacterial symbiont of *L. japonicus*. However, an important distinction between *L. japonicus* and *M. truncatula* is that they represent the major classes of nodule ontogeny, determinate and indeterminate development, respectively (Cook *et al.*, 1997).

1.3. Systematics of *Rhizobiaceae*

Bacteria belonging to the family *Rhizobiaceae* are Gram negative, chemolithotrophic or chemoorganotrophic and obligate aerobes. They are rod-shaped, non-spore forming and motile with one polar or subpolar flagellum, and between two and six peritrichous flagella.

The 1984 edition of Bergey's Manual of Systematic Bacteriology classified four genera in the family *Rhizobiaceae* viz., *Rhizobium*, *Bradyrhizobium*, *Agrobacterium* and *Phyllobacterium* (Werner, 1992). More recently, there have been proposals for new genera namely, *Azorhizobium*, *Sinorhizobium* and ~~Phyllobacterium~~^{Meso} *rhizobium* (Werner, 1992). Rhizobia, which encompasses *Rhizobium*, *Azorhizobium* and *Bradyrhizobium* form nodules on the root or stem of legumes where they fix nitrogen. The plant pathogen ~~Agrobacterium~~^{Mesorhizobium} and *Phyllobacterium* form ~~nodules~~^{galls} on leaves of species of plants in the families, *Myrsinaceae* and *Rubiaceae* on which they cause hypertrophies of leaves (Werner, 1992).

Generally, bacterial taxonomy relies on morphological, biochemical and physiological characteristics. Rhizobia however, cannot be distinguished from other soil microorganisms on the basis of cultural and biochemical characteristics, but by their ability to form nodules on leguminous plants. Classification within the genera is still evolving and ~~is~~^{is} primarily based on the host plant infection spectrum. The

classification assumes that each rhizobial species nodulates plants within a particular “cross inoculation” group. Table 1.1 shows some identified nodule forming bacterial species and their host legumes. However, rhizobial taxonomy based on cross inoculation groups was found to be impracticable because of a large genetic diversity.

Table 1.1 *Rhizobium*, *Sinorhizobium*, *Bradyrhizobium* and *Azorhizobium*

Species and their legume hosts (Marínez-Romero and Cabllero-Mellado, 1996)

Species	Principal legume host
<i>S. meliloti</i>	<i>Medicago</i> , <i>Melilotus</i> , <i>Trigonella</i>
<i>R. fredii</i>	<i>Glycine max</i> and <i>G. soja</i>
<i>S. saheli</i>	<i>Acacia</i> and <i>Sesbania</i>
<i>S. teranga</i>	<i>Acacia</i> and <i>Sesbania</i>
<i>R. leguminosarum</i>	
bv <i>viciae</i>	<i>Vicia</i> and <i>Pisum</i>
bv <i>trifolii</i>	<i>Trifolium</i>
bv <i>phaseoli</i>	<i>Phaseolus vulgaris</i>
<i>R. tropici</i>	<i>Phaseolus vulgaris</i> and <i>Leucaena</i>
<i>R. etli</i>	<i>Phaseolus vulgaris</i>
<i>R. galegae</i>	<i>Galega officinalis</i> , <i>G. orientalis</i>
<i>M. ~~R.~~ loti</i>	<i>Lotus</i> spp.
<i>R. huakuii</i>	<i>Astragalus sinucus</i>
<i>R. ciceri</i>	<i>Cicer arietinum</i>
<i>R. tianshanense</i>	<i>Glycine max</i>
<i>B. japonicum</i>	<i>Glycine max</i>
<i>B. elkanii</i>	<i>G. max</i>
<i>A. caulinodans</i>	<i>Sesbsnia rostrata</i>

For example, *R. tropici*, *R. elti* and *R. leguminosarum* bv. *phaseoli* nodulate *Phaseolus vulgaris*; *Acacia* is nodulated by *Bradyrhizobium*, *R. loti*, *R. huakuii* and *Sinorhizobium saheli* (Martínez-Romero and Caballero-Mellado, 1996), and *R.*

meliloti strain 04213 forms nodule on soybean and alfalfa (Gao and Yang, 1995). A majority of *Bradyrhizobium* from Amazonia are promiscuous and have no clear specificity for legumes (Moreira *et al.*, 1993).

Rhizobia are mainly isolated from surface sterilised nodules (Vincent, 1970). Other techniques for direct isolation of rhizobia from the soil demonstrate the existence of non-nodule forming *Rhizobium* strains (Segovia *et al.*, 1991; Gault and Schwinghamer, 1993; Laguerre *et al.*, 1993a), hence the need for a comprehensive approach to rhizobial taxonomy. Recent rhizobia classification tends to consider factors which include host plant, growth rate, bacteriophage susceptibility, DNA-DNA hybridisation, DNA base ratios, serology, multi locus enzyme electrophoresis (MLEE), extracellular polysaccharide composition, RNA cataloguing, 23S and 16S rRNA sequencing, nutritional and metabolic characteristics (Jordan, 1984; Elkan, 1992), analysis of fatty acid profiles (Jarvis and Tighe, 1994), REP- and ERIC-PCR techniques (de Bruijn, 1992).

Analysis of 16S ribosomal sequences reveal that *Rhizobium* and *Bradyrhizobium* are distantly related, but that each relates closely to other groups of bacteria that are not plant symbionts (Sawada *et al.*, 1993; Willems and Collins, 1993; Yanagi and Yamasato, 1993; Young *et al.*, 1991). Fig. 1.1 illustrates the phylogenetic tree of rhizobia and some related bacteria. *Bradyrhizobium* spp., including the phototrophic strains, BTail, are more related to *Rhodopseudomonas*, to *Azotobacter* and to *Blastobacter denitrificans* (Willems and Collins, 1992; Young *et al.*, 1991). *Rhizobium* is related to *Agrobacterium*, to *Brucella*, to *Rochalimea* and to *Bartonella*. *Phyllobacterium* appears to be related to *R. huakuii* and to *R. loti* (Yanagi and Yamasato, 1993). *Agrobacterium* and *Rhizobium* species are very closely related (Sawada *et al.*, 1993; Willems and Collins, 1993; Yanagi and Yamasato, 1993). For example, *R. tropici* and *A. rhizogenes* are remarkably related. However, *R. tropici* has been divided into two subgroups based on phenotypic differences and differences in ribosomal RNA genes (Martínez *et al.*, 1991).

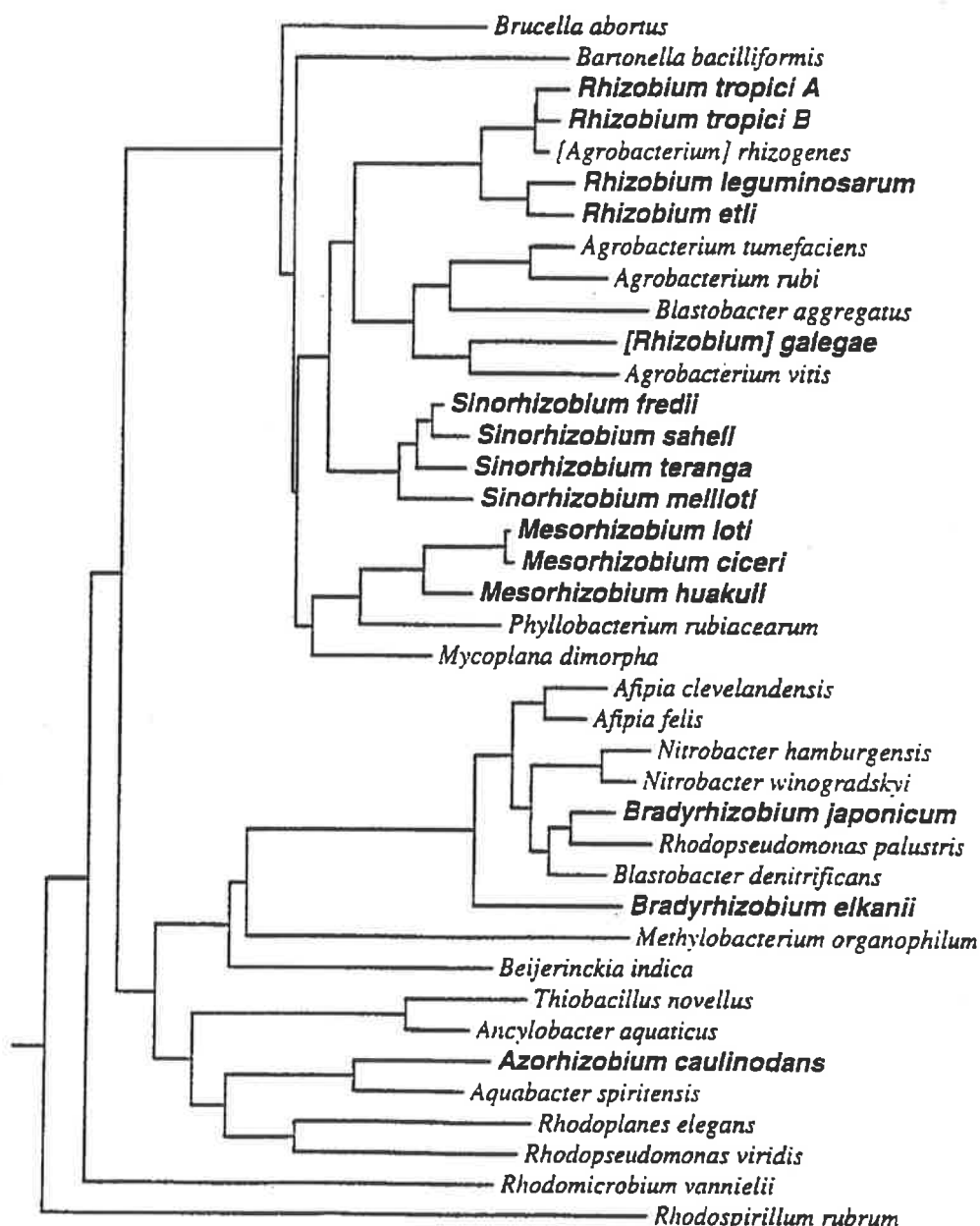


Fig. 1.1 Phylogenetic tree of rhizobia and some related bacteria in the alpha subdivision of the *Proteobacteria* based on 16S rRNA sequence (Young and Haukka, 1996).

DNA-DNA and DNA-rRNA hybridisation, and 16S rRNA sequence comparison reveal a large genetic diversity between strains of *R. fredii* and *R. meliloti* taxonomic groups and other rhizobia. Thus, a new genus, *Sinorhizobium*, was described with four species, *S. meliloti*, *S. fredii*, *S. saheli* and *S. teranga* (de Lajudie *et al.*, 1994;

Rome *et al.*, 1996; Martínez-Romero and Caballero-Mellado, 1996). The use of biochemical and serological differences for the classification of rhizobia have limitations as reflected by the variability of results obtained (Martínez-Romero and Caballero-Mellado, 1996). Classification based only on phenotypic characteristics may distort relationships since few genetic changes could have profound phenotypic effect (Martínez-Romero and Caballero-Mellado, 1996).

Nevertheless, the limitations of the present trend of adopting a polyphasic approach to rhizobial classification, as reflected by mismatches between phenotypic- and genetic-based groupings, seem to cast doubt on its usefulness in classifying rhizobia (Martínez-Romero and Caballero-Mellado, 1996).

1.4. Rhizobial Plasmids

Extra-chromosomal DNA forms an important proportion of the *Rhizobium* genome. Plasmids seem to have more genetic diversity than the chromosomes and range in size from 150 to 1700 Kb (Martínez *et al.*, 1990; Martínez-Romero, 1994; Fischer, 1994). The same types of chromosome are known to harbour different types of plasmids (Laguette *et al.*, 1993). Hence, the symbiotic or pathogenic behaviour of closely related bacteria may depend on the plasmid content. Most symbiotic information in *Rhizobium* is contained on the extra-chromosomal symbiotic (Sym) plasmids (Martínez *et al.*, 1990). However, other plasmids have been shown to be involved in symbiosis (Brom *et al.*, 1992; Hynes and McGregor, 1990; Martínez and Rosenblueth, 1990). In *Agrobacterium*, plasmids are responsible for tumorigenesis. Non-symbiotic rhizobia (Segovia *et al.*, 1991) and non-pathogenic *Agrobacterium* (Kerstens and de Ley, 1984) lack symbiotic and pathogenic plasmids respectively, but are otherwise similar to their symbiotic and pathogenic strains. This may suggest that Sym plasmids are continuously lost and gained among rhizobia (Segovia *et al.*, 1991). Restriction fragment length polymorphism (RFLP) analysis has shown plasmid transfer within two geographically distant populations of *R. galegae* (Kaijalainen and Lindström, 1989). *Rhizobium* and *Agrobacterium* strains readily interchange plasmids under laboratory conditions. Different *Rhizobium* species containing Ti (tumor-inducing) plasmids from *Agrobacterium tumefaciens*

are tumour inducing, although the tumours are relatively small in size (Hooykaas *et al.*, 1977). *Agrobacterium tumefaciens* with symbiotic plasmids from *Rhizobium* form nodules on the corresponding host legume (Hooykaas *et al.*, 1982, 1985; Truchet *et al.*, 1984). The transfer of *R. tropici* Sym plasmid into *A. tumefaciens* lacking this plasmid produced *A. tumefaciens* transconjugants capable of nodulation and fixing nitrogen in bean, although at a reduced level. The transconjugants could also nodulate *Leucaena* (Martínez *et al.*, 1987).

Except for ^{M.}~~B.~~ *loti*, *B. japonicum* and *A. caulinodans*, the genetic information for nodulation and nitrogen fixation in *Rhizobium* spp. is on the plasmids (Fischer, 1994). *S. meliloti* contains two megaplasmids, one carries genes for exopolysaccharide and lipopolysaccharide syntheses (Finan *et al.*, 1986) and the other contains the *nod* and *nif* genes (Mercado-Blanco and Toro, 1996). Similarly, the *nod* and *nif* genes of *R. etli* strain and *R. leguminosarum* bv. *viciae* strain VF39 are located on a different Sym plasmid from those carrying the LPS genes (Brom *et al.*, 1992; Hynes and McGregor, 1990).

The Sym plasmids also carry the non-symbiotic genes for the metabolism of secondary products of the symbiotic relationship that serve as sources of carbon and nitrogen. For example, catabolism genes for trigonelline (Boivin *et al.*, 1990a; b), proline, betaine and stachydrine (Goldman *et al.*, 1991) are located on Sym plasmids. Moreover, the Sym plasmid contains genes for the catabolism and synthesis of rhizopines, which is essential for the survival and competitiveness of rhizopine producing strains (Murphy and Saint, 1992). Other genes identified on Sym plasmids include those for the utilisation of lactose (*lac*), dulcitol (*dul*), and β -hydroxybutyrate (*bhb*), and for the syntheses of thiamine (*thi*), and melanin (Mercado-Blanco and Toro, 1996). Recently, in a landmark study Freiberg and coworkers sequenced the entire Sym plasmid of rhizobia strain NGR234 (Freiberg *et al.*, 1997).

1.5. Legume-Rhizobia Symbiosis

The symbiotic relationship between legumes and rhizobia usually leads to nodule formation. Nodulation involves stages which include, rhizobial attraction and adhesion to the plant root, development of the infection thread and nodule

development (see reviews such as Nap & Bisseling, 1990; Fisher and Long, 1992; Hirsch, 1992; van Rhijn and vanderleyden, 1995; Niner and Hirsch, 1998; Schultze and Kondorosi, 1998). These stages are briefly discussed below.

1.5.1 Rhizobial attraction to plants

Symbiotic association between rhizobia and legumes is mediated by the exchange of diffusible chemical signals such as flavonoids from seedling (Philips, 1992) and Nod factors from the bacteria (Dénarié *et al.*, 1992). Flavonoids are constitutively produced in higher plants, especially legumes under nitrogen-limiting conditions (Coronado *et al.*, 1995), and belong to a group of oxygen-containing heterocyclic compounds with a common 1,3-diphenylpropane ring system (Wong, 1970). Levels of flavonoids, particularly isoflavonoid phytoalexins, can be modulated by biological and environmental stress, including microbial challenge (Van Etten and Pueppke, 1976; Koes *et al.*, 1994). Low concentrations of flavonoids are known to influence polysaccharide production (Dunn *et al.*, 1992; Reuhs *et al.*, 1994), growth (Hartwig *et al.*, 1991), and metabolism in bacteria (Prinsen *et al.*, 1991; Krishnan and Pueppke, 1993). Also, flavonoids act as chemoattractants for rhizobia (Aguilar *et al.*, 1988; Caetano-Anollés *et al.*, 1988; Kape *et al.*, 1991), and induce the expression of nodulation or *nod* genes, which make up the *nod* regulon. *Nod* genes encode proteins whose catalytic activities apparently lead to the production of Nod factors, lipo- oligosaccharides (Carlson *et al.*, 1994; Spaink *et al.*, 1994), which seem to serve as return signals (Pueppke, 1996; Fig. 1.2). Nod factors modulate flavonoid release (Schmidt *et al.*, 1994) and are able to substitute for bacterial cells in inducing certain morphological changes associated with nodulation (Spaink *et al.*, 1991; Truchet *et al.*, 1991; Stokkermans and Peters, 1994). Figure 1.2 shows a schematic representation of signal exchanges between lucerne and *S. meliloti*.

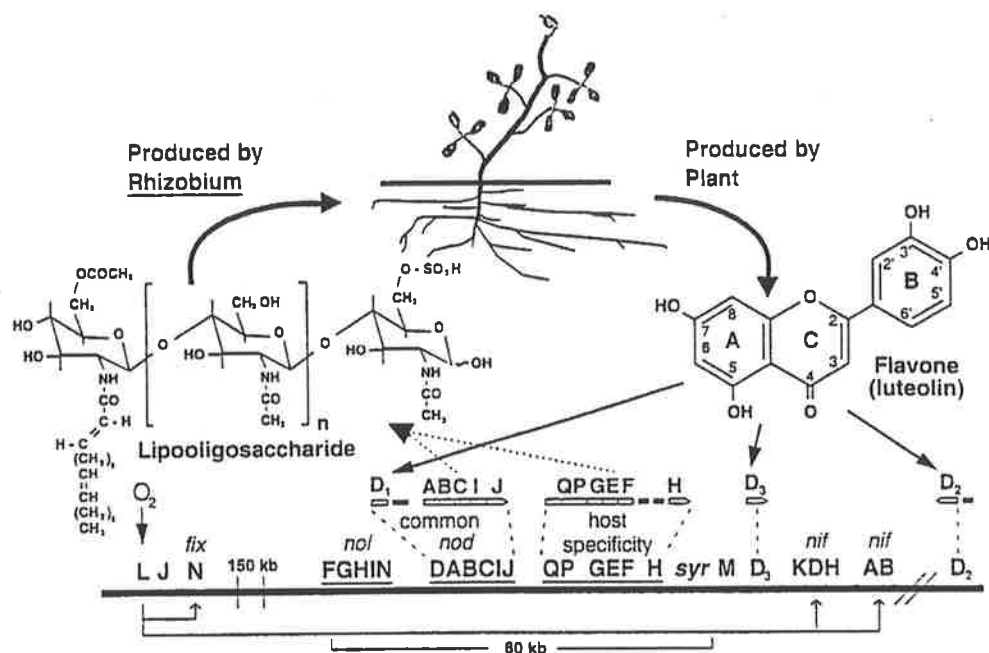


Fig. 1.2 Symbiotic signalling between lucerne and *S. meliloti*. Phenolic flavonoids and isoflavonoids exuded from plant roots bind to the rhizobial *nodD* gene product, which in turn activates transcription of the *nod* genes. The *nod* gene products catalyse the synthesis of lipo-chitooligosaccharides (LCOs), which act to induce root hair curling and plant cell division in root nodule initiation (Vance, 1997).

1.5.2 Rhizobial attachment to root hairs

Rhizobia, in response to the chemoattractants from leguminous plants adhere to the root hair surface. Factors which mediate rhizobial-root interaction are not properly understood (Smit *et al.*, 1992). However, bacteria mostly adsorb to immature root hairs, at points beyond the root tip within the zone of differentiation, apparently with the involvement of bacterial fimbriae and lectins (Vesper and Bouer, 1986; Ho *et al.*, 1990), cellular fibrils (Smit *et al.*, 1987), the calcium binding protein rhicadhesin (Swart *et al.*, 1994), and plant lectins (Dazzo and Truchet, 1983). Subsequently, hydrolysis of the plant cell wall begins at the point of bacterial attachment and the legume root undergoes morphological changes which consist of various types of bending, twinning, twisting, branching and curling (Haack, 1964).

1.5.3 Infection thread development

Rhizobial attempts to penetrate the root hair cell wall lead to the formation of an infection^u thread, which is roughly tubular in form and of plant origin (Callaham and Torrey, 1981). The bacteria multiply within the infection thread which is bound by host cell wall-like material and surrounded by a membrane. The thread conducts rhizobia from the site of penetration to the nodule meristem (Brewin, 1991), which is derived from zones of localized cell division in the cortex (Pueppke, 1996). These zones of cell division in determinate nodules (that lack meristems at maturity) are located in the outer cell layers of legumes whereas in indeterminate nodules, they are found in the deeper layers adjacent to the pericycle of species with persistent meristem activity (Spaink *et al.*, 1991; Truchet *et al.*, 1991; Stokkermans and Peters, 1994), and hence influence cells that form infection^u threads as well as those that receive bacteria from these threads (Pueppke, 1996). However, little is known about the physiology and genetic basis for infection^u thread biogenesis and development (Rae *et al.*, 1992; van Spronson *et al.*, 1994).

1.5.4 Nodulation

Two main types of nodule result from the rhizobial-legume interaction viz., determinate and indeterminate. Determinate nodules such as those formed on bean and soybean are spherical with transient internal meristemic development. Indeterminate nodules formed in legumes such as pea, vetch, alfalfa and clover are elongated, branched, and have protracted meristematic development (Pueppke, 1996). Determinate nodules reach their final size by cell expansion after an initial period of meristematic activity instead of by continued cell division from a distal meristem. In contrast, indeterminate nodules recapitulate in space, the temporal phases of nodule development, yielding four main zones of gradient differentiation from the distal to the proximal end of the nodule (Niner and Hirsch, 1998). Zone I is the actively dividing meristem containing small isodiametric cells free from bacteria. Adjacent to the meristem is zone II made up of cells filled with infection threads containing the invading bacteria. Between zone II and zone III is the interzone II/III filled with amyloplasts. The bacteria are subsequently released into the interzone to colonise expanding host cells after transiting from vegetative to elongated

bacteroids (Vasse *et al.*, 1990). Nitrogen fixation occurs within zone III which is at the central portion of the nodule and contains cells filled with bacteroids and highly vacuolated plant cells devoid of bacteroids. Zone IV, the senescence section is proximal to the site of attachment to the plant root and consists of the oldest tissues of the nodule. The cells in zone IV contain amyloplasts and vegetative rhizobia which are surrounded by plant membrane and are returned to the soil at the death of the nodule (Vasse *et al.*, 1990; Niner and Hirsch, 1998).

Nodule development starts with flavonoids inducing the transcription of rhizobial *nod* genes which code for the proteins that synthesise Nod factors (Fig. 1.3). Usually, NodD binds to the conserved *nod* box sequence and in the presence of the flavonoid inducer, activates *nod* gene expression (Schlaman *et al.*, 1992). However, in *Bradyrhizobium japonicum*, the regulation of *nod* gene expression requires additional two-component regulatory system, NodVW, as well as a MerR-type regulator, NodA (Fig. 1.3).

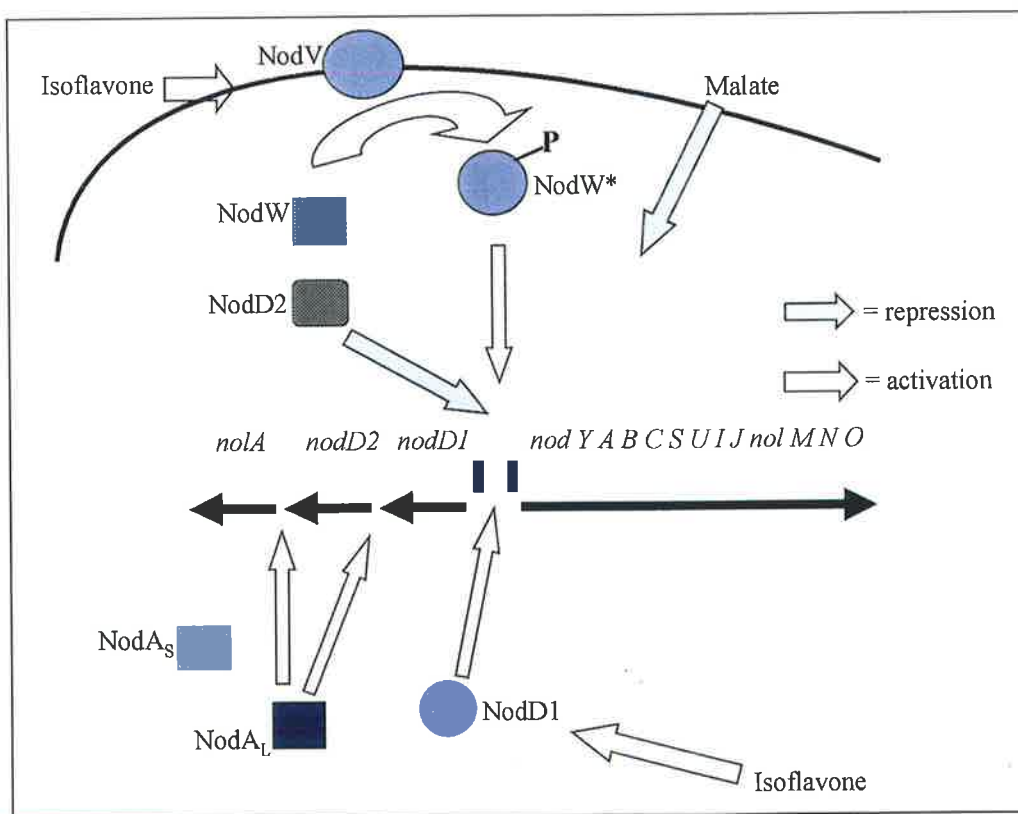


Fig. 1.3 Schematic drawing of the components controlling *nod* gene expression (Loh *et al.*, 1996). In *B. japonicum*

The initiation of infection threads is followed by the division of cortical cells. Rhizobia, such as *S. meliloti* and *R. leguminosarum*, which form indeterminate nodules induce division in inner cortical layers using Nod factors with highly unsaturated fatty acyl groups, unlike *Bradyrhizobium japonicum* which forms determinate nodules (Mylona *et al.*, 1995). Apparently, Nod factors alter the balance of plant hormones, especially the auxin/cytokinin balance that regulates nodule initiation (Spaink, 1995). This may involve inhibiting auxin transport and increasing the cytokinin level, thus eliciting cell division in the infected roots (Pawlowski and Bisseling, 1996; Dénarié *et al.*, 1996). Figure 1.4 describes the working model for indeterminate nodule initiation.

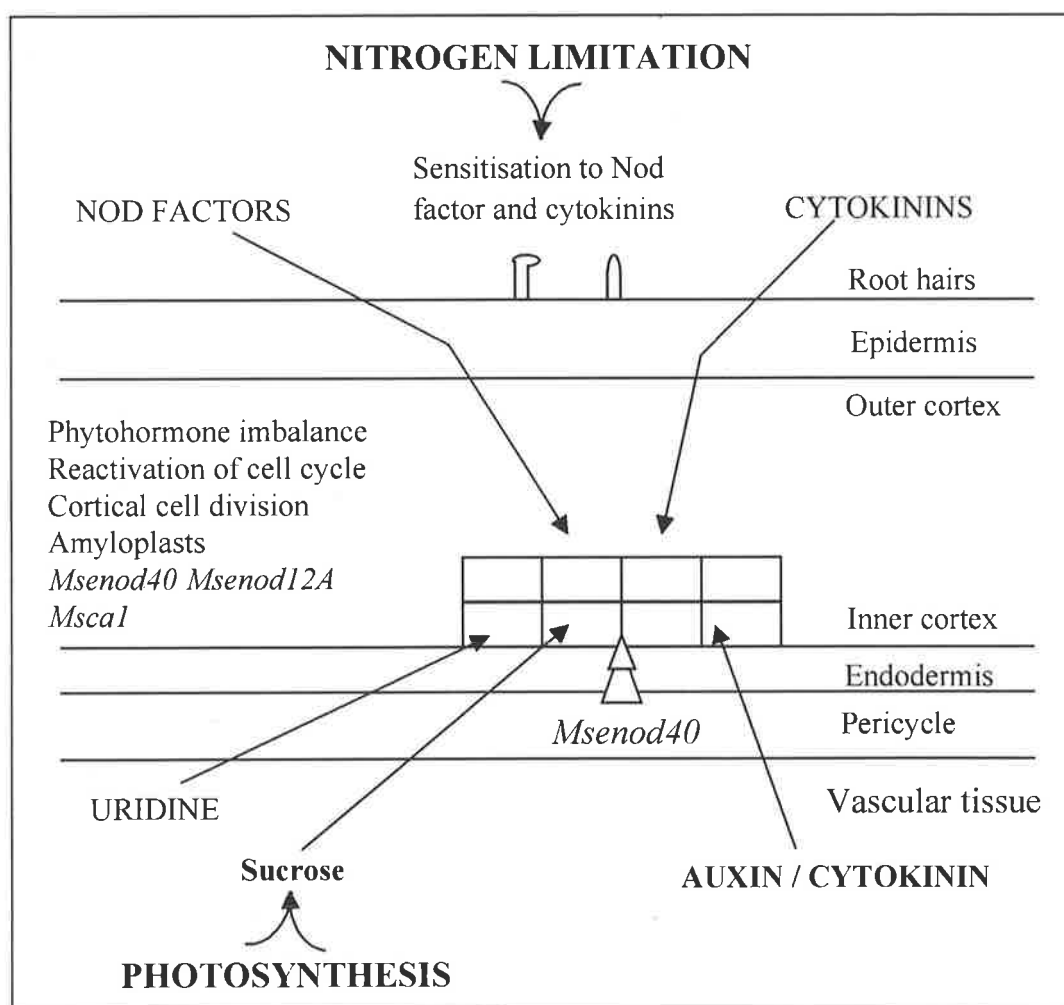


Fig. 1.4 Working model for the control of indeterminate initiation in *Medicago sativa* (Crespi *et al.*, 1996).

It is known that the induction of Nod factors leads to the expression of nodulin genes. For example, the nodulin gene, *enod40* which encodes a protein homologous to a peptide that changes the response of plant cells to auxin in non-legumes, may be involved in inducing nodule primordium formation (Pawlowski and Bisseling, 1996).

When the infection thread reaches the nodule primordium in the inner cortex, the bacteria are released into the primordial cells and rhizobia are surrounded by a plant-derived membrane, the peribacterioid membrane (Pawlowski and Bisseling, 1996). The peribacterioid membrane is important in the control of metabolic exchanges and contains phospholipids and proteins such as nodulins (Mylona *et al.*, 1995). Bacterial propagation ceases within the peribacterioid membrane and they become pleiomorphic, and are referred to as “bacteroids”. The bacteroids differ in shape and size from the free-living rhizobia and possess a modified cell wall which facilitate the exchange of metabolites (Gallon and Chaplin, 1987). When the bacteroids are released into the host cytoplasm, they increase in size and engage actively in nitrogen fixation (Paau *et al.*, 1978).

Nodule organogenesis begins with the formation of nodule primordia in the cortex. The primordia are induced by lipo-oligosaccharides, at positions where the young root emerges, opposite or virtually opposite the protoxylem poles of the central cylinder. This occurs in both the determinate nodule forming plants such as *Glycine soja* and *Lotus prestii* (Lopez-Lare *et al.*, 1995), and indeterminate nodule-forming plants like *Vicia sativa* and *Medicago sativa* (Spaink *et al.*, 1991; Truchet *et al.*, 1991). In *Medicago* and *Glycine soja*, the nodule primordia can develop into full-grown nodules with appropriate anatomical and histological features, but in other plants such as *Vicia*, the nodules stop developing at a stage when small outgrowths are externally visible on the roots (Spaink, 1995). However, little is known about the chemical signals which determine the position of root nodulation. Nevertheless, the presence of uridine, a hormone, has been shown to be essential for lipo-oligosaccharide induction of root cortical cell division (Spaink *et al.*, 1994). Other

compounds such as the auxin transport inhibitor, 2,3,5-triiodobenzoic acid (TIBA) are known to act like lipo-oligosaccharides in inducing the formation of nodule primordia in *Medicago* (Franssen *et al.*, 1992; Hirsch, 1992). Moreover, trans-zeatin, a product of the *tze* gene from *Agrobacterium*, restores nodulation phenotype to mutant strains of *S. meliloti* that are unable to produce lipo-oligosaccharides (Cooper and Long, 1994).

Nodule senescence is apparently an intrinsic progression of indeterminate nodule development (Vasse *et al.*, 1990). Usually, undifferentiated bacteria within the infection thread are recovered from a senescent nodule. Nodule senescence eventually results in host cell and bacteroid lyses, and autophagy of the cytoplasm and organelles (Brewin, 1991). During senescence, the symbiosome (organelle-like structure which encloses rhizobia bacteroids within dividing cortical cells) undergoes physiological stress manifested by the presence of stress proteins in bacteroids and the alteration of lipopolysaccharides (LPS) due to low oxygen or pH (Werner, 1992). Moreover, the bacteroid space contains dicarboxylic acids and protons (generated via an ATPase transport system) while acid hydrolases are present in the peribacteroid fluid (Brewin, 1991). However, active nitrogen-fixing bacteroids in nodules usually increase pH because of their utilisation of dicarboxylic acids for respiration and the excretion of ammonia (Kannenberg and Brewin, 1989).

1.5.5 Nodule specific proteins (nodulins) and leghaemoglobins

Nodule specific proteins or nodulins are proteins that accumulate particularly in nodules (Sanchez *et al.*, 1991). Over 30 nodule specific proteins have been identified by comparing gels of root mRNA and nodule mRNA. About 15 days post infection, 15 nodulins and the four leghaemoglobins (Lba, Lbc, Lbc2, and Lbc3) appear with very high expression of nodulins, N-68 (68 KDa) and N-21. Nodulin expression is induced by a variety of factors. For example, the activity of the *nod* genes and the *hsn* genes of *R. leguminosarum* induce ENOD12 (Scheres *et al.*, 1990). Some other nodulin genes are induced by auxin transport inhibitors (Hirsch *et al.*, 1989). Nodule extracts induce the transcription of nodulin 23, nodulin 24 and leghaemoglobin C3 from nucleic acid of uninfected soybean tissues. The induction

of nodulin 35 occurs at the later stages of nodule development (Mauro and Verma, 1988). However, the functions of only a few nodulins are known. For example, nodulin N-35 is a nodule specific uricase (urate oxidase, uricase II) located in the peroxisomes of uninfected cells. Also present are nodule specific forms of sucrose synthetase and choline kinase (CKII) which are involved in carbon metabolism and architectural restructuring of infected cells, respectively (Verma *et al.*, 1991; Mellor *et al.*, 1986). Nodulin 26 was previously thought to be involved in the transport of dicarboxylic acids (Day, 1992; Brewin *et al.* 1993) but has recently been shown to be an aquaporin (Dean *et al.*, 1999)"

However, some non-nodule specific proteins are enhanced in nodules e.g. glutamine synthetase which plays a role in ammonium assimilation is 10 to 100 times more active in the nodules than in the roots, while phosphoenol pyruvate carboxylase involved in carbon metabolism demonstrates sevenfold more enzyme activity in the nodule than in root tissue (Cullimore *et al.*, 1990; Egli *et al.*, 1991a, b; Cullimore and Bennet, 1992).

Leghaemoglobin is localized in the cytoplasm of the host cell (Werner, 1992). The activation of leghaemoglobin genes in soybean occurs 7 to 8 days post infection, with an apparent increase 4 days later. All four leghaemoglobin genes of soybean contain three introns in codons 32, 68-69 and 103-104, as is the case in other plant globin genes. The order of leghaemoglobin genes activation is Lbc3, Lbc2, Lbc and Lba (Werner, 1992).

1.6. Regulation of *nod* genes

Nodulation genes, *nod*, *nol* and *noe*, seem to primarily encode enzymes involved in the synthesis of Nod factors i.e. lipo-oligosaccharides (also called chito-lipo-oligosaccharides or lipo-chito-oligosaccharides with an oligomeric backbone of β -1,4-linked N-acetylglucosamine [GlcNAc] residues [Fig. 1.5a, 1.5b]) molecules which function as plant growth regulating components (Carlson *et al.*, 1994; Long, 1996). Bacterial *nod* genes encode for proteins involved at the early stages of

nodulation. The *nodD* gene is the only constitutively transcribed *nod* gene in free-living cells. Both *R. leguminosarum* bv. *viciae* and *R. leguminosarum* bv. *trifolii*, have a single copy of *nodD* while *S. meliloti* has four allelic forms, *nodD1*, *nodD2*, *nodD3* and *syrM* (Schlaman *et al.*, 1992). *Nod* genes in rhizobia can be induced by *nodD* gene product and a specific host produced flavonoid (Peters and Verma, 1990).

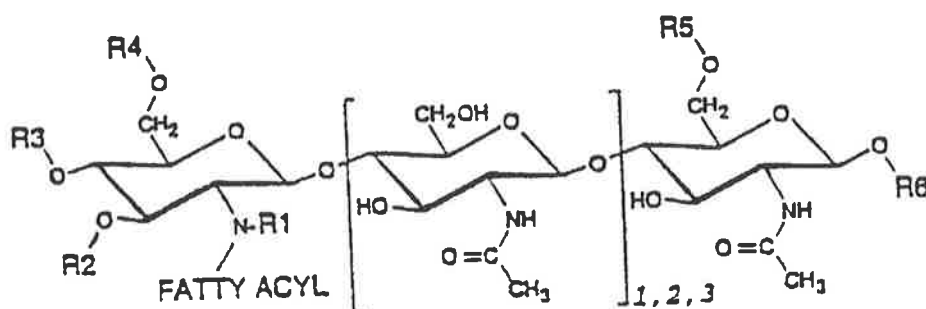


Fig. 1.5a General structure of lipo-chitin oligosaccharide (LCO). The nature of substituents, R, is shown on Fig. 1.5b. In the absence of host specific modification, R1-6 stand for hydrogen groups (Spaink, 1995).

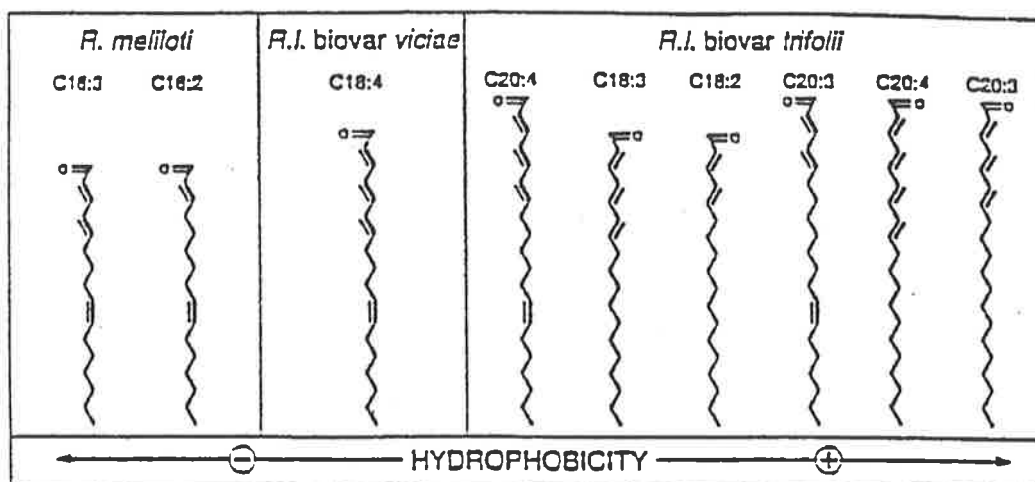


Fig 1.5b Chemical structure of polyunsaturated fatty acyl moieties of LCOs of *S. meliloti* and *R. leguminosarum* biovars *viciae* and *trifolii* (Spaink, 1995).

However, the *nodD* gene products of various bacterial species appear to activate *nod* genes differentially in response to the specific mixtures of inducers derived from the roots of each host plant (Long, 1996). Examples of such inducers are luteolin,

methoxychalcone, naringenin and daidzein, as well as nonflavonoid inducers like trigonelline and stachydrine (Philips *et al.*, 1994). Usually, NodD, a DNA binding protein, binds to the conserved *nod* box sequences preceding most *nod* operons and activates *nod* gene expression in the presence of flavonoid inducers, flavones and flavanones (Schlaman *et al.*, 1992; Perret *et al.*, 1999), while isoflavones and coumarins act as anti-inducers for *Rhizobium* species (Schlaman *et al.*, 1991). The NodA, NodB and NodC proteins are sufficient to produce the basic lipo-oligosaccharide structure. NodC and NodB function as chitin synthase and a chitin deacetylase respectively, producing NodBC intermediate metabolite in lipo-oligosaccharide synthesis (Spaink, 1995) while NodA adds the fatty acyl moiety to the NodBC metabolite (Schmidt *et al.*, 1994; Spaink *et al.*, 1994). The NodE-determined polyunsaturated fatty acyl moieties of *R. leguminosarum* bvs. *viciae* and *trifolii* and *S. meliloti* are important effectors or determinants of host specificity. The NodH-determined sulphate moiety of the lipo-oligosaccharides from *S. meliloti* is required to induce responses in host plants and to prevent activity ^{on} individual non-host plants (Roche *et al.*, 1991; Lerouge *et al.*, 1990). However, the regulation of *nod* gene expression in *B. japonicum*, requires in addition to NodD, a two component regulatory system, NodVW and a Mer-R-type regulator, Nola, [Fig. 1.3] (Loh *et al.*, 1996). In *B. japonicum*, there are two *nodD* genes, *nodD1* and *nodD2*, arranged in tandem (Göttfert *et al.*, 1989). Apparently, *nodD1* is the only functional copy, as a *nodD1D2* deletion mutant is phenotypically complemented by *nodD1* but not by *nodD2* (Göttfert *et al.*, 1992). Unlike the NodD proteins of most *Rhizobium* species, *nodD1* expression in *B. japonicum* requires isoflavones and positively regulates its own expression (Banfalvi *et al.*, 1988). Moreover, glucosylated forms of the isoflavones, genistein and daidzein are known to induce *nodD1* expression (Hungria and Stacey, 1997; Pueppcke, 1996). However, these derivatives do not affect the expression of the common *nod* genes, *nodYABC* in *B. japonicum*. The inducible *nodABC* and *nodFEL* genes are involved in the early stages of nodulation. The products of these genes function in root hair curling, infection thread formation and initiation of cortical cell division. The *nodO* gene from *R. leguminosarum* bv. *viciae* encodes an excreted calcium binding protein which forms ion channels in membranes that allow the movement of monovalent cations across the membrane

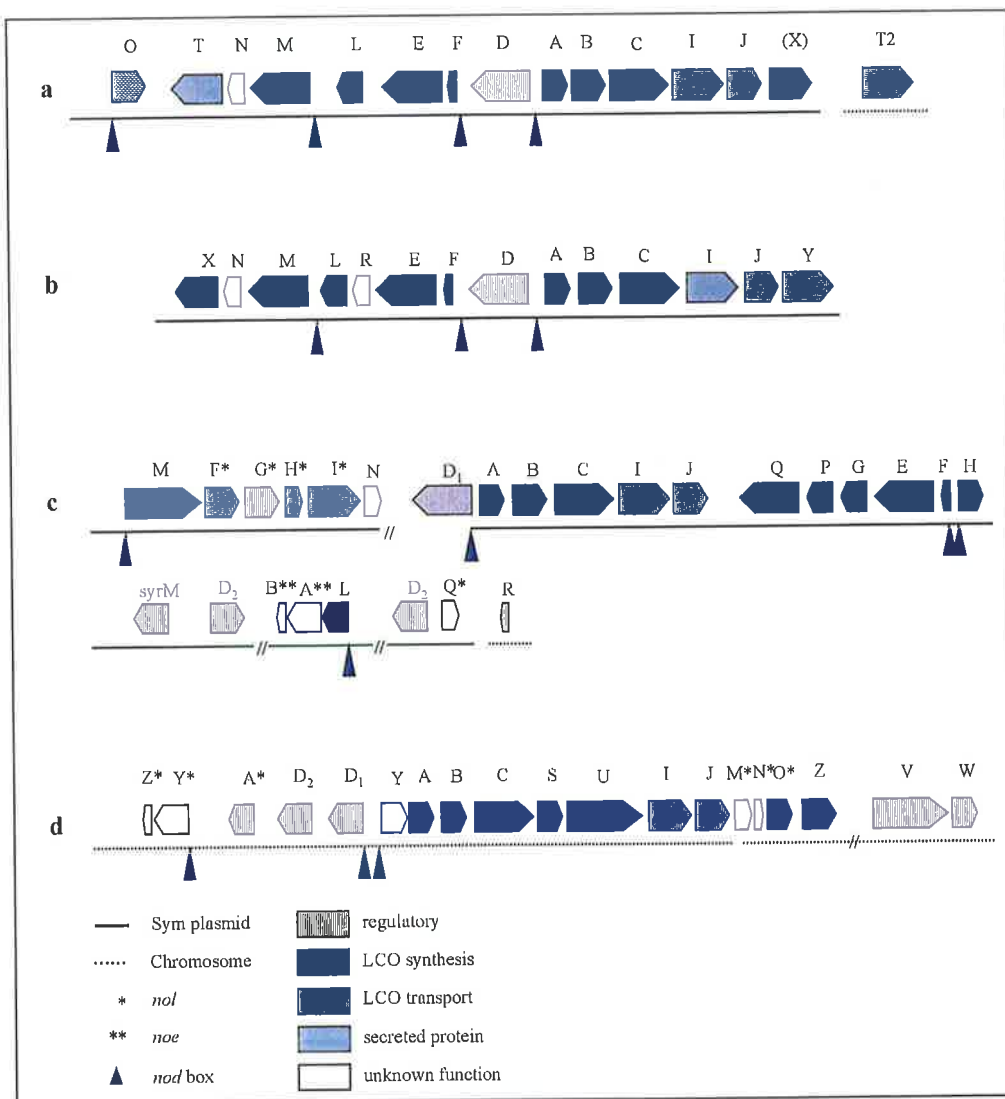


Fig. 1.6 Genetic map and proposed role of *nod*, *nol* and *noe* genes of rhizobia. Arrows depict the open reading frames of the genes of (a) *R. leguminosarum* biovar *viciae* (strain 248, except *nodX*, which is present in strain TOM), (b) *R. leguminosarum* biovar *trifolii* (strain ANU843), (c) *S. meliloti* (strain 2011 and AK41), and (d) *B. japonicum* (strain USDA110). Other rhizobial genes include *nolB*, *nolE*, *nolJ*, *nolU*, *nolV*, *nolW* and *nolX* of *R. fredii* which have been indicated to determine cultivar specificity, *nodK* of *B. elkani* (homologous to *nodY*), *nolC* of *R. leguminosarum* biovar *phaseoli* and *nolP* and *nolK* of *A. caulinodans* (Fisher, 1994).

(Spaink, 1995; Sutton *et al.*, 1994; Downie, 1994; Economou *et al.*, 1994). Other *nod* genes identified in *R. leguminosarum* bv. *viciae* are *nodIJ* and *nodMNT*. Mutations in these genes affect nodulation depending on the host plant (Schlaman *et*

al., 1991). Table 1.2 shows a summary of nodulation genes and their functions while Fig. 1.6 depicts the organisation of nodulation genes in *S. meliloti*, *R. leguminosarum* bv. *viciae*, *R. leguminosarum* bv. *trifolii* and *B. japonicum*.

Table 1.2 Summary of the functions of nodulation genes (Werner, 1992; Carlson *et al.*, 1994; Dénarie *et al.*, 1996; Long, 1996).

Gene	Proposed function of gene products
<i>nodABC</i>	Synthesis of Nod factor backbone, deacetylation, acylation
<i>nodD</i>	Transcriptional regulator of <i>nod</i> genes
<i>nodE</i>	Synthesis of polysaccharide - host specificity determinant
<i>nodF</i>	Synthesis of special acyl moiety - host specificity determinant
<i>nodG</i>	Similar sequence to dehydrogenase - host specificity determinant
<i>nodH</i>	Similar sequence to sulphotransferase - host specificity determinant
<i>nodI</i>	ATP binding protein, transport Nod factors across cell membrane
<i>nodJ</i>	Membrane protein, export of signals
<i>nodK</i>	Unknown function, only present in <i>Bradyrhizobium</i> sp.
<i>nodL</i>	Similar sequence to acetyl transferase, host specificity determinant for <i>R. leguminosarum</i> .
<i>nodM</i>	D-glucosamine synthase for NodRm-1 synthesis
<i>nodN</i>	Biosynthesis of hair deformation factor
<i>nodO</i>	Calcium binding protein
<i>nodP</i>	Similar sequence to ATP sulphurylase - host specificity determinant
<i>nodQ</i>	Similar sequence to ATP sulphurylase, APS kinase - host specificity determinant
<i>nodR</i>	Present only in <i>R. leguminosarum</i> bv. <i>trifolii</i> , host range
<i>nodS</i>	Similar sequence to S-adenosyl methionine methyl transferase - possible host specificity determinant
<i>nodT</i>	Nod factor export, outer membrane protein
<i>nodU</i>	Synthesis of 6-O-carbamyltransferase

<i>nodV</i>	Similar to two-component family sensor - only in <i>B. japonicum</i>
<i>nodW</i>	Similar to two-component family regulator - only in <i>B. japonicum</i>
<i>nodX</i>	TOM sequence similar to acetyl transferase - <i>R. leguminosarum</i> host specificity determinant.
<i>nodY</i>	Unknown function
<i>nodZ</i>	Possibly fucosyl transferase - host specificity determinant
<i>nolA</i>	Similar sequence to MerR-type regulator - only in <i>B. japonicum</i>
<i>nolBC</i>	Present in only certain <i>B. japonicum</i> strains
<i>nolFGHI</i>	Membrane proteins, only in <i>S. meliloti</i>
<i>nolK</i>	Similar sequence to sugar epimerase
<i>syrM</i>	Transcriptional regulator in <i>S. meliloti</i>

1.7. Nitrogen Fixation

Nitrogen fixation by rhizobia takes place in conditions of limiting ammonium and oxygen. Symbiotic nitrogen fixation in the nodule occurs within bacteroids, which reduce atmospheric nitrogen by a series of reactions catalysed by nitrogenase enzymes (the nitrogenase complex) to form ammonia. The nitrogenase complex is present in only prokaryotes and consist of two oxygen sensitive proteins, the MoFe-protein (dinitrogenase or molybdoferrodoxin) and the Fe-protein (dinitrogenase reductase or azoferrodoxin) (Gallon and Chaplin, 1987). Dinitrogenase has α - and β -subunits encoded by *nifD* and *nifK* genes, respectively while dinitrogenase reductase has two identical polypeptides encoded by the *nifH* gene (Ligon, 1990). Dinitrogenase reductase catalyses the transfer of electrons from ferredoxin to dinitrogenase, which has the active site for nitrogen fixation (Kennedy, 1989). The reduction of a molecule of nitrogen to ammonia requires the transfer of electrons by Fe-protein to the active site of nitrogenase and the utilisation of up to 16 ATP molecules (Kennedy, 1989; Hill, 1992).

Because rhizobia require oxygen as terminal electron acceptor (via oxidative phosphorylation) for the synthesis of ATP which provides the high energy essential for nitrogen fixation, legumes and rhizobia evolved certain structural and

biochemical modifications to overcome the sensitivity of the nitrogenase complex to oxygen. The tightly packed cortical cells in the central zone of the nodule act as oxygen diffusion barrier while leghaemoglobin present in nodules bind oxygen to reduce the level of free oxygen. The bound oxygen is subsequently released at controlled low concentration but high flux to bacteroids. Further, rhizobia synthesise a specialised high affinity terminal oxidase essential for aerobic respiratory metabolism in a microaerobic environment (Fischer, 1996).

The rhizobia-legume symbiosis results in the derepression of dinitrogenase, a process apparently optimised by the differentiation of nodule cells (Werner, 1992). Nitrogenase synthesis closely follows the release of bacteria from infection threads although nitrogenase mRNA may be detected earlier (Paau and Brill, 1982). The heme group of leghaemoglobin facilitates the reversible binding of oxygen, thus preventing damage to the oxygen sensitive nitrogenase complex (Hardison, 1996). Haemoglobin production in host cells causes active nodules to turn pink in colour and occurs almost at the same time as nitrogen fixation. Nevertheless, nitrogenase development in *Rhizobium* and *Bradyrhizobium* is accompanied by the repression of glutamine synthetase (GS) thus, enabling bacteroids to reduce nitrogen to NH_4^+ without assimilating the NH_4^+ into organic compounds (Sprent, 1989). Nitrogen compounds synthesised by the plant from NH_4^+ in the nodule are usually exported into the plant cytoplasm and assimilated through the glutamine synthase-glutamate oxoglutarate amino transferase (GS-GOGAT) biosynthetic pathways in roots, leaves and in seeds to produce glutamine and ~~glutamate~~ ^{glutamate} (Gober and Kashket, 1987). The enzyme aspartate aminotransferase (AAT) and asparagine synthase (AS) control the flow of carbon between amino acids and organic acids and catalyse the synthesis of aspartate and asparagine [Fig. 1.7]. However, the form of nitrogen transported within the host plant depends upon the type of plant. For example, nitrogen is transported out of the nodule in the xylem stream in the form of asparagine to other plant organs for nitrogen metabolism in *Pisum*, *Medicago* and *Lupinus* (Mifflin and Cullimore, 1984; Cullimore and Bennett, 1992), and as purine-based compounds in tropical legumes.

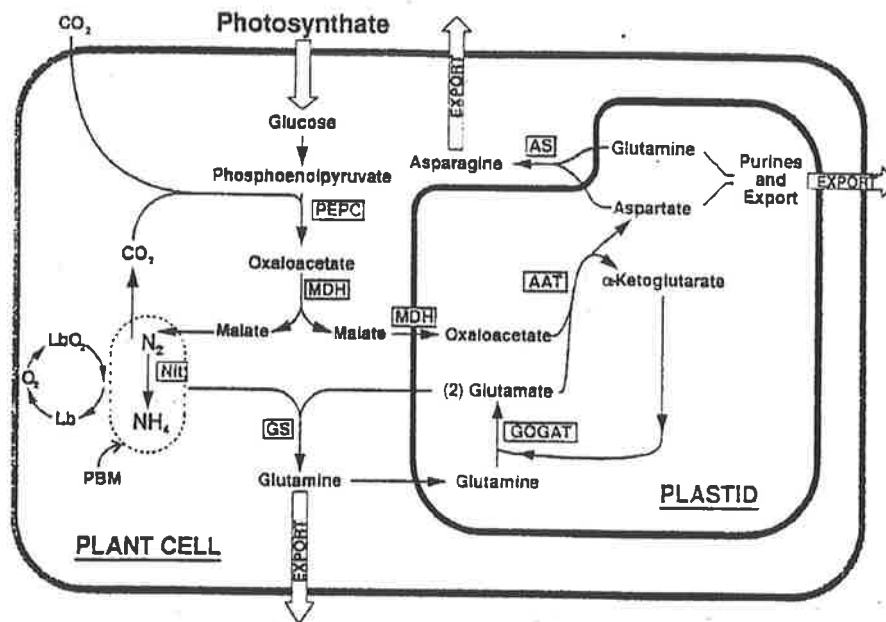
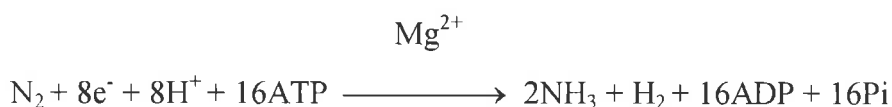


Fig. 1.7 Schematic presentation of general nitrogen assimilation in legume root nodules and the GS/GOGAT/AAT/AS pathway. Participating enzymes are enclosed in blocks: AAT, aspartate aminotransferase; AS, asparagine synthase; GOGAT, glutamate oxoglutarate amino transferase; GS, glutamine synthetase, MDH, malate dehydrogenase, PEPC, phosphoenolpyruvate carboxylase; Nit, nitrogenase. Lb is leghaemoglobin (Vance, 1997).

The general equation for nitrogen fixation in legumes is:



Nitrogen fixation involves mainly the *nif* and *fix* genes. In *Klebsiella pneumoniae*, environmental nitrogen and oxygen influence gene expression mediated via the nitrogen regulatory (*ntr*) system and the *nifLA* operon. The transcription of *nifLA* is activated by *ntrA* gene product, the sigma factor σ^{54} and the *ntrC* gene product (a transcriptional activator) under anaerobic nitrogen-limiting^{ed} conditions. The *nifA* gene product, NifA, activates all other *nif* operons resulting in the synthesis and assembly of functional nitrogenase. However, the presence of oxygen and/or excess nitrogen, alter *nifL* gene product and inactivates NifA, thereby repressing the *nif* operons. This indicates that NifA is a positive regulator that allows the expression

of *nif* genes whereas NifL acts as a negative regulator in repressing nitrogenase synthesis (Vance, 1997). However, the control of nitrogenase system is different in root nodules. The expression of the *nifA* gene product, NifA and its activity depend largely on cellular oxygen conditions and are not controlled by the *ntr* system but are rather regulated by a two component system involving FixL and FixJ ^{in *S. meliloti*}. These function as a sensor-transducer of low oxygen potential in root nodules and activate the transcription of *nifA* and *fixK*, which subsequently activate the transcription of the other *nif* and *fix* operons (Vance, 1997).

Rhizobial *nif* genes and the 20 *K. pneumoniae* ²⁰*nif* genes are structurally homologous and may perform similar functions in both bacterial species (Arnold *et al.*, 1988). About nine different rhizobial *nif* genes are present in *S. meliloti*, *B. japonicum* and *A. caulinodans*. The *fix* genes do not share any homology with *K. pneumoniae* and include genes involved in the development and metabolism of bacteroids (Fischer, 1994). The *nif* and *fix* genes of *S. meliloti*, *B. japonicum* and *A. caulinodans* are in clusters differing in their structure and genomic location within these species. In *K. pneumoniae*, the 20 *nif* genes are in eight operons within ca 24 kb DNA (Arnold *et al.*, 1988). There are two *nif* and *fix* gene clusters, cluster I (*nifHDKE*, *nifN*, *fixABCX*, *nifA*, *nifB*, *frdX*), and cluster II (*fixLJ*, *fixK*, *fixNOQP*, *fixGHIS*) in *S. meliloti* and both are located on the 1400 kb pSym-a or megaplasmid 1 (Fig. 1.6) (Fischer, 1994). Cluster II genes map about 220 kb down stream of *nifHDKE* operon and are transcribed in opposite orientation to it. However, the *fix* and *nif* genes of *B. japonicum* and *A. caulinodans* are located in three (plus *fixK*₁ at a fourth locus) and four (plus *nifB* and *nfrA* in two additional loci) clusters respectively, on the chromosome as both bacteria lack plasmids (Fischer, 1994). Figure 1.8 depicts the organization of *nif* and *fix* gene clusters in *S. meliloti*, *B. japonicum* and *A. caulinodans*.

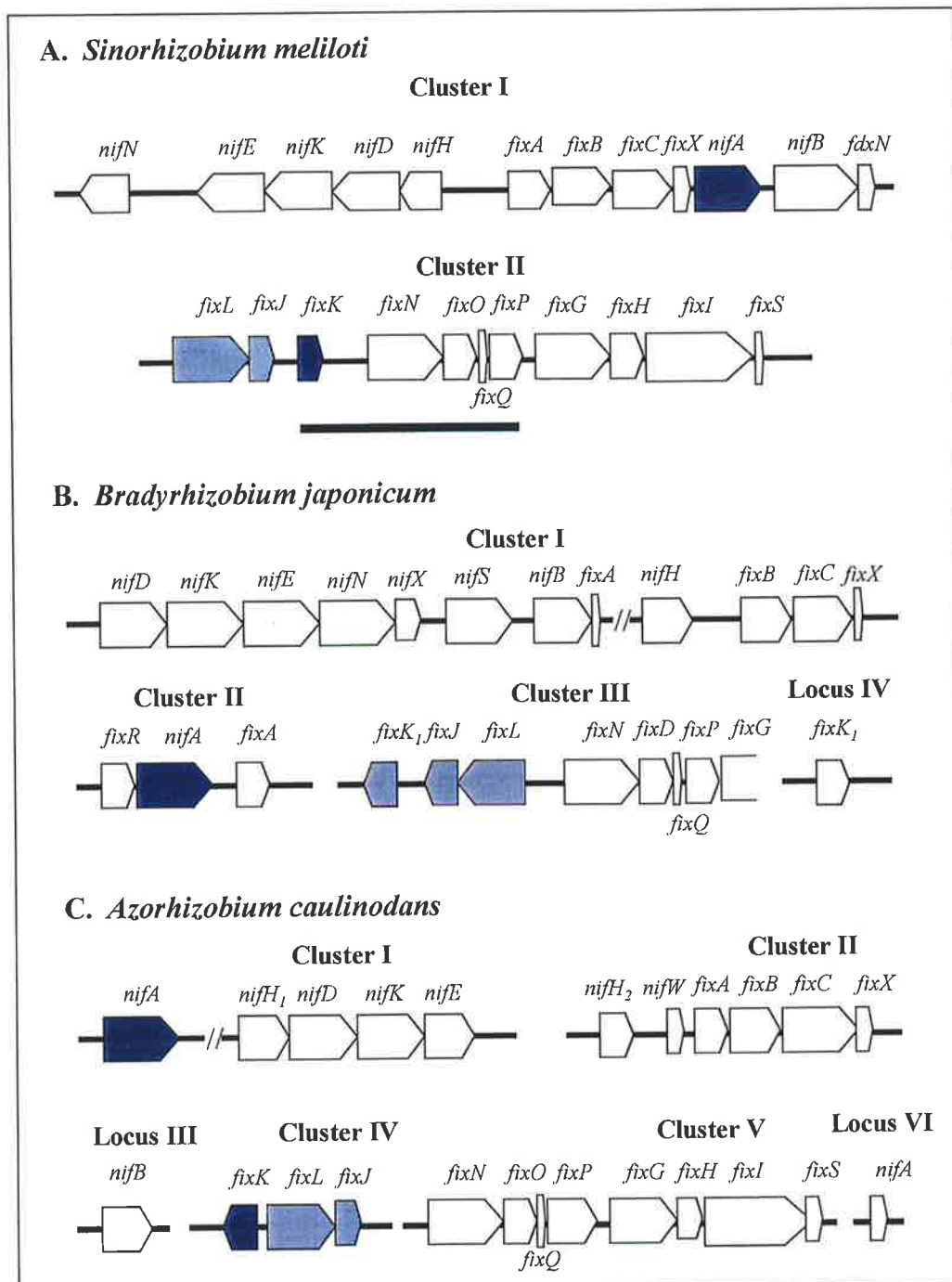


Fig. 1.8 Organisation of *nif* and *fix* gene clusters in *S. meliloti* (A), *B. japonicum* (B), and *A. caulinodans* (C). Several known ORFs associated with the same *nif* or *fix* genes are not shown. Homologous regulatory genes are marked by identical patterns. The hatch bar below cluster II of *S. meliloti* refers to a region which is duplicated on the megaplasmid (Fisher, 1994).

The *nifB*, *nifE* and *nifN* genes encode proteins involved in the biosynthesis of Fe-Mo factors required for nitrogen fixation (Allen *et al.*, 1994). At about 40 kb upstream of the *nifHDK* operon is a functional duplication of the region spanning *fixK* and *fixNOQP* (Renalier *et al.*, 1987). Within the 30 kb region between *nifE* and *nifN*, is a cluster of *nod* genes including *nodABC* (Long, 1989). The 1700 kb pSym-b or megaplasmid 2 and the chromosome contain the additional genes required for symbiosis (Dylan *et al.*, 1986; Honeycutt *et al.*, 1993). The *fixL* and *fixJ* genes activate *nifA* and *fixK* at low oxygen level (David *et al.*, 1988; Batut *et al.*, 1989). The *nifA* gene is a positive regulatory transcription factor (Allen *et al.*, 1994), located upstream of the *nifHDK* gene cluster, that activates the structural *nif* gene operon, *nifHDK* as well as *fixABC*, *nifN* and *nifB* (Mulligan *et al.*, 1988). The *fixK* regulates the transcription of *fixNOQP* (de Phillip *et al.*, 1990). However, NifA and FixK only function at a low oxygen level because of their sensitivity to oxygen (Ligon, 1990). The *nif* and *fix* genes and the functions of their products are summarised in Table 1.3.

Table 1.3 Rhizobia *nif* and *fix* genes and gene product functions (Werner, 1992)

Gene	Gene product and functions
<i>nifA</i>	DNA binding regulatory protein for <i>nif</i> genes and some <i>fix</i> genes transcription.
<i>nifH, D, K</i>	Structural genes for dinitrogenase reductase (<i>nifH</i>) and dinitrogenase (<i>nifD</i> , <i>nifK</i>)
<i>nifB, E, N</i>	Synthesis of the FeMo cofactor of dinitrogenase
<i>nifS</i>	maturation of dinitrogenase
<i>fixL, J</i>	Transcriptional regulation of <i>nifA</i> gene; <i>fixL</i> shares sequence homology with a membrane sensor family, <i>fixJ</i> has similar sequence with transcriptional regulator protein.
<i>fixA, B, C, X</i>	May function in electron transport to nitrogenase.
<i>fixG, H, I, S</i>	May function as a membrane-bound complex coupling a cation pump to redox processes.
<i>fixK, N, R</i>	Unknown function (<i>FixK</i> is regulatory)

1.8. Competition Among Rhizobia For Nodulation

The diversity of rhizobia creates a great potential for competition in legume nodulation. Invariably, rhizobia in the rhizosphere that can rapidly and efficiently initiate infection and nodulation will have competitive advantage over slower ones. Factors such as high-affinity receptors eg. lectin, and host specificity genes enhance the ability of some rhizobia to initiate nodulation (Downing and Broughton, 1986). Inoculation ^{with} efficient commercial nodule eliciting rhizobia often fail to establish effective symbiosis with legumes apparently because of environmental factors such as ^{soil} type, nutrient, moisture, acidity and temperature (Downing and Broughton, 1986). Moreover, competitiveness of the inoculant for nodulation is influenced by the indigenous rhizobial population. When there is a low number of indigenous *Rhizobium* present, the inoculant can establish efficient nodulation but large numbers of a well adapted population usually outcompete the introduced strain in nodule formation (Brookwell and Bottomley, 1995; Triplett and Sandowsky, 1992). For example, introduced *Bradyrhizobium* isolates formed 90% of the nodules when the indigenous population was 5 rhizobia per gram but produced 18% of the nodules in the presence of 700 indigenous rhizobia per gram (Lochner *et al.*, 1989). Biotic factors such as motility (Bauer and Caetano-Anollés, 1992), predatory protozoa (Postma *et al.*, 1990), susceptibility of bacteriophages (Abebe *et al.*, 1992) and toxin production affect the persistence and competitiveness of rhizobia. For example, *R. leguminosarum* bv. *trifolii* strain T24 which produces a bacteriostatic trifolitoxin against some *Rhizobium* strains is highly competitive (Triplett, 1990a, b). The ability of certain strains of rhizobia to utilise specific nutrients such as calystegins, trigonelline, flavones and homoserine also enhance their competitiveness in the rhizosphere (van Egeraat, 1975; Tepfer *et al.*, 1988a, b; Boivin *et al.*, 1991; Hartwig *et al.*, 1991). Also, the catabolism of rhizopines by some rhizobial strains apparently enhance ^stheir competitiveness for nodulation in the rhizosphere.

1.9 Rhizopines in Rhizobia-Plant Interaction

1.9.1 Rhizopines

Rhizopines are inositol-based compounds synthesised in root nodules by bacteroids, the differentiated symbiotic form of infecting rhizobia (Murphy *et al.*, 1987). Two structurally similar rhizopines, L-3-O-methyl-scylo-inosamine (3-O-MSI) and *scyllo*-inosamine (SI) have been identified. They are distinguishable by the presence of a methoxyl group at the C3 position in 3-O-MSI in place of a hydroxyl group in SI (Fig 1.9).

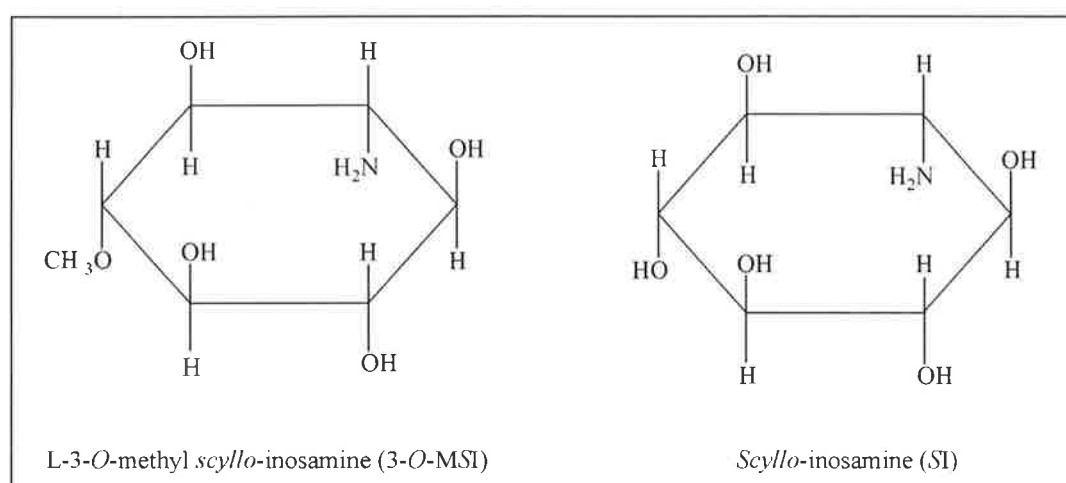


Fig. 1.9 Structure of rhizopines.

Approximately 11% of *S. meliloti* and 12% *R. leguminosarum* bv. *viciae* from 330 *Rhizobium* isolates investigated produced rhizopines. To date, no rhizopine producing strains of *R. elti*, *R. tropici*, *R. leguminosarum* bvs. *trifolii* and *phaseoli* have been identified (Wexler *et al.*, 1995). Except for *S. meliloti* strain Rm220-3 which produces SI, the nodules of other rhizopine strains contain 3-O-MSI (Saint *et al.*, 1993; Rao *et al.*, 1995).

Rhizopines were initially isolated by analogy with opines in the plant pathogenic *Agrobacterium*-crown gall interaction (Tempé and Petit, 1983), and were predicted, as is the case for opines, to provide a selective food substrate for rhizopine producing strains in the rhizosphere and thereby enhance their capacity to nodulate roots of legumes (Murphy and Saint, 1982; Murphy *et al.*, 1995). Rhizopines are

catabolised mainly by the free-living rhizobial population (i.e. those within the nodule infection threads and in the rhizosphere), and as shown recently, by a few other soil inhabiting bacteria (McSpadden Gardener *et al.*, 1998). These soil bacteria belong to five distinct genera, *Aeromonas*, *Alcaligenes*, *Arthrobacter*, *Pseudomonas* and *Sinorhizobium*, and interestingly lack known *mocA*, *mocB* and *mocC* genes of *S. meliloti* L5-30 ^{required} for rhizopine catabolism (McSpadden Gardener *et al.*, 1998). However, rhizopines play a role in intra-species competition for nodulation and seem not to serve solely as a selective growth substrate since they appear very early in the rhizobia-legume interaction (Heinrich, 1999). Intra-species competition studies have shown that rhizopine metabolising strains occupy up to 90% of root nodules (Murphy and Saint, 1992; Murphy *et al.*, 1995; Gordon *et al.*, 1996). Nevertheless, the mode of action of rhizopines has not yet been fully established, but appears not to be involved in direct bacteria-bacteria competition, and may be linked to a response from the plant root-rhizobia interaction during the early stages of nodule initiation (Murphy *et al.*, 1995).

The rhizopine gene complex consists of two components, the synthesis (*mos*) and catabolic (*moc*) genes which are located on the Sym plasmid and seem to have evolved as a functional cassette (Murphy *et al.*, 1987; Saint *et al.*, 1993; Wexler *et al.*, 1995; Wexler *et al.*, 1996). The *mos* genes do not enhance nodulation on their own but work together with the *moc* genes to confer an advantage in competition for nodulation (Gordon *et al.*, 1996). Although the rhizopine *mos* and *moc* genes act in concert, they are expressed at different stages of development (i.e. bacteroids and free-living bacterial states, respectively) of the symbiotic relationship.

1.9.2 Rhizopine synthesis genes

The *mos* locus in *S. meliloti* strain L5-30 contains four open reading frames (ORFs), ORF1, *mosA*, *mosB* and *mosC* arranged in an operon structure (Murphy and Saint, 1992). Fig. 1.10 shows an illustration of the mosaic structure of the *S. meliloti* L5-30 *mos* operon.

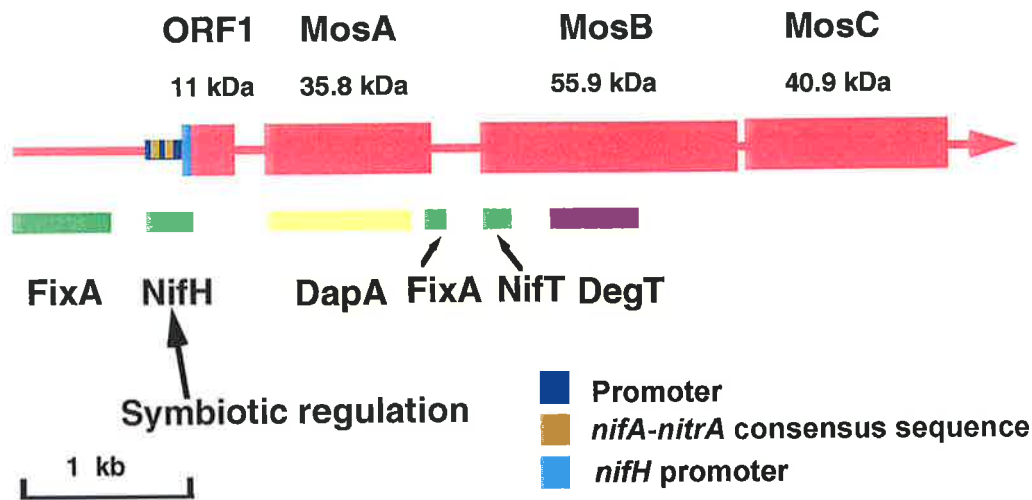


Fig. 1.10 The *mos* locus of *S. meliloti* L5-30 (Murphy and Saint, 1992).

Interspersed within the *mos* locus are regions homologous to *nifT*, *fixU* and *fixA*, thus suggesting a high frequency of rearrangement with symbiotic genes (Murphy *et al.*, 1993; Petes and Hill, 1988; Romero *et al.*, 1991). The *mos* genes are regulated via a promoter with motifs recognised by NtrA and NifA proteins of the NifA/NtrA regulatory system which controls symbiotic genes such as *nif* and *fix* operons. The symbiotic regulation of *mos* genes suggests that they are involved in the symbiotic interaction and function in active nitrogen fixing bacteroids (Murphy *et al.*, 1988; Saint *et al.*, 1993). The ORF1 does not express a functional protein product for rhizopine production and a frame shift mutation in this locus has no effect on 3-*O*-MSI production (Murphy and Saint, 1992; Murphy *et al.*, 1993). However, ORF1 shares a region of homology with the *nfeI* gene of *S. meliloti* GR4, that is part of the *nfe* regulon which is involved in nodulation competition (Soto *et al.*, 1993). Nevertheless, the lack of expression of ORF1 in *S. meliloti* L5-30 creates uncertainty about the role of this region (Murphy *et al.*, 1995). The MosA protein, responsible for 3-*O*-MSI methylation, shows significant homology with DapA, which is involved in the lysine biosynthetic pathway (Rao *et al.*, 1995). MosB was previously considered to have a regulatory role in 3-*O*-MSI biosynthesis, since it shares considerable homology with regulatory proteins such as DegT from *Bacillus stearothermophilus* and DnrJ from *Streptomyces typhimurium* (Murphy *et al.*, 1993). However, recent database analysis suggests that it is an amino transferase

and shares homology with *stsC* of the *sts* gene cluster of *Streptomyces griseus* (Heinrich, 1999). The *mosC* gene has two in-frame start codons coding for *mosC1* and *mosC2*. The amino terminal end of MosC1 contains a 19 amino acid hydrophilic region not present in MosC2. The remainder of the MosC protein is very hydrophobic with 12 membrane spanning domains, and is proposed to be involved in transporting a rhizopine precursor into, and/or rhizopines out of the bacteroid. A point mutation at the start of either *mosC1* or *mosC2* seems not to affect rhizopine production, thus indicating that either of the proteins can be used to synthesise rhizopine (Murphy *et al.*, 1993; Grzowski, 1994).

1.9.3 Rhizopine catabolic genes

The ability to catabolise rhizopines has been shown to confer a nodulation advantage to rhizopine strains over wild type strains lacking in this property (Murphy *et al.*, 1995; Gordon *et al.*, 1996). Early studies reported four rhizopine catabolic genes, *mocA*, *B*, *C* and *R* encompassed within a 15.1 kb DNA fragment in *S. meliloti* (Murphy *et al.*, 1987; Rossbach *et al.*, 1994). The *mocABC* genes are involved in the uptake and breakdown of rhizopine while *mocR* has a regulatory function (Rossbach *et al.*, 1994). Figure 1.11 schematically represents *S. meliloti* MocABCR.

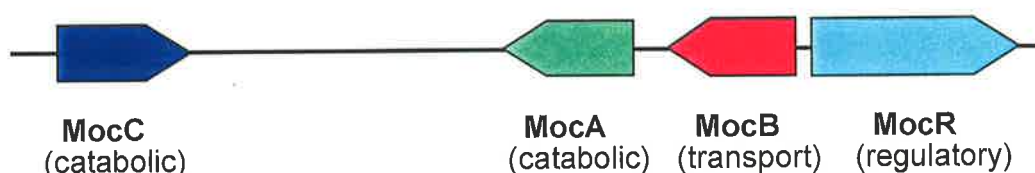


Fig. 1.11 An illustration of *S. meliloti* L5-30 *mocABCR* gene cluster (Rossbach *et al.*, 1994)

MocA contains a $\beta\alpha\beta$ -fold motif that is able to bind NAD(H) factors (Thompson and Donkersloot, 1992) and may therefore be involved in a NADH-dependent dehydrogenase reaction during rhizopine catabolism (Rossbach *et al.*, 1994). The MocB protein is likely a periplasmic binding protein important in rhizopine transport into the bacterial cytoplasm. However, the role of the MocC protein in rhizopine degradation is unknown (Rossbach *et al.*, 1994). The *moc* loci are neither regulated by the symbiotic gene *nifA* nor the common free-living metabolic regulatory genes *ntrC* or *ntrA* (Saint *et al.*, 1993). More recently, a further cluster,

mocDE(F) has been shown to be required for rhizopine catabolism in *R. leguminosarum* bv. *viciae* strain 1a (Bahar *et al.*, 1998)

The *R. leguminosarum* bv. *viciae* *moc* genes are homologous to equivalent genes in *S. meliloti* (Murphy *et al.*, 1995). The catabolism of 3-*O*-MSI in *R. leguminosarum* bv. *viciae* strain 1a requires the catabolic *moc* gene cluster, *mocCABRDE(F)* located on the Sym plasmid (Bahar *et al.*, 1998). The gene products of *mocCABR* resemble those of *S. meliloti*. MocD and MocE contain motifs which correspond to a TOL-like oxygenase and a [2Fe-2S] Rieske-like ferredoxin, respectively. The *mocF* gene encodes a ferredoxin reductase which completes the oxygenase system, although it is not essential for rhizopine catabolism (Bahar *et al.*, 1998). Bahar and co-workers proposed a rhizopine catabolic model whereby MocB transports rhizopine into the cell while MocDE and MocF (or a similar protein encoded by the chromosome), under MocR regulation, act together to form a ferredoxin oxygenase system which demethylates 3-*O*-MSI to *scyllo*-inosamine. The catabolic process is continued by MocA, an NAD(H)-dependent dehydrogenase and MocC, yielding compounds which enter the inositol catabolic pathway (Bahar *et al.*, 1998).

1.10 Project Aims and Objectives

Genes responsible for rhizopine synthesis in *S. meliloti* strain L5-30 have previously been isolated and characterised and have four open reading frames, ORF1, *mosA*, *mosB*, and *mosC* (Murphy and Saint, 1992; Murphy *et al.*, 1993). In the rhizopine producing *R. leguminosarum* bv. *viciae* strain 1a, *mosA* has been characterised and a *mosB* hybridising region identified but attempts to identify regions which hybridise to *mosC* have so far proved unsuccessful.

The work of Bahar (1997) showed that the *mosA* gene from *R. leguminosarum* bv. *viciae* strain 1a has 47% homology to the *dapA* gene (which encodes dihydrodipicolinate synthase which catalyses the condensation of pyruvate and L-aspartate to form dihydrodipicolinate, a precursor in the lysine biosynthetic pathway)

in *E. coli*. This raised a question about the function of the *mosA* homologue in *R. leguminosarum*.

The hydrophobicity pattern of *S. meliloti* MosC suggests that it is a membrane protein (Murphy *et al.*, 1993), however, the location of this protein (i.e. whether it is on the peribacteroid or bacteroid membrane) is still not certain.

Therefore, the aims and objectives of this work are to:

- (a) isolate, sequence and analyse *mosB* and *mosC* homologues in *R. leguminosarum* bv. *viciae* strain 1a,
- (b) establish whether *mosA*, *mosB* and *mosC* homologues are present in only rhizopine strains, and determine their location in the genome of *R. leguminosarum* bv. *viciae*,
- (c) determine whether *mosA* has a similar biochemical function to *dapA* by complementing a *dapA*⁻ *E. coli* with *mosA* from either *R. leguminosarum* bv. *viciae* strain 1a or *S. meliloti* strain L5-30,
- (d) demonstrate whether *mosA* in *R. leguminosarum* bv. *viciae* strain 1a is involved in rhizopine synthesis by mutagenising this gene,
- (e) demonstrate the effect of rhizopine application to roots on nodulation, and
- (f) identify whether *S. meliloti* strain L5-30 MosC is located in the bacteroid and/or peribacteroid membrane in *S. meliloti* strain L5-30 induced lucerne nodule.

Chapter 2

Materials and Methods



THE UNIVERSITY OF ADELAIDE

CHAPTER 2

Materials and Methods

2.1 Bacterial strains, Plasmids and Media

Bacterial strains and plasmids used in this work are described in Tables 2.1. and 2.2 respectively. *Escherichia coli* strains were grown at 37 °C on LB or TY (Sambrook *et al.* 1989; see Appendix 1.1) and when necessary supplemented with appropriate concentrations (in µg/ml) of antibiotics; kanamycin 25, tetracycline 10, spectinomycin 100 and ampicillin 100. Solid media contained 1.5% agar No. 2 (Difco). *Rhizobium* cultures were grown on TY (Beringer, 1974) and when required antibiotic supplements were included at the following concentration (in µg/ml); rifampicin 100, kanamycin 200, streptomycin 200 and tetracycline 10.

Table 2.1 Bacterial strains used.

Bacterial strains	Characteristics	References
<i>Escherichia coli</i> JM109	F' <i>traD</i> 36 <i>lacIq</i> $\Delta(lacZ)M15$ <i>proA⁻B⁻lel4</i> (McrA ⁻) $\Delta(lacproAB)$ <i>thi</i> <i>gyrA96</i> (Nal ^r) <i>endA1</i> <i>hsd17</i> (r ⁻ km ⁺ k) <i>relA1</i> <i>supE44</i> <i>recA1</i>	Yanisch-Perron, <i>et al.</i> 1985
<i>E. coli</i> DH5 α	<i>supE44</i> $\Delta lacU169$ (f80 <i>lacZ</i> Δ M15) <i>hsdR17</i> <i>recA1</i> <i>endA1</i> <i>gyrA96thi-1</i> <i>relA1</i>	Sambrook <i>et al.</i> 1989
<i>E. coli</i> HB101	F' $\Delta(gpt-proA)62$ <i>leu</i> <i>supE44</i> <i>ara14</i> <i>galK2</i> <i>lacY1</i> $\Delta(mcrC-$ <i>mrr)rpsL20</i> (Str ^r) <i>syl-5</i> <i>mtl-1</i> <i>recA13</i>	Boyer and Roulland- Dussoix, 1969
<i>S. meliloti</i> L5-30 Rm1021	Wild type strain, Mos ⁺ , Moc ⁺ Wild type strain, Mos ⁻ , Moc ⁻	Kowalski, 1967 Meade, <i>et al.</i> , 1982

Table 2.2 List of plasmids and cosmids used

Plasmids/ Cosmids	Description	References
pGEM7Zf(+)	Amp ^r cloning and sequencing vector	Promega
pBluescript® KS (-)	Amp ^r cloning and sequencing vector	Stratagene
pUCD2608	Km ^r , broad host range Inc W cosmid	Rogowski (unpublished)
pKT231	Km ^r , broad host range Inc P-1 cosmid	~~Bahr et al., 1998~~ Bagdasarian et al., 1981

2.2 General DNA Techniques

2.2.1 Isolation of plasmid DNA

2.2.1.1 Alkaline lysis method:

Small scale plasmid preparation was done by the alkaline lysis method (Sambrook *et al.*, 1989) as described in the Promega Technical Manual. For plasmid preparations used in sequencing, the DNA was treated with 13% polyethylene glycol (PEG 8000) as described in the manual. Usually 10 ml of LB medium containing the appropriate antibiotic was inoculated with a single bacterial colony and incubated at 37°C overnight with vigorous shaking. 1.5 ml of the overnight culture was spun at 12,000 × g for 3 min and the pellet resuspended in 100 µl lysis buffer (25 mM Tris-HCl, pH 8.0, 10 mM EDTA, 50 mM glucose) and maintained at room temperature for 5 min. 200 µl freshly prepared 0.1 N NaOH, 1% SDS, was added to the suspension, mixed by inversion and incubated in ice for 5 min. Then, 150 µl 3 M potassium acetate, pH 4.8, was added, inverted several times to mix and incubated in ice for 5 min. The suspension was centrifuged at 12,000 × g for 5 min to sediment the precipitate. The supernatant was transferred to a fresh tube and DNase-free RNase A was added to a final concentration of 20 µg/ml and then incubated at 37°C for 45 min. Thereafter, the solution was extracted once with phenol/chloroform/isoamyl mixture (ratio 25:24:1), and once with chloroform/isoamyl mixture (ratio 24:1). The plasmid DNA was precipitated by adding cold (−20°C) absolute ethanol, incubated

at -70°C for 10 min and the spun for 5 min at $12,000 \times g$. The DNA precipitate was washed once with 1 ml cold (-20°C) 70% ethanol, dried in SpeedVac SVC 100 (Selby Scientific Instruments), resuspended in 10-20 μl MilliQ H_2O and stored at -20°C . However, for DNA preparations for sequencing, the dry DNA pellet was resuspended in 16 μl sterile MilliQ H_2O , 4 μl 4 M NaCl was added, mixed, then 20 μl 13% PEG 8000 was added, mixed well and incubated on ice for 20 min. The mixture was centrifuged for 7 min at $12,000 \times g$, the pellet washed once with prechilled 70% ethanol, dried under vacuum, resuspended in 10-20 μl sterile MilliQ H_2O and stored at -20°C .

2.2.1.2 Plasmid DNA adsorption to matrix method

Quantum Prep® plasmid miniprep kit (BIO-RAD) was used in accordance with the instructions of the manufacturer to prepare plasmids. The kit utilises the silicon dioxide exoskeleton of diatoms as the matrix for binding DNA.

2.2.1.3 Small scale Sym plasmid preparation

The Sym plasmid from *Rhizobium leguminosarum* species was prepared by a modified method for Ti plasmid purification from *Agrobacterium* described by Zhang *et al.* (1993). Briefly, 10 ml bacteria^e/culture grown in TY shaking overnight at 28°C was centrifuged at $12,000 \times g$ for 2 min and the pellet resuspended in 1 ml TE buffer pH 8. 100 μl 5M NaCl and 20 μl 10% sodium sarkosyl were added and gently mixed by inversion and spun at $12,000 \times g$ for 2 min. The sediment was resuspended in 200 μl 2 mg/ml lysozyme in solution I (25 mM Tris, 10 mM EDTA, 50 mM glucose pH 8.0), 400 μl of solution II (0.2 M NaOH, 1% SDS) added, mixed by inversion and maintained at 37°C for 10 min. 100 μl 2 M Tris pH 7.0 was added, mixed by inversion and left at room temperature for 30 min. 100 μl 5 M NaCl and 800 μl of phenol/chloroform/isoamyl plus 8-hydroxyquinoline (Mainiatis *et al.* 1982) were added, mixed by inversion for 5 min and centrifuged for 10 min at $12,000 \times g$. The aqueous phase was transferred into a 2 ml reaction tube and equal volume of isopropanol added, mixed by inversion and spun at $12,000 \times g$ for 15 min. The pellet was washed with cold 70% ethanol and air dried for 2 min. The precipitate was then purified by dissolving the pellet in 100 μl solution I, 170 μl

MilliQ H₂O and 2 M Tris pH 7.0, mixed gently and left at 50°C for 5 min. Then 20 µl 2 M NaOH and 150 µl ice cold 3 M potassium acetate solution were added, mixed and kept in ice for 5 min. 40 µl 1% ethidium bromide was added, gently mixed, kept in ice for 30 min and spun at 12,000 × g for 5 min. The supernatant was transferred to another tube and extracted twice with phenol/chloroform/isoamyl and chloroform/isoamyl. The aqueous phase was transferred to another tube, precipitated with equal volume of isopropanol, washed with 1 ml 70% ethanol, dried in vacuum and resuspended in 10-20 µl MilliQ H₂O.

2.2.2 Genomic DNA preparation

Total DNA from *Rhizobium* spp was extracted by slightly modifying the method of Wilson (1991). Ten ml of two day old broth culture grown at 28°C in an orbital shaker was pelleted and washed with 2.5 ml saline. The pellet was resuspended in 500 µl 25% sucrose in 50 mM Tris-HCl pH 8.0. 250 µl lysozyme (10 mg/ml in 0.25 M EDTA pH 8.0) was added and mixed by inversion. The mixture was kept in ice for 20 min. Then, 750 µl Tris-EDTA (10 mM Tris-HCl, 1 mM EDTA pH 8), 250 µl lysis solution (5% sarkosyl, 0.05 M Tris-HCL pH 8, 0.0625 M EDTA) and 20 µl proteinase K (20 mg/ml) were added and mixed by inversion. The mixture was maintained at 60°C until it became clear and then extracted twice with equal volume of Tris saturated phenol. Subsequent to phenol extraction, two phenol/chloroform/isoamyl extractions and two chloroform/isoamyl extractions were done. The DNA was obtained by ethanol precipitation.

2.2.3 Transformation

2.2.3.1 Preparation of electro-competent *Escherichia coli* cells and electroporation

A single colony of *E. coli* culture was inoculated into 10 ml broth of appropriate growth medium (LB or TY supplemented with diaminopimelic acid [DAPA]) and incubated with shaking overnight at 37°C. 5 ml of overnight culture was inoculated into 100 ml of growth medium and incubated with shaking at 37°C for approximately 2 h to an optical density of 0.5-1 at OD₆₀₀. Cells were chilled for 30 min in ice and spun at 10,000 rpm in a JA-14 Beckman rotor for 10 min at 4°C. The

pellet was washed with 100 ml sterile cold MilliQ H₂O and 50 ml cold MilliQ H₂O and then resuspended in 20 ml cold sterile 10% glycerol. The suspension was transferred into an SS34 tube and centrifuged in a JA20 rotor at 15,000 rpm for 15 min at 4°C. The pellet was resuspended in 1-2 ml cold sterile 10% glycerol and aliquots of 100 µl were transferred into 1.5 ml micro-centrifuge tubes, snap frozen in dry ice/ethanol or liquid nitrogen and stored at -70°C until required.

2.2.3.2 Electroporation and plating of transformants.

One hundred micro litres of electro-competent cells were transferred into an ice cold Gene Pulser^R 0.2 cm cuvette (BIORAD), DNA vector or ligation mixture was added and mixed thoroughly while the cuvette was still in ice. The mixture was transferred into an *E. coli* Pulser (BIORAD) and pulsed at 2.00 volts. Immediately after that, 1 ml of SOC medium (2% bactotryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄ and 20 mM glucose) was added, mixed thoroughly and maintained at 37°C for at least 1 h before plating 100-200 µl per agar medium (LB or TY) supplemented with appropriate antibiotics.

2.2.4 Restriction digests and electrophoresis of DNA

DNA was digested in reactions containing appropriate restriction enzymes (2 units enzyme per µg DNA) and corresponding buffer adjusted to appropriate volume with MilliQ H₂O, and incubated for 2-4 h at 37°C. At the end of the reaction, loading buffer (0.25% bromophenol blue and 40% [w/v] sucrose) was added at a ratio of 1:10, and the reaction solution loaded into wells of agarose, 0.7% to 1% (Seakem ME), gel in TAE buffer (40 mM Tris-acetate, 2 mM EDTA, pH 8.0). Following electrophoresis at 90-100 volts for 1 to 4 h, the gel was stained in ethidium bromide (10 µg/ml) solution and visualised under UV light (254 nm).

2.2.5 Extraction of DNA from agarose gel

Geneclean kit (BIO 101, Inc., CA, USA) or the JETquick spin column technique (GENOME Inc. USA) was used to extract required DNA fragments from agarose gels in accordance with the manufacturer's instructions. Briefly, in the Geneclean method, excised gel slice with the target DNA band was placed in a 1.5 ml plastic

reaction tube containing 3× its volume of 5M sodium iodide and incubated at 45-55°C to melt. 5 µl of thoroughly resuspended GLASSMILK (silica matrix) was added to solutions containing 5 µg or less DNA and thoroughly mixed. The resulting GLASSMILK/DNA complex was pelleted by spinning for 5 sec at 12,000 × g. The pellet was then washed 3 times with 400 µl ice cold NEW WASH solution (100 mM NaCl, 1 mM EDTA, 10 mM Tris-HCl pH 7.5, 50% ethanol) and resuspended in 20 µl MilliQ H₂O. The suspension was centrifuged to elute the DNA. The DNA within the supernatant was recovered by ethanol precipitation as in section 2.2.1.

2.2.6 Cloning of DNA fragments into plasmid

Subcloning of purified DNA from agarose gels into a vector was based on the protocols outlined in Sambrook, (1989) and Promega Corp., Protocols and Application Guide (1995). Briefly, the purified vector DNA was digested with appropriate restriction enzyme for 2 h, dephosphorylated with calf intestinal alkaline phosphatase (CIAP) to prevent self religation, phenol/chloroform extracted and recovered by ethanol precipitation. For the ligation reactions appropriate vector to insert ratios were added into a 0.5 ml reaction tube then 1 µl of 10× ligase buffer and 1 µl T4 DNA ligase were added and the solution made up to 10 µl with MilliQ H₂O and thoroughly mixed. Blunt ended ligations were incubated at 15°C for 4-16 h while sticky ended ligations were maintained at 22°C for 3 h or 4°C overnight.

2.3 Eckhardt Gel Electrophoresis

Large plasmids (>100 kb) in bacteria were separated, after *in situ* lysis, by electrophoresis in agarose using a modified method of Eckhardt (1978). Rhizobial cells were grown in an orbital incubator overnight in TY broth at 28°C while *E.coli* cells were grown in LB broth at 37°C. The optical density (OD₅₈₀) of the cells was measured using a 1cm diameter tube in a spectrophotometer (Turner Model 330). The appropriate volume of broth (see Table 2.3) was transferred to 1.5 ml plastic reaction tube and spun for 1 min in a microcentrifuge, the supernatant was removed with a microtip pipette and discarded. The pellet was kept at -20°C for at least 10 min. A TBE agarose gel (0.65%) was prepared in a midi (14×11 cm) gel tank using

two adjacent combs, the outer one with a single slot and the other with 8-14 slots. After the gel had set, the single slot comb was removed and 5 ml 0.5% agarose containing 1% SDS and 0.25% bromophenol blue was added into the slot. Prior to gel loading, the pellet was resuspended in 20 μ l lysis buffer (50 μ g/ml lysozyme, 125 μ g/ml RNase [in 10 mM Tris HCl pH 7.5, 15 mM NaCl, boiled in water for 15 min to inactivate nucleases] in 12.5% sucrose). The gel was run in TBE buffer at 7 mA for 15-30 min to ensure lysis, then 100 mA for approximately 3 hours and thereafter, was stained in ethidium bromide (10 μ g/ml) solution and visualized under UV light (254 nm).

Table 2.3 Volume of bacterial broth culture in relation to optical density (OD_{580}) that was used in Eckhardt gels

OD_{580}	Volume of bacterial broth (μ l)
0.10	750
0.15	500
0.20	375
0.25	300
0.30	250
0.35	215
0.40	188
0.45	167
0.50	150
0.55	136
0.60	125
0.70	110
0.80	90
0.90	80
1.00	70

2.4 High Voltage Paper Electrophoresis (HVPE)

HVPE was carried out as described by Murphy *et al.* (1987) and Wexler (1994). 1.0-2.0 μ l nodule extract was spotted in the centre of a sheet of 1.0 mm Whatman paper and air dried. The paper was saturated with formic acid/acetic acid buffer (28.4 ml 98% formic acid, 59.2 ml glacial acetic acid in 1 L water, pH 1.75) and electrophoresed at 3000 V for 15-30 min in formic/acetic acid buffer with distilled perchloroethylene as inert buffer. The electrophoretogram was dried and stained with silver nitrate reagent (0.4% AgNO_3 in acetone). Subsequently, the paper was air dried and submerged in 2.0% NaOH in 90% ethanol. Following the appearance of spots, the paper was then fixed in thiosulphate solution (100.0 g sodium thiosulphate, 15.0 g sodium metabisulphite in 1 litre H_2O), washed in water for 2-5 h, then air dried and documented by photography.

2.5 Purification of Bacteroids from Nodules

Bacteroids were purified by using the modified method of Robertson *et al.* (1978). Nodules (5.0-8.0 g) were crushed in a porcelain mortar containing 10 ml of homogenisation buffer (0.5 M sucrose in 10.0 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ pH 7.2). The homogenate was filtered through Miracloth and spun at $10,000 \times g$ (in an Eppendorf 5415C centrifuge) for 1 min. The upper layer of the pellet containing bacteroids was resuspended in 0.5 ml homogenisation buffer and layered on top of 0.6 ml of 1.2 M sucrose in 10.0 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ pH 7.2, while the lower portion of the pellet containing starch granules was discarded. The gradient was centrifuged at $10,000 \times g$ for 30 sec, the upper layer and the interface were collected and further spun for 1.5 min at $10,000 \times g$. The pellet was resuspended in 0.5 ml of homogenisation buffer and layered on top of 1.0 M sucrose in 10.0 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ pH 7.2. The gradient was spun at $12,000 \times g$ for 5 min and the pellet containing bacteroids enclosed in peribacteroid membranes was collected, washed 3 times with the homogenisation buffer by resuspending and spinning at $12,000 \times g$ for 1 min each time. The purified bacteroids were resuspended in homogenisation buffer and stored at -20°C until required.

2.6 Preparation of Peribacteroid Membrane

Peribacteroid membranes were prepared by modifying the method of Robertson *et al.* (1978). Peribacteroid membrane-enclosed bacteroids were osmotically shocked by resuspending in 140 μ l suspension buffer (10.0 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ pH 7.2), then frozen in liquid nitrogen and thawed twice to ensure the rupture of the peribacteroid membrane. This suspension was layered over 200 μ l of 1 M sucrose, 10.0 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ pH 7.2 and centrifuged at 16,000 g (14,000 rpm) for 15 min in an Eppendorf centrifuge. The peribacteroid membrane remained on top of the 1 M sucrose layer and 200 μ l of this was removed. The pellet contained the osmotically shocked bacteroids. The supernatant (i.e. the 200 μ l containing the peribacteroid membrane) was adjusted to 0.25 M sucrose in homogenisation buffer (10.0 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ pH 7.2) and layered over 0.5 M sucrose in the same buffer, then centrifuged at 60,000 g (29,500 rpm) for 30 min at 4°C using^a Beckman 70Ti rotor. The pellet consisted of peribacteroid membranes and the supernatant contained soluble proteins released during the osmotic shock of the bacteroids. The pellet was washed 3 times with the homogenisation buffer, resuspended in the same buffer and stored at -20°C until needed.

2.7 Preparation of Bacteroid Total Envelope Fraction and Cytoplasmic Contents

The bacteroid total envelope fraction and cytoplasmic contents were prepared by modifying the method of Robertson *et al.* (1978). The osmotically shocked bacteroid pellet in Section 2.6 above, was washed 3 times with the 0.5 M sucrose in homogenisation buffer, resuspended in 140 μ l of homogenisation buffer, frozen in liquid nitrogen and thawed three times, and then sonicated (using a Branson Sonifier Cell Disruptor B15) three times for 10 sec each with a microtip sonicator at 50 cycles per min with a power setting of 30W. Clearing of the cell suspension indicated lysis. The lysate was spun at 16,000 g in an Eppendorf centrifuge for 5 min. The supernatant collected contained bacteroids cytoplasmic contents, and was

concentrated in SpeedVac SVC 100 (Selby Scientific Instruments) to 1/8 of the original volume and then stored at -20°C. The pellet which contained the total envelop^e/fraction was washed twice with suspension buffer, resuspended in 50 µl suspension buffer and stored at -20°C.

2.8 Southern Hybridisation

Southern hybridisation was carried out using DIG (Digoxigenin) labeled probes and detection was by both colorimetric and chemiluminescent methods as recommended by the manufacturer, Boehringer-Mannheim (see Sections 2.8.7 and 2.8.8)

2.8.1 Southern blotting

Southern blotting was essentially the method of Southern (1975) as described in Boehringer Mannheim DIG System User's Guide for Filter Hybridisation. After gel electrophoresis, the gel was submerged in 250 mM HCl for 10 min with shaking at room temperature. The gel was then rinsed with RO water and submerged in denaturation solution for 2× 15 min at room temperature with gentle shaking. The gel was rinsed with RO water and submerged in neutralisation solution 2× 15 min at room temperature. Overnight capillary transfer of DNA from gel to Boehringer-Mannheim positively charged nylon membranes was done in 20× SSC at room temperature, then the DNA was fixed to the membrane by exposure to UV light for 2 min at room temperature. Prehybridisation and hybridisation were carried out as described in Section 2.8.5.

2.8.2 Dot blotting

DNA dilutions were denatured for 10 min at 100°C and chilled immediately on ice. Spot marks for identification were made on the membrane (Boehringer-Mannheim Nylon membrane positively charged) and 1µl DNA dilution was dotted on each spot. The DNA was fixed to the membrane by UV crosslinking for 2 min. Prehybridisation and hybridisation were done following the description in Section 2.8.5.

2.8.3 Colony hybridisation

Colony blotting was carried out using DIG (Digoxigenin) labeled probes as described in Boehringer-Mannheim DIG System User's Guide for Filter hybridisation. 2 layers of Whatma 3MM paper were soaked for each of the following solutions; denaturation solution, neutralisation solution and 2× SSC. Membrane disc was carefully placed over pre-cooled colonies (4°C for 30 min) on agar plate, left for 1 min and marks made on the membrane to identify the orientation to the plate. The membrane was carefully removed from the plate and blotted briefly on dry 3MM paper. The membrane disc (with colonies side up) was placed for 15 min on filter paper soaked with denaturation solution, blotted briefly on 3MM paper and then placed on filter paper soaked with neutralisation solution for 15 min. Then, it was briefly blotted on 3MM paper and place on 2× SSC for 10 min. The transferred DNA was crosslinked with UV light for 2 min. The membrane disc was placed on clean aluminum foil and 0.5 ml 2 mg/ml Proteinase K pipetted onto the membrane disc (of 82 mm in diameter), distributed evenly and incubated for 1 h at 37°C. Cellular debris were removed by blotting the membrane between wetted filter paper with the application of gentle pressure over the area and removing the filter paper. Prehybridisation and hybridisation were as described in Section 2.8.5. Detection was by colorimetric method (see Section 2.8.7) as recommended by the manufacturer, Boehringer-Mannheim.

2.8.4 Digoxigenin (DIG) labeling of DNA

The labeling of probes with DIG was by random primed incorporation of DIG-labeled dUTP (a thymidine analogue) according to the instruction of the manufacturer, Boehringer-Mannheim.

2.8.5 Prehybridisation, hybridisation and washing conditions

The standard hybridisation solution was used in Southern hybridisation (as described by Boehringer-Mannheim). The nylon membrane (Boehringer-Mannheim) containing fixed DNA was prehybridised for 4 h at 65°C. DIG labeled probes for hybridisation were denatured by boiling for 5 min at 95-100°C and cooled rapidly on ice. Hybridisation at 65°C was overnight in a shaking hybridisation water bath (Belly

Dancer, Stovall Life Science Inc., USA). Membranes were then washed twice in $2 \times$ SSC, 0.1% SDS at room temperature for 15 minutes, then twice in $0.5 \times$ SSC, 0.1% SDS for 20 min at 68°C. Approximately, 300 ml of solution was used to wash the membrane in each case.

2.8.6 Detection of Southern hybridisation using the DIG system

Detection of DIG labeled DNA was carried out using both the colorimetric and the chemiluminescent methods (Boehringer-Mannheim).

2.8.7 Colorimetric detection of DIG labeled DNA

DIG labeled DNA was detected using NBT (75 mg/ml nitroblue tetrazolium salt in 70% [v/v] dimethylformamide) and BCIP (50 mg/ml 5-bromom-4-chloro-3-indolyl phosphate, toluidinium salt in 100% dimethylformamide) after post hybridization washes, blocking and treatment with anti-dioxigenin, Fab fragments conjugated to alkaline phosphatase, following the manufacturer's (Boehringer-Mannheim) Instruction manual.

2.8.8 Chemiluminescent detection of DIG labeled DNA

DIG labeled DNA was detected using chemiluminescent substrate CSPD, 25 mM disodium 3-(4-methoxyspiro{1,2-dioxetane-3,2'-(5'-chloro)tricyclo[3.3.1.1^{3,7}]decan}-4-yl) phenyl phosphate, after post hybridization washes, blocking and treatment with anti-dioxigenin, Fab fragments conjugated to alkaline phosphatase, according to the protocol of the manufacturer, Boehringer-Mannheim.

2.8.9 Autoradiography

Autoradiography was done by exposing the treated membrane to X-ray film (Fuji RX) between two intensifying screens. X-ray cassette was kept at room temperature for 45 to 90 minutes. The X-ray film was developed (with developer G153) and fixed (with fixer G353) in automated CURIX 60 (AGFA) instrument. The film was then washed thoroughly with RO water and left to dry at room temperature.

2.8.10 Stripping and reprobing of membranes

DIG labeled DNA was removed from nylon membranes and ^{the membranes} rehybridized to a different probe following the instruction of the manufacturer, Boehringer Mannheim. A brief description of the procedure is as follows: The membrane was washed in water for 1 minute, incubated twice for 10 minutes in alkaline probe-stripping solution (0.2N NaOH, 0.1% SDS) at 37°C, and then rinsed thoroughly in $2 \times$ SSC and thereafter reprobbed commencing with the prehybridization step.

2.9 Sequencing of DNA

2.9.1 Sequencing Reactions

DNA clones into pGEM7Zf(+) was sequenced with M13 forward and reverse primers, following the dideoxy method described by Sanger *et al.* (1977). Sequence reactions were performed using a PRISM® Ready Reaction Dye Primer Cycle Sequencing Kit (Applied Biosystem Inc.) and analyzed in an Applied Biosystems 373A automated DNA sequencer, in accordance with the manufacturer's instruction manual. In order to complete DNA sequencing in both directions, crude preparations of oligonucleotides (15-20 bp) designed using the programme, OLIGO 4.0-s, and synthesized by GeneWorks (South Australia), were used as primers for DyeDeoxy terminator sequencing reaction, following the instruction of the manufacturer.

2.9.2 Analysis of DNA sequences

Data obtained from sequencing were compiled with the programmes, Sequence Editor v1.03 (Applied Biosystems Inc.) and DNA Strider personal computer. DNA and amino acid similarities between pairs of sequences were analyzed with GCG (Genetics Computer Group, Madison, Wis., USA) version 7.0 suite of programmes (Devereux *et al.*, 1984). Possible open reading frames were determined with the CODONPREFERENCE program. BLAST and BESTFIT programs were used for DNA and protein searches in data base (Altschul *et al.*, 1990). Alignment of sequences was done with the program PILEUP and homologous regions emphasised with PRETTYBOX.

2.10 Protein Analysis

2.10.1 Antibody preparation

Cultures of PM1203 (*E. coli* strain carrying pPM1203) or PM1204 (*E. coli* strain carrying pPM1204) [see Table 3.1] on LB agar supplemented with 100 µg/ml ampicillin was grown as described by Grezmski (1994). Culture of PM1203 and PM1204 were inoculated into 50 ml LB plus 100 µg/ml ampicillin and incubated in an orbital shaker for 3 h at 37°C. 50 µl 0.4M IPTG was added (to achieve a final concentration of 0.4 mM) to induce T7 polymerase and the broth culture was incubated for 3 h. 50 µl of 100 mg/ml rifampicin was added to a final concentration of 100 µg/ml to arrest cell growth and then further incubated for 1 h. Cells were sedimented by spinning at $12,000 \times g$ for 10 min (Sorval RC2-B centrifuge, SS-34 rotor), resuspended in 10 mM Tris-HCl (pH 7.5), 25 mM EDTA and lysed by freezing and thawing once. The lysate was centrifuged for 10 min at $12,000 \times g$, the sediment containing fusion proteins was resuspended in the same buffer. DNase and RNase were added to a final concentration of 1.0 µg/ml each and incubated at 37°C for 1 h, to remove DNA and RNA contaminants. Fusion proteins were collected by centrifuging at $12,000 \times g$ for 10 min and washed 3 times with 10.0 mM Tris-HCl (pH 7.5), 25 mM EDTA, and stored at -20°C in the same buffer. The cell lysates containing fusion proteins were digested with LUG buffer (50.0 mM Tris-HCl [pH 6.8], 10.0% [v/v] glycerol, 2.0% SDS, 5.0% [v/v] β-mercaptoethanol, 0.00125% [w/v] bromophenol blue) (Maniatis, 1982), and run on PAGE (Laemmli, 1970). Bands corresponding to MosB (55 kDa) and MosC (43 kDa) (Murphy *et al.*, 1993) were extracted with a scalpel and lyophilized. The lyophilized samples were then submitted to Institute of Medical and Veterinary Research (IMVS), South Australia for antibody preparation as described by Harlow and Lane (1988).

2.10.2 Absorption of antibodies

Escherichia coli BL2(DE3) was inoculated into LB broth and incubated in an orbital shaker overnight at 37°C to a cell population of greater than 10^9 cells/ml. 1.5 ml of broth culture was transferred into each of 8 reaction tubes and pelleted at $12,000 \times g$

for 3 min. The cells were washed once with 1 ml PBS (4.3 mM $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$, 137 mM NaCl, 2.7 mM KCl, 1.4 mM KH_2PO_4) containing 0.05% sodium azide. The cells in the first tube were resuspended in 0.5 ml PBS/sodium azide solution and an equal volume of antiserum was added and mixed thoroughly. The mixture was incubated rotating (on a Rotor-Torgue EXTCH Equipment Ltd.) at 37°C for 1 h. The cells were then pelleted and the supernatant transferred to the second reaction tube containing washed *E. coli* cells. The cells were again resuspended and treated as in the first tube. The cycle of pelleting and resuspension was repeated until the 5th tube. After resuspension in the 6th tube, cells were left overnight at 4°C after which the cells were sedimented. The supernatant was further absorbed twice at 37°C for 1 h and then filter sterilised and stored at 4°C until required.

2.10.3 Polyacrylamide gel electrophoresis (PAGE)

Whole cells or sediments of cell lysates were digested and separated on SDS-PAGE as described by Laemmli (1970). 12% (w/v) polyacrylamide gels were used. Pellet cells were resuspended in 50 µl PBS and equal volume of 2× LUG buffer (see section 2.10.1) was added and the mixture heated for 5 min at 95-100°C. The gels were run for 60-90 min at 120V on min on a mini-Protean II Dual Slab Cell (Bio-Rad) and stained with 0.1% Coomassie Blue stain dissolved in 40.0% methanol and 10.0% acetic acid. Bio-Rad low molecular weight standards in Dalton^s phosphorylase b (97,400), BSA (66,200), ovalbumin (45,000), carbonic anhydrase (31,000), soybean trypsin inhibitor (21,500) and lysozyme (14,400) were used. The bands were photographed for documentation.

2.10.4 Western blotting

Protein bands separated by SDS-PAGE (Section 2.10.3) were transferred to nitrocellulose membrane (NitroBind MSI membrane, GeneWorks) at 15V for 45 min using a Trans-Blot[®] Semi-dry apparatus following the instructions of the manufacturers (BioRad Laboratories USA). The membrane was rinsed for 1 min in TTBS (20 mM Tris-HCl [pH 8.0], 0.9% [w/v] NaCl, 0.05% [v/v] Tween 20 in MilliQ H₂O) in clean Petri dish and then suspended in antiserum solution (3 ml 0.5% sodium azide, 3 ml 10× TTBS, 24 ml MilliQ H₂O, 30 µl 0.02% skim milk in

TTBS, 6µl absorbed antiserum [i.e. 1:5,000 dilution]) and incubated for 4-18 h. The membrane was washed 3× for 10 min each with TTBS and resuspended in peroxidase solution (15µl goat antirabbit IgG {whole molecule} peroxidase conjugate [Sigma], 30 µl 0.02% skim milk in 30 ml TTBS) and incubated at room temperature with gentle agitation for 3.5 h (Sambrook *et al.*, 1989). The membrane was then washed 4 times for 5 min with TTBS, twice with TBS (20 mM Tris-HCl [pH 8.0], 0.9% [w/v] NaCl) and then immersed in freshly prepared detection solution (15 µl H₂O₂ [4°C] in 16.5 ml TBS [room temperature] mixed with 9.9 mg 4-chloro-1-naphtol in 3.3 ml methanol [−20°C]). Colour development occurred after 5 to 60 min and the reaction was stopped by washing once in water. The membrane was dried at room temperature and then photographed.

2.11 Electron Microscopy

Nodules obtained from lucerne plants inoculated with either *Sinorhizobium meliloti* strains L5-30 or Rm1021 were suspended overnight at 4°C in 200 µl of fixative (4% paraformaldehyde, 0.25% glutaraldehyde, 4% sucrose in PBS) in an eppendorf tube. The nodules were then washed twice with washing buffer (4% sucrose in PBS) and suspended in 2.3 M (78.7%) sucrose in PBS. Using Cryo-microtome (Leica Reichert Ultracuts), thin sections of 200nm were cut at −110°C to −100°C and picked up on carbon coated nickel grids and labeled at RT with AuroProbeTM, following the manufacturer's (Amersham) instructions, as follows: Grids were incubated in 0.02M glycine in PBSG buffer (0.5% BSA, 0.1% gelatin in PBS) for 10 min, rinsed twice for 1 min with PBSG buffer and placed on buffer supplemented with 5% normal rabbit serum for 15 min. The grids were incubated with absorbed 1:20 dilution of rabbit anti-MosB or 1:50 dilution of anti-MosC for 60 min. Grids were rinsed twice for 5 min with PBSG buffer and incubated with 1:25 dilution of 5nm Auroprobe EM goat anti-rabbit IgG (H+L) G5 (Amersham) for 30 min. Grids were rinsed twice for 5 min with drops of 250 µl PBSG buffer and then twice with 2% glutaraldehyde in PBS. The grids were then washed three times with PBS for five minutes, and two times five minutes, then in 1.5% uranyl acetate in MilliQ H₂O for 5 min, and again washed three times with MilliQ H₂O for 5 min. The grids were

then placed in 1% methyl cellulose plus 0.5% uranyl acetate in MilliQ H₂O (w/v) for 1 second (Hayat, 1988). The stained samples were examined and photographed with transmission electron microscope (Philips CM100 TEM) at the Centre for Electron Microscopy of South Australia (CEMMSA).

2.12 Rhizopine Extraction

Nodules harvested from 4-6 weeks old plants inoculated with rhizopine producing *Rhizobium* strain were crushed in water (2.5 g/ml). The extract was purified by passing it through Dowex 50W-X8 (H) (BDH, Poole, England) cation exchange columns. Columns were left overnight in 0.1 M HCl. Before loading the extract, the pH was increased to approximately 2.5 by running water through the column. The extract was loaded on the column at 1 ml per min and left overnight. The column was washed with 5 bed volumes of water which was also added at 1 ml per min. The negatively charged compounds were eluted slowly (15 ml/h) with 5 bed volumes of 2 M ammonia solution. The eluate was collected when the pH reached 9. Higher yields were obtained when columns were left overnight.

2.13 Rhizopine Catabolism Tests

The rhizopine, 3-*O*-methyl-*Scyllo* inosamine (3-*O*-MSI), was catabolised with rhizobial strains grown on TY agar at 28°C for 36 h. 500 µl of bacterial cell suspension of optical density 0.5 at OD₅₈₀ made from the agar culture, was washed with saline (0.85% w/v NaCl) and resuspended in an equal volume of Bergersen's minimal medium (Bergersen 1961, Appendix 1). 200 µl of this bacterial suspension was mixed with 5 µl of 3-*O*-MSI (in a sterile polystyrene tubes (Elkay^(R) products, INC. Shrewsbury, MA., USA) and the culture was incubated with shaking at 28°C for 5 days under atmospheric or microaerophilic condition. 4 µl of the suspension was subjected to HVPE analysis.

2.14 Seed Sterilization

Lucerne seeds were immersed in 2% sodium hypochlorite for 6 min and transferred into 70% ethanol for 2 min. The seeds were then rinsed 6 times in MilliQ H₂O.

2.15 Root Extract

Lucerne roots (2 g) were cut into small pieces and ground in a mortar with a pestle. 1 ml of MilliQ water was added and the aqueous extract was concentrated in a SpeedVac to 1/10 of its original volume.

2.16. Mating of bacteria

Mating was done by mixing the suspensions of saline washed bacterial cells grown overnight on agar medium. The Pellet from the centrifuged mixture was resuspended in appropriate broth medium and spread on a nitrocellulose membrane placed on LB agar (for *E. coli*) or TY agar (for *Rhizobium* sp.) in a Petri dish. This was then incubated at 30°C for 4-6 h. The cells were then washed into eppendorf tube with a broth medium supplemented with appropriate antibiotics and spread on a similar agar medium. The transconjugants were isolated after incubation under appropriate conditions.

Chapter 3

Analysis of *mosB* and *mosC* homologue gene sequences from *R. leguminosarum* bv. *viciae* strain 1a



THE UNIVERSITY OF ADELAIDE

CHAPTER 3

Analysis of *mosB* and *mosC* homologue gene sequences from *R. leguminosarum* bv. *viciae* strain 1a

3.1 Introduction

In *S. meliloti* rhizopine strains such as L5-30 the metabolism of rhizopine (L-3-*O*-methyl *scyllo*-inosamine [3-*O*-MSI]), confers a competitive advantage in nodulation (Gordon *et al.*, 1996). The rhizopine catabolic and synthetic genes in *S. meliloti* strain L5-30 are closely linked and located on the symbiotic plasmid (pSym) (Murphy *et al.*, 1987). The *mos* locus, which contains the rhizopine biosynthetic genes, has four open reading frames (ORFs), ORF1, *mosA*, *mosB* and *mosC* arranged in an operon (Murphy *et al.*, 1993). Interspersed within the *mos* locus are regions homologous to *nifT*, *fixU* and *fixA*, thus suggesting a high frequency of rearrangement with symbiotic genes (Murphy *et al.*, 1993). The *mos* locus is regulated via a promoter with motifs recognised by NtrA and NifA proteins of the NifA/NtrA regulatory system which controls symbiotic genes such as those in *nif* and *fix* operons (Murphy *et al.*, 1988). The symbiotic regulation of the *mos* locus indicates that it functions in active nitrogen fixing bacteroids. ORF1 appears not to be required for rhizopine production since a frame shift mutation in this region did not affect 3-*O*-MSI production and no protein product was detected with appropriate antibodies (Murphy *et al.*, 1993). However, ORF1 curiously shares a region of homology with the *nfe1* gene of *S. meliloti* GR4, that is part of the *nfe* regulon which is involved in nodulation competition (Soto *et al.*, 1993).

The MosA protein, responsible for 3-*O*-MSI methylation (Rao *et al.*, 1995), has significant homology with DapA (dihydrodipicolinate synthase) (Bahar, 1997), which is involved in the lysine biosynthetic pathway (Yugari and Gilvergi, 1965). MosB which was previously thought to have a regulatory role in 3-*O*-MSI biosynthesis, due to its considerable homology with regulatory proteins such as DegT from *Bacillus stearothermophilus* and DnrJ from *Streptomyces typhimurium* (Murphy *et al.*, 1993) is now considered more likely to be an amino transferase (see

discussion). The *mosC* gene has two in-frame start codons encoding *mosC1* and *mosC2*. The amino terminal end of MosC1 contains a 19 amino acid hydrophilic region whereas MosC2 does not. The remainder of the MosC protein is very hydrophobic with 12 membrane spanning domains, and is postulated to be involved in transporting a rhizopine precursor into bacteroids, or alternatively, transporting rhizopine out. A point mutation at the start of either *mosC1* or *mosC2* does not affect rhizopine production, thus indicating that either of the proteins can be used to synthesise rhizopines (Murphy *et al.*, 1993; Grzowski, 1994).

In previous studies, a *mosB* hybridising region was isolated from a subclone of a cosmid library from *R. leguminosarum* bv. *viciae* strain 1a (Wexler, 1994). This clone was termed pPM1302. In the same study no hybridisation to a *mosC* probe was observed (Wexler, 1994). The objectives of this study are to further subclone and sequence the *mosB* hybridising region and re-investigate whether the *mosC* homologue can be found in *R. leguminosarum* bv. *viciae* strain 1a.

3.2 Results

Table 3.1 Plasmids used in this chapter

Plasmid	Description	Source
pPM1203	1.2 kb <i>Nsi</i> I- <i>Hind</i> III fragment internal to <i>mosB</i> fragment from <i>S. meliloti</i> strain L5-30 cloned in pGEMEX	Grzinski, 1994
pPM1204	1.4 kb <i>Mlu</i> I fragment containing 3' region, including untranslated region of <i>mosC</i> from <i>S. meliloti</i> strain L5-30 cloned in pGEMEX	Grzinski, 1994
pPM1205	30.6 kb partial <i>Hind</i> III fragment from <i>R. leguminosarum</i> bv <i>viciae</i> strain 1a homologous to <i>mos</i> genes in cosmid pVK102, Km ^r , Tet ^r .	Wexler, 1994
pPM1302	5.3 kb <i>Hind</i> III DNA <i>mosB</i> hybridising fragment from <i>R. leguminosarum</i> bv <i>viciae</i> strain 1a pSym, cloned into pUCD2608, Km ^r .	Wexler, 1994
pPM1361	0.8 kb <i>Bam</i> H I- <i>Nsi</i> I <i>mosB</i> homologous DNA fragment from pPM1302 in pGEM-7Zf(+), Amp ^r .	This study

3.2.1 DNA purification and plasmid mapping

Plasmid DNA, pPM1302 [Table 3.1] (Wexler, 1994), containing a 5.3 kb *mosB* homologous region from *R. leguminosarum* bv. *viciae* strain 1a Sym plasmid, was extracted by the alkaline lysis method (Section 2.2.1.1). A restriction map of this DNA fragment was made by a series of single, double and multiple restriction enzyme digests using *Bam*H I, *Nsi* I and *Hind* III (Fig. 3.1) and the fragment sizes estimated from agarose gel photograph.

3.2.2 Southern blot with DIG labeling *mosB* and sub-cloning.

Southern blot analysis of *Bam*H I and *Bam*H I/*Nsi* I digests of pPM1302 (Wexler, 1994; Wexler *et al.*, 1996) (schematically represented in Fig. 3.1), was done using a

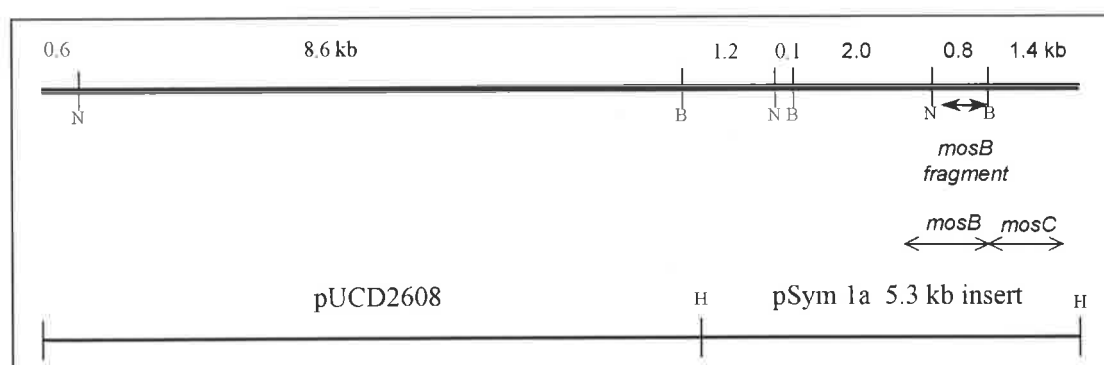


Fig. 3.1 Schematic restriction map of pPM1302 (Table 3.1) using *Nsi* I (N), *Bam*H I (B), and *Hind* III (H).

DIG labeled (Section 2.8) *S. meliloti* strain L5-30 *mosB* gene specific probe. The *mosB* probe (from pPM1203) [Table 3.1] was a 1.2 kb *Nsi* I-*Hind* III fragment from the internal region of *mosB* from *S. meliloti* strain L5-30 (Grzemski, 1994; Murphy *et al.*, 1993). Figure 3.2 shows strong hybridising bands of 2.8 kb in the *Bam*H I digested DNA and 0.8 kb in the *Bam*H I-*Nsi* I DNA digest. This 0.8 kb *Bam*H I-*Nsi* I DNA band was extracted, purified and the fragment cloned into pGEM7Zf(+) (Section 2.2) to yield pPM1361 (Table 3.1).

3.2.3 Southern hybridisation of *Bam*H I digested *R. leguminosarum* bv. *viciae* strain 1a pSym with DIG labeled *mosC*

Extracted *Rhizobium leguminosarum* bv. *viciae* strain 1a pSym DNA (see Section 2.2.1.3) digested with *Bam*H I was probed with a DIG labeled (see Section 2.8) specific *S. meliloti* strain L5-30 *mosC* DNA. The *mosC* probe (pPM1204) [Table 3.1] was a 1.4 kb *Mlu* I-*Apa* I fragment cloned into a blunt ended *Eco*R I-*Apa* I site in pGEMEX. The probe was a 1 kb fragment released from this plasmid by digesting with *Pst* I-*Apa* I. This contains approximately 800bp of the 3' end of *mosC* and 200bp of the untranslated region. The southern blot revealed a very faint hybridising band (data not shown).

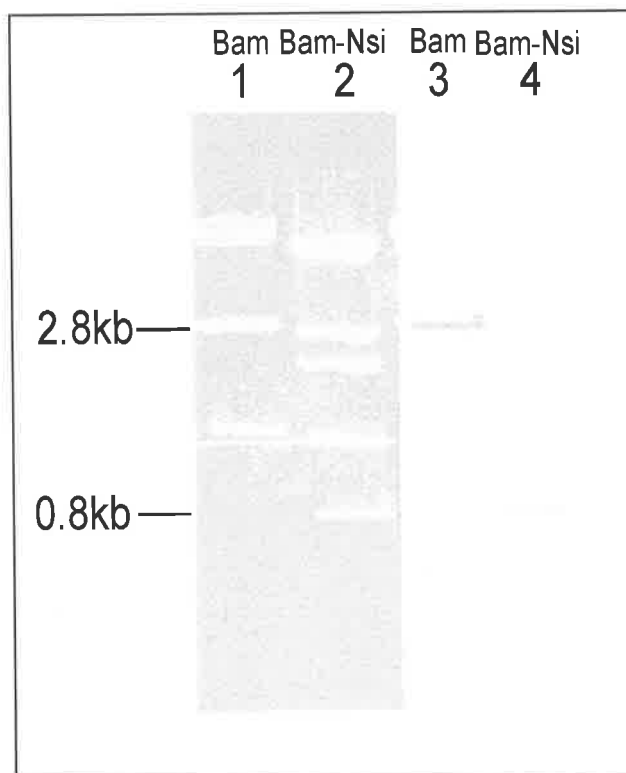


Fig. 3.2 Agarose gel (lanes 1 and 2) of *Bam*H I/*Nsi* I digests of pPM1302 and corresponding Southern blot (lanes 3 and 4), probed with DIG labeled *mosB* fragment from *S. meliloti* strain L5-30. Lane 1 corresponds with lane 3 (*Bam*H I digested), while lane 2 (*Bam*H I-*Nsi* I digested) corresponds with lane 4.

3.2.4 Sequencing

pPM1361 (Table 3.1) was partially sequenced with M13 forward and reverse primers (Section 2.10). The resulting sequences were progressively used to construct oligonucleotides (designed with the program, OLIGO 4.0-s) to complete the sequencing of the 0.8 kb insert. Since the *mosB* and *mosC* genes in *S. meliloti* strain L5-30 are linked (Murphy *et al.*, 1993), it was assumed that this may also be the case with *R. leguminosarum* bv. *viciae* strain 1a. Therefore, further sequencing downstream of the *mosB* homologue was carried out in pPM1302 and a *mosC* homologue was found in this region (Fig. 3.3). As well, the very faint *mosC* hybridisation (Section 3.2.3) observed suggested the presence of *mosC* in *R. leguminosarum* bv. *viciae* strain 1a. The *mosC* homologous region terminated 8bp before the *Hind* III site of the 5.3 kb insert in pPM1302, therefore an overlapping clone, pPM1205 (Table 3.1) was used to extend the sequence.

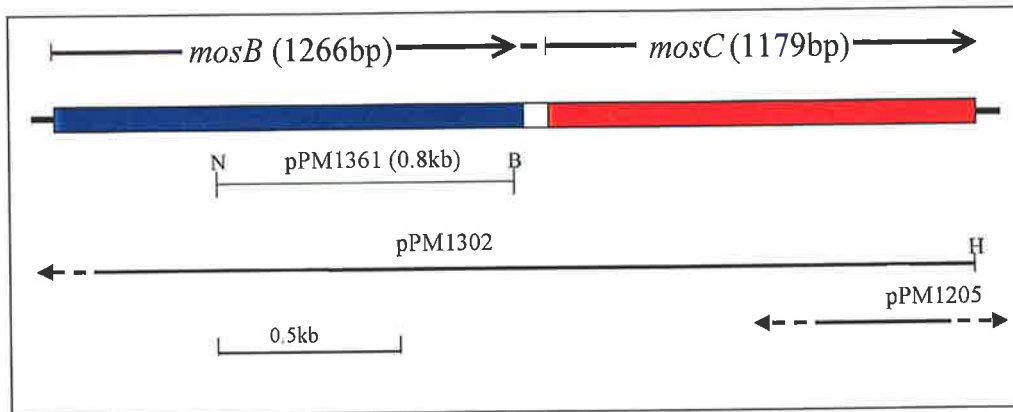


Fig. 3.3 Schematic representation of *mosB* and *mosC* homologous genes of *R. leguminosarum* strain 1a. The *mosB* fragment (Fig. 3.1) was cloned into pGEM-7Zf(+) yielding pPM1361 which was initially sequenced. With progressively constructed oligonucleotides, the sequence of the *mosB* homologue in pPM1302, and *mosC* in pPM1302 and pPM1205 were completed (see text for details). The shaded areas show the open reading frames and between them is a clear portion showing a 79bp intergenic region. The direction of transcription is shown by arrows flanking the homologous *mosB* and *mosC* labels. pPM1361 contains the 0.8 kb *Nsi* I (N)/*Bam*H I (B) insert and the *Hind* III (H) at the right end of pPM1302 indicates one end of 5.3 kb *Hind* III insert (see Table 3.1)

Using the programs MAP and CODONPREFERENCE, two open reading frames (ORFs) of 1266bp and 1179bp, separated by an intervening sequence of 79bp were found in the *R. leguminosarum* bv. *viciae* strain 1a sequence (Fig. 3.4).

BLAST analysis indicated that the first of these had homology with *mosB* and the second with *mosC*. ~~Calculated~~ ^{Predicted} molecular weights of the MosB and MosC homologues ~~were~~ ^{were} 46135 and 41252 Dalton respectively (Table 3.2). A Shine-Dalgarno sequence for the *R. leguminosarum* bv. *viciae* strain 1a *mosB* homologue was not identified but that for *mosC* (GGAG) was separated from the start codon, ATG, by 6 bp. Comparison of MosB and MosC from *S. meliloti* L5-30 with their homologues in *R. leguminosarum* bv. *viciae* 1a using PILEUP/PRETTYBOX

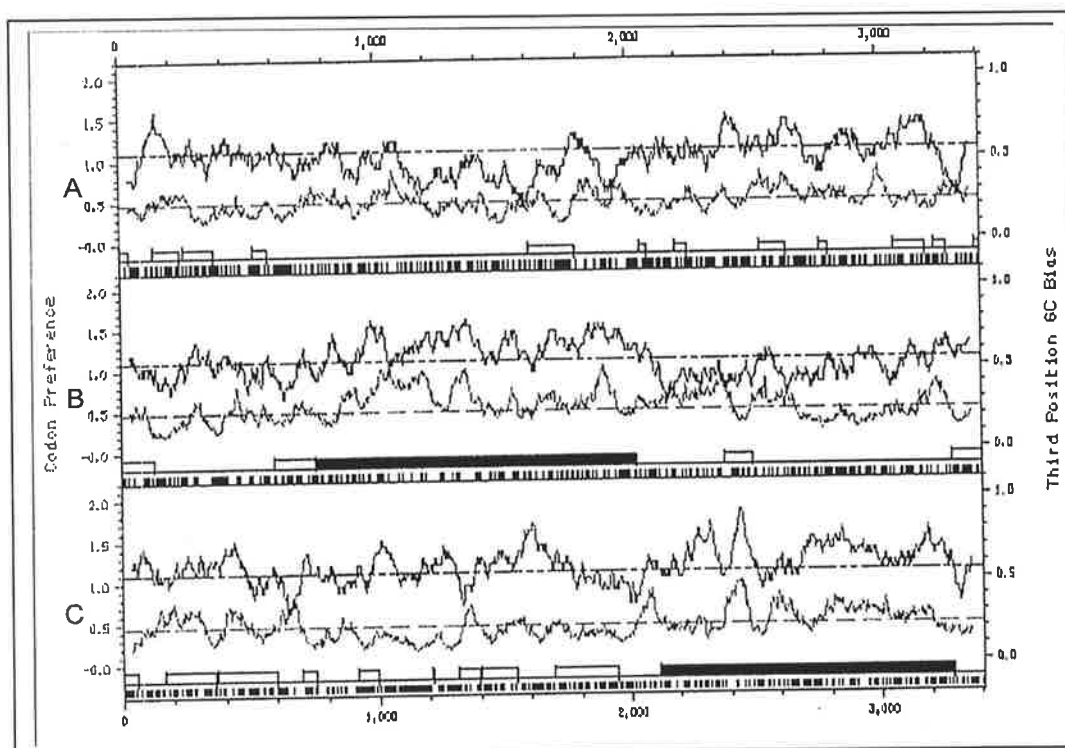


Fig. 3.4 CODONPREFERENCE of the *mosBC* homologue showing two open reading frames at panels B and C. Panels A, B and C are three forward reading frames. The shaded portion in panel B represents the MosB homologue ORF while that in panel C is the MosC homologue ORF. Codon table: GenRun Data ecohigh.cod. PrefWindow: 25 Rare Codon Threshold 0.10 BlastWindow: 25 Density: 111.6.

Table 3.2 Characterisation of open reading frames from *mosB* and *mosC* homologous regions of *R. leguminosarum* bv. *viciae* strain 1a.

ORFs	Shine-Dalgarno sequence/Intervening bases/ start codon/number of residues/stop codon	~~Calculated~~ Predicted molecular weight (Dalton)
ORFb	?/ATG/214/TAA	8384
ORFa	GAGG/10/ATG/225/TGA	8529
MosB homologue	?/ATG/1260/TAG	46135
MosC homologue	GGAG/6/ATG/1173/TAA	41252

revealed 69% identity with MosB (Fig. 3.5) and 49% with MosC (Fig. 3.6), respectively. Upstream of the *mosB* homologue (in the putative promoter region) are two putative

RNA polymerase binding sites with *nifA*/ σ^{54} (*ntrA/rpoN*^{*glnF*}~~*glnF*~~) recognition motifs, which are separated by 297bp. In the distal (i.e. farther from the *mosB* homologue) "promoter", the *nifA* and σ^{54} recognition motifs are separated by 44bp and in the proximal "promoter" these motifs are separated by 51bp. The arrangement of the two *nifA*/ σ^{54} sites is therefore as follows; TGT-N₁₀-ACA {*nifA*} —44bp— TGGCAC-N₅-GTGC { σ^{54} } —297bp—TGT-N₁₀-ACA {*nifA*}—51bp—TGGCAC-N₅-GTGC { σ^{54} } (see Fig. 3.7 and Fig. 3.8). The distal *nifA*/ σ^{54} recognition sequence precedes a Shine-Dalgarno sequence (GAGG) and an ORF encoding 77 amino acids (henceforth referred to as ORFa). BLAST analysis of DNA sequence of ORFa (a 231bp region, i.e. 369-599bp in Fig. 3.8) revealed extensive homology with *mosB* from *S. meliloti* strain L5-30. BESTFIT analysis of DNA sequence from 242 to 611 bp with the DNA sequences between 621 and 1007 bp showed 82% similarity and identity, suggesting sequence duplication has occurred (Fig. 3.9).

Upstream and overlapping the distal *nifA* recognition sequence is a 219bp ORF (ORFb) encoding 73 amino acids transcribed in the opposite direction to the *mosBC* homologues. The deduced protein sequence from ORFb has two short regions of homology (varying from 41-53% identity and 64-69% similarity) with a hypothetical 28.8 kDa protein in the NifR3-like 5' region of *Azospirillum brasilense*. This covers 2 regions of a total of 30 amino acids (i.e. 7/17 and 7/13 identity, and 11/17 and 9/13

similarity). The *ntrBC* genes of *A. brasilense* and *R. leguminosarum* are part of a negatively regulated *nifR3*-like-*ntrB*-*ntrC* operon (Machado *et al.*, 1995; Patriarca *et al.*, 1993). The *nifR3*-like gene shows similarity with a region termed ORF1 in *R. leguminosarum* upstream of the *ntrB* gene in a complex operon (Machado *et al.*, 1995).

mosbrle	..TIRINGD	VLSAVIPKKD	LEETPISVAN	EDLWGGSIIL	0
mosbrme	MKV TIRINGD	VLSAVIPKKD	LEETPISVAN	EDLWGGSIIL	40
mosbrle	..RRLALPH	LPQEARLPVT	VEARRMFGRL	DRGAHEK?DR	0
mosbrme	RNGWRLALPH	LPQEARLPVT	VEARRMFGRL	DRGAHEK?DR	80
mosbrle	...MPSKPNF	QVPSAEPNGT	RTLHRWPILT	GANYRKVIAC	37
mosbrme	MTRIGEISSS	QDAAMLAEEN	QKMHPWPALT	RTAYEDVAAC	120
mosbrle	ISSGELSGSG	LEVINEFETT	VENWIGGGHV	VSTNSGTAL	77
mosbrme	ISSGELSGSG	LGIIINAFERR	MEEWIGGGYV	VSASSGTAL	160
mosbrle	MVALTLGVG	PEDVVLVPSY	TWSATAFAAT	LIGAIPRFVD	117
mosbrme	TVALIALGIQ	PGDVVLVPSY	TWATAALAPL	LIGAIPRFVD	200
mosbrle	IDS DTYNLS	AVAKAITPD	VKAIIVVHMH	GLSCDMDEIT	157
mosbrme	IDPNSYNISP	TALAAAITPD	VKAIIVVHMH	GLSCDMDEIT	240
mosbrle	ATAREKGVPI	IEDCAQAHGA	LYKKQYVGT	SDIGCFSMQK	197
mosbrme	CHAREQGIIV	IEDCAQAHGA	LYKGOHVGL	SDIGCFSMQK	290
mosbrle	SKHFSAGDGG	FLVTRNAALA	QKARDICNFG	LPTPKHNYHF	327
mosbrme	SKHLSAGDGG	FMVTRDPTLA	QKMRDICNFG	LPTPKANYRF	320
mosbrle	E E G L R E G Y A V	F R E C E Q I G G M	F R L N P L S A A L	V I D Q L E H L D Q	277
mosbrme	D E V V R D G Y A V	F R E C E Q I G G M	F R L Q P M S A A L	V M H O L E H L R Q	360
mosbrle	R E W L Q E A M K	P L V A E A A Q V P	F L K I T Q P G S H	R T H V W H K I R V	317
mosbrme	R I A W L Q T A M E	P L V E E S A K I P	F F K I T R S H V D	R T H V W H K I R V	400
mosbrle	G I D Y G A V D H Y	G R S M A D I R Q G	L R N T L A R S G I	P S T L W T A P I L	357
mosbrme	G I D Y A A V D Y F	G R S M A E I R R G	L R N S L A E R G I	S S T L W T A P I L	440
mosbrle	P L Q K V F R P Y A	G T I D W T K T A G	H L A I E N S F I V	F D E K Y P L I A Q	397
mosbrme	P L Q T A F R P Y A	G V V E W T K S A G	H L A I E N S F I V	F D E N Y P L I A Q	460
mosbrle	D P S T M T A L V S	K L R Q V W S D Y F	A S L E * . . .	421	
mosbrme	E P A K M A E L V V	K L Q Q A W D E H F	T S L A R G K * 507		

Fig. 3.5 PRETTYBOX from PILEUP showing extensive homology between deduced amino acids of *S. meliloti* strain L5-30 MosB and *R. leguminosarum* bv. *viciae* strain 1a MosB homologue.

BESTFIT analysis of the deduced amino acids of the MosB homologue in *R. leguminosarum* bv. *viciae* strain 1a with *S. meliloti* strain L5-30 MosB showed 72% identity and 83% similarity over 401 amino acids. The MosC homologue has 49% identity and 64-68% similarity with *S. meliloti* strain L5-30 MosC over 205 amino acids. However, *S. meliloti* strain L5-30 MosB and MosC are 86 and 15 amino acids respectively, longer than their homologues in *R. leguminosarum* bv. *viciae*

strain 1a. *S. meliloti* strain L5-30 MosB differs from the homologue in *R. leguminosarum* bv. *viciae* strain 1a by having an extra 83 amino acids in the 5' region and 3 additional amino acids at the 3' region. *S. meliloti* strain L5-30 MosC has 19 additional amino acids in the 5' region and 9 less amino acids in the 3' region compared to the MosC homologue. There are also a number of small gaps when comparing the central region of the sequence (see Fig. 3.6).

1a__mosc	M	SNARIFVPQK	.RVMAVAFQF	20
1530mosc	MTRTSPRHHA	PSETKRRVPM	GGVHTNPGKK	LTITAVAFQF 40
1a__mosc	FINGVVYGTW	ATNILHIRRS	FGLSEFGLSI	VFLTMGAHAV 60
1530mosc	FINGLVLAAW	ATSIIPHVKNA	YSFNDAELGV	LLIIMAAAGAL 80
1a__mosc	LIMSLTGYLV	QIFGSRRISI	LSCLEFFPAAL	CFVFWAPNHI 100
1530mosc	LFMSLAGYFS	HVFGSRRMSI	QSALLFPSSAL	VLIFAAPNCM 120
1a__mosc	ALLVGVGFLFG	AANGSMDVSM	NDQGTNLGKR	GSDPVISLHL 140
1530mosc	TLFCSI VLFGL	AANGAMDVL M	NHQAKALEEN	GFPRIMAFHL 160
1a__mosc	GFSSTIGALAG	SLLSYW WIGR	GYHPLVQAIA	LLVGTLMLSW 180
1530mosc	GCSSTGILAG	.IMTFGVIGD	G.HYVARSVT	LLTGILIVAR 199
1a__mosc	RLFPHLDLDA	DRSRISLELG	ALGNRGLILL	GFLSFLTMMS 220
1530mosc	WLFPHLLDDV	RSGEHRLAIG	ELRNCKLLMF	GILSFLTMVT 238
1a__mosc	DGAIA DWSTI	YLEDYWPTSP	QTAVLGYLAF	AVFMIVGRIS 260
1530mosc	DGAIAEWSKL	YLI RVEQVTD	QVGS LGYVAF	TLLMIAGRIS 278
1a__mosc	GDRIGSVIGD	RAMIAISGGL	MSAGMTIALS	IPFFAAALAG 300
1530mosc	GDRVKDAIGC	RALIAISGSL	ASAGMTTALF	MPSFAGKLAG 319
1a__mosc	FSLLGIGMAN	LVPIIFRNAS	EV..VPDGLG	TA FVSVC GYS 338
1530mosc	FALLGLGMAN	LVPIIFSEAA	SMNTVSKTVG	LT FVSVC GYS 358
1a__mosc	GFLVGPPIIG	GAAGAVCLNK	ALLIVSAVGA	TLAYVIGPFS 378
1530mosc	GFLVGPPIIG	ASRRPLGSGE	LCSSSF AVGV	IVACASVFRD 398
1a__mosc	RTQT PARFTL	MDLGMIQQ *396		
1530mosc	RHRSGQPEP *	407		

Fig. 3.6 PRETTYBOX from PILEUP showing extensive homology between deduced amino acids of *S. meliloti* strain L5-30 MosC and *R. leguminosarum* bv. *viciae* strain 1a MosC homologue.

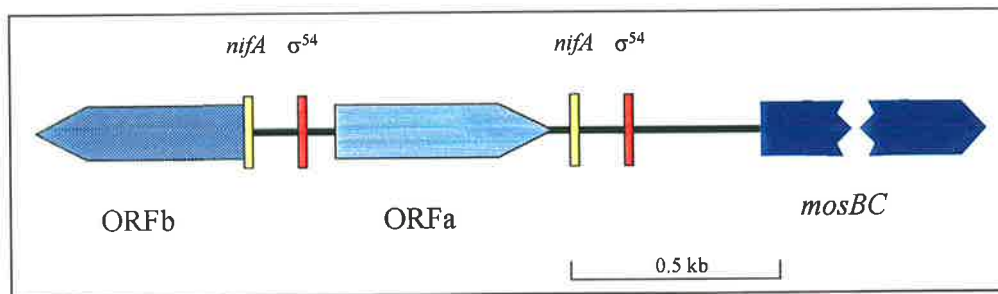


Fig. 3.7 Schematic representation of the ORFa, ORFb and the *nifA*/σ⁵⁴ recognition sites in the proximal and distal putative promoters upstream of the *mosBC* homologues.

Interestingly, the start of the MosC homologue corresponds with that of MosC2 of *S. meliloti* strain L5-30. In *S. meliloti* strain L5-30, MosC1 and MosC2 represent 2 inframe start sites. Either start site can be used for rhizopine biosynthesis (Grzanski, 1994). The characteristics of the ORFa, ORFb, MosB homologue and MosC homologue are shown in Table 3.2.

1 GGC GGATAACGGT GCGGATAACGCCACATCAGGCAGATTAACAAACGTACGGAGATAGCATTTTCGCGAGCACGGTGAACGTG 80
 81 TGCGGCAGAGTAATAAACAGCTCGCGGGGAAGACCGCCGTCATACAAATGAGGCCGAACGGGGCGAAAACCTCTCCAGA 160
 161 GATGGAAACCGCAGAACTCAAGATGGGTTTCGGACTTACGTGTGTGCGGGCAACTGAAGACAACAATCTGGACATGAGCA 240
nifA consensus IHF consensus σ consensus
 241 CGTCCAAT**TGTCGGACATCCGAC**ACCGCCAAGCGCGAAACTCCCCTTCATTTCCGATGCGGGTCTGGTT**TGGCAGCAGCAGT** 320
 321 **GCTAAGAGTAGGCCGAGGATGCACCAACCAGACAGAGGATAGATAGAGATGCCCTCAAGCCCGAACTTTCGGGTTCTCTC** 400
 401 TGCGGACCCAAAGCGAAACCCGGACGCTGCATCCGTGGCCGATTCTGACGGGAGCGATCTACCGGAAAGTCATCGTGTGTAT 480
 481 TTTTCGCTGGTGAATATCCGGCTCGGGTCGGGAAGTCATTAATGAGTTTGAAAAGAAGTGGAGAATTGGATCGGCAGTG 560
nifA consensus
 561 GGCATGTCGTCTCGACAAATAGCGGTACTGCCGCGCCGATGATAGCTCTACACGATCAT**TGTCCAATGTCCGACATAAGA** 640
 IHF consensus σ consensus
 641 CGCAGCTCAGTTCGAATTTCTCTTTATTT**TTCAATGGAAATCTGGT****TGGCAGCAGATGTGCT**CAGTATGGAAGCGGGCAA 720
 721 CATTGACGCCCGGGATGCACCAACCCACACGGGGTAAATTAATGCCATCAAAGCCAAATTTTCAGGTTCCCTCTGCGG 800
 M P S K P N F Q V P S A E
MosB homologue ➔
 801 AGCCAAACGGAACCCGGACGCTGCATCGGTGGCCGATTTTAACGGGAGCGAACTACCGGAAAGTCATCGCGTGTATTTCG 880
 P N G T R T L H R W P I L T G A N Y R K V I A C I S
 881 TCTGTGAACTTTCGGCTCGGGTCTGGAAGTCATTAATGAGTTTGAAACGACGGTGGAGAATTGGATCGGCGCGGTCA 960
 S G E L S G S G L E V I N E F E T T V E N W I G G G H
 961 TGTCTCTCGACAAATAGCGGCACTACCGCGCTGATGGTGGCCCTCTGACGCTCGGCGTTGGGCCGGAAGACGTTGTTC 1040
 V V S T N S G T T A L M V A L L T L G V G P E D V V L
 1041 TTGTGCTTCTACACCTGGAGCGCACTGCATTTGCAGGCACTTTATCGGCGCCATTCCCCGCTTCGTTGACATTGAT 1120
 V P S Y T W S A T A F A A T L I G A I P R F V D I D
 1121 AGCGACACCTACAATCTCTTACCGAGGCACTGGCTAAAGCGATTACCGCGGACGTCAGGCAATCATCGTCTTCATAT 1200
 S D T Y N L S T E A V A K A I T P D V K A I I V V H M
 1201 GCATGGCCTGAGCTGCGAAATGGACGAGATTACCGCCATTGCGCGGGAGAAAGGAGTTCCCATCATGAGGACTGCGCCC 1280
 H G L S C E M D E I T A I A R E K G V P I I E D C A Q
 1281 AGGCCCAACGGCGCACTCTATTAAGAAACAGTATGTGCGGACTCTCTCTGACATCGGCTGCTTCAGCATGCAGAAAGCAAG 1360
 A H G A L Y K K Q Y V G T L S D I G C F S M Q K S K
 1361 CATTTCTCCGCGCGGCGACGTTGGCTTCCTTGTGACGAGAAATGCCGCCCTTGCCCAAAGGCACGCGATATCTGCAACTT 1440
 H F S A G D G G F L V T R N A A L A Q K A R D I C N F
 1441 CGGCCTTCTACGCCAAAGCATAATATCACTTTGAGGAGGGGCTGCGGGAAGGATATGCCGTCTTTCGTGAGTGCAGGC 1520
 G L P T P K H N Y H F E E G L R E G Y A V F R E C E Q
 1521 AGATCGGCGGCAATTTAGGTAAACCCGCTTTTCGGCGGCTCTGGTGATCGATCAGTAGAGCATCTCGATCAACGGATA 1600
 I G G M F R L N P L S A A L V I D Q L E H L D Q R I
 1601 GAATGGCTCCAGGAGGCTATGAAGCCACTCGTTGCAGAGGCGGCACAAGTTCCCTTTCTCAAATCACCAGCCCGGCTC 1680
 E W L Q E A M K P L V A E A A Q V P F L K I T Q P G S
 1681 CCACCGAACGCATGTTTGGCACAAGATCCGCGTTCGATCTGACTATGGGCGAGTCGATCACTACGGGAGATCAATGGCGG 1760
 H R T H V W H K I R V G I D Y G A V D H Y G R S M A D
 1761 ACATCAGACAAGGGCTGCGCAACACTCTTGCTAGGAGCGGCATTCCGAGCACGCTGTGGACAGCACCTATACTCCGCTT 1840
 I R Q G L R N T L A R S G I P S T L W T A P I L P L
 1841 CAGAAGGTGTTACGGCCGATGCGGGCAGCATGACTGGACCAAACTGCGGGGCACCTTGCAATCGAAAACCTCGTTTCT 1920
 Q K V F R P Y A G T I D W T K T A G H L A I E N S F I
 1921 CGTTTTTCGACGAGAAGTACCCATTGATTGCCAGGATCCCTCAACCATGACGGCTCTCGTTTCGAAGCTCAGGCAAGTCT 2000
 V F D E K Y P L I A Q D P S T M T A L V S K L R Q V W
 2001 GGAGCGATTACTTTGCCTCATTAGAATAGACAAGCAACCCACGGGCTGATGGTTCTGGCAGCTCTACGCAGTCTCTGAT 2080
 S D Y F A S L E *
 2081 CATCTCGACAACGGAAGCGGAGTCTCCAATGAGTAACGCTCGCATATTTGTTCCCCAAAGCGAGTCATGGCGGTGGCTT 2160
 M S N A R I F V P Q K R V M A V A F
MosC homologue ➔
 2161 TCCAGTTCTTTATAACCGAGTGGTATATGGCACCTGGGCGACAAACATACTTCATATCCGACGGTCTGTTGGTCTGAGC 2240
 Q F F I N G V V Y G T W A T N I L H I R R S F G L S
 2241 GAGTTCCGGGCTCAGCATCGTGTTCCTAACAATGGGAGCCGACGGGTGTTGATCATGTGCTCACGGGTTATCTCGTGCA 2320
 E F G L S I V F L T M G A A A V L I M S L T G Y L V Q

```

2321  GATTTTCGGGAGCCGGCGGATATCGATCCTGTCCTATTCTTTCCAGCTGCGCTATGTTTGTGTTGTAGCTCCAA 2400
      I F G S R R I S I L S C L F F P A A L C F V F V A P N
2401  ACCACATCGCCCTCCTGGTCGGCGTCGGCCTGTTTCGGCGCCGCGAACGGCTCAATGGACGTTTCGATGAACGACCAGGGC 2480
      H I A L L V G V G L F G A A N G S M D V S M N D Q G
2481  ACCAATCTTGGTAAGAGAGGAAGCGATCCCGTCATATCGCTACTGCGATGGTTTCTCCAGCATTGGAGCGCTTGCGGGTTCT 2560
      T N L G K R G S D P V I S L L H G F S S I G A L A G S
2561  ACTTCTGTCCTACTGGTGGATTGGTAGAGGCTATCATCCGCTGGTCCAGGCCATTGCCTTGCTGGTTGGCACTCTAATGC 2640
      L L S Y W W I G R G Y H P L V Q A I A L L V G T L M L
2641  TTTCTTGGCGACTCTTCCACACCTTGATCTGGACGCTGATCGGAGCAGGATCAGTCTGGAGTTAGGGGCATTGGGCAAC 2720
      S W R L F P H L D L D A D R S R I S L E L G A L G N
2721  GGCAGGCTCATTTTGTCTCGGATTCCTGTCTTCTCACCATGATGAGCGATGGAGCAATCGCCGATTGGAGCAGGATCTA 2800
      G R L I L L G F L S F L T M M S D G A I A D W S T I Y
2801  TCTAGAGGACTACTGGCCGACGTCGCTCAGACTGCCGCTCTGGGTTACCTGGCCTTTGCGGTCTTTATGATCGTCGGCC 2880
      L E D Y W P T S P Q T A V L G Y L A F A V F M I V G R
2881  GGATCTCAGGAGACCGCATCGGAAGCGTTATCGGTGATCGCGCCATGATTGCGATTAGCGGCGGCCCTCATGTGCGCGGGC 2960
      I S G D R I G S V I G D R A M I A I S G G L M S A G
2961  ATGACAATAGCATTGTCAATCCCGTTTTTCGCAGCTGCGCTCGCGGGCTTTTCTTACTGGGCATAGGCATGGCTAACCT 3040
      M T I A L S I P F F A A A L A G F S L L G I G M A N L
3041  AGTGCCTATCATCTTCAGAAATGCAAGTGAGGTGGTGCCAGACGGTCTCGGGACCGCATTTGTGTGCGGTCTGCGGCTATT 3120
      V P I I F R N A S E V V P D G L G T A F V S V C G Y S
3121  CCGGATTTCTCGTCGGTCCACCCATCATTGGTGGCGCGGCCGCGCAGTGGGGCTCAACAAGCCTTGCTGATCGTCAGT 3200
      G F L V G P P I I G G A A G A V G L N K A L L I V S
3201  GCGGTGCGGAGCGACGCTGGCGTATGTCATCGGTCCTTTTCAAGGACGCAACGCCCGCCGCTTCACGCTGATGGCCTT 3280
      A V G A T L A Y V I G P F S R T Q T P A R F T L M A L
3281  GGGATGAGCCGCAATAAGCTTTTGTGCAAAAGGGAATATTCTATATGAAAAGGAGAGACGTACTAACCCTCATCGGGGC 3360
      G *
3361  GCTCGTTGCCGCCGGTCCGATCGGATGCAATCCCGCCTGTGGC 3403

```

Fig. 3.8 DNA sequence of *R. leguminosarum* bv. *viciae* strain 1a *mosBC* homologues with deduced amino acid sequence. Upstream of the *mosBC* homologue are two consensus sequences for *nifA* σ^{54} (in bold) and Integration host factor, IHF (underlined).

```

242 GTCCAATGTCGGACATCCGACACCGCCAAGCGCGAAACTCCCCTTCATTT 291
    |||||
621 GTCCAATGTCGGACATAAGACGCAGCTCAGTTCGAATTTCCCTCTTTATTT 670

292 CCGATGCGGGTCTGGTTGGCACGAGCAGTGCTAAGAGT..... 329
    |||
671 TCAATGGAATCTGGTTGGCACGAGATGTGCTCAGTATGGAAGCGGGCAA 720

330 .....AGGCCGAGGATGCACCAACCAGACAGAGGATAGATAGAGATGCCC 374
    |||
721 CATTGACGCCCGGGATGCACCAACCCACACGGGGGTAAAT.TAAATGCCA 769

375 TCAAGCCCGAACTTTTCGGGTCTCTCTGCGGACCCAAGCGAAACCCGGAC 424
    |||
770 TCAAAGCCAAATTTTCAGGTTCCCTCTGCGGAGCCAAACGGAACCCGGAC 819

425 GCTGCATCCGTGGCCGATTCTGACGGGAGCGATCTACCGGAAAGTCATCG 474
    |||||
820 GCTGCATCGGTGGCCGATTTTAACGGGAGCGAACTACCGGAAAGTCATCG 869

475 TGTGTATTTTCGCTGGTGAACATCCGGCTCGGGTCGGGAAGTCATTAAT 524
    |||||
870 CGTGTATTTTCGTCTGGTGAACTTTCCGGCTCGGGTCTGGAAGTCATTAAT 919

525 GAGTTTGAAAAGA.AGTGGAGAATTGGATCGGCAGTGGGCATGTCGTCTC 573
    |||||
920 GAGTTTGAAACGACGGTGGAGAATTGGATCGGCGGCGGTCATGTCGTCTC 969

574 GACAAATAGCGGTACTGCCGCGCCGATGATAGCTCTAC 611
    |||||
970 GACAAATAGCGGCACTACCGCGCTGATGGTGGCCCTCC 1007

```

Fig. 3.9 Sequence duplication in the upstream region of the *mosB* homologue between 242-611bp and 621-1007bp.

Hydrophobicity analysis of the deduced protein from the *mosC* homologue in *R. leguminosarum* bv. *viciae* strain 1a, using the ^{Kyte Doolittle program (1982)} DNA STRIDER program, showed a protein with 12 membrane spanning regions (Fig. 3.10) thus, suggesting a membrane transport protein. A comparison with the hydrophobicity pattern of other hypothetical transport proteins, for example, MosC from *S. meliloti* strain L5-30 and y4WD of *Rhizobium* species NGR234 reveal similar 12 membrane spanning regions (Fig. 3.10).

Comparing the proximal and distal putative promoters upstream of the *mosBC* homologous region in *R. leguminosarum* bv. *viciae* strain 1a with those from other

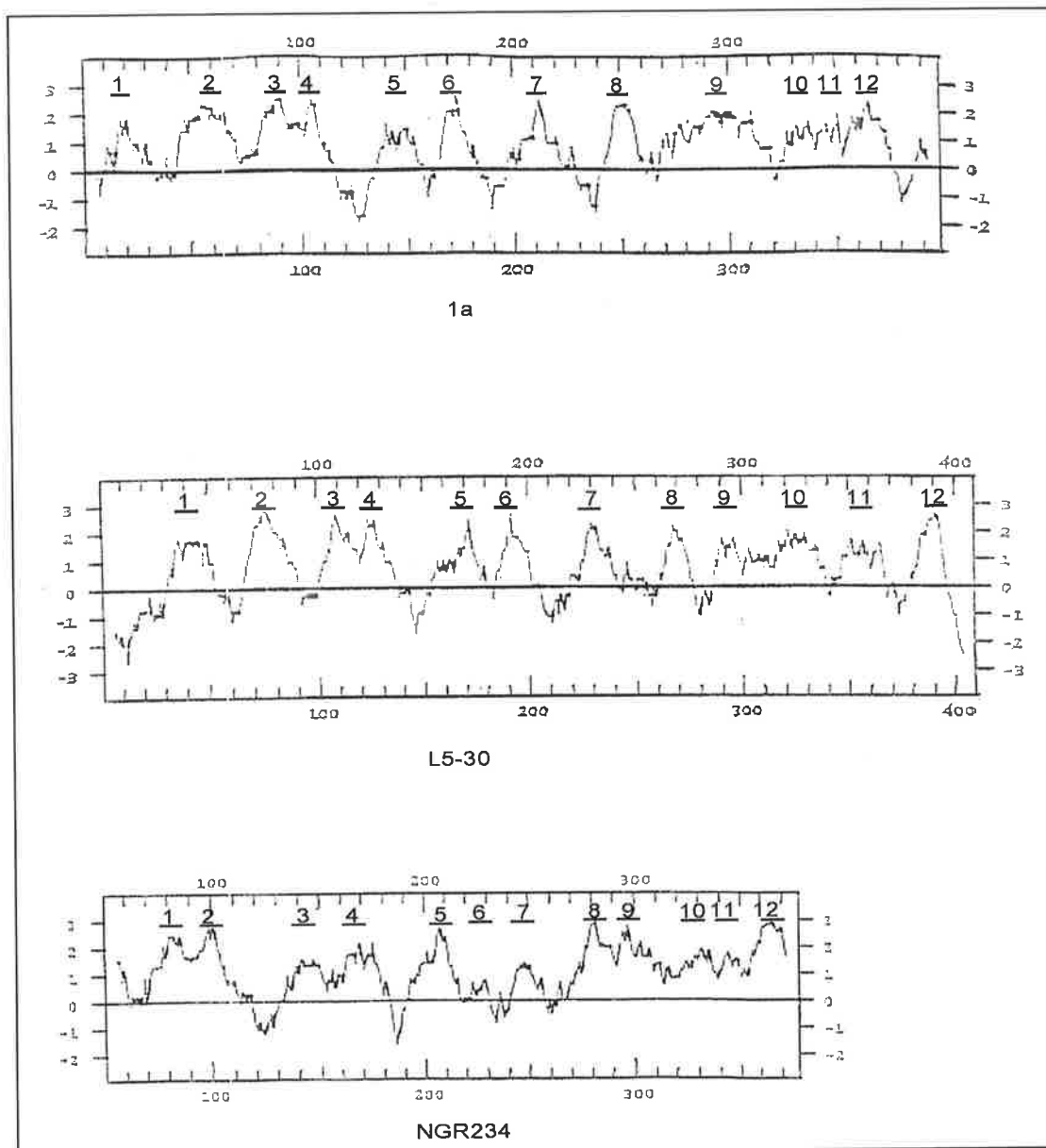


Fig. 3.10 Hydrophobicity plot of deduced protein sequence of the MosC homologue from *R. leguminosarum* strain 1a (1a), MosC from *S. meliloti* strain L5-30 (L5-30) and the hypothetical transport protein y4WD (NGR234) from *Rhizobium* species strain NGR234, derived from the program DNA Strider by the procedure of Kyte and Doolittle (1982)

bacterial species containing *nifA*/ σ^{54} recognition sequences revealed similar *nifA* motifs in all the motifs shown in Fig. 3.11. However, the conserved σ^{54} region varied slightly in the putative distal promoter of *R. leguminosarum* bv. *viciae* strain 1a *mosBC* homologue, showing **G**TGCT as opposed to TTGCT or CTGCT in the 5 base-conserved segment (see underlined bold base in Fig. 3.11). Moreover, in *Sm*-

mos and *Rm-nifH*, there are 79 to 84 bases between the *nifA* and σ^{54} motifs with leader sequences of about 70 bases, but in the *R. leguminosarum mosBC* homologue 5' region, the proximal "promoter" has only 51 bases between the *nifA* and σ^{54} motifs and a likely leader sequence of about 50 bases, while 44 bases separate the *nifA* and σ^{54} motifs in the distal "promoter" with a leader sequence of about 35 bases.

	<i>nifA</i>			σ^{54}			ATS	ALS		
<i>Kp-nifH</i>	TGT	N ₁₀	ACA	N ₉₄	CTGG T ATG	N ₄	CTGCA	N ₁₀	N ₃₀	ATG
<i>Sm-mos</i> (p)	TGT	N ₁₀	ACA	N ₈₄	CTGGC A CG	N ₄	TTGCA	N ₁₀	N ₆₉	ATG
<i>Sm-mos</i> (d)	TGT	N ₁₀	ACA	N ₈₇	CTGGC A CG	N ₄	TTGCA	N ₁₀	N ₁₂₁	ATG
<i>Rm-nifH</i>	TGT	N ₁₀	ACA	N ₇₉	CTGGC A CG	N ₄	TTGCA	N ₁₀	N ₇₀	ATG
<i>Tf-nifH</i>	TGT	N ₁₀	ACA	N ₁₀₁	CTGGC A CG	N ₄	TTGCA	N ₁₀	N ₆₆	ATG
<i>Bj-nifH</i>	TGT	N ₁₀	ACA	N ₆₉	TTGGC A CG	N ₄	TTGCT	N ₁₀	N ₄₈	ATG
<i>Rl-mosBC</i> (p)	TGT	N ₁₀	ACA	N ₅₁	TTGGC A CG	N ₄	CTGCT	N ₁₀	N ₅₀	ATG
<i>Rl-mosBC</i> (d)	TGT	N ₁₀	ACA	N ₄₃	TTGGC A CG	N ₄	G TGCT	N ₁₀	N ₃₅	ATG

Fig. 3.11 DNA sequence comparison of the *nifA* and σ^{54} motifs in promoters of *Klebsiella pneumonia nifH* (*Kp-nifH*), *S. meliloti* strain L5-30 *mos* proximal (*Sm-mos* [p]) and distal (*Sm-mos* [d]) promoters, Rm102F34 *nifH* (*Rm-nifH*), *Thiobacillus ferrooxidans nifH* (*Tf-nifH*), *Bradyrhizobium japonicum nifH* (*Bj-nifH*), and putative proximal (*Rl-mosBC*[p]) and distal (*Rl-mosBC*[d]) promoters of *R. leguminosarum* bv. *viciae* strain 1a *mosBC* homologues. The bold letters indicate the bases which show variation in the *nifA* and σ^{54} motifs of the bacterial species. ATS is the assumed transcription starting site. ALS is the assumed leader sequence.

In the promoters of *Klebsiella pneumonia nifH* (*Kp-nifH*) and *Bradyrhizobium japonicum nifH* (*Bj-nifH*), the IHF (integration host factor) sites [IHF is one of the regulatory proteins binding the minor groove of target DNA sequence, whose role includes site specific recombination, transcription, initiation of plasmid replication and gene expression (Friedman, 1988)] (see discussion) are in the forward direction but in the reverse direction in Rm102F34 *nifH* (*Rm-nifH*), *S. meliloti* strain L5-30 *mos*, (*Sm-mos*) and *R. leguminosarum* bv. *viciae* strain 1a homologous *mosBC* (*Rl-mosBC*) (Fig. 3.12).

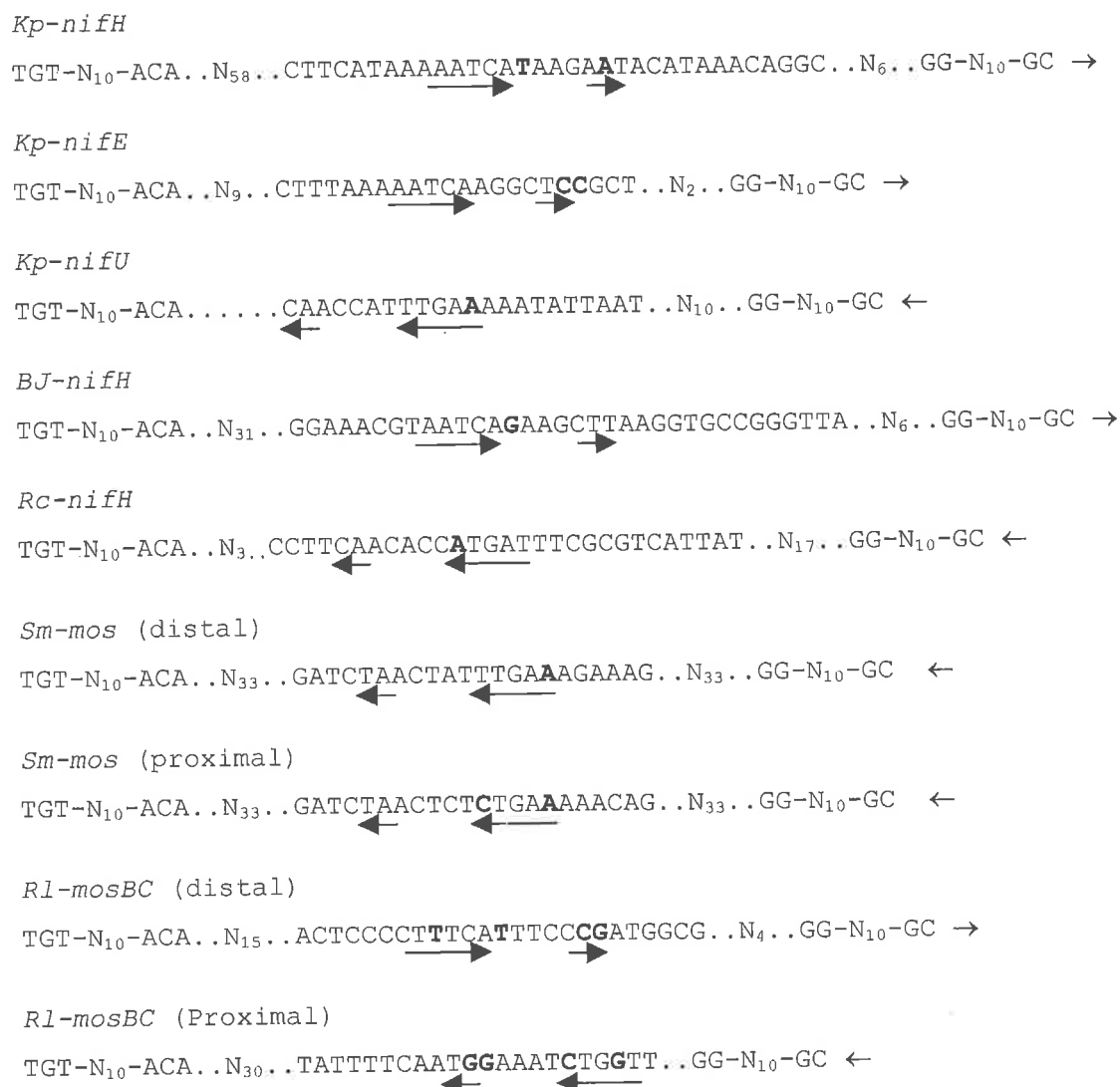


Fig. 3.12 Comparison of integration host factor (IHF) consensus DNA sequence (underlined) in the promoters of *Klebsiella pneumonia nifH* (*Kp-nifH*), *Klebsiella pneumonia nifE* (*Kp-nifE*), *Klebsiella pneumonia nifU* (*Kp-nifU*), *Bradyrhizobium japonicum nifH* (*Bj-nifH*), *Rhodobacter capsulatus nifH* (*Rc-nifH*), *S. meliloti* strain L5-30 *mos* (*Sm-mos* [distal and proximal]) and the putative promoters of *R. leguminosarum* bv. *viciae* strain 1a *mosBC* homologues (*Rl-mosBC* [distal and proximal]). The arrows indicate the 5'-3' direction of the IHF recognition sequence. The bold bases indicate variations from standard consensus sequence. Certain deviations and mismatches in IHF consensus sequence, 5'-W[a/t]ATCAA-N4-TTR[a/g] (Craig and Nash, 1984; Leong *et al.*, 1985) are not important for IHF binding. Binding is reduced by changes of W to C, of C at position 4 to T, G or A, of A at position 5 to T, G or C, of A at position 6 to C, or of T at position 12 to A (Hoover *et al.*, 1990).

3.3 Discussion

The RNA polymerase sigma factor (σ^{54} / σ^N /rpoN/NtrA/~~glhF~~^{glhF}) is involved in nitrogen regulation and bacterial differentiation (Gussin *et al.*, 1986) and is required in both free-living and symbiotic rhizobia. The σ^{54} regulated genes are characterized by a conserved motif, 5'-CTGGYAYR-N₄-TTGCA-3' located at -24 to -12 (or 5'-TGGCAC-N5-TTGCa/t-3' located between -26 to -11 { ± 1 bp}) from the start of the gene (Gussin *et al.*, 1986; Merrick, 1993; Wang *et al.*, 1998). The minimal σ^{54} consensus sequence is regarded as GG-N₁₀-GC (Drummond *et al.*, 1983). The characteristic minimal conserved *nifA* sequence, TGT-N₁₀-ACA (Cannon *et al.*, 1991) and σ^{54} recognition sequences were identified in the two putative promoters in the 5' region of the *mosBC* homologue *R. leguminosarum* bv. *viciae* strain 1a (Fig. 3.11), thus suggesting the *mosBC* homologue may be regulated by σ^{54} and NifA. This regulatory system is known to control symbiotic genes, such as the *nif* and *fix* operons. Similarly, the *mosABC* operon in *S. meliloti* L5-30 is regulated in the same manner (Murphy *et al.*, 1988).

The *mosA* homologue in *R. leguminosarum* bv. *viciae* strain 1a is not closely associated with the *mosBC* homologue^S and is located at least 34 kb away (Wexler, 1994). In *S. meliloti* strain L5-30, *mosABC* are arranged in an operon (Murphy *et al.*, 1993). This variation in arrangement of L5-30 *mos* genes and their 1a homologues may reflect regulatory differences in the 2 genera probably due to variation in evolutionary and genetic development.

BESTFIT analysis of the DNA sequence in the region surrounding ORFa with that of the *mosB* homologue region showed 82% similarity and identity (cf. regions from 242 to 611 bp with 621 and 1007 bp) (Fig. 3.9). In other words, ORFa represents a duplication of the 5' region and part of the ORF of the *mosB* homologue. The DNA homology between the two extends further than the amino acid homology as the absence of a base causes a frame shift in ORFa (Fig. 3.13). Further work would be required to establish the significance of this apparently reiterated promoter sequence and ORF. Rearrangements giving rise to deletion and duplication of large fragments

and single genes in pSym are common in rhizobia (Romero *et al.*, 1991). For example, duplicate copies of *nodD* (Honma and Ausubel, 1987; Rodriguez-Quinones *et al.*, 1987) and *nifH* (Quinto *et al.*, 1985) have been observed. Other reports indicated the presence of duplicated segments of symbiotic genes and promoter elements within the vicinity of functional symbiotic genes of *S. meliloti* (Better *et al.*, 1983).

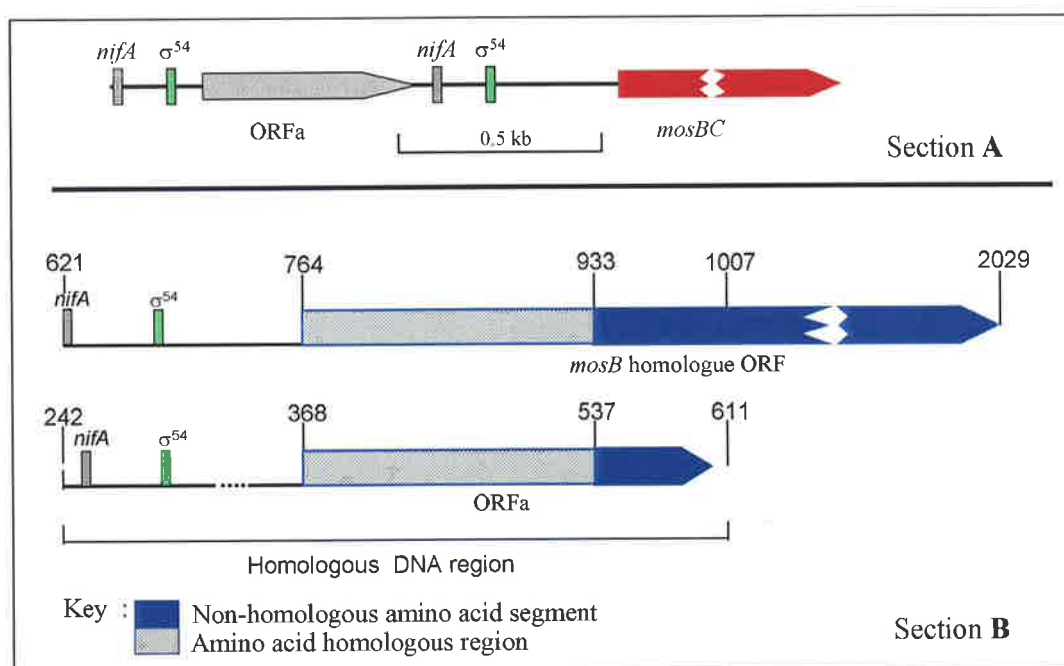


Fig. 3.13 The figure in the upper Section A is an overview of the expanded Section B below. Section B is a schematic representation of the DNA duplication in the upstream region (242-611 bp) and *R. leguminosarum* bv. *viciae* strain 1a *mosB* homologue (621-1007 bp). The dotted line represents a gap (between 329 and 330) in the DNA homology in the 5' region. DNA positions 764-2029bp and 368-611bp mark the ORFs (with large arrows) for *R. leguminosarum* bv. *viciae* strain 1a *mosB* homologue and ORFa respectively. The *nifA* and σ^{54} recognition motifs are shown upstream the ORFs. A frame shift occurred beyond the 537bp position on ORFa thereby terminating the amino acid homology. The thin line is the 5' region.

Within the *S. meliloti* *mos* region, there are reiterated fragments of *fixA*, *nifH*, *nifT* and *fixU*. Also, there is a duplication of the *nifA* and σ^{54} element upstream of the *mos* promoter (Murphy *et al.*, 1993). It is also likely that ORFa is a reiterated DNA segment of *nifR3* present in the *mosBC* homologue region.

Possible upstream IHF (integration host factor) binding sites have been identified in the putative promoters of *R. leguminosarum* bv. *viciae* strain 1a *mosBC* homologue [this work] (Fig. 3.12). Examination of previously published *S. meliloti* strain L5-30 *mosABC* proximal and distal promoters (Murphy *et al.*, 1993) also revealed likely IHF motifs (Fig. 3.12). The role of IHF in *E. coli* includes site specific recombination, transcription, initiation of plasmid replication and gene expression (Friedman, 1988). IHF is a small heterodimeric protein of two peptides encoded by the *himA* and *himD* (also called *hip*) genes (Nash and Robertson, 1981; Friedman, 1988) and is one of the regulatory proteins binding the minor groove of target DNA sequence, resulting in a sharp bend ($>140^\circ$) at the cognate site (Thompson and Landy, 1988; Robertson and Nash, 1988; Kosturko *et al.*, 1989). IHF binding sequences in *nif* promoter regulatory regions are similar to the defined consensus, 5'-W[a/t]ATCAA-N₄-TTR[a/g] (Craig and Nash, 1984; Leong *et al.*, 1985). However, binding sequences deviate in some instances eg. *Klebsiella nifH* promoter regulatory region has a 2bp deviation (Fig. 3.12). Nevertheless, it has been shown that certain mismatches are not important for IHF binding. Binding is greatly reduced by changes of W to C; of C at position 4 to T, G or A; of A at position 5 to T, G or C; of A at position 6 to C; or of T at position 12 to A (Hoover *et al.*, 1990). IHF binds to *nif* promoter regulatory regions in a variety of nitrogen fixing bacteria at a site between the promoter and the upstream binding site for NifA, bends the DNA adjacent to the promoter and greatly stimulates NifA-mediated activation of transcription from the *nifH* promoter eg. in *Klebsiella pneumoniae*, by facilitating the interaction between NifA and σ^{54} holoenzyme. (Hoover *et al.*, 1990; Santero *et al.*, 1992). IHF also stimulates NifA-mediated activation of transcription from the *nifH* and *nifD* promoters of *Bradyrhizobium japonicum* but has less of an effect on the *nifH* promoters of ~~*Rhizobium*~~^{*Sinorhizobium*} *meliloti* and *Thiobacillus ferrooxidans*, as stimulation is greatest at promoters that have weak binding sites for σ^{54} -holoenzyme in closed complexes (Santero *et al.*, 1992). Further, IHF has been shown to be directly involved in the transcriptional activation of the *p_L* promoter of phage λ and binds to two sites upstream of *p_L*, one proximal at ca. position -80 and one distal at about -180 (Giladi *et al.*, 1990). Promoter activity *in vivo* is increased four fold in IHF-plus

compared with the IHF-minus strains (Giladi *et al.*, 1992). In λ phage, IHF could stimulate p_L by bringing farther-upstream sequences into contact with the RNAP and/or by directly interacting with the enzyme that might stimulate formation of a closed complex (Pérez-Martín *et al.*, 1994). Activation of p_L requires an intact α subunit of RNA polymerase (RNAP) suggesting the requirement of additional protein-protein or protein-DNA interaction (Giladi *et al.*, 1992). Stimulation of p_L by IHF may involve the formation of a distorted DNA loop in the region between the sites bound by IHF and RNAP (Pérez-Martín *et al.*, 1994). The *E. coli* *ilvPg2* promoter is also regulated by IHF. Two upstream regions independently enhance *ilvPg2* transcription (Pagel *et al.*, 1992), the proximal (at ca. -50) spanning an intrinsic DNA bend and the more distal (at -90) containing an IHF site. The proximal activating sequence stimulates promoter activity 7-fold whereas the distal sequence stimulates activity about 4-fold, together both cause an increase of about 28-fold in promoter activity (Pagel *et al.*, 1992). In the *S. meliloti* strain L5-50 *mos* region the IHF sequences in the proximal and distal promoters fit the conserved sequences with slight variations (see bold letters in Fig. 3.12). However, these differences are reported to have no effect on IHF binding (Craig & Nash, 1984; Leong *et al.*, 1985). For *R. leguminosarum* bv. *viciae* strain 1a *mosBC*, the IHF sequences in the putative proximal and distal promoters also fit the consensus sequence with variations which also do not reduce binding of IHF (Fig. 3.12). However, the start of the IHF consensus sequence in the putative proximal promoter of *R. leguminosarum* bv. *viciae* strain 1a *mosBC* is adjacent to the σ^{54} conserved recognition motif and whether this will affect the binding and function of IHF is not known. Nevertheless, some identified IHFs are close to either the σ^{54} or *nifA* motif. For example, IHF consensus sequence in *Kp-nifE* is 5bp away from the σ^{54} recognition motif while in *Kp-nifU*, the IHF consensus sequence is adjacent to the *nifA* motif (see Fig. 3.12). The intervening sequence between the σ^{54} and *nifA* recognition sites of *Kp-nifE* (33bp) and *Kp-nifU* (35bp) are adequate for the bending required for transcriptional activation. It is therefore likely that the intervening sequence between σ^{54} and *nifA* motifs in *Rl-mosBC* [proximal] (43bp) and *Rl-mosBC* [distal] (51bp) (see Fig. 3.11) would permit the necessary bending for transcriptional activation in the putative promoters. Invariably, demonstration of physiological activity based on IHF binding to the putative IHF

consensus sequence in *Rl-mosBC* promoter region will be required to confirm their functionality.

The high level (89%) of DNA homology, between *S. meliloti* L5-30 *mosB* (Murphy *et al.*, 1993) and the *R. leguminosarum* bv. *viciae* strain 1a *mosB* homologue, and 72% for the deduced amino acid sequences as shown by BESTFIT analysis, suggest a close evolutionary relationship and possibly the production of a related protein product. BLAST analysis of the DNA sequence of the *R. leguminosarum* bv. *viciae* strain 1a *mosB* homologue reveals 52% identity and similarity (over 452bp) to *S. griseus* *strO* from the *sts* gene cluster (Ahlert *et al.*, 1997). Moreover, BLAST analysis of the amino acid translation of the *mosB* homologue (consisting of 422 amino acids) shares homology with *Zymomonas mobilis* histidinol phosphate aminotransferase (30-44% identity and 56-59% similarity over 57 amino acids) and *Bacillus stearothermophilus* aspartate aminotransferase (28-43% identity and 52-69% similarity over 77 amino acids). It is, thus, likely that the *mosB* homologue encodes an aminotransferase as it also shares appreciable homology with the *stsC* gene (see Fig. 3.15), which encodes a pyridoxal-phosphate-dependent aminotransferase in *Streptomyces griseus*. This protein catalyses the transfer of an amino group from L-glutamine to *scyllo*-inosose to produce *scyllo*-inosamine (Ahlert *et al.*, 1997), a likely step in rhizopine synthesis (see proposed rhizopine synthetic pathway in Fig. 3.14).

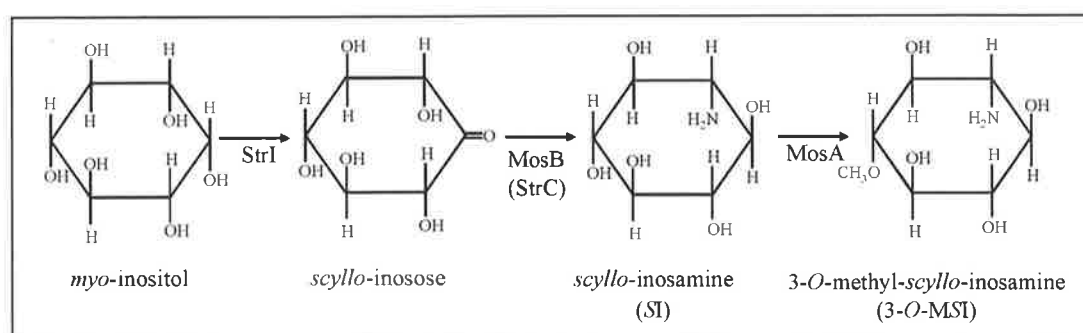


Fig. 3.14 Proposed rhizopine synthetic pathway in *S. meliloti* strain L5-30. StrI in *Streptomyces griseus* shares homology with dehydrogenases (Mansouri and Wolfgang, 1991) and is expected to have a homologous enzyme in *S. meliloti* (Murphy, pers. commun.).

Comparison of *S. meliloti* strain L5-30 MosB with the protein data bases indicated two domains; the first domain in the N-terminal end showed 58.6% identity over 70 amino acids with the *fixU* gene product, which has no known function (Murphy *et al.*, 1993). The second domain shares extensive homology with StsC, an aminotransferase from *Streptomyces griseus* (see below, Fig. 3.15). It is probable that the extra 83 amino acids in the 5' region of *S. meliloti* strain L5-30 MosB relative to *R. leguminosarum* bv. *viciae* strain 1a MosB homologue (Fig. 3.15) is non-functional while the second *S. meliloti* strain L5-30 MosB domain which aligns with the *R. leguminosarum* bv. *viciae* strain 1a MosB homologue is the functional domain.

During the synthesis of *scyllo*-inosamine in *Streptomyces griseus*, the α -amino group transfer from glutamine to *scyllo*-inosose has been shown to require pyridoxal phosphate which is apparently hydrolysed by StrO, thus releasing energy for the amino transfer reaction and the synthesis of *scyllo*-inosamine (Walker and Walker, 1969). The actual amino transfer reaction is performed by L-glutamine:*scyllo* aminotransferase encoded by *stsC* (Ahlert *et al.*, 1997). *S. meliloti* strain L5-30 MosB and the *R. leguminosarum* bv. *viciae* strain 1a MosB homologue have extensive homology to StsC (see PILEUP analysis in Fig. 3.15). StsC belongs to a super family of secondary metabolic aminotransferases (SMATS), which have the consensus sequence G-(D,V,T)-(E,V,Q)- $\times$ 76-84-(E,Q,R)-D- $\times$ 10-12-(G,D,Q)- $\times$ 3-5-G- $\times$ 8-10-S- $\times$ 4-K- $\times$ 4-5-(G,A,L,F)-(E,D,Q)-G-G (Ahlert, *et al.*, 1997). *S. meliloti* strain L5-30 MosB has this consensus, [G-D-V- $\times$ 77-E-D- $\times$ 10-G- $\times$ 3-G- $\times$ 8-S- $\times$ 4-K- $\times$ 4-G-D-G-G]. This together with amino transfer being a likely step in rhizopine biosynthesis strongly suggests that L5-30 MosB is a SMATS protein. The *R. leguminosarum* bv. *viciae* strain 1a *mosB* homologue consensus, [E-D-V- $\times$ 77-E-D- $\times$ 11-Q- $\times$ 2-G- $\times$ 8-S- $\times$ 4-K- $\times$ 4-G-D-G-G] is a closely related sequence. Slight alterations in the conserved region exist. These are shown as underlined and in bold letter; i.e. E for G, and the 2 amino acids between Q and G instead of 3-5 amino acids in the above sequence. Although, the *R. leguminosarum* bv. *viciae* strain 1a *mosB* homologue does not have exact SMATS consensus sequence, the close sequence similarity suggests that this protein may also have an aminotransferase function.

mosbrle	..VTIRINGD	VLSAYIPKED	LEETIISVAN	EDLWGGSELL	0
mosbrme	MKV TIRINGD	VLSAYIPKED	LEETIISVAN	EDLWGGSELL	40
stscpep					0
mosbrle	..NGWRLLALPH	LPQEARLPVT	VEARRNFGRL	DRGAHEKPD	0
mosbrme	..NGWRLLALPH	LPQEARLPVT	VEARRNFGRL	DRGAHEKPD	30
stscpep					0
mosbrle	..MPSKPNF	QVPSAEPNGT	RTLHRWPILT	GANYRKVIAC	37
mosbrme	MTRIGETISS	QDAAMLAEEN	QKMHPWPALT	RTAYEDVAAC	120
stscpep	..MDSSL	ATSGGPRLSN	REWPRWPQPG	DRALKSLEDV	35
mosbrle	ISSGELS.GS	GLEVTNEFET	T...VENWI	GGGHVVSTNS	72
mosbrme	ISSGELS.GS	GLGIINAFER	R...MEEWI	GGGYVVSASS	155
stscpep	LTSGRWTISC	AYQGRDSYER	QFASAFADYC	GSAMCVPIST	75
mosbrle	GTTALMVALL	TLGVGPEDVV	LVPSYTW SAT	AFATLLIGAI	112
mosbrme	GTAALTVALI	ALGTEOPGDVV	LVPSYTW SAT	ALAPLLIGAI	195
stscpep	GTAALTALE	ACGVGAGDEV	IVPGLSWVAS	ASAVLGINAV	115
mosbrle	PRFVDIDSDT	YNLSTEAFAK	AITPDVKAIT	VVHMHGLSCB	152
mosbrme	PRFVDIDPNS	YNISPTALAA	AITPDVKAIT	VVHMHGLSCD	235
stscpep	PVLVDVDPAT	YCLDPAATEA	AITERTRAIT	VVHAYSAAVD	155
mosbrle	MDEITATARE	KGVPIIEDCA	QAHGALYKKQ	YVGTLSIDIC	192
mosbrme	MDEITICHARE	QGLAVIEDCA	QAHGALYKGO	HVGLLSIDIC	275
stscpep	LDALLDIARR	HGLEPLIEDCA	HAHGAGFERGR	PVGAHGAGAGV	195
mosbrle	FSMQSKKHLS	AGDGGFLVTR	NAALAQKARD	ICNFGLPTPK	232
mosbrme	FSMQSKKHLS	AGDGGFMVTR	DPTLAQKMRD	ICNFGLPTPK	315
stscpep	FSMQGSKLLT	CGEGGALVTD	DADVALLRAEH	LRADGRVVR	235
mosbrle	HNYHFEEGLR	EGYAVFRECE	OIGGMFRLNP	LSAALVIDQL	272
mosbrme	ANYRFDEEVV	DGYAVFRECE	OIGGMFRLQP	MSAALVMHQL	355
stscpep	EPVGVGEMEL	DETGRNMGSN	AC...LSE	PHAAVLLDQL	270
mosbrle	EHLDRQRIEWL	QEAMKPLVAE	AAQVPFL..K	ITQPGSHRTH	310
mosbrme	EHLDRQRIEWL	QAMEPLVEE	SAKIPFF..K	ITRSHVDRTH	393
stscpep	ELLDGQNAARR	TRAADHLLTDR	LSELGMTAQA	TAPGTTARAY	310
mosbrle	VWHKIRVGID	YGAVDHYGRS	MADIRQGLRN	TLARSGTIPST	350
mosbrme	VWHKIRVGID	YAAVDYFGRS	MAEIRRGRLN	SLAERGISTST	433
stscpep	YRYLVRLPDE	VLAAPVERF	AHALTAE...	GEAVTQTHRP	348
mosbrle	LWTAPILPLQ	KVFRPYAGTI	DW.....	...TKTAG	377
mosbrme	LWTAPILPLQ	TAFRPYAGVV	EW.....	...TKSAG	460
stscpep	LNDNP.LNRP	SSRRKREATDA	RYLERVDP SR	FDLPAAKRAH	387
mosbrle	HLAIENSFIV	FDEKYPLIAQ	DPSTMTALVS	KLRQVWSDYF	417
mosbrme	HLAIENSFIV	FDENYPLIAQ	EPAKMAELVV	KLQQAWEHF	500
stscpep	ESVVSFSHEV	LILAPLDAIDD	IARAFRKVL D	NVRSVSR...	424
mosbrle	ASLE*...421				
mosbrme	TSLARGK*507				
stscpep	...424				

Fig. 3.15 PILEUP/PRETTYBOX of the deduced amino acids of *R. leguminosarum* bv. *viciae* strain 1a MosB homologue (mosbrle), *S. meliloti* strain L5-30 MosB (mosbrme) and *Streptomyces* sp. StsC (stscpep). Underlined portion covers the consensus sequence for secondary metabolic aminotransferases (SMATS). The region of FixU homology in *S. meliloti* strain L5-30 MosB is shown (1-70 amino acids).

Hydrophobicity analysis of the amino acid translation of the *mosC* homologue (393 amino acids) in *R. leguminosarum* bv. *viciae* 1a revealed 12 membrane spanning regions. Comparing the hydrophobicity patterns of the *R. leguminosarum* strain 1a MosC homologue with those of other hypothetical transport proteins such as *S. meliloti* MosC (Murphy *et al.*, 1993) and *Rhizobium* species strain NGR234 y4WD (Freiberg *et al.*, 1997), showed the same number of membrane spanning regions (Fig. 3.10). Such regions are common in sugar transporting proteins. In addition, the *R. leguminosarum* strain 1a MosC homologue shares 53% identity and 68% similarity with *S. meliloti* strain L5-30 MosC and 24-38% identity and 43-65% similarity over 7 regions of 384 amino acids with NGR234 y4WD. The *R. leguminosarum* bv. *viciae* strain 1a MosC homologue also shares homology with other transporter proteins. These include, the *Salmonella typhimurium* hexose phosphate transport protein (25-53% identity and 42-84% similarity over 5 regions of 116 amino acids) and hypothetical metabolite transport protein in *E. coli* [23% identity and 43% similarity over 71 amino acids] (Blattner *et al.*, 1997; Island *et al.*, 1992). Although the amino acid homologies are apparently not high, the hydrophobicity nature of many of the transport proteins are similar.

Examination of the consensus sequence of sugar transport proteins (STPS), [LIVMF]-x-G-[LIVMFA]-x(2)-G-x(8)-[LIFY]-x(2)-[EQ]-x(6)-[RK], revealed the presence of a closely related amino acid sequence in *R. leguminosarum* bv. *viciae* strain 1a MosC homologue, I-x-G-V-x(2)-G-x(8)-I-x(2)-**S**-x(6)-**G**. The variations from the STPS consensus are underlined bold letters, i.e. S for E or Q, and G for R or K. Although the sequence present in the *R. leguminosarum* bv. *viciae* strain 1a MosC homologue deviates slightly from the exact STPS consensus (which are known to have some exceptions), it may be similar enough to suggest that the MosC homologue belongs to the large STPS family. However, it would require a demonstration of the physiological function of *R. leguminosarum* bv. *viciae* strain 1a MosC to confirm that it belongs to this group of proteins.

Chapter 4

**Analysis of *R. leguminosarum* bv. *viciae*
strains for *mos* gene homologues and
complementation of DapA⁻ *E. coli* with MosA**



THE UNIVERSITY OF ADELAIDE

CHAPTER 4

Analysis of *R. leguminosarum* bv. *viciae* strains for *mos* gene homologues and complementation of DapA⁻ *E. coli* with MosA

4.1 Introduction

Rhizobium leguminosarum bv. *viciae* nodulates legumes within the cultivars *Pisium sativa*, *Lathyrus* sp., *Lens esculenta*, *Vicia faba*, *Vicia sativa* and some other *Vicia* spp. Two strains of *R. leguminosarum* bv. *viciae* were previously shown to produce unusual cyclitols in pea nodules (Skøt and Egsgaard, 1984). One of these, *R. leguminosarum* bv. *viciae* strain 1a produced rhizopine (Wexler *et al.*, 1995). Moreover, further work identified other rhizopine producing *R. leguminosarum* bv. *viciae* strains. Rhizopine producing strains of *R. leguminosarum* bv. *viciae* identified so far are 1a, CC310, M1201, M1325, P314, P444, SA1, Sp18, and Sp75 (Wexler *et al.*, 1995).

Ecological and genetic studies of *S. meliloti* rhizopine strains suggest that rhizopines are important in the *Rhizobium*-legume symbiosis in both annual and perennial *Medicago sativa* legumes (Wexler *et al.*, 1995). A survey of 332 rhizobia from around the world for rhizopine producing strains revealed that 11% of *S. meliloti* and 12% of *R. leguminosarum* bv. *viciae* isolates synthesise rhizopines as well as catabolise them. Distinct from these strains are 2 out of 9 *Rhizobium* sp. (Phaseolus) that catabolised rhizopine but did not synthesise it. Interestingly, no rhizopine strains were found among *R. leguminosarum* bv. *trifoli* and bv. *phaseoli*, *R. elti* and *R. tropicii* (Wexler *et al.*, 1995). With the exception of the 2 *Rhizobium* sp. (Phaseolus) strains, all rhizobial strains which synthesise the rhizopine, 3-*O*-methyl-scyllo-inosamine (3-*O*-MSI), (including strain Rm2203 which produces scyllo-inosamine [SI]) can catabolise both 3-*O*-MSI and SI (Saint *et al.*, 1993, Wexler *et al.*, 1995). Other bacteria capable of catabolising but not synthesising rhizopines were recently identified within the genera, *Arthrobacter*, *Sinorhizobium*, *Pseudomonas*, *Aeromonas* and *Alcaligenes* (McSpadden Gardener *et al.*, 1998). Amongst rhizopine strains of *S. meliloti* and *R. leguminosarum* bv. *viciae*, rhizopine

synthesis (*mos*) genes show greater sequence divergence than rhizopine catabolism (*moc*) genes, although less divergence was observed in the *mos* and *moc* genes of *R. leguminosarum* bv. *viciae* than in *S. meliloti* (Wexler *et al.*, 1995).

Studies with *S. meliloti* strains show that rhizopine strains, unlike non-rhizopine strains, catabolise rhizopines and have a nodulation advantage over non-rhizopine strains (Murphy & Saint, 1992). However, the mechanism of rhizopine action in enhancing nodulation is not yet known. The rhizopine synthesis genes in *S. meliloti*, the *mos* locus, consists of four open reading frames (ORFs), viz., ORF1, *mosA*, *mosB* and *mosC* arranged in an operon structure (Murphy and Saint, 1992). In *R. leguminosarum* bv. *viciae* 1a, the *mosA* homologue is at least 34 kb away from the *mosB* homologue (Wexler, 1994) and only *mosB* and *mosC* homologues form a likely operon (this work, Chapter 3). The *mosA* homologue in *R. leguminosarum* bv. *viciae* strain 1a shares over 87% homology with *mosA* in *S. meliloti* strain L5-30 and extensive homology with *dapA*, which encodes dihydropicolinate synthetase (DHDPS) in *E. coli* (Bahar, 1997). The *dapA* gene is a unique enzyme in the lysine biosynthetic pathway of procaryotes, some *Phycomycetes* and higher plants, that catalyses the condensation of pyruvate and L-aspartate β -semialdehyde (L-ASA) (Yugari and Gilvergi, 1965), to L-2,3,-dihydrodipicolinate, which is then converted by dihydrodipicolinate reductase to L-2,3,4,5-tetrahydrodipicolinic acid (THDA) (Blingkling *et al.*, 1997). The amino acid, *meso*-diaminopimelate (*m*-DAP) which is synthesised from THDA via one of three different pathways in bacteria, is a precursor for the synthesis of lysine or peptidoglycan (Scapin *et al.*, 1995 [see Fig. 4.1]; Wehrmann *et al.*, 1998). DAP is an essential constituent of the peptidoglycan layer within bacterial cell walls which are necessary for the survival of vegetative bacteria. Nevertheless, the significance of the extensive homology between *dapA* and *mosA* is not clear, particularly as they seem to be involved in different physiological reactions.

Previously, the *mosA* and *mosBC* homologous regions of *R. leguminosarum* bv. *viciae* strain 1a (identified from a total DNA library) were cloned into separate compatible plasmids and both transferred to a non-rhizopine strain (although the

mosC homologue was part of the transfer, it had not been located at the time but was shown to be linked to the *mosB* homologue in this work [see Chapter 3]), but no rhizopine was produced. This, together with the lack of linkage of *mosA* to *mosBC*

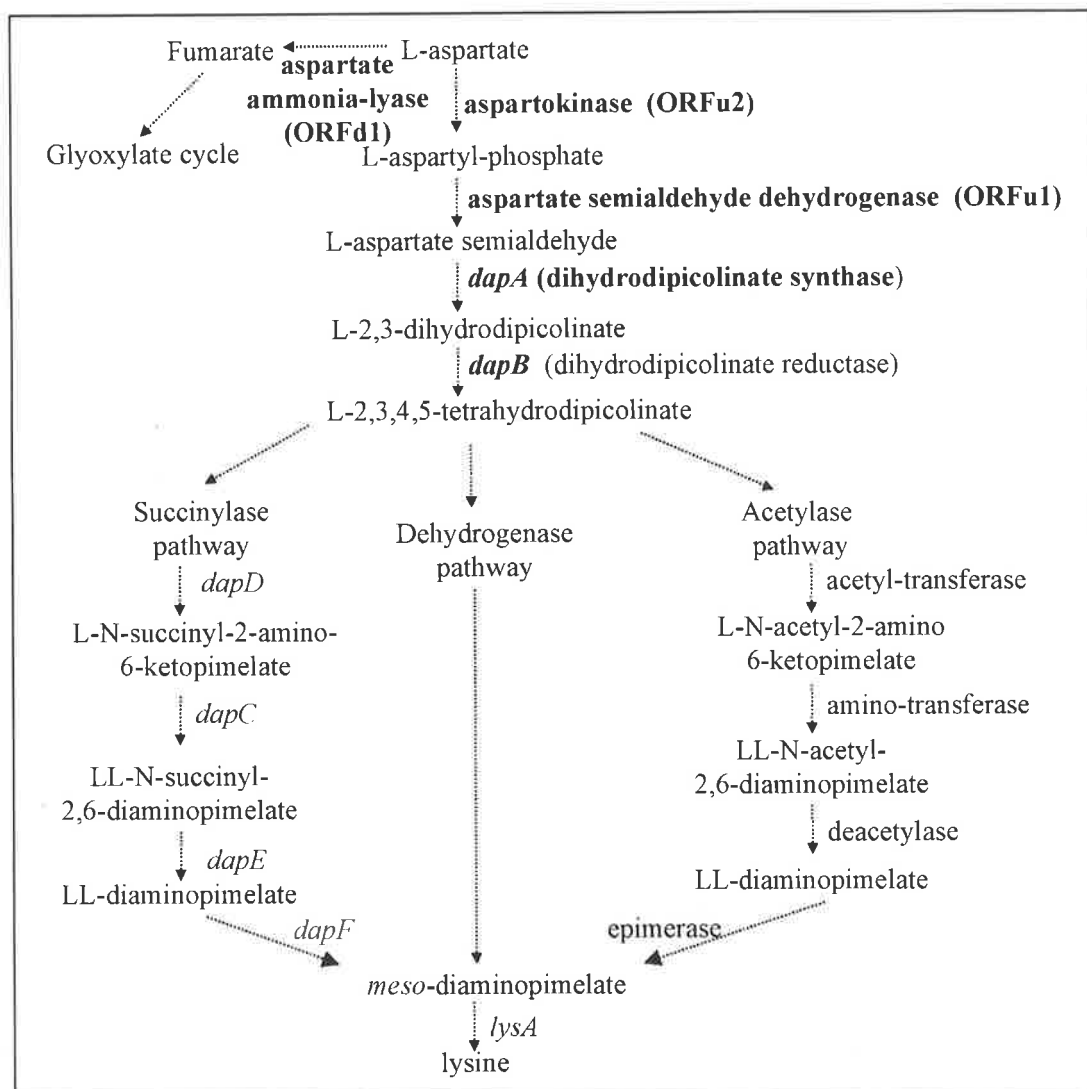


Fig. 4.1 Three pathways (viz., succinylase pathway, acetylase pathway and dehydrogenase pathway) for the biosynthesis of *meso*-diaminopimelate and lysine in bacteria (Scapin *et al.*, 1995). The enzymes in bold share homology with adjacent ORFs in parenthesis (see Fig. 4.11).

homologues in this strain and the similarity between *mosA* and *dapA* DNA sequence, raise the possibility that the *mosA* homologue may indeed be a *dapA* gene in *R. leguminosarum* bv. *viciae* strain 1a. Hence, the presence of the *mosA* homologue in

rhizopine and non-rhizopine strains was examined. As well, the DNA sequence of the regions upstream and downstream of the *mosA* homologue were determined. Further aims of this Chapter are to test the ability of MosA to complement a DapA⁻ *E. coli* strain and comparatively analyse the *R. leguminosarum* bv. *viciae* strain 1a *mosA* homologue and *S. meliloti* strain L5-30 *mosA*. Also, site directed mutagenesis of the *mosA* homologue was initiated to determine the physiological function(s) of the *mosA* homologue.

Moreover, previous work by Wexler (1994) identified a *mosB* homologous region (using a non-specific 5.3 kb fragment probe) in rhizopine strains but did not identify a *mosC* homologue. Therefore, the presence of the *mosB* and *mosC* homologues (using specific probes), in rhizopine and non-rhizopine strains were also investigated.

4.2 Results

Table 4.1 Bacterial strains used

Bacteria	Description	Source
<i>Agrobacterium radiobacter</i> strain K749	C58 pAt ⁻ , pTi ⁻ , Rif ^r , Sr ^r , Sp ^r	Donner <i>et al.</i> , 1993.
<i>Agrobacterium radiobacter</i> strain K749pSym::Tn5	pSym1a::Tn5-mob	M. Wexler, unpublished
<i>Rhizobium leguminosarum</i> strain 1a	Mos ⁺ , Moc ⁺ , produces 3-O-MSI	Jensen, 1987.
<i>R. leguminosarum</i> strain 8401	bv. <i>phaseoli</i> strain 8002 cured of its pSym, Sr ^r , Nod ⁻	Lamb <i>et al.</i> , 1982.
<i>R. leguminosarum</i> strain 8401(pRL1)	Mos ⁻ , Moc ⁻ , Sr ^r , Rif ^r .	Downie <i>et al.</i> , 1983.
<i>R. leguminosarum</i> strain L165	Mos ⁻ , Moc ⁻	See Wexler <i>et al.</i> , 1995.
<i>R. leguminosarum</i> strain M1201	Mos ⁺ , Moc ⁺ , (3-O-MSI)	See Wexler <i>et al.</i> , 1995.
<i>R. leguminosarum</i> strain M1325	Mos ⁺ , Moc ⁺ , (3-O-MSI)	See Wexler, 1994.
<i>R. leguminosarum</i> strain NA533	Mos ⁻ , Moc ⁻	See Wexler <i>et al.</i> , 1995.
<i>R. leguminosarum</i> strain P135	Mos ⁻ , Moc ⁻	G. Laguerre, Institut National de la Recherche Agronomique, Montpellier Cedex France.
<i>R. leguminosarum</i> strain P314	Mos ⁺ , Moc ⁺ , (3-O-MSI)	See Wexler <i>et al.</i> , 1995.
<i>R. leguminosarum</i> strain P342	Mos ⁻ , Moc ⁻	G. Laguerre, Institut National de la Recherche Agronomique, Montpellier Cedex France.
<i>R. leguminosarum</i> strain SA1	Mos ⁺ , Moc ⁺ , (3-O-MSI)	See Wexler <i>et al.</i> , 1995
<i>R. leguminosarum</i> strain Sp18	Mos ⁺ , Moc ⁺ , (3-O-MSI)	See Wexler, 1994.
<i>R. leguminosarum</i> strain Sp59	Mos ⁻ , Moc ⁻	J.P.W. Young, Dept of Biol. Univ of York, Heslington, York, UK.
<i>R. leguminosarum</i> strain Sp45	Mos ⁻ , Moc ⁻	See Wexler <i>et al.</i> , 1995.

<i>R. leguminosarum</i> strain Sp75	Mos ⁺ , Moc ⁺ , (3-O-MST)	See Wexler, 1994.
<i>R. leguminosarum</i> strain WSM1015	Mos ⁻ , Moc ⁻	See Wexler <i>et al.</i> , 1995.
<i>E. coli</i> CGSG 4548	<i>dapA16</i> , <i>thi-1</i> , <i>relA1</i> , λ^- , <i>spoT1</i>	Bukhari and Taylor, 1971. Obtained from Nicholas Toro, Granada, Spain.
<i>E. coli</i> DH5 α	F ⁻ (ϕ 80 <i>dlacZ</i> λ M15) <i>recA</i> <i>endA1</i> <i>gyrA96</i> <i>thi-1</i> <i>hsdR17</i> (r _k ⁻ m _k ⁺) <i>supE44</i> <i>relA1</i> <i>deoR</i> λ (<i>lacZYA-argF</i>) U169	Sambrook <i>et al.</i> , 1989
<i>E. coli</i> JM109	F' <i>traD</i> 36 <i>lacI</i> ^q Δ (<i>lacZ</i>)M15 <i>proA</i> ⁺ <i>B</i> ⁺ /el4 ⁻ (McrA ⁻) Δ (<i>lacproAB</i>) <i>thi</i> <i>gyr96</i> (Nal ^r) <i>endA1</i> <i>hsd17</i> (r ⁻ km ⁺ k) <i>relA1</i> <i>supE44</i> <i>recA1</i>	Yanisch-Perron <i>et al.</i> , 1985
<i>E. coli</i> S17-1	RP4 2-Tc::Mu-Km::Tn7; 294 <i>recA</i> , Tp ^r , Sm ^r .	Simon <i>et al.</i> , 1983
PM2335	<i>R. leguminosarum</i> strain 1a, Rif ^r	Wexler, 1994
PM2617	<i>R. leguminosarum</i> strain 1a transconjugant carrying pSUP202-0.73 kb 1a <i>mosA</i> , Rif ^r , Cm ^r	This work

Table 4.2 Plasmids used

Plasmids	Description	Source
pPM1089	3.4 kb <i>EcoR</i> I <i>mos</i> fragment in pACYC 184, containing <i>mosA</i> from <i>S. meliloti</i> strain L5-30.	Murphy <i>et al.</i> , 1993
pPM1161	22.9 kb partial <i>Hind</i> III fragment containing a region homologous to L5-30 <i>mosA</i> gene in pVK102, Tc ^r	Wexler, 1994
pPM1316	11 kb <i>Eco</i> RI fragment from pPM1161, homologous to L5-30 <i>mosA</i> in pUCD2608, Km ^r	Bahar, 1997
pPM1260	1.4 kb <i>Hind</i> III <i>mos</i> hybridising fragment in pGEM7Zf(+), Amp ^r	Wexler, 1994
pPM1324	5 kb <i>Bgl</i> II fragment from pPM1316, homologous to <i>S. meliloti</i> strain L5-30 <i>mosA</i> , in pKT231, Km ^r	Bahar, 1997
pPM1339	5.4 kb <i>Sac</i> I fragment from pPM1316, homologous to L5-30 <i>mosA</i> in pGEM-5Zf(+), Amp ^r	Bahar, 1997
PM1373	0.73 kb <i>Bam</i> H I <i>mosA</i> homologue fragment from <i>R. leguminosarum</i> bv. <i>viciae</i> strain 1a cloned into pSUP202, in <i>E.coli</i> JM109.	This work
pSUP202	pBR325 Mob Ap ^r , Tc ^r , Cm ^r	Simon <i>et al.</i> , 1983

4.2.1 Screening for a *mosA* homologue in *R. leguminosarum* bv. *viciae* strains.

Total DNA extracted from *R. leguminosarum* bv. *viciae* strains (Section 2.2.2) was digested with *Nco* I (which cuts within 7bp of the 5' region of the *Hind* III 1.4 kb probe fragment and outside of this region at the 3' end, see Fig. 4.2) and probed with a DIG labeled 1.4 kb *mosA Hind* III fragment (containing the *mosA* homologue and part of the upstream ORF, see Fig. 4.2). This revealed strong single hybridising bands of 3.0 kb in the rhizopine strains (1a, M1201, P314 and Sp18). In the non-rhizopine strains, a 6.0 kb band was observed in strain Sp45 and 5.7 kb bands in strains 8401(pRL1) and NA533 (Fig. 4.3).

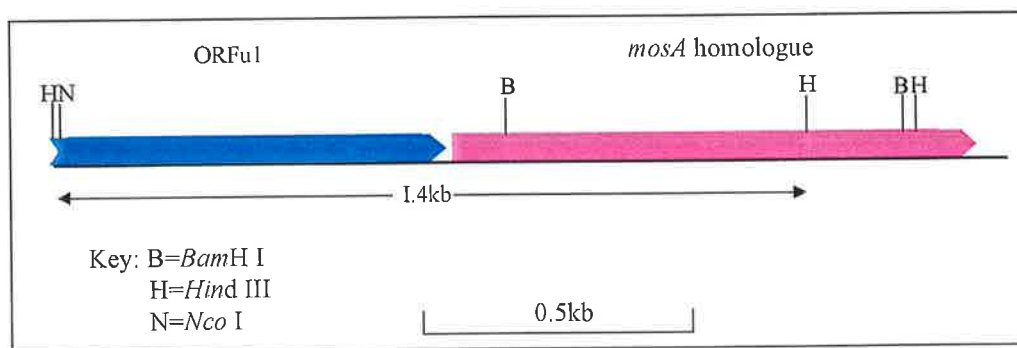


Fig. 4.2 Selected restriction enzyme map of *R. leguminosarum* strain 1a *mosA* and 0.74 kb upstream region (an incomplete ORFu1).

In order to establish whether, (a) the hybridisation observed with the non-specific 1.4 kb *mosA* homologue probe in non-rhizopine strains was due to the 0.74 kb upstream sequence [see Fig. 4.2] and (b) the *mosA* homologue is located on pSym as is the case with the *mos* genes in *S. meliloti* strain L5-30 (Murphy *et al.*, 1987), a DIG labeled internal *Bam*H I 0.73 kb *mosA* homologue fragment (Fig. 4.2) was used as a probe. Total DNA of *R. leguminosarum* rhizopine strains (1a, M1201, M1325, P314, SA1, Sp18 and Sp75), non-rhizopine strains (8401, 8401(pRL1), L165, NA533, P135 and P342), *Agrobacterium radiobacter* strain K749pSym::Tn5 and *A. radiobacter* strain K749 were digested with *Pst* I (which has no restriction site within the 0.73 kb strain 1a *mosA* homologous region) and probed with the labeled 0.73 kb *Bam*H I fragment (see Fig. 4.2), using the Southern blotting technique.

Examination of the DIG autoradiogram revealed no hybridisation to non-rhizopine strains 8401, 8401(pRL1),

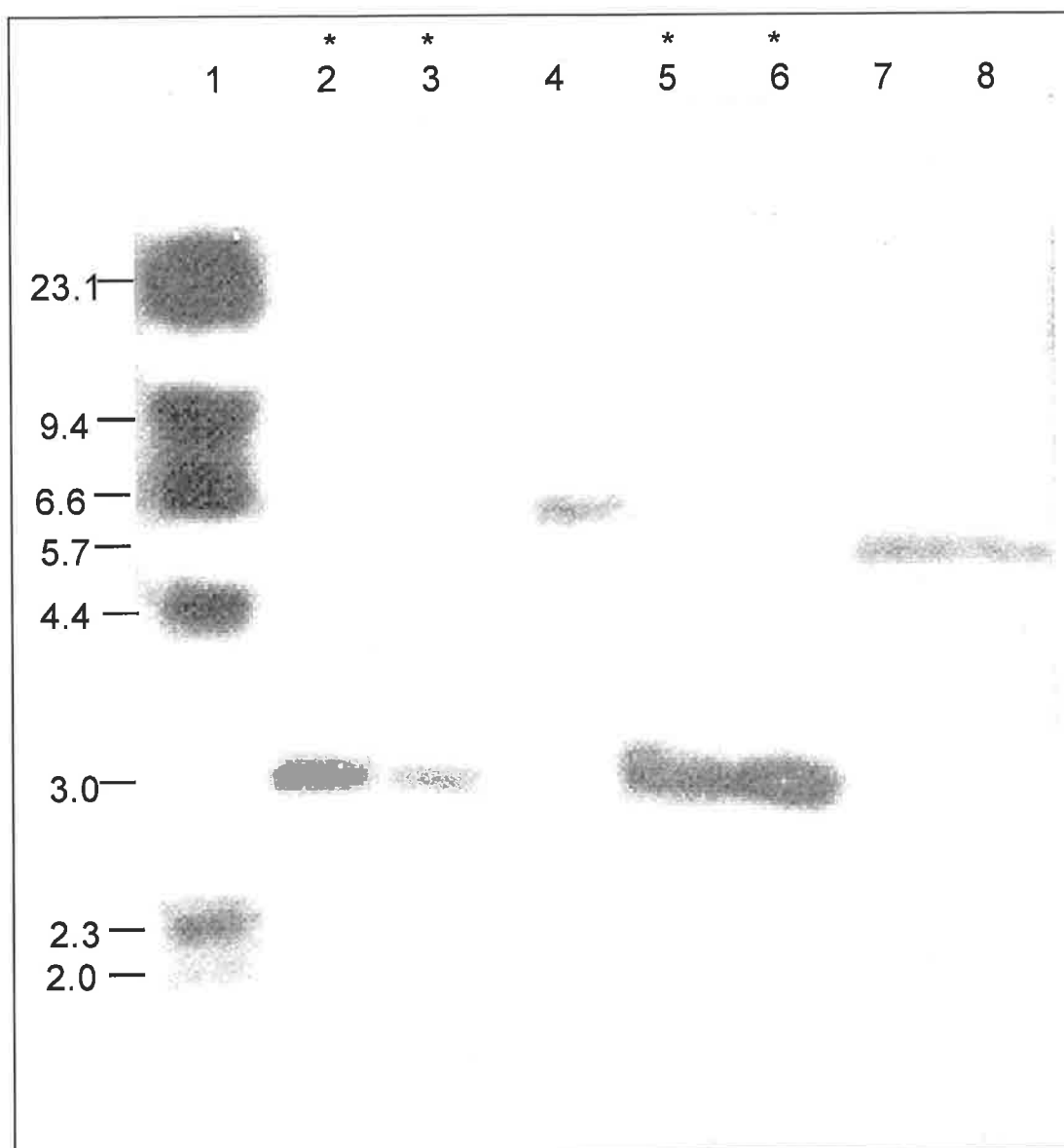


Fig 4.3 *R. leguminosarum* bv. *viciae* strains total DNA digested with *Nco* I and probed with a DIG labeled 1.4 kb *Hind* III strain 1a *mosA* region (also containing 0.74 kb upstream sequence). 1, Marker; 2, 1a; 3, M1201; 4, Sp45; 5, P314; 6, Sp18; 7, 8401(pRL1); 8, NA533. Lanes with asterisk are rhizopine strains.

P342, NA533, and P135, but strong hybridization to single bands of 3.5 kb or 3.6 kb in the rhizopine strains, 1a, M1325, M1201, Sp75, P314, SA1, Sp18 and the *A. radiobacter* strain K749pSym::Tn5. Additional faintly hybridising bands were observed in M1201 (6.2 kb) and Sp75 (4.4 and 6.2 kb) (Fig. 4.4a and 4.4b). However, there was no hybridising band in *A. radiobacter* strain K749. The above result indicates that only rhizopine strains hybridise to a specific *mosA* probe.

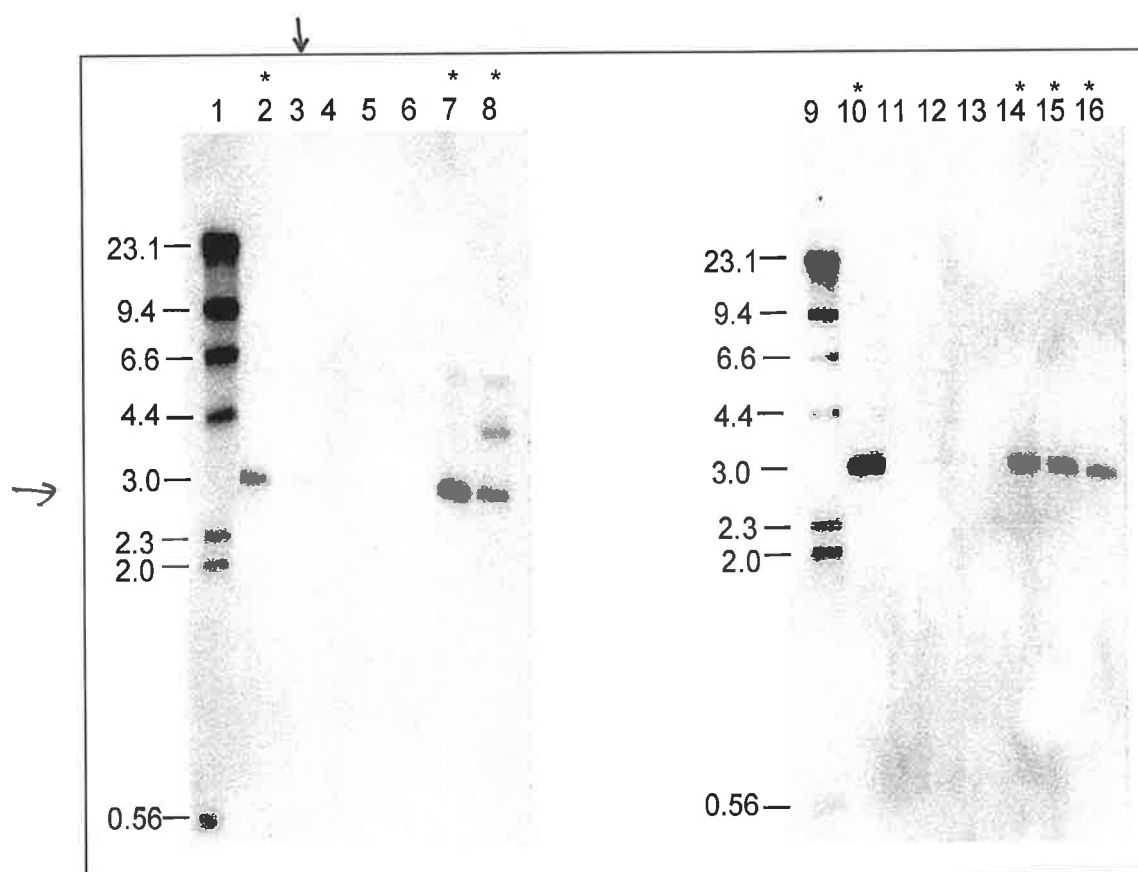


Fig. 4.4a

Fig. 4.4b

Bacterial total DNA digested with *Pst* I and probed with a DIG labeled 0.73 kb probe internal to 1a *mosA* homologue. 1, Marker, 2, *R. leguminosarum* bv. *viciae* strain 1a; 3, *Agrobacterium radiobacter* strain K749pSym::Tn5; 4, *A. radiobacter* strain K749; 5, *R. leguminosarum* bv. *phaseoli* strain 8401(pRL1); 6, *R. leguminosarum* bv. *phaseoli* strain 8401; 7, *R. leguminosarum* bv. *viciae* strain M1201; 8, *R. leguminosarum* bv. *viciae* strain Sp75; 9, Marker; 10, *R. leguminosarum* bv. *viciae* strain Sp18; 11, *R. leguminosarum* bv. *viciae* strain P342; 12, *R. leguminosarum* bv. *viciae* strain NA533; 13, *R. leguminosarum* bv. *viciae* strain P135; 14, *R. leguminosarum* bv. *viciae* strain SA1; 15, *R. leguminosarum* bv. *viciae* strain P314; 16, *R. leguminosarum* bv. *viciae* strain M1325. Lanes with asterisk are rhizopine strains.

Note. Fig 4.4a lane 3 has a very faint band at 3.0 kb.

4.2.2 Screening of *R. leguminosarum* bv. *viciae* strains for *mosB* and *mosC* homologues.

Using a DIG labeled *Bam*H I-*Nsi* I internal 0.8 kb *mosB* homologue fragment (Fig. 4.5) to probe total DNA of *R. leguminosarum* bv. *viciae* strains digested with *Bam*H I, revealed hybridization to single bands in rhizopine strains, 1a (3.7 kb), M1201 (3.8 kb) and SA1 (3.8 kb), and also in M1325 (2.8 kb), P314 (2.9 kb) and Sp18 (3.0 kb). However, no hybridising bands were observed in non-rhizopine strains, 8401(pRL1), L165, NA533, P135, P342, Sp45, Sp59, and WSM1015 (Fig. 4.6a and 4.6b).

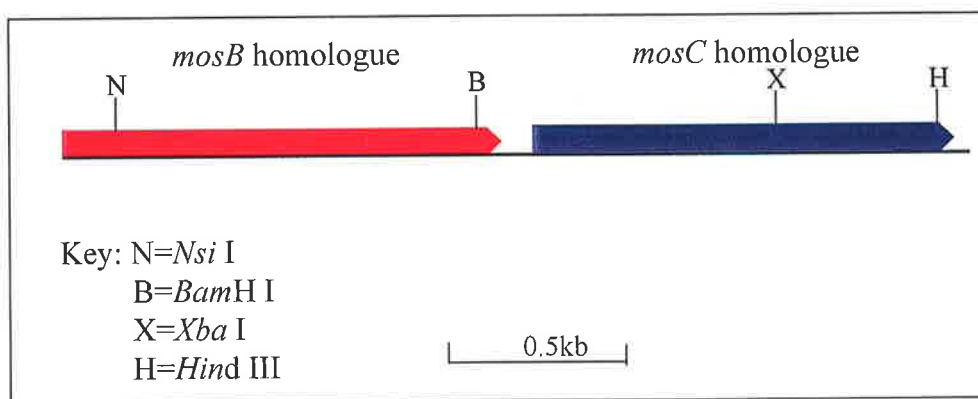


Fig. 4.5 Restriction map of *R. leguminosarum* bv. *viciae* strain 1a *mosBC* homologues.

To establish whether the *mosB* homologue (and by association the *mosC* homologue) is located on pSym of strain 1a, *Pst* I digests of total DNA from *R. leguminosarum* bv. *viciae* strain 1a, *Agrobacterium radiobacter* strain K749(pSym1a::Tn5) and *A. radiobacter* strain K749, were probed by Southern hybridization with a DIG labeled *Bam*H I-*Nsi* I 0.8 kb *mosB* homologue internal probe (see Fig. 4.5). The autoradiogram revealed single hybridising bands of approximately 9 kb for strains 1a and *A. radiobacter* strain K749(pSym::Tn5) but no hybridising band for *A. radiobacter* strain K749 (Fig. 4.7). This result indicates that the *mosB* homologue (and by implication, the *mosC* homologue) is located on pSym1a.

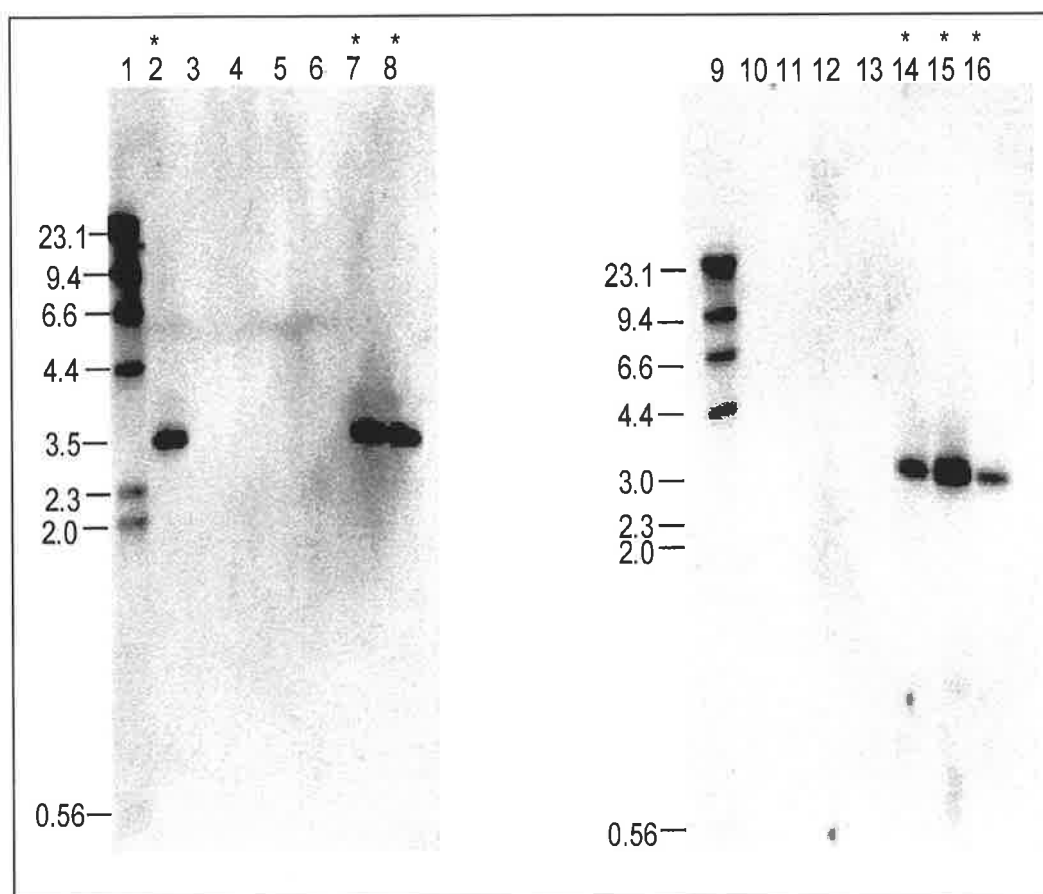


Fig. 4.6a

Fig. 4.6b

R. leguminosarum bv. *viciae* total DNA digested with *Bam*H I and probed with a DIG labeled *Nsi* I-*Bam*H I 0.8 kb 1a *mosB* homologue DNA fragment. 1, Marker; 2, strain 1a; 3, strain 8401(pRL1); 4, strain P342; 5, strain NA533; 6, strain WSM1015; 7, strain SA1; 8, strain M1201; 9, Marker; 10, strain L165; 11, strain Sp59; 12, strain Sp45; 13, strain P 135; 14, strain P314; 15, strain Sp18; 16, strain M1325. Lanes with asterisk are rhizopine strains.

Using a DIG labeled *Xba* I-*Hind* III 0.5 kb probe internal to the *mosC* homologue (Fig. 4.5) to screen total DNA of *R. leguminosarum* bv. *viciae* strains digested with *Bam*H I (which does not cut the *Xba* I-*Hind* III fragment) also revealed single hybridising bands to rhizopine strains, 1a (3.4 kb), M1201 (3.2 kb), M1325 (3.9 kb), P314 (3.9 kb), Sp18 (3.9 kb), and SA1 (3.1 kb). No bands were observed in non-rhizopine strains, 8401(pRL1), L165, NA533, P135, P342, Sp45, Sp59 and WSM1015 (Fig. 4.8a and 4.8b). The results indicate that *mosC* homologue is specific to rhizopine strains.

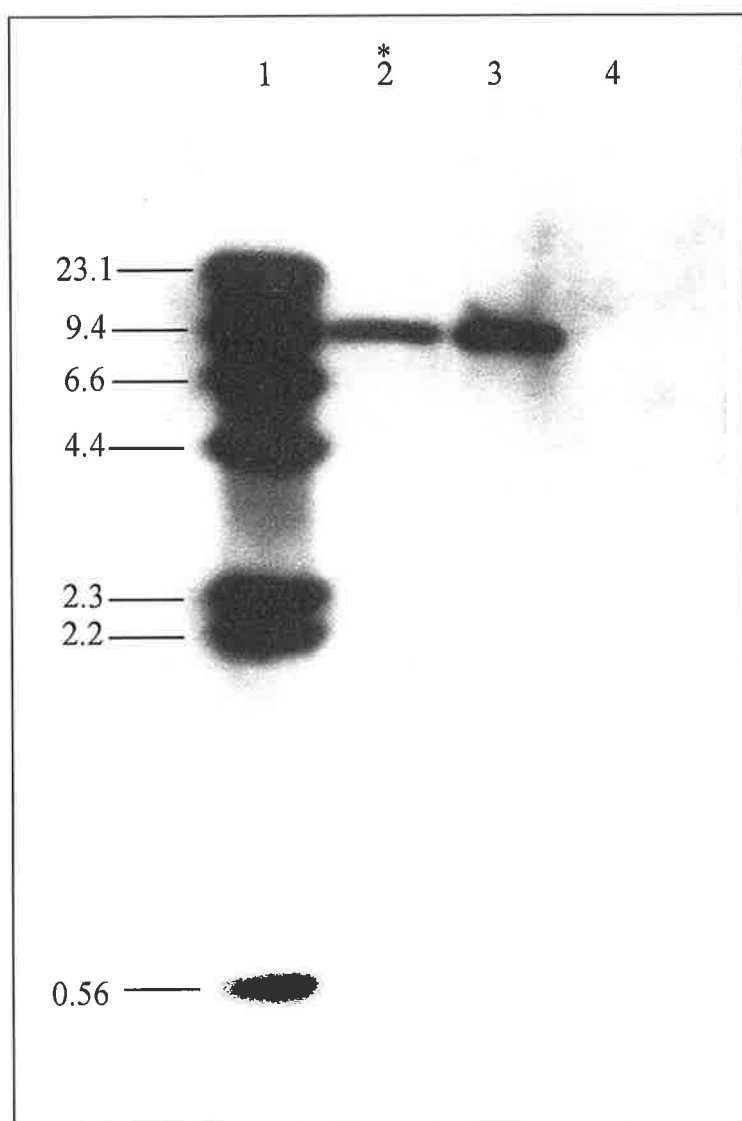


Fig. 4.7 Bacterial total DNA digested with a *Pst* I, probed with a DIG labeled 0.8 kb probe internal to the *mosB* homologue from *R. leguminosarum* bv. *viciae* strain 1a. 1, Marker, 2, *R. leguminosarum* bv. *viciae* strain 1a; 3, *Agrobacterium radiobacter* strain K749pSym::Tn5; 4, *A. radiobacter* strain K749. Lane with asterisk is the rhizopine strain.

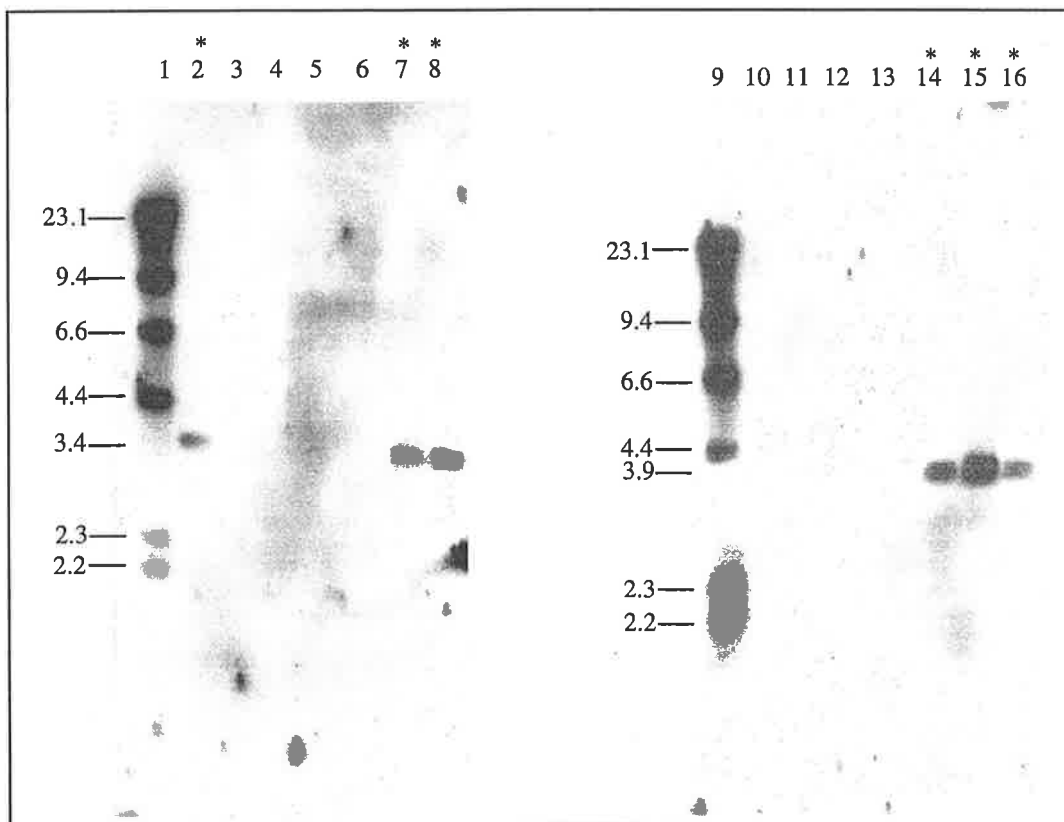


Fig. 4.8a

Fig. 4.8b

R. leguminosarum bv *viciae* total DNA digested with *Bam*H I and probed with a DIG labeled *Xba*I-*Hind* III 0.5 kb 1a *mosC* homologue DNA fragment. 1, Marker; 2, strain 1a; 3, strain 8401(pRL1); 4, strain P342; 5, strain NA533; 6, strain WSM1015; 7, strain SA1; 8, strain M1201; 9, Marker; 10, strain L165; 11, strain Sp59; 12, strain Sp45; 13, strain P135; 14, strain P314; 15, strain Sp18; 16, strain M1325. Lanes with asterisk are rhizopine strains.

4.2.3 Complementation of DapA⁻ *E. coli* with MosA.

A DapA⁻ *E. coli* strain (Bukhari and Taylor, 1971) was transformed by electroporation (Section 2.2.3.2) with plasmids, pPM1339 (carrying the *mosA* homologue from *R. leguminosarum* bv. *viciae* strain 1a) or pPM1089 (containing the *mosA* gene from *S. meliloti* strain L5-30) (Table 4.5). The transformants were selected on TY-DAPA (TY containing diaminopimelic acid) supplemented with ampicillin [Amp] (for pPM1339 transformants) and tetracycline [Tc] (for pPM1089 transformants). To test for complementation, the transformants were inoculated onto TY agar and incubated at 37°C overnight.

Complementation of DapA⁻ *E. coli* with vectors carrying the *mosA* gene from *S. meliloti* strain L5-30 (pPM1089) or the *mosA* gene homologue for *R. leguminosarum* bv. *viciae* strain 1a (pPM1339) was not successful, as transformants showed no growth on TY but grew on TY supplemented with DAPA (diaminopimelic acid) and appropriate antibiotics (Tc for pPM1089 and Amp for pPM1339).

4.2.4 Sequencing regions upstream and downstream of the *mosA* homologue

Examination of the upstream region of the previously sequenced *mosA* homologue from *R. leguminosarum* bv. *viciae* strain 1a (Wexler, 1994), revealed an incomplete ORF. Hence, the upstream region as well as the downstream region of the *mosA* homologue were further sequenced using clone pPM1339 with progressively constructed oligonucleotides. Oligonucleotides were designed with the programme OLIGO 4.0. Upstream of the *mosA* homologue is an ORF (ORFu1) of 1380bp (see Fig. 4.10; Fig. 4.11), with a molecular weight of 48,629 Daltons. There is a gap of 82bp between ORFu1 and the *mosA* homologue ORF. BLAST analysis of this upstream ORF showed extensive homology (67% identity and 79% similarity over 447 amino acids) with a hypothetical aldehyde dehydrogenase, y4uC in *Rhizobium* species strain NGR234 (Frelberg *et al.*, 1997). ORFu1 also has homology with succinate semialdehyde dehydrogenase in *E. coli* (30-53% identity and 48-69% similarity over 404 amino acids) [Niegemann *et al.*, 1993], NADP-dependent glyceraldehyde-3-phosphate in *Streptococcus mutans* (22-35% identity and 41-53% similarity over 471 amino acids) [Boyd *et al.*, 1995], probable succinate semialdehyde dehydrogenase, GabD in *Rhizobium* species strain NGR234 (Frelberg *et al.*, 1997), and benzaldehyde dehydrogenase in *Pseudomonas putida* (28-42% identity and 48-62% similarity over 318 amino acids) [Inoue *et al.*, 1995]. Upstream of ORFu1 is an incomplete ORF termed ORFu2, the sequence of which was not completed due to time constraints. There is no gap between ORFu1 and ORFu2 (Fig. 4.10; Fig. 4.11).

1 TGGCAAACGAACTGCTTGGTGCAAATCTCAAGCCCGAGGATGTAAGCTGCAAGGGTATCTCCGGTCTTTACCTTCCGAC 80
 A N E L L G A N L K P E D V S C K G I S G L S P S D
 → ORFu2 →

81 ATCGAGGATGCAGCACAGGCAAACAGCCGCTGGAAGCTGGTGGATCGGCTGTCCGCAATGACGATGGCAGCGTGACCGG 160
 I E D A A Q A N S R W K L V G S A V R N D D G S V T G

161 CAGTGTTCACCGAGGCGGATTAGCCTCGATCATCCCTCGCAGGCGTCAATGGTGCCACCAATGCGGTGTGCTCGACA 240
 S V A P R R I S L D H P L A G V N G A T N A V S L D T

241 CCGCACTGCTGGGCGCGGTGACGATCACCGGACCTGGCGCCGCGCGCATCGAAACCGCTATGCGCTTCTTTCCGATATC 320
 A L L G A V T I T G P G A G R I E T A Y A L L S D I

321 GTCGCCATTTCATCCGCGGCGCTCGGCCATCGTTAAGGAGGCTGCATGATGCACAACCTTGCACTCACCCAAACGAAC 400
 V A I H S A G A S A I V K E A A * M M H N L A L T Q T N
 ORFu1 →

401 GCTCCTGCAGAAATTGCTGTTACCAGCCCGTTCGACGGTACGCTGGTCGGAACGGTCTGGAACCTGTGCTACAAGCGT 480
 A P A E I A V T S P F D G T L V G T V V E T C A T S V

481 GAACGTGGCTCTGGAGCGGGCAAGCGCGGGCGCAGATTGCCCGAACGCTGCCGCGTCACAAGCGAGCGGTGATCCTTG 560
 N V A L E R A R R G A Q I A R T L P R H K R A V I L E

561 AAAAGGCGCTGCGGCCATCGAAGCCGCGCGGAGGACTTCGCGCTTCTCATCGTCAAAGAGGCGGGCAAGACGATTACG 640
 K A A A A I E A A R E D F A L L I V K E A G K T I T

641 CAGGCCCCGAAAGAGACAACGCGGTGCGTCAATACGCTCAAGCTGTGAGCCGACGAGGCCAAACGCAATGCGGGCGAGGT 720
 Q A R K E T T R C V N T L K L S A D E A K R N A G E V

721 CATCCCGTTCGATGCTACGAGGCTCGGAGGCTCGCCAAGGCTGGTACACGAGAGAACCCTGGGGATTATTGCCGCCA 800
 I P F D A Y A G S E A R Q G W Y T R E P L G I I A A I

801 TTACGCCCTATAACGATCCCTCAATCTCGTCGCCCCAACAATTGGTCCGGCAATTGCCGGCGGCAACGCAAGTCTTTG 880
 T P Y N D P L N L V A H K L G P A I A G G N A V L L

881 AAGCCGTCGGAACCTTACCCCTCTCTGTCATCAAGCTCGTTCGAGGCTCGTTCGATGCGGTCTACCCGAGGAAATCAT 960
 K P S E L T P L S A I K L V E V L V E C G L P E E I I

961 AACGGTCGCGATTGGCGGCCCTGAGCTTGGAAAGGCTCTGGTTCGAGCGAGAGCGTTCGCATGATCTCCTTACCGGTG 1040
 T V A I G G P E L G K A L V A A R D V R M I S F T G G

1041 GCTTTGCCACCGCGGAGGCGATTGCCAAAACAGCAGGCCTCAAGAAGCTTGCCATGGACCTTGGCGGCAATGCTCCTGTC 1120
 F A T G E A I A K T A G A V E A C V S G A Y W A A G Q N C

1121 ATCGTCATGGAGAATTGCAACTTCGATGACCGCCGCTCGAGGCTCGGTATCTGGAGCCTACTGGCGGCCCGGTGAGAACTG 1200
 I V M E N C N F D A A V E A T G C V S G A Y W A A G Q N C

1201 TATTGGCAGCGAGCGAATTTTGATCCAGCGGCCGATCTATGAGCAATTAAAGCGACGTTTCGTGCGAGAAACGAGGAAGC 1280
 I G T Q R I L I Q R P I Y E Q L K A T F V A E T R K L

1281 TGAAGACCGGCAATCTTTGGAAGCCGACACTGATGTCGCCCCGATGATCTCACAAGGCGGTGCAAGAGCGGCCGCA 1360
 K T G N P L E A D T D V G P M I S Q K A V E R A A A

1361 ATGGTCGAACGGGCGCTCGCTGCGCGTGCACATTGCTTTGGGGCCATAAGCCGACTGGCAACCTCTATCCGCCGACGGT 1440
 M V E R A L A A G A T L L W G H K P T G N L Y P P T V

1441 CCTGGAGAATGTTCCAGCCTCTGTGACGCTGGGGCGAGGAAGTGTTCGCGCAATCGTCATTCTTCAACCTTCGATG 1520
 L E N V P A S C D A W G E E V F A P I V I L Q P F D G

1521 GGCTGGCGGATGCGATCGATCTGGCAACAGCCCGAATACAGACTTCACGCCGCAATTTTACAAACGACCTGGAAGGG 1600
 L A D A I D L A N S P E Y R L H A G I F T N D L E G

1601 GCGCTTGAGGCTGCAAGCAAGCTCGAGGCTGGAGGCTGATGATCAACGACTCCTCCGACTATCGCTTCGATGCAATGCC 1680
 A L E A A S K L E A G G V M I N D S S D Y R F D A M P

1681 GTTTGGCGGTTTCAAATATTTCCACGCATGGGCCGCGAGGGCGTCCGGTTTGCCTATGAGGACATGACCCAGCCGAAAGT 1760
 F G G F K Y F H A W A A R A S G L P M R T *

1761 CGTGTGCATCAACCGGCTGAAGACGTAACGTCAGGCCGGCTCAAATCCATCGGCAACTCAGGAGTAAAAATGTTCAAGGG 1840
 M F K G
 Mosa →

1841 TTCTATTACGCGCTTGTACGCTTTTCTGATGGTTCGTTTCGATGAGGTTGCATTGCGCGACCTCGTCGAGTGGCAGA 1920
 S I T A L V T P F P D G R F D E V A L R D L V E W Q I

1921 TCGAAGAAGGATCTCCGGTCTCGTTCCATGCGGGACAACGGGGGAAAGTCCGACGCTGAGCCATTCCGAGCATAAGCAG 2000
 E E G S S G L V P C G T T G E S P T L S H S E H K Q

2001 GTGGTCGAGATCACGATCGCGGCCCGAAGGGCCGTGTGCCGTCATCGCGGGCGCGGATCAAACAGCACGGCGGAAGC 2080
 V V E I T I A A A K G R V P V I A G A G S N S T A E A

2081 AATCGACTTCGTTCCGATCGCGAGGACGCCGCGCGGCTGTGCTGATTGTGCGGCCCTATTACAACAAGCCAACCC 2160
 I D F V R H A Q D A R A D G V L I V A P Y Y N K P T Q

2161 AGGAAGGGATCTACCAGCATTTCAAGGCGATCAATGCTGCGGCCACCGTTCCGATTATCGTCTACAATATCCCCGGCCGC 2240
E G I Y Q H F K A I N A A A T V P I I V Y N I P G R

2241 AGCGTCATCGATATTCAGGTCGAAACGTTGGCCCGGATATTTGAAGACTGCCCAAACGTGAAGGCGTCAAGGATGCGAC 2320
S V I D I Q V E T L A R I F E D C P N V K G V K D A T

2321 GGGAAACCTGTTGCGCCCGTCCCTCGAGCGTATGGCCTGCGGGCACTCGTTCAATCTCCTGACCGGTGAAGACGGCACTG 2400
G N L L R P S L E R M A C G H S F N L L T G E D G T A

2401 CGCTCGGCTACATGGCGCATGGTGGACATGGCTGCATCTCGGTACGGCGAATGTGCGACCCGGCCTTTGCGCTGACTTC 2480
L G Y M A H G G H G C I S V T A N V A P G L C A D F

2481 CAGCAGGCTGCCTGAAACGCGACTTCGCCGCCGCTTGAAGCTTCAGGACCGTCTGATGCCCCTGCACCGGGCGCTTTT 2560
Q Q A C L K R D F A A A L K L Q D R L M P L H R A L F

2561 TCTCGAAACCAACCCTGCTGGGCCGAAATATGCCTTGACGCTCTCGGTGCGATACGTGGGGATCTGCGGCTACCACTCG 2640
L E T N P A G P K Y A L Q R L G R I R G D L R L P L V

2641 TTACGATCAGTCTTCCGTTCAAGAAGAGATCGATAGCGCCATGCGCCATGCGGGGATCCTGGTTTGATGGACAAGTCAA 2720
T I S P S V Q E E I D S A M R H A G I L V * M D K S K

ORFd1 =

2721 AGCTTGCTGAGCGTATCGAGTCAGATCTTATCGGGCCGCTCGCTATTCCCGACACGTCTTTACGGCGTCCATACACGG 2800
L A E R I E S D L I G P L A I P R H V L Y G V H T R

2801 CGCGCGAACAGAAATTTGACGTATCGGGCCTGCGCCTGAGGGATTTTCTGAGCTCATCCAGTCGATGGCGATGGTGAA 2880
R A E Q N F D V S G L R L R D F P E L I Q S M A M V K

2881 AAAGGCTTCTGCTCTCGCCAACATGGAGCTCGGGCTTCTTTCTGCCGAAAAGGCCGAAGCGATCTCTCAAGCATCCGACG 2960
K A S A L A N M E L G L L S A E K A E A I S Q A S D E

2961 AACTCATCGCGCTCAATGATATCGATGAGAACTTTCCCGTCGACATGATGCAGGGAGGGGCCGGCACCTCGACCAACATG 3040
L I A L N D I D E N F P V D M M Q G G A G T S T N M

3041 AACGTGAACGAGGTTCTGACAAACCTCGCTCTTTTGAAGATGGGCGCCGACGTGCGTGAGTATTCGAGGCTGCACCCAAA 3120
N V N E V L T N L A L L K M G A D V G E Y S R L H P N

3121 TGACGACGTCAATCTGTCCAGTCGACGAACGACGTCTACCCAACAGCTATCCGGCTGACGATCTTGCGAGCGTGCGAAG 3200
D D V N L S Q S T N D V Y P T A I R L T I L R A C E G

3201 GGCTTATCGCAGCTGAGGTGCGCTTCGCGACGCCTTTCGCGCAAGGCCATGCAATATGCAAGCGCTCTAAAGGTC 3277
L I A A E V A L R D A F R G K A M Q Y A S A L K V ➡

Fig. 4.7 DNA sequence with deduced amino acid sequence of incomplete ORFu2, ORFu1, *mosA*, homologue and incomplete ORFd1.

BESTFIT analysis shows that the incomplete ORFu2 shares homology (31% identity and 51% similarity over 121 amino acids) with homoserine dehydrogenase from *Lactococcus lactis* (Madsen *et al.*, 1996), and with *E. coli* bicomponent aspartokinase-homoserine dehydrogenase II (28% identity and 47% similarity over 109 amino acids) [Zakin *et al.*, 1983]. Aspartokinase precedes aspartate semialdehyde dehydrogenase in the diaminopimelate and lysine biosynthetic pathway (Fig. 4.1), just as ORFu2 precedes ORFu1. The *mosA* gene is located amongst genes related to those which function in the diaminopimelate pathway and suggests that the *mosA* homologue possibly has a *dapA*-like function.

The downstream region of the *mosA* homologue revealed ORFd11 (partially sequenced, due to time constraints) which shares homology with aspartate ammonia-lyase (aspartase) in certain bacteria, such as, *Corynebacterium glutamicum* (51% identity and 64% similarity over 182 amino acids), *Pseudomonas fluorescens* (48-49% identity and 64-65% similarity over 197 amino acids), *Serratia marcescens* (47% identity and 63% similarity over 182 amino acids), *Haemophilus influenza* (46% identity and 64% similarity over 182 amino acids) and *E. coli* (47% identity and 61% similarity over 182 amino acids).

Table 4.6 Characterisation of open reading frames from upstream (ORFu1 and ORFu2) and downstream sequence (incomplete ORFd11) of *R. leguminosarum* bv. *viciae* strain 1a *mosA* homologue.

ORFs	Shine-Dalgarno sequence/ intervening bases/start codon/number of residues /stop codon.	Calculated molecular weight (Dalton).	Source
ORFu1	GGAGG/4/ATG/1374/TGA	48,629	This work
<i>mosA</i> homologue	GGAG/5/ATG/873/TGA	31,284	Wexler, 1994; Bahar, 1997
Incomplete ORFu2	??/?/+368/TGA	?	This work
Incomplete ORFd11	??/?/ATG/567	?	This work

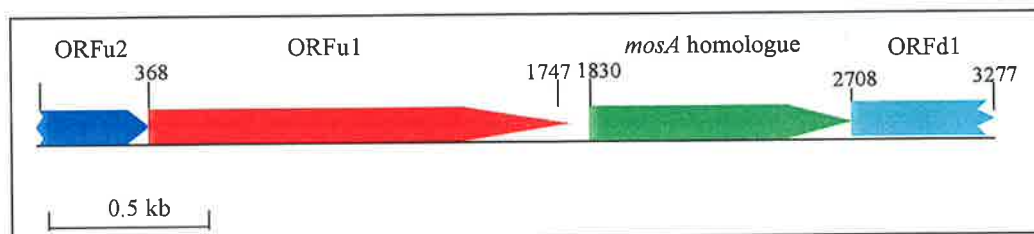


Fig. 4.11 Schematic representation of upstream ORF (ORFu1 and incompletely sequenced ORFu2), *mosA* homologue and incompletely sequenced downstream region, ORFd1, of *R. leguminosarum* bv. *viciae* strain 1a.

4.2.5 Comparison of aligned sequences of *R. leguminosarum* bv. *viciae* strain 1a *mosA* homologue and *S. meliloti* strain L5-30 *mosA*

There is over 87% homology between the DNA sequence of the *R. leguminosarum* bv. *viciae* strain 1a *mosA* homologue and the *S. meliloti* strain L5-30 *mosA* (Bahar, 1997). BESTFIT analysis of these sequences indicates a deletion of a base, in the 1a *mosA* homologue (at position 892) compared to that in strain L5-30 *mosA* (Fig. 4.12). This apparent deletion results in a frame shift and termination of the MosA homologue protein in strain 1a earlier than is the case with strain L5-30 MosA protein (Fig. 4.12). Hence, this section in strains L5-30 *mosA* and the 1a *mosA*

			M	R	H	A	G	I	L	V	*					
1a <i>mosA</i>	881		ATGCGCCATGC.	GGGG	ATCCTGGTTTGATGGACAAGTCA							918				
L530 <i>mosA</i>	881		ATGCGCCACGCG	GGGGG	ATCCTTTTATGATGGACAACGCA							919				
			M	R	H	A	G	D	P	F	M	M	D	N	A	→ 36 *

Fig. 4.12 Alignment of a section of the *R. leguminosarum* bv. *viciae* *mosA* homologue (1a *mosA*, showing the corresponding amino acids above) with that of *S. meliloti* strain L5-30 *mosA* (L5-30 *mosA*, with the corresponding amino acids below); →36*, indicates 36 amino acids before the termination codon (*) in *S. meliloti* MosA.

homologue was re-sequenced (section 2.9.1) in pPM1089 (containing the L5-30 *mosA* region) and pPM1339 (containing 1a *mosA* homologue region) respectively, using oligonucleotides. The 4Gs (bold), in strain 1a *mosA* homologue (Bahar, 1997)

and the published sequence of strain L5-30 *mosA* with 5Gs [bold] (Murphy *et al.*, 1993), were found to be correct.

4.2.6 Screening the R primes (R')

DNA from six R primes, R'1, R'2, R'4, R'5, R'8 and R'9 (Heinrich, 1999) prepared from the *R. leguminosarum* bv. *viciae* strain 1a Sym plasmid were screened with DIG labeled *mosA*, *mosB* and *mosC* homologue specific probes using the dot blotting technique (Section 2.8). The *mos A*, *B* and *C* homologue probes only hybridised to R'4 DNA digest (data not shown) while all R' except R'1 hybridised to the 1.4 kb *mosA* homologue probe (see Section 4.2.1). To confirm the presence of *mos* homologues in R'4, the total DNA was restricted with *Nco* I and probed with DIG labeled *mosA* homologue. Also, R'4 total DNA digested with *Bam*H I was probed with DIG labeled *mosB* or *mosC* homologue. All the three homologue probes hybridised to R'4 (data not shown). Unfortunately, it was not possible to establish R'4 phenotypic characteristics, due to several unsuccessful attempts to recover R'4 from glycerol stock at -80°C and mate it into a non-rhizopine strain. It was not clear why this occurred, apparently, the *E. coli* host lost the R'4 and the culture was dead.

Further attempts to obtain a rhizopine synthesis gene phenotype using R's were unsuccessful.

4.2.7 Mutagenesis strategy

The plasmid pSUP202 was used as a mutagen. This plasmid replicates in *E. coli* but cannot replicate in *Rhizobium* unless it integrates into the host genome (Simon *et al.*, 1983). If a region internal to the gene of interest is cloned into pSUP202 and the construct is mated into *Rhizobium*, a single cross-over event will incorporate pSUP202 into the gene, thus disrupting it. As this mutagenesis strategy is dependent on a single cross-over, it is often more efficient than strategies relying on double cross-overs.

4.2.8 Mutagenesis of the *mosA* homologue and selection of mutant

Shown below is a restriction enzyme map of the *R. leguminosarum* bv. *viciae* strain 1a *mosA* homologue with *Bam*H I showing a 730bp fragment (see Fig. 4.12). The 0.73 kb *Bam*H I fragment from the 1a *mosA* homologue (see Fig. 4.12) was ligated into *Bam*H I restricted pSUP202 to yield pPM1373. The pPM1373 plasmid was electroporated into S17-1 (see Section 2.2.3.2) and mated into rifampicin resistant *R. leguminosarum* bv. *viciae* strain 1a (PM2335) (see Section 2.16).

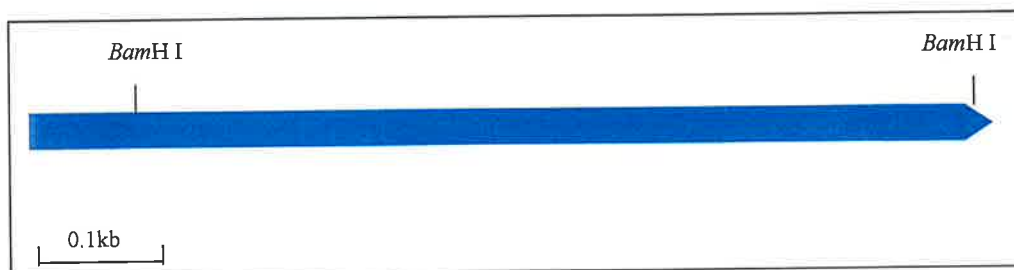


Fig. 4.12 Schematic representation of *Bam*H I restriction map of the *R. leguminosarum* bv. *viciae* strain 1a *mosA* homologue.

After mating, colonies were selected on TY (supplemented with 50 µg/ml rifampicin and 50 µg/ml chloramphenicol) and PM2617 (*R. leguminosarum* bv. *viciae* strain 1a [Rif^r] {pSUP202-*mosA*}) isolated.

4.2.9 Screening of the *mosA* homologue mutant

To confirm that the *mosA* homologue has been mutagenised, total DNA from the transconjugant and the parent strain were digested with *Pst* I (which cuts outside the cloned region used for mutagenesis but cuts once in pSUP202) and probed with a DIG labeled 0.73 kb *mosA* homologue fragment. The transconjugant DNA showed a 9.4 kb band (and a much weaker higher molecular weight band which is presumably uncut DNA) compared to a 3.0 kb band in the parent *R. leguminosarum* bv. *viciae* strain 1a Rif^r strain (Fig. 4.13). The shift in banding pattern suggesting that *mosA* homologue has been disrupted in the transconjugant strain by the insertion of pSUP202. The transconjugant, PM2617 was subsequently referred to as the MosA⁻ mutant. The MosA⁻ mutant showed abnormally rapid growth at 28°C and 37°C. Under the suspicion that it may be an *E. coli* carrying pSUP202:*mosA*, pPM1373, that mutated to rifampicin resistance (such a strain would show a different banding

profile to strain 1a when probed with *mosA* as was shown in Fig. 4.13), the strain was grown on Colilert medium (Edberg *et al.*, 1988). This medium is used to distinguish *E. coli* strains which appear yellow and fluoresce (Edberg *et al.*, 1988). The *MosA*⁻ mutant appeared yellow and fluoresced suggesting that it was *E. coli*. To confirm this, DNA from the mutant would need to be probed with a *nif* probe. If this is *E. coli*, it indicates the need for stronger selection (by the incorporation of another selective marker) during mating. Unfortunately, due to time constraint, further work could not be done on this.

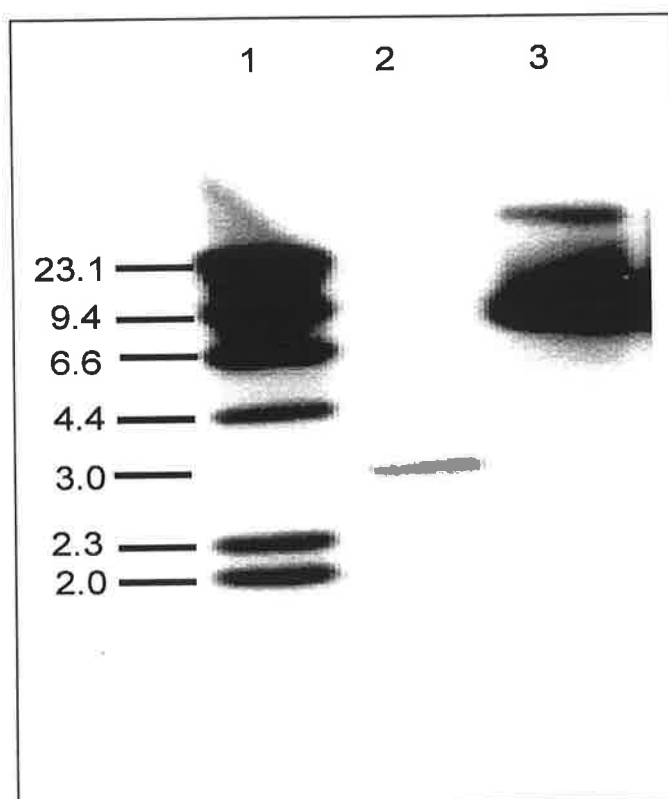


Fig. 4.13 Southern blot of *R. leguminosarum* strain 1a (Rif^r) and *R. leguminosarum* strain 1a transconjugant (PM2617) probed with 0.73 kb *mosA* homologue fragment. (1) λ /Hind III Marker; (2) *R. leguminosarum* strain 1a (Rif^r) total DNA digested with *Pst* I; and (3) PM2617 total DNA digested with *Pst* I.

4.3 Discussion

4.3.1 Presence and location of *mos* homologues in *R. leguminosarum*.

Hybridisation studies using probes prepared from the internal regions of the *mosA*, *B* and *C* homologues from *R. leguminosarum* strain 1a only hybridised to rhizopine strains. This strongly suggests that the *mos* homologues are required for rhizopine production. It is therefore surprising that the introduction *in trans* of the *mosA* and *mosBC* homologues from *R. leguminosarum* strain 1a into a non-rhizopine *R. leguminosarum* bv. *viciae* strain, failed to induce rhizopine synthesis (Bahar, 1997). The inability of this transformant to produce rhizopine may suggest the involvement of other gene(s) besides *mosA* and *mosBC* that are important in rhizopine biosynthesis in this species. With few exceptions each of the *mosA*, *mosB* and *mosC* homologue probes hybridised to single bands in the rhizopine strains. The exceptions where there was hybridisation to more than one band were in M1201 and Sp75 when probed with the *mosA* homologue probe which showed two and three hybridising bands respectively. More than one band in M1201 and Sp75 may be a reflection of incomplete digestion of the DNA, the presence of *Pst* I restriction site(s) within the *mosA* homologues in the two strains (apparently arising from evolutionary difference or genetic rearrangement), and/or homology to a completely different region on the pSym or chromosomal DNA.

There was little variation in band size for each of the probes. For example, rhizopine strains hybridised to a 3.1/3.2 kb fragment using the *mosA* homologue as a probe and 2.9/3.7 kb and 3.1/3.9 kb fragments when *mosB* and *mosC*, respectively, were used as probes. This suggests that there is little polymorphism in the homologues in *R. leguminosarum*. Wexler *et al.* (1996) made similar observations.

Distinct from *S. meliloti* where the *mos* genes are in an operon, the *R. leguminosarum* *mos* homologues are in two regions; the *mosA* region and the *mosBC* region, which is at least 34 kb away. Hybridisation studies to the plasmidless *Agrobacterium* strain and a transconjugant containing the pSym plasmid from strain 1a indicate that both *mosA* and the *mosBC* homologues are on the pSym plasmid.

This adds further support that the *mosA/mosBC* homologues are important in rhizopine synthesis as rhizopine production is encoded by the symbiotic plasmid in *R. leguminosarum* strain 1a (Heinrich, 1999).

4.3.2 MosA/DapA homology

The *R. leguminosarum mosA* homologue has very extensive homology to *mosA* in *S. meliloti* (87% at the amino acid level). There is also extensive homology with *dapA* (37%-47% at the amino acid level) encoding dihydrodipicolinate synthase. The *dapA* gene has a vital role in lysine and peptidoglycan biosynthesis (Wehrmann, 1998; Scapin *et al.*, 1995). This extensive homology and the fact that the *mosA* homologue was not linked to *mosBC* raised the possibility that the *mosA* homologue is in fact a *dapA* gene. Further pursuing this possibility, attempts were made to complement a DapA⁻ mutant of *E. coli* with plasmid borne *mosA* from *S. meliloti* and the *mosA* homologue from *R. leguminosarum* bv. *viciae*. Neither *mosA* nor the *mosA* homologue complemented the mutant. This can be explained by the fact that the genes do not encode a DapA function or the promoters do not express in *E. coli*.

4.3.3 Comparison of *mosA* and *mosA* homologue sequences

A point deletion in the *R. leguminosarum* bv. *viciae* strain 1a *mosA* homologue sequence relative to the *S. meliloti mosA* sequence (Fig. 4.12), resulted in a frame shift and early termination of the deduced protein. However, the DNA homology with strain L5-30 *mosA* continued beyond the point at which the deduced protein of the 1a MosA homologue terminated. The occurrence of this frame shift suggested that there may have been a sequencing error in either of the sequences. However, resequencing the DNA sequences showed them to be correct.

The single base pair deletion results in the shortening of the predicted amino acid sequence of the *R. leguminosarum* MosA homologue by 44 amino acids. In this region of the hypothetical extended protein, 13 of the next 20 amino acids are identical between the two MosAs. The amino acid sequence then diverges at a point near where the *S. meliloti* MosA has homology to FixA (see Fig. 4.13).

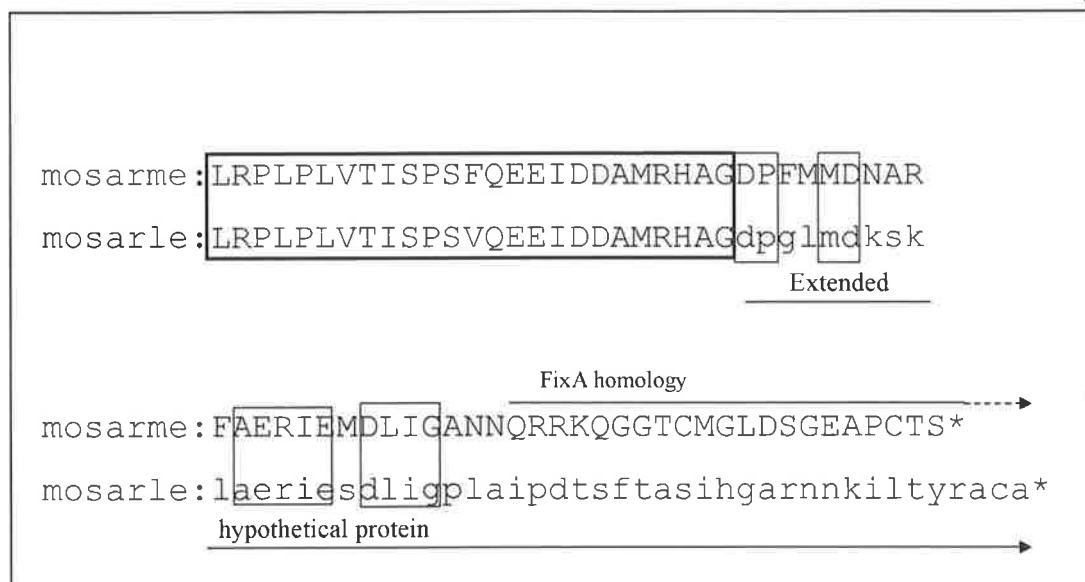


Fig. 4.13 Aligned *S. meliloti* MosA (mosarme) and *R. leguminosarum* MosA homologue (mosarle) at the carboxylic acid end, as shown in thick lined box. The amino acids shown in lower case in **mosarle** sequence represent the hypothetical extended protein, if an additional base was added to the gap in 1a *mosA*, as indicated in Fig. 4.12. The boxed segments are areas of homology. The FixA homology is to **mosarme**.

These results might suggest that the carboxyl-terminal end of the MosA protein, the region with FixA homology, has no important function but rather the result of a re-arrangement of symbiotic genes as has been previously suggested (Murphy *et al.*, 1993). However, it is curious as to why a point deletion that shortens the *R. leguminosarum* MosA homologue, eliminating part of the amino acid sequence in common with *S. meliloti* MosA, has occurred. The shortened MosA homologue has a size of 292 amino acids which is a size very similar to *E. coli* DapA. Comparison of a number of DapA sequences indicates they range in size from 289-302 amino acids (Bahar, 1997).

Amino acid sequence comparison shows that MosA from *S. meliloti* and the homologue from *R. leguminosarum* are much more similar to each other (87%) than to any of the *dapA* genes or than the *dapA* genes are to each other (see phenogram, Fig. 4.14). However, from the sequence alone it is not possible to predict which parts of the sequence might be important for DapA or MosA function. Mutagenesis of this region and determination of DapA and MosA phenotype is required.

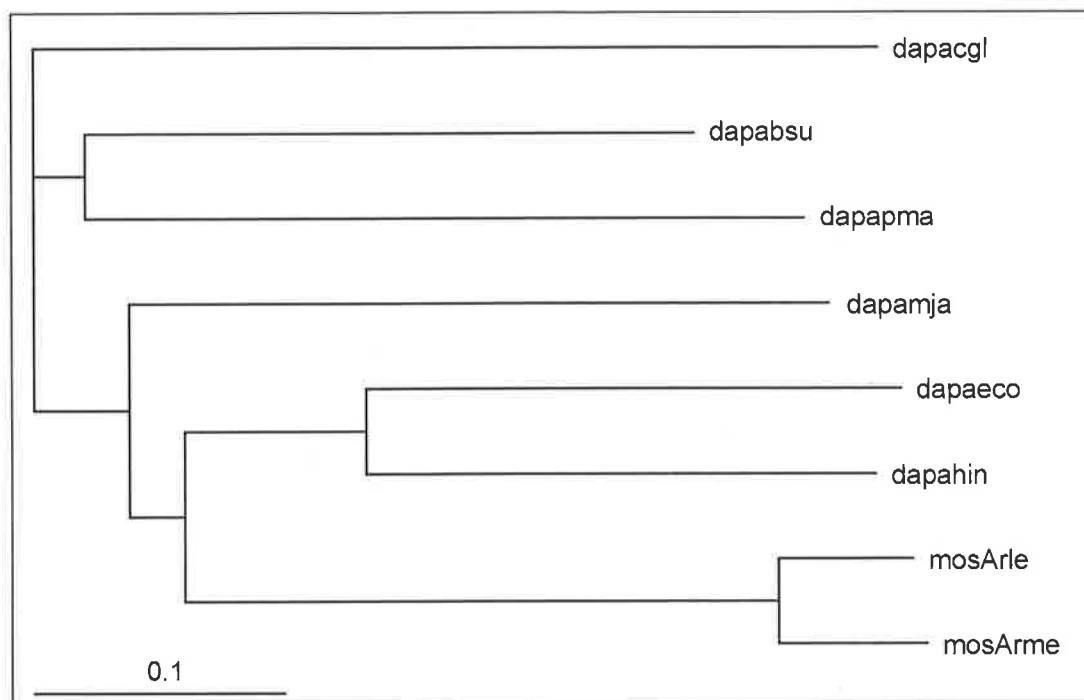


Fig. 4.14 Relative phenograms (generated with the programme, PILEUP) of the genetic relationships among *dapA* and *mosA* from some bacterial species; *Haemophilus influenzae* (*dapahin*), *E. coli* (*dapaeco*), *Methanococcus jannaschii* (*dapamja*), *Prochlorococcus marinus* (*dapapma*), *Bacillus subtilis* (*dapabsu*), *Corynebacterium glutamicum* (*dapacgl*), *S. meliloti mosA* (*mosArme*) and *R. leguminosarum mosA* homologue (*mosArle*).

Since there is considerable homology between *mosA* and *dapA* it might have been expected that the *mosA* homologue probe would detect the presence of both genes in a strain. However, this probe only detected a single band in rhizopine strains, which likely corresponds to *mosA*. There were no additional hybridising bands in rhizopine strains that might correspond to a *dapA* gene. The exception to this was strains M1201 and Sp75, the presence of several bands most likely explained by partial digestion. Similarly, there was no hybridising band in non-rhizopine strains that might correspond to a *dapA* gene. DapA is an ~~important~~^{essential} enzyme involved in lysine biosynthesis and is found in many bacterial species such as *E. coli*, *Bacillus subtilis* and *Haemophilus influenzae*. It is therefore likely that it is present in *Rhizobium*.

It is possible that although there is considerable homology between *mosA* and *dapA* (39-47%) it is not sufficient to detect a *dapA* gene in *Rhizobium*. Recently, by

random sequencing, a *dapA* gene from a cryptic plasmid in *S. meliloti* GR4 was isolated and shares 25% identity with *mosA* in *S. meliloti* (Nicolas Toro, person. comm.). This level of homology, especially if it is dispersed within the sequence, would make detection difficult. Difficulties in using heterologous probes was shown in the present study when a *mosC* probe from *S. meliloti*, which was later shown to have 49% identity to the *R. leguminosarum* *mosC* homologue, barely detected the gene in total DNA from *R. leguminosarum*.

4.3.4 DNA sequence adjacent to the *mosA* homologue.

From the above discussion it is not clear whether the *mosA* homologue has *mosA* function and/or *dapA* function. To help clarify this, the region around the *mosA* homologue was sequenced to see if this gave an indication as to its function. Two ORFs were found upstream of the *mosA* homologue. These are ORFu1 and further upstream ORFu2, for which the complete sequence was not obtained due to time constraints. Downstream of the *mosA* homologue, an incompletely sequenced ORF (ORFd1) was identified.

The pathway in which DapA is involved during lysine biosynthesis involves the following enzymatic steps; aspartate kinase → aspartate β-semialdehyde dehydrogenase → dihydrolipicolinate synthase (DapA) (see Fig. 4.1). ORFu1 has extensive homology with certain aldehyde dehydrogenase^S such as the hypothetical aldehyde dehydrogenase (y4uC) from *Rhizobium* species NGR234 and *E. coli* succinate semialdehyde dehydrogenase. As well, there is homology to glyceraldehyde-3-phosphate dehydrogenase from *Streptomyces mutans* which has a similar mode of action to aspartate β-semialdehyde dehydrogenase (Black and Wright, 1955). The above 3 proteins have the aldehyde dehydrogenase consensus pattern, (FYLVA)-x(3)-G-(QE)-x-C-(LIVMGSTANC)-(AGCN)-x-(GSTADNEKR). ORFu1 has a motif very similar to this, Y-x(3)-G-Q-x-C-I-G-**x(2)**-R. The only difference [bold and underlined] is that ORFu1 has **x(2)** instead of x in the consensus protein. Interestingly, ORFu2 has homology to aspartate kinase, the enzyme preceding aspartate β-semialdehyde dehydrogenase in the DAP pathway. The ORFd1 shares homology with aspartate ammonia-lyase (aspartase) which is located upstream of aspartate kinase in the lysine biosynthetic pathway involving

DapA (Fig. 4.1). Aspartate ammonia-lyase catalyses the conversion of aspartate to fumarate, an intermediate metabolite in the glyoxylate pathway.

Taken together, this indicates that the *mosA* homologue is located amongst genes which could be part of a cluster of genes (of which *dapA* is one) involved in lysine biosynthesis. This appears to conflict with the role of the *mosA* homologue in rhizopine biosynthesis.

When *R. leguminosarum* rhizopine and non-rhizopine strains were probed with a probe containing the *mosA* homologue, ORFu1 and part of ORFu2, the 3 non-rhizopine strains showed single hybridising bands. As these strains did not hybridise with the *mosA* homologue probe alone, the hybridisation is presumably due to ORFu1 and/or ORFu2. The single hybridising band at 3.0 kb in rhizopine strains may contain ORFu1 and ORFu2 as well as the *mosA* homologue. It is therefore possible that all the strains have homology to ORFu1 and ORFu2. This could be tested with a specific probe to ORFu1/ORFu2.

Data to date are consistent with the *mosA* homologue having a role in rhizopine biosynthesis and possibly in a DAP-type pathway. It seems surprising that such diverse functions could be performed by one enzyme. Mutagenesis of this region is required to clarify this.

Chapter 5

Immunological studies with *Sinorhizobium meliloti* anti-MosC serum



THE UNIVERSITY OF ADELAIDE

CHAPTER 5

Immunological studies with *Sinorhizobium meliloti* anti-MosC serum

5.1 Introduction

In *Sinorhizobium meliloti*, the *mos* locus contains four open reading frames (ORFs), *mosORF1*, *mosA*, *mosB* and *mosC* in an operon (Murphy and Saint, 1992). The MosA protein is responsible for 3-*O*-MSI methylation and shares homology with DapA, which is involved in the lysine biosynthetic pathway (Rao *et al.*, 1995). MosB, appears to be an amino transferase involved in 3-*O*-MSI biosynthesis (Chapter 3). The MosC protein is very hydrophobic with 12 membrane spanning domains, and is likely a membrane protein which may be involved in transporting rhizopine precursor into, and/or rhizopines out of the bacteroid (Murphy *et al.*, 1993; Grzemski, 1994). Analysis of deduced amino acid sequences of MosA and MosB suggest that they are likely soluble cytoplasmic proteins (Grzemski, 1994).

Immunogold electron microscopy is useful in *in situ* localization of cellular molecules in animals, plants, fungi and bacteria. The procedure is based on the detection of electron dense colloidal gold particles in transmission electron microscopy (TEM), at a relatively low magnification (Hayat, 1989). The gold particles adsorb to macromolecules based on complex electrochemical interactions and bind to macromolecules such as enzymes while the antibodies still remain biologically active. Relatively high labeling efficiency is achieved with colloidal gold particles of small size (Holgate *et al.*, 1983)

Previously, Grzemski (1994) over-expressed a MosC fusion protein in *E. coli* and prepared antibodies to the purified protein. Attempts to locate MosC in nodule sections by immunogold labeling were inconclusive as the level of gold labeling was low due to difficulties in raising antibodies to the MosC fusion protein. The objective of this chapter is therefore, to identify the location of MosC proteins in nodules induced by *S. meliloti* strain L5-30.

5.2 Results

Table 5.1

Bacterial strains	Description	Source
PM 1204	MosC1-protein 10 fusion construct in pGEMEX-1	Grzmski (1994)

5.2.1 Growing of plants, bacterial inoculation and harvesting of nodules

Prior to germination, lucerne (alfalfa) seeds were surface sterilised (see Section 2.14). The seeds were allowed to germinate on water agar at 28°C for 2-3 days. Rhizobial cultures were inoculated into TY broth and incubated by shaking at 28°C for 48 h. The rhizobia were harvested by centrifugation, washed once in saline and resuspended in Hoagland's solution. The radicles from the germinating seeds were dipped into resuspended rhizobial cells in Hoagland's solution and then transferred onto Hoagland's agar slant (see appendix III) in large test tubes (20cm long, 2.7cm diameter). Two plants were put in each tube. Four replicates of each treatment were made. The plants were then kept in the glass house at 25°C for 5-6 weeks. The nodules were harvested and used for rhizopine analysis and immunological studies.

5.2.2 Preparation of antibodies

A fusion protein containing the ~~hydrophobic~~^{hydrophilic} region of the MosC protein was over-expressed in *E. coli*. The construct was previously prepared using pGEMEX and MosC1 (pPM1204) by Grezmski (1994). This construct expresses poorly, likely due to the hydrophilic nature of the MosC portion of the protein. A 43 kD protein band corresponding to the predicted size of the MosC fusion protein was extracted from the gel and antibodies commercially prepared (Section 2.10.1). The antiserum was adsorbed to protein from the host *E. coli* to remove non-specific antibodies. When this antiserum was used as a probe in a Western blot, it hybridised strongly to a 43 kD band from the *E. coli* strain over-producing the fusion protein and also to a similarly sized band in the host *E. coli* strain not containing the plasmid over-

expressing the MosC fusion protein. However, it did not hybridise to extracts from free-living *S. meliloti* strain L5-30 or Rm1021 (Fig. 5.1).

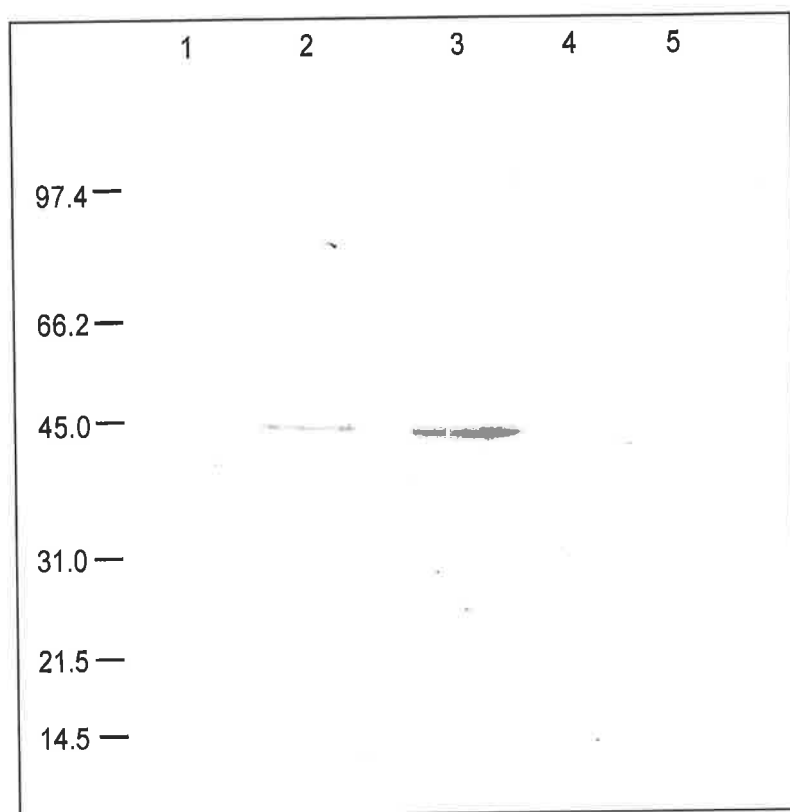


Fig. 5.1 Western blot probed with *S. meliloti* strain L5-30 MosC antibodies. 1, Marker; 2, *E. coli* strain BL21/pLysE; 3, *E. coli* strain BL21/pLysE containing pPM1204; 4, *S. meliloti* strain L5-30; 5, *S. meliloti* strain Rm1021

5.2.3 Western blot analysis of bacteroid fractions

Purified bacteroids were separated into three fractions; (i) peribacteroid membrane (PBM), (ii) bacteroid envelope (BE) [which contains the bacteroid membrane] and (iii) the cytoplasmic contents, as described in sections 2.6 and 2.7. The fractions were digested with LUG buffer to release the protein and subjected to PAGE (see Section 2.10.3)

Western blot analysis (see Section 2.10.4) of bacteroid extracts from *S. meliloti* L5-30 and Rm1021 induced nodules indicated that *mosC* antibodies hybridised to three bands (sizes of 43 kD, 50 kD and 52 kD) in the PBM extract. However, only one band of 52 kD was observed in the BE fraction (Fig. 5.2). No bands were observed in the cytoplasmic contents of bacteroids from either strain.

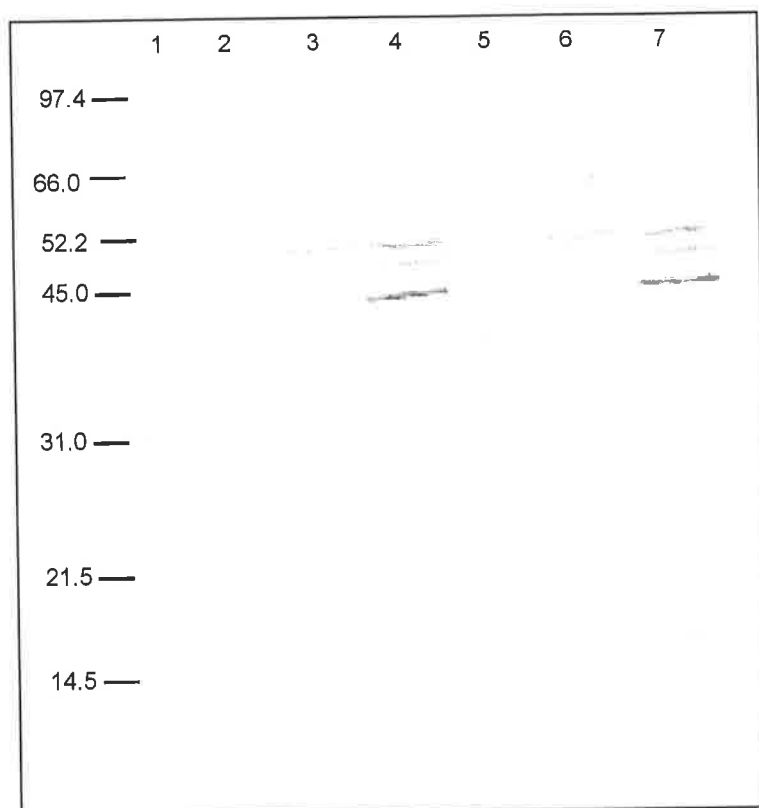


Fig. 5.2 Western blot of bacteroid fractions from *S. meliloti* strain L5-30 and Rm1021 induced nodules probed with MosC antibodies. 1) Marker (not visible); 2) L5-30 bacteroid cytoplasm; 3) L5-30 bacteroid envelope (BE); 4, L5-30 peribacteroid membrane (PBM); 5) Rm1021 bacteroid cytoplasm; 6) Rm1021 BE; 7) Rm1021 PBM.

To test whether the MosC antibodies cross-react with proteins outside of the nodule, lucerne root extract (Section 2.15) was included with *S. meliloti* L5-30 and Rm1021 bacteroid PBM extracts on a Western blot and probed with MosC antibodies (Fig 5.3). The lucerne root extract showed a faint single band of 50 kD whereas three bands were observed in the bacteroids from *S. meliloti* strains L5-30 and Rm1021 as shown previously (one strong band of 43 kD and two faint bands of 52 kD and 50 kD).

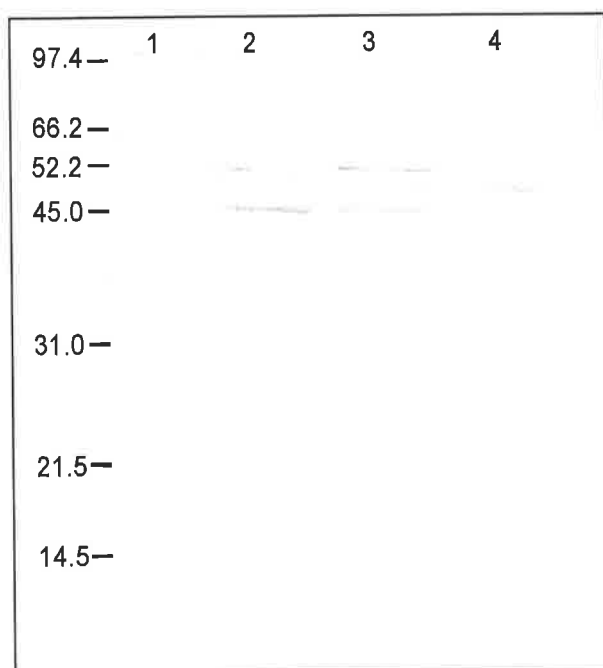


Fig. 5.3 Western blot of root extracts and bacteroid extracts probed with MosC antiserum. 1) Marker (not visible); 2) L5-30 bacteroid peribacteroid membrane (PBM); 3) Rm1021 PBM; 4) Lucerne root extract.

5.2.4 Immunogold labeling and electron microscopy

Very thin cryo-sections of nodules induced by *S. meliloti* strain L5-30 and *R. leguminosarum* bv. *viciae* strain 1a were made with a Cryo-microtome (Lieca Reichert Ultracuts). The nodule sections were immunogold labeled (see Section 2.11) using AuroProbe EM goat anti-rabbit IgG (H+L) G5 (Amersham) and 1:50 dilution of anti-MosC serum. The labeled sections were examined with an electron transmission microscope (Philips CM100 TEM). The anti-MosC labeled nodule sections induced by strain L5-30, revealed gold particles mainly on the peribacteroid membrane, some on the bacteroid membrane and in the cytoplasm (Fig. 5.4). There were also gold particles on what appears to be inter-connecting strands between the bacteroid membrane and peribacteroid membrane (Fig. 5.4). However, in Rm1021 bacteroids, there were virtually no gold particles in the peribacteroid membrane, bacteroid membrane or cytoplasm (Fig. 5.5).

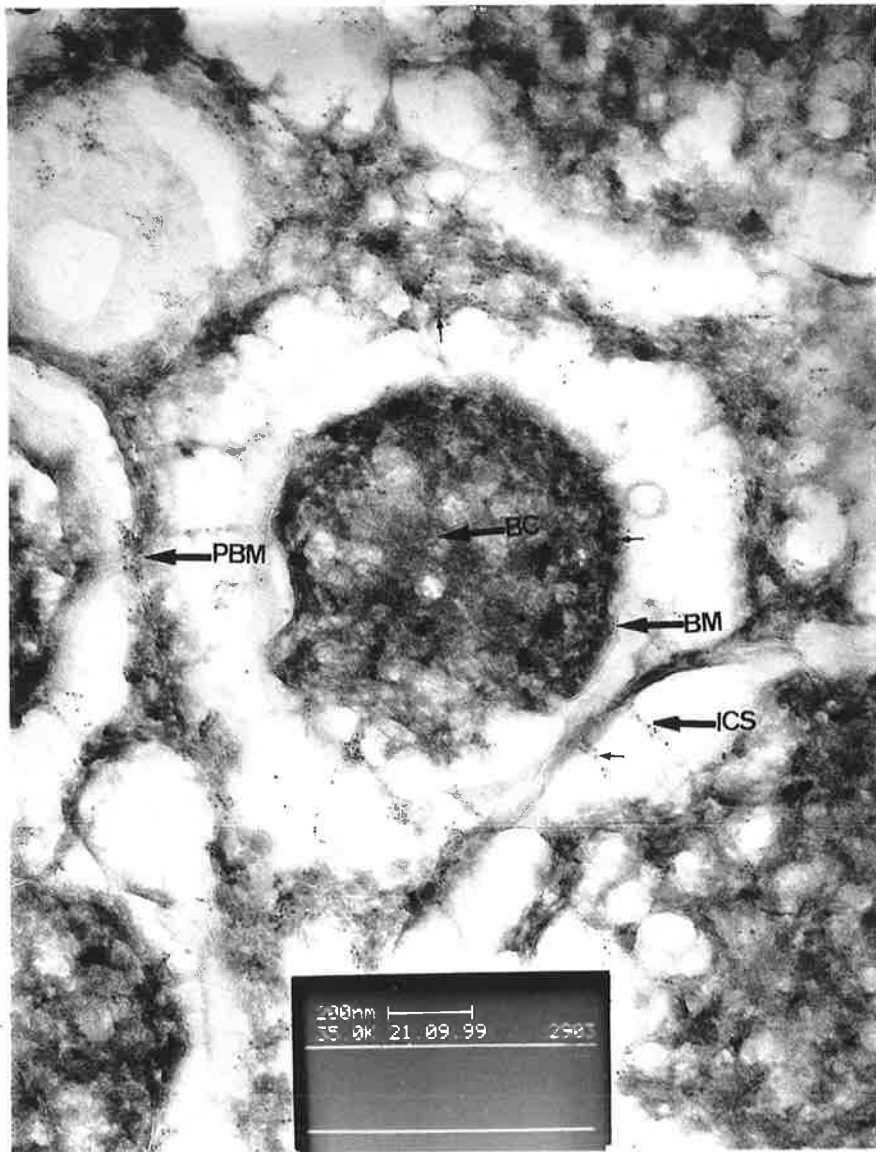


Fig. 5.4 Immunogold labeled section of *S. meliloti* strain L5-30 induced nodule treated with anti-MosC serum. The peribacteroid membrane (PBM), bacteroid membrane (BM), bacteroid cytoplasm (BC), and inter-connecting strands (ICS) are shown with large arrows. The gold particles in the PBM, BM, BC and ICS are shown with small arrows.

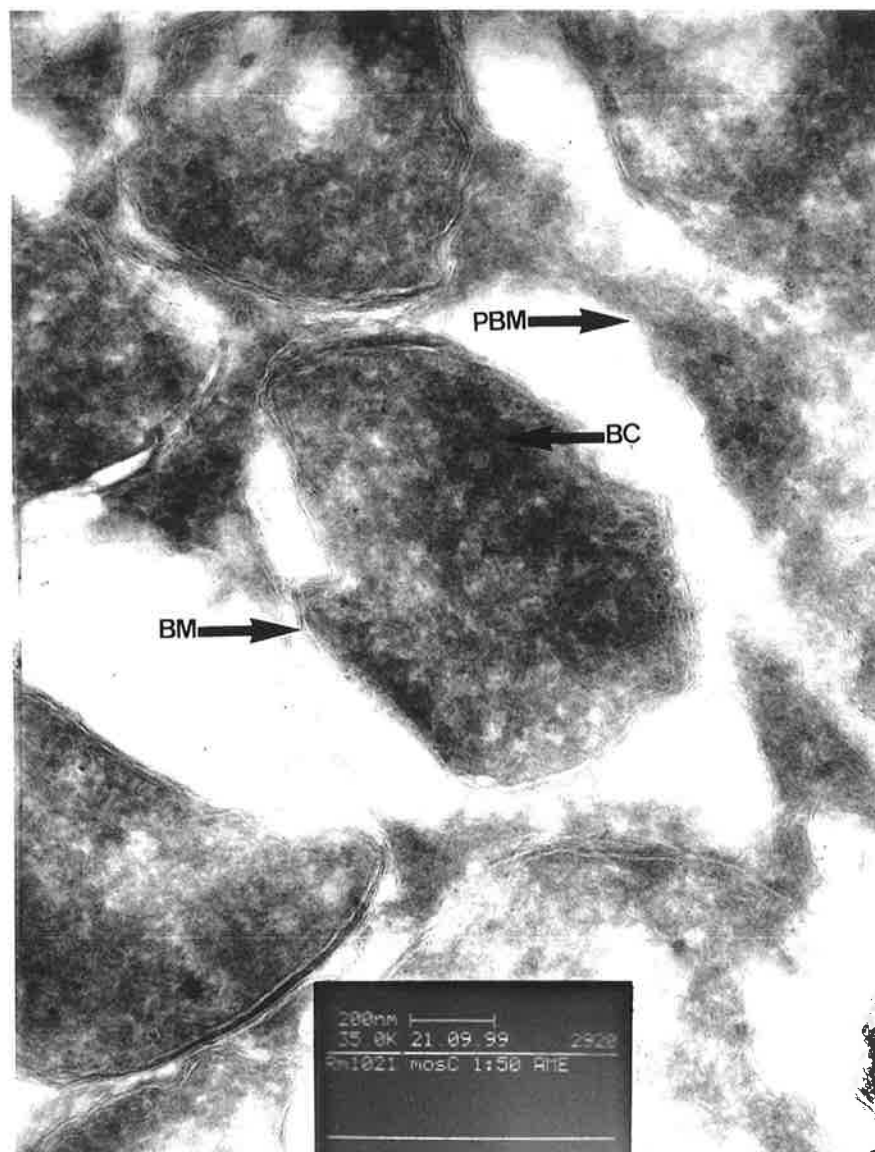


Fig. 5.5 Immunogold labeled section of *S. meliloti* strain Rm1021 induced nodule treated with anti-MosC serum. The peribacteroid membrane (PBM), bacteroid membrane (BM), and the bacteroid cytoplasm (BC) are shown with arrows.

5.3 Discussion

The antibodies prepared to the MosC fusion protein were not pure. They hybridised to a component in the *E. coli* host strain used to over-produce the MosC protein. This hybridising band was the same size as the MosC fusion protein (Fig. 5.1). The most likely explanation for this is that the extracted band contained several proteins to which antibodies were raised. Despite the absorption of the antiserum with *E. coli* protein, some *E. coli* protein antibodies appeared to have remained. The antibodies also non-specifically hybridised to a component of the roots which did not have nodules (Fig. 5.3).

The antibodies showed no selectivity between bacteroid membrane fragments of *S. meliloti* strain Rm1021, which has no rhizopine genes, and *S. meliloti* strain L5-30 which has rhizopine genes. In both cases three hybridising bands were observed in the peribacteroid membrane (PBM) and one corresponding to the largest of these was also present in the bacteroid envelope (BE) [which contains the bacteroid membrane]. No hybridisation was observed to the cytoplasmic fraction. The above results indicate that the antibodies are not specific for the MosC protein.

The lack of specificity could be due to a number of proteins being present in the extract used to make the antibodies as is suggested by the experiments with *E. coli* (Fig. 5.1). As well, cross-reactivity of the MosC antibodies with a similar epitope on other proteins is also likely to be occurring as is evident from the hybridisation to proteins extracted from roots and to Rm1021 bacteroids which do not have *mos* genes. There may, however, be some degree of specificity with the antibodies as they bind to PBM and BE fractions of the nodules formed by strain Rm1021 and L5-30 but not to the free living forms of these bacteria.

The banding pattern observed in the bacteroid membrane extracts cannot be due to MosC alone. It is possible a MosC protein is masked by the other bands. That the hybridising bands are membrane bound is consistent with a protein like MosC, which has 12 membrane binding domains and is predicted to be located in the membrane (Murphy *et al.*, 1993). It is curious that the hybridisation pattern

observed in the bacteroid membranes from the Mos⁻ strain Rm1021 is the same as that for the Mos⁺ strain L5-30. Possibly, the antibodies are reacting with an epitope to a bacteroid membrane protein other than MosC. MosC, which is proposed to transport rhizopine, has a consensus sequence for sugar transport proteins (Chapter 3). Therefore, it is possible the antibodies are detecting a more general epitope for membrane bound proteins.

The immunogold labeling studies provided interesting results. The gold particles bound strongly to *S. meliloti* L5-30 induced nodule sections. This was mainly to the PBM and BM (Fig. 5.4) suggesting that MosC is a membrane protein. This is in agreement with the hydrophobicity profile of MosC that shows a 12 membrane-spanning region (Murphy *et al.*, 1993). However, the attachment of gold particles to what appears to be inter-connecting strands (ICS), suggests a link between the BM and PBM (Fig. 5.4) and possibly a channel through which MosC is transported from the BM to the PBM. This would be necessary since the synthesis of MosC occurs in the bacteroid. The few gold particles observed binding to the bacteroid cytoplasm may be attached to MosC synthesised in the cytoplasm prior to its translocation to the BM.

Distinct from the study with L5-30 nodules, the MosC antibodies did not bind strongly to sections of Rm1021 nodules (Fig. 5.5). This is surprising given that the Western blot of bacteroid fractions showed a similar hybridisation pattern for both Rm1021 and L5-30 (Fig. 5.4). Possibly this is due to a difference in sensitivity between Western blotting and immunogold labeling. The cross-reactivity with Rm1021 proteins may be due to a limited similarity in their epitopes. Hence, this limited binding to cross-reacting protein may not be strong enough to withstand the various treatments during the preparation of immunogold labeling to keep the gold particles bound to the proteins.

Clearly, there has been a problem with the lack of specificity of the MosC antibodies. This is likely due to the poor expression of the fusion protein. Future work should be directed at choosing a portion of the MosC protein which can be

used to prepare a fusion protein which expresses at a much higher level, thus improving the possibility of obtaining specific antibodies.

Chapter 6

Physiological and biochemical studies of rhizopine



THE UNIVERSITY OF ADELAIDE

CHAPTER 6

Physiological and biochemical studies of rhizopine

6.1 Introduction

6.1.1 Effect of exogenously applied rhizopine on root nodulation

Rhizopine strains demonstrate an intra-species nodulation advantage over non-rhizopine strains. This phenomenon seems to require the presence of both the *mos* and *moc* genes (Gordon *et al.*, 1996). However, the mode and site of rhizopine action are not known.

Previously, competition experiments have shown that in the presence of rhizopine those strains which can catabolise it have a nodulation advantage. All of these studies were conducted with rhizopine being produced by the nodulating strain of rhizobia. Therefore, one of the objectives of this Chapter is to determine whether the direct application of rhizopine (3-*O*-MSI) to lucerne rootlets affects nodulation.

6.1.2 Antibiotic activity of rhizopine catabolic products

Rhizopines are catabolised by free-living rhizopine strains of rhizobia within the infection thread and the rhizosphere, and are utilised as a growth substrate (Murphy *et al.*, 1987; Murphy and Saint, 1992; Murphy *et al.*, 1995). Furthermore, rhizopine catabolising bacteria other than rhizobia, were recently identified within the genera, *Arthrobacter*, *Pseudomonas*, *Aeromonas* and *Alcaligenes* (McSpadden Gardener *et al.*, 1998).

The rhizopine, *scyllo*-inosamine, is a metabolic intermediate in the synthesis of the antibiotic streptomycin by *Streptomyces griseus* (Walker and Walker, 1967; Ahlert *et al.*, 1997). It was therefore considered possible that rhizopine may have an antibiotic effect. However, Heinrich (1999), found that 3-*O*-MSI was not inhibitory to the growth of rhizobia or *E. coli*, although the by-products from the catabolism of rhizopine by rhizobia were not tested for antibiotic activity. Therefore, another

objective of this chapter is to determine whether rhizopine catabolic products produced by *S. meliloti* strain L5-30 and *R. leguminosarum* bv. *viciae* strain 1a have antibiotic activity.

6.1.3 Microaerophilic induction of rhizopine

Rhizopines are produced in legume root nodules by the bacteroids of rhizopine producing rhizobia (Murphy *et al.*, 1987; Wexler *et al.*, 1995). In *S. meliloti*, it has been shown that rhizopine synthesis is controlled by a NifA/NtrA regulated symbiotic promoter (Murphy *et al.*, 1988; Murphy *et al.*, 1993). Consequently, rhizopine synthesis is under the same regulatory control as nitrogen fixation, which occurs in microaerophilic conditions and a nitrogen deficient environment (Fischer, 1996). It was therefore investigated whether rhizopine could be produced by free-living bacteria under microaerophilic conditions.

6.2 Results

Table 6.1 Strains of rhizobia used

Rhizobia strains	Description	Source/Reference
<i>S. meliloti</i> strain L5-30	Mos ⁺ Moc ⁺	Kowalski, 1967
<i>S. meliloti</i> strain Rm1021	Mos ⁻ Moc ⁻	Meade <i>et al.</i> , 1982
<i>S. meliloti</i> strain Rm2129	Mos ⁺ Moc ⁻ , L5-30::Tn5	Gordon <i>et al.</i> , 1996
<i>R. leguminosarum</i> bv. <i>viciae</i> strain 1a	Mos ⁺ Moc ⁺	Jensen, 1984

6.2.1 Effect of addition of 3-*O*-MSI to lucerne roots upon nodulation by *S. meliloti* strains L5-30, Rm1021 and Rm2129

Lucerne (*Medicago sativa* var. Hunter River) seeds were sterilised (Section 2.14) and placed on water agar in a Petri dish, to germinate for 2 days, in the dark. The rootlets in the 2-day old seedlings were dipped into *S. meliloti* strains L5-30 (Mos⁺Moc⁺), Rm1021 (Mos⁻Moc⁻) or Rm2129 (Mos⁺Moc⁻) suspension, prepared in 0.5 ml Hoagland's solution (see Appendix III), and then transferred onto a reduced nitrogen (2 mM KNO₃) Hoagland's agar slant in large test tubes (20 cm long, 2.7 cm diameter). 2.5 µl 3-*O*-MSI [2.5 mg/ml] (previously extracted as described in Section 2.12, was sterilised by adding absolute ethanol to achieve 70% ethanol concentration, maintained at room temperature for 30 min and spun in a SpeedVac to remove the alcohol by vacuum evaporation), was added to each seedling rootlet on the agar slants. Two seedlings were placed in each tube. The seedlings were allowed to grow at 22-24°C for about 5 weeks in a temperature regulated glass house. The number of nodules formed on the 3-*O*-MSI treated and untreated plant roots at intervals of 2 days was determined.

Two separate experiments were performed. In the first experiment, there were three replicates in each experiment. The shoots of growing plants were manually guided through the cotton wool used to stopper the tubes, two weeks after rhizobial inoculation (Fig. 6.1). One ml of the nutrient (Hoagland's solution without nitrogen)

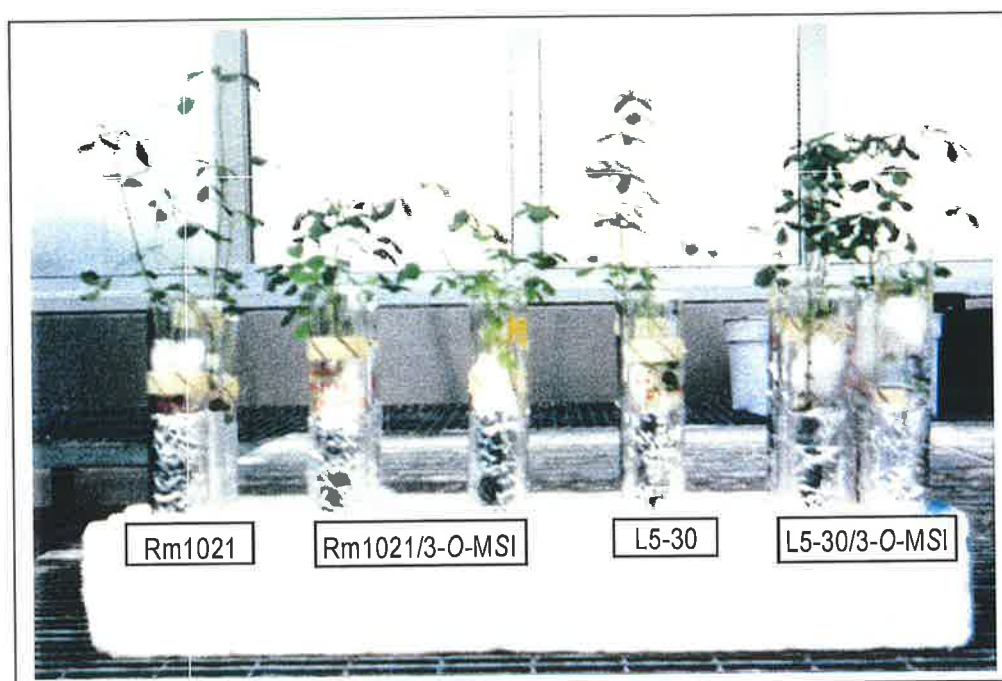


Fig. 6.1 The first experiment showing lucerne plants grown in large tubes. The shoots were guided out of the cotton wool. Rm1021/3-O-MSI and Rm1021 indicate plant roots treated with and without 3-O-MSI respectively, and inoculated with *S. meliloti* strain Rm1021. L5-30/3-O-MSI and L5-30 refer to plant roots treated with and without 3-O-MSI respectively, and inoculated with *S. meliloti* strain L5-30.

was added to the agar slant at the time of rhizobia inoculation. After three weeks, 1-2 ml of nutrient solution was added at intervals of 2-4 days as required, to prevent the agar from drying. The control was not inoculated with rhizobia.

In the second experiment, there were six replicates. The plants were not allowed to grow outside of the cotton wool plugged test tube for the entire duration of the experiment (Fig. 6.2). An average of one white nodule was formed in the non-inoculated control. No more nutrients were added to the tubes throughout the duration of the experiment, as sufficient nutrient (5 ml Hoagland's solution without nitrogen) was introduced into the agar slant, at the beginning of the experiment, i.e. during the inoculation of rhizobia strains. The intention of keeping the tubes stoppered with cotton wool was to eliminate contamination and reduce transpiration,

although the plants may have been stressed as the growing shoots pressed against the cotton wool for generally 18 days, before the termination of the experiment.



Fig. 6.2 The second experiment showing lucerne plants grown in large tubes. The leaves were restricted as they pressed against the cotton wool used to stopper the tubes. Rm1021, Rm2129 and L5-30 refer to treatments of lucerne plants with and without 3-*O*-MSI and inoculated with *S. meliloti* strains Rm1021, Rm2129 and L5-30 respectively. The pea plants were inoculated with *R. leguminosarum* strain 1a and represent a trial experiment to grow pea plants on agar, but are not discussed in this work.

6.2.1.1 Results of the first experiment

In the first experiment, three replicates were used in each experiment. Nodules were first observed after six days on plants inoculated with only strains Rm1021 or L5-30, and 8 days in roots treated with 3-*O*-MSI. The average number of nodules observed on lucerne roots treated with 3-*O*-MSI and inoculated with L5-30 (L5-30/3-*O*-MSI) (Fig. 6.3) or Rm1021 (Rm1021/3-*O*-MSI) (Fig. 6.4), was higher than the number of nodules present on plants not treated with 3-*O*-MSI. However, after 3 weeks when the plants were guided out of the tubes, nodules started appearing on plants in the control experiment, i.e. uninoculated roots. About 50% of these were pink nodules while the rest were white nodules.

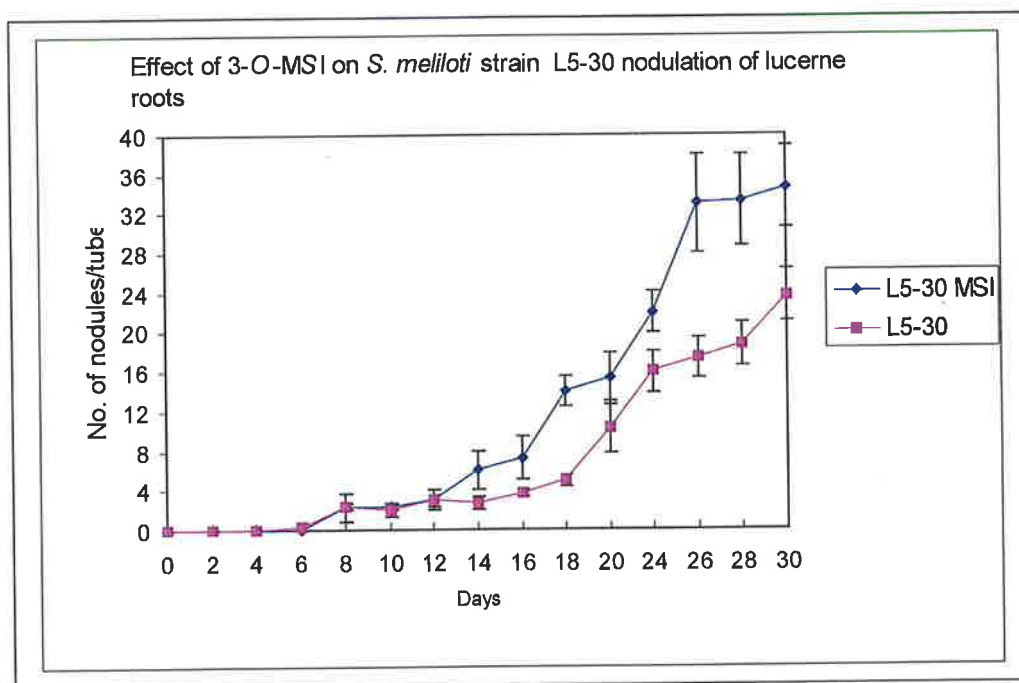


Fig. 6.3 Experiment 1. Number of nodules formed over time on rhizopine treated and untreated lucerne seedling rootlets inoculated with L5-30. Error bars are for standard error.

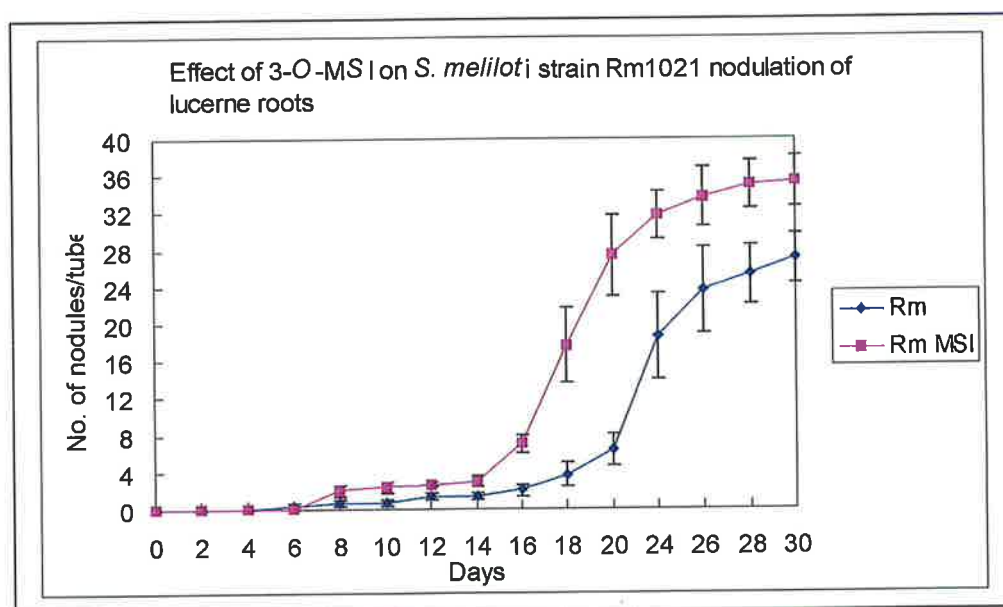


Fig. 6.4 Experiment 1. Number of nodules formed on 3-O-MSI treated and untreated roots inoculated with Rm1021. Error bars indicate standard error.

Regression analysis was used to determine whether there were significant differences between the number of nodules formed on roots treated with 3-O-MSI and the

untreated ones. The square root of the mean number of nodules formed was used in the analysis so as to create a constant variation or make the residuals more even. Using this transformation, the variation in the number of nodules on roots treated with L5-30 and L5-30/3-*O*-MSI was significant at $p < 0.05$. Similarly, there was a significant difference ($p < 0.05$) in nodule number between roots treated with Rm1021 and Rm1021/3-*O*-MSI. The results indicate that the application of the rhizopine, 3-*O*-MSI enhances nodulation on lucerne roots.

6.2.1.2 Results of the second experiment

In an attempt to prevent the late contamination of the uninoculated control that was observed in the first experiment, the second experiment was conducted under sealed conditions. Six replicates were used for each treatment. Nodules were observed six days after rhizobial inoculation.

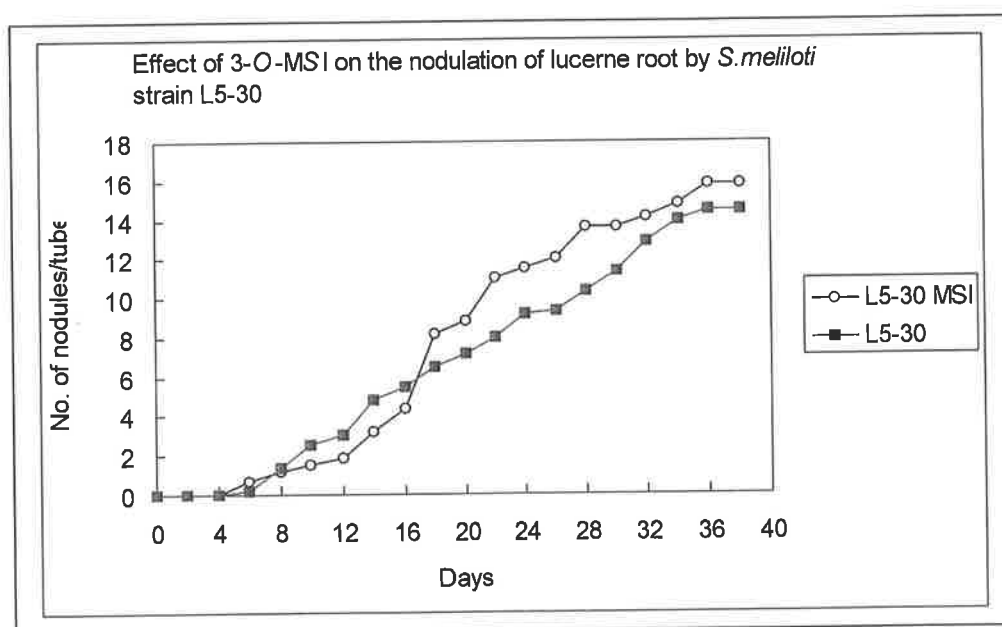


Fig. 6.5 Experiment 2-Number of nodules formed on roots inoculated with *S. meliloti* strain L5-30 and treated with rhizopine (L5-30 MSI) or without rhizopine (L5-30). Standard error bars are not included because they overlap and clutter the graph (see values in Appendix 4, Table 6.2)

The lucerne roots treated with 3-*O*-MSI and inoculated with either L5-30 (L5-30/3-*O*-MSI) (Fig. 6.5) or Rm1021 (Rm1021/3-*O*-MSI) (Fig. 6.6) showed an initial delay in nodulation but later produced more nodules than those not treated with rhizopine i.e. between 14 and 30 days. Towards the end of the experiment, there was little difference in the average number of nodules formed by L5-30 or Rm1021 in both 3-*O*-MSI treated and untreated rootlets (Fig. 6.5 and Fig.6.6). Moreover, the number of nodules on roots treated with L5-30/3-*O*-MSI were virtually the same with those treated with Rm1021/3-*O*-MSI, at the end of the growth period (Fig. 6.5; Fig. 6.6).

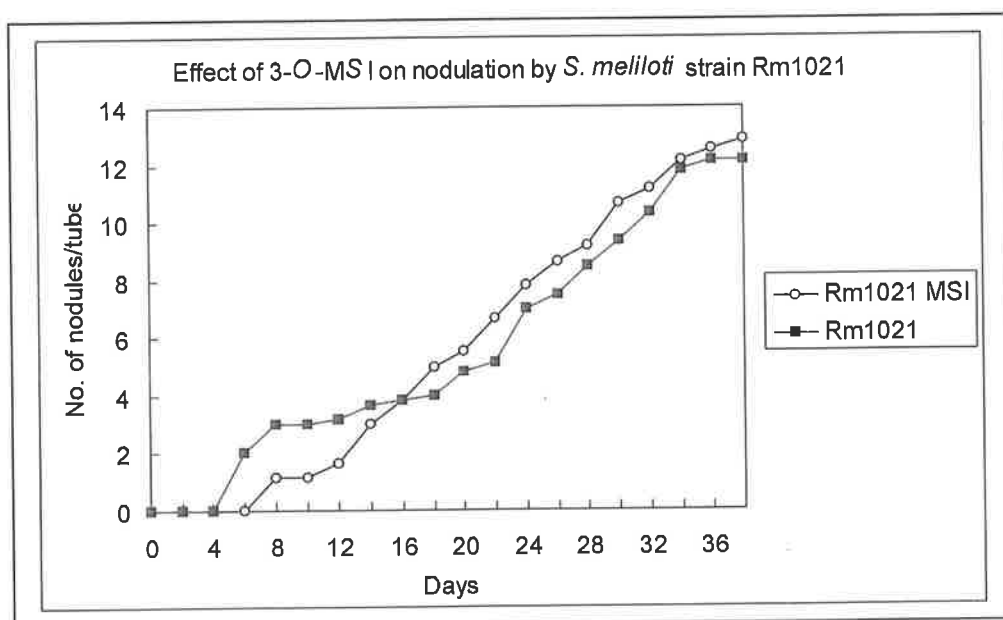


Fig. 6.6 Experiment 2-Number of nodules formed on roots inoculated with Rm1021 and treated with rhizopine (Rm1021 MSI) or without rhizopine (Rm1021). Standard error bars are not included because some overlap and clutter the graph (see values in Appendix 4, Table 6.3)

For the lucerne roots treated with Rm2129/3-*O*-MSI, there was a greater delay in initiating nodulation, however, the number of nodules formed with time was greater than in those inoculated with only Rm2129. Again the difference in the number of nodules formed between 3-*O*-MSI treated and untreated roots decreased towards the end of the experiment (Fig. 6.7). Generally, nodules appeared after 12 days, in roots treated with Rm2129/3-*O*-MSI, but after 6 days in those inoculated with Rm2129. As with the previous experiment, using regression analysis on the square root of the mean number of nodules showed a significant difference at $p < 0.05$ between the L5-

30 and L5-30/3-*O*-MSI treated lucerne roots. Significant difference ($p < 0.05$) were also observed in roots treated with Rm1021 and Rm1021/3-*O*-MSI, as well as those treated with Rm2129/3-*O*-MSI and Rm2129. The significant differences observed in the various treatments indicate that the application of rhizopine to the root surface increases the number of nodules formed.

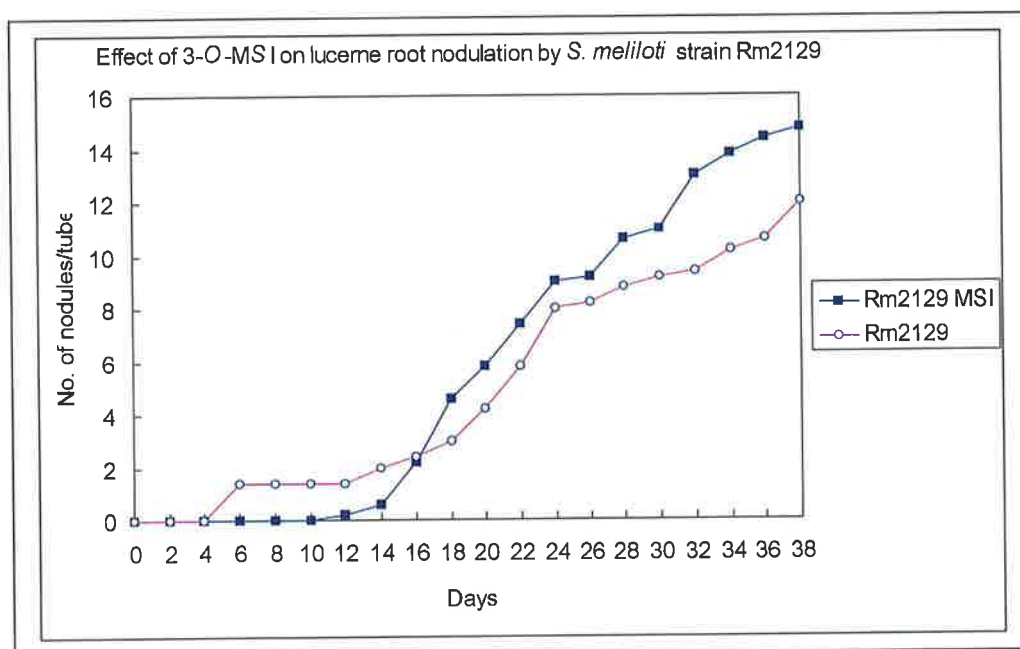


Fig. 6.7 Experiment 2- Number of nodules formed on roots inoculated with Rm2129 and treated with rhizopine (Rm2129 MSI) or without rhizopine (Rm2129). Standard error bars are not included because some overlap and clutter the graph (see values in Appendix 4, Table 6.4)

6.2.2 Antibiotic activity of rhizopine catabolic products

Previously, Heinrich (1999) showed that 3-*O*-MSI did not inhibit the growth of *E. coli* and rhizobia. This work examines whether a breakdown product from the catabolism of rhizopine has antibiotic activity. Catabolic tests were conducted with *S. meliloti* strain L5-30 or *R. leguminosarum* bv. *viciae* strain 1a, in minimal media supplemented with rhizopine (Section 2.13). Discs (5 mm in diameter) made from Whatman 3MM paper and sterilised, were impregnated with 5 μ l of 8 $\times$ concentration of the cell free extract [concentrated by vacuum evaporation in SpeedVac svc100 {Selby scientific}] and then placed on M9 minimal medium (supplemented with 40 μ g/ml thiamine HCl) agar plates uniformly streaked with either *E. coli* DH5 α or *S.*

meliloti Rm1021. The extracts had no inhibitory effect on the growth of either of these bacteria.

6.2.3 Microaerophilic induction of rhizopine

An experiment was conducted to determine whether rhizopine could be produced by free-living rhizopine strains under microaerophilic and atmospheric conditions. The rhizopine strains, *S. meliloti* strain L5-30 and *R. leguminosarum* bv. *viciae* strain 1a were grown shaking at 28°C in TY broth under microaerophilic and atmospheric conditions. The microaerophilic environment was created by burning to extinction a lit candle in a plastic desiccator (Nalgene), sealed with paraffin wax. The exact gaseous content of this microaerophilic environment has not yet been determined. The cell-free extracts were obtained by centrifuging the broth culture at 15,000g for 4 min and the supernatant filter sterilised through a 0.45µ filter (Schleicher & Schuell). The extracts were then analysed for the presence of rhizopine by HVPE (Section 2.4). There was no rhizopine observed in extracts from *R. leguminosarum* bv. *viciae* strain 1a and *S. meliloti* strains L5-30 grown under the microaerophilic condition outlined.

6.3 Discussion

6.3.1 Effect of exogenous rhizopine on root nodulation

Two independent experiments were conducted to investigate the effect of exogenously added rhizopine on nodulation by single isolates of rhizobia. As the rhizobia were added at levels vastly in excess (4×10^8) of that required for nodulation and the experiments did not involve competition between strains, it is unlikely that differences in the number of bacteria outside of the nodule were responsible for the effects observed.

In the first experiment there was an enhanced nodulation when rhizopine was added to rootlets inoculated with strains L5-30 and Rm1021. As Rm1021 can not catabolise the rhizopine, this suggests that rhizopine catabolic activity is not responsible for the enhanced nodulation by exogenously applied rhizopine.

In the second experiment, there was less overall nodulation than the first experiment and the increased nodulation due to the addition of rhizopine, although significant, was considerably less than in the first experiment. As with the first experiment the effect was observed with the Moc⁺ strain L5-30, the Moc⁻ strain Rm1021 and additionally in this experiment, the Moc⁻ L5-30 mutant strain Rm2129. In this experiment the addition of rhizopine initially resulted in a delay in nodulation between days 4-12. This was also not dependent on the ability to catabolise rhizopine as it occurred in all strains regardless of whether they were Moc⁺ or Moc⁻. This effect was not observed in the first experiment. It is possible that the delay in nodule initiation on roots treated with rhizopine may be due to rhizopine modulating the plant signals such as flavonoids and isoflavonoids which induce the expression of nodulation genes in rhizobia and production of Nod factors. The stress to which plants in the second experiment were exposed may have also influenced the plant signals. It has been shown that the levels of flavonoids and isoflavonoids produced by plants can be modulated by biological and environmental stress (Van Etten and Pueppke, 1976; Koes *et al.*, 1994). However, due to the variation between the experiments and in consideration that the delay in nodulation occurs when nodule

numbers are low, it would be well worth repeating this experiment to confirm this observation.

A rhizopine induced delay in the early phases of nodulation would be consistent with previous observations that a Moc^- strain is at a competitive disadvantage in competition (Gordon *et al.*, 1996). A model has been proposed whereby rhizopine inhibits the early phases of nodulation and in conjunction with this, a Moc^+ strain catabolises the rhizopine giving these strains an advantage. Competing Moc^- strains would have a delayed nodulation. When two strains are competing this initial delay in nodulation would be important in determining which strain nodulates as once nodulation is initiated feedback inhibition would restrict subsequent nodulation events. However, no effect would be observed when bacteria are not competing as the number of initiation events greatly exceed the number of nodules that form (Caetano-Anollés and Gresshoff, 1991).

The later effects of enhanced nodulation by exogenous addition of rhizopine appear to be different from that proposed in the early events in nodulation. When rhizopine is produced by bacteria during nodulation (as opposed to its exogenous addition outlined here) there is no difference between the level of nodulation for individual isolates (Gordon *et al.*, 1996). The differences are only observed when the bacteria are competing for nodulation.

Any explanation of the enhanced nodulation by externally applied rhizopine can only be speculative without further experimentation. Nevertheless, some possible aspects where rhizopine could be having an effect are:-

- (a) The rate of nodule formation has been reported to decline as the number of nodule per plant increased (Caetano-Anollés and Gresshoff, 1991). This is due to a feedback inhibition mechanism which regulates the infection events in nodule formation and limits the number of nodules formed on a root, so as to permit the metabolic efficiency of the plant (Calvert *et al.*, 1984). A diffusible shoot factor has been postulated to be important here (Caetano-Anollés *et al.*, 1990). Caetano-Anollés *et al.* (1990) suggested that the diffusible signal may be modified or

transmuted into another signal, which represses the formation of new nodule premordia in the roots. Rhizopine may be suppressing the effect of the diffusible signal from the shoot to the root.

(b) Rhizopine may be resulting in changes in the root hairs and/or interacting with the Nod factor. Microinjection of Nod factors from *R. elti* into living root hairs of bean (*Phaseolus vulgaris*) have been shown to alter the organisation of actin microfilaments in root hair cells, a likely prelude to the formation of infection threads (Cardenas *et al.*, 1998). Also, preinfection threads differentiation (a stage necessary for successful infection) in alfalfa, requires in addition to Nod factors, complementary signalling (Timmers *et al.*, 1999).

The appearance of nodules in the uninoculated control in the first experiment appeared to cast some doubts on the results. However, nodules were not observed until after 21 days. It is likely bacterial contamination occurred when or after the plants were directed out of the top of the tubes at around day 12, or when Hoagland's solution was added. Nevertheless, it would seem unlikely that all the increased nodulation seen on roots treated with rhizopine was due to contamination. Contamination at around day 12 would not be manifest as nodules until after 20 days when the effects of the addition of rhizopine were already observed. To avoid a re-occurrence of the contamination that occurred in the first experiment the plant shoots were kept contained within the tubes throughout the second experiment.

Generally, the first experiment produced about twice the number of nodules per plant as in the second experiment. This could be explained by the increased growth and metabolic activity of the plants in the first experiment due to them being freely growing outside the top of the tube and being exposed to more light. An increase in metabolic activity would invariably affect the infection events leading to nodulation (Calvert *et al.*, 1984). The slightly longer day light conditions for the first experiment (12 h) compared with the second experiment (10.5 h) may have also had an effect. It would be interesting to see if these different plant growth conditions were in any way responsible for the differing response to rhizopine in the two experiments.

6.3.2 Antibiotic activity of catabolised rhizopine

The products of rhizopine (3-*O*-MST) catabolism by *R. leguminosarum* bv. *viciae* strain 1a and *S. meliloti* L5-30 showed no inhibitory activity against *E. coli* DH5 α and *S. meliloti* strain Rm1021. Thus, suggesting that these rhizobia do not produce antibiotic(s) as the end product of rhizopine catabolism. The lack of inhibitory activity by the by-products of rhizopine catabolism suggests that, the competitive advantage enjoyed by rhizopine strains is not associated with inhibitory activity of the by-products of rhizopine catabolism. Nevertheless, further study is required to elucidate the nature and mechanism of rhizopine-root interaction.

6.3.3 Microaerophilic induction of rhizopine synthesis

No rhizopine was detected in microaerobically induced free-living bacteria. The experiment is therefore inconclusive. A precedent exists for *nifA* regulated gene to be activated in free-living cells. Pankhurst and Craig (1978) reported that free-living rhizobia developed nitrogenase activity *in vitro* under reduced oxygen condition in soft agar. As well, although genes regulated by NifA/NtrA are usually expressed in bacteroids, there are examples where such genes are regulated outside this environment. *glnII*, *groESL* and *mela* have been activated in free-living rhizobia (Hawkins *et al.*, 1991; Carlson *et al.*, 1993; Fisher *et al.*, 1993). A large number of reasons could account for no rhizopine being detected. The microaerophilic condition may not be correct, induction of the *mos* genes may not be sufficient under microaerophilic condition, a plant input either as a substrate or perhaps an inducer may be required. It is possible that rhizopine was produced but not released since bacteroids which release rhizopine from the nodules have modified cell wall, unlike their free-living forms (Paau *et al.*, 1978). Whether this morphological difference between the bacteroid and free living form affect the transport of rhizopine out of the cell is not known. This could be investigated by analysing lysed cells of L5-30 grown under microaerophilic environment for rhizopine. A further experiment to test for the synthesis of rhizopine *in vitro* may involve the growth of a rhizopine strain in a minimal medium supplemented with plant extract under microaerophilic condition.

Chapter 7

General discussion



THE UNIVERSITY OF ADELAIDE

CHAPTER 7

General Discussion

7.1 Introduction

Many rhizobia survive in the soil as saprophytes and engage in a symbiotic relationship in the presence of appropriate legume hosts by inducing the formation of nodules on the roots. Within the nodules the bacteroids fix nitrogen and in some rhizobia, bacteroids also produce rhizopine. Ecological and genetic studies of *S. meliloti* rhizopine producing strains suggest that rhizopines are important in the *Rhizobium*-legume symbiosis in both annual and perennial *Medicago* species (Wexler *et al.*, 1995).

The rhizopine genes confer a competitive nodulation advantage on rhizopine strains, however, this phenomenon seems to require the presence of both the rhizopine *mos* (synthesis) and *moc* (catabolic) genes (Gordon *et al.*, 1996).

The *mos* locus in *S. meliloti* consists of the *mosABC* genes in a single operon structure. The *mosA* gene shares homology with *dapA* (Bahar, 1997) which is a unique enzyme in the diaminopimelic acid (DAPA) biosynthetic pathway of prokaryotes, some *Phycomycetes* and higher plants. The unusual amino acid, *meso*-diaminopimelate (*m*-DAP) is an essential constituent of the peptidoglycan layer within bacterial cell walls (Wehrmann *et al.*, 1998). The MosB protein was previously thought to be a regulatory protein (Murphy *et al.*, 1993), but recent data suggest that it is an amino transferase (Chapter 3). The MosC protein is possibly a transport protein as it is very hydrophobic with 12 membrane spanning domains and may be involved in transporting rhizopine and/or rhizopine precursors (Murphy *et al.*, 1993).

7.2 The *mosBC* Gene Homologues in *R. leguminosarum* strain 1a

In this study, the *mosBC* homologues in *R. leguminosarum* bv. *viciae* strain 1a were sequenced. The arrangement of the *mos* gene homologues in *R. leguminosarum* bv.

viciae strain 1a is such that *mosA* is not co-located with *mosBC*, unlike in *S. meliloti* strain L5-30 where *mosABC* form an operon. This variation in arrangement of the *mos* loci suggests different regulatory mechanisms for *mosA* in the two species. The promoter regulating the *mosA* homologue in *R. leguminosarum* bv. *viciae* strain 1a was not identified in this work due to time limitation. It will be of interest to ascertain the mechanism of regulation of this gene. The promoter controlling the *mosBC* homologues appears to be regulated by the NifA/NtrA regulatory system, as is the case with the *mosABC* operon in *S. meliloti*. Future work using NifA⁻/NtrA⁻ mutant strains could confirm the regulatory role of these proteins. The role of the integration host factor (IHF) in regulating *mosBC* homologues would also be of interest.

7.3 MosA-DapA Hypothesis

During this study it was shown that the *mosA* homologue, along with the *mosB* and *mosC* homologues, only hybridised to rhizopine strains. This, together with the high level of homology (87%) to the *S. meliloti* *mosA* gene suggest that the *mosA* homologue is in fact a *mosA* gene. However, other evidence suggest that the homologue may have an alternative or separate function. Firstly, the *mosA* homologue has 47% homology with *dapA*. Secondly, the DNA sequence of the upstream and downstream regions of the *R. leguminosarum* bv. *viciae* strain 1a *mosA* homologue indicated the presence of genes which are co-located with the *dapA* gene in *E. coli* (Chapter 4). The *dapA* gene is vital in the synthesis of diaminopimelate (Wehrmann *et al.*, 1998; see Fig 4.1), which is required for the formation of bacterial cell walls (Scapin *et al.*, 1995). Thus, *dapA* is expected to be present in both rhizopine and non-rhizopine strains. A probe containing the *mosA* homologue and upstream DAPA pathway genes hybridised to non-rhizopine strains. It is likely that this hybridisation was due to the regions upstream of *mosA* present on the probe. However, due to time constraints, it was not possible to confirm this. Nevertheless, it is important to demonstrate that the pathway in which *dapA* is involved is present in other rhizobia and also to isolate *dapA* in these rhizobia, perhaps by using a more specific probe than the *mosA* homologue. Isolation of this

gene would then allow for comparison of the *mosA* homologue and *dapA* at the genetic and possibly physiological levels.

It is possible that in *R. leguminosarum* bv. *viciae* strain 1a, the *mosA* homologue functions both as a *mosA* gene and a *dapA* gene. An example of a protein with two functions is the yeast protein CDC6, which is required for the initiation of DNA replication and also participates in the suppression of nuclear division (Bueno & Russel, 1992). If the *mosA* homologue has a dual function, it raises some interesting questions. For example, what are the implications of this dual functionality and what are the environmental factors which may influence the transcription of this gene? It is noteworthy that MosA functions in the nodule under reduced nitrogen and microaerophilic conditions (Murphy *et al.*, 1993), while DapA is involved in diaminopimelate synthesis under normal aerobic condition (Yugari and Gilvergi, 1965). Whether dual functionality of DapA/MosA is environmentally dependent, will be an interesting topic for further study. However, it is first essential to determine the physiological function(s) of the MosA homologue mutant in the nodule and in free-living bacteria.

7.4 Rhizopine and Root Nodulation

Chapter 6 reports that lucerne roots treated with rhizopine and inoculated with rhizobia initially developed nodules later than the non-treated roots but this was followed by an enhanced level of nodulation. Two independent experiments were conducted. The level of enhanced nodulation was considerably different in these two experiments, therefore the results need to be confirmed.

The observed effects of exogenously applied rhizopine appear to be different from those that occur when rhizopine is produced by the nodulating bacteria, as the effect is not dependent on whether or not the nodulating strain can catabolise the rhizopine. This is distinct from competition experiment where the rhizopine is produced by the nodulating strain and where the non-catabolic strains are at a distinct competitive disadvantage (Gordon *et al.*, 1996).

To influence competition within two weeks (Gordon *et al.*, 1996), rhizopine must be affecting the early events of nodulation. At this point, rhizopine levels would be low. Heinrich (1999), using gas chromatography-mass spectrophotometry (GC-MS) found trace amounts of rhizopine at 4 days after the inoculation of seedlings. The level of rhizopine equated to approximately 5×10^{-9} M at this time. This is comparable to levels (10^{-11} M) at which Nod factor works. Picomolar concentrations of Nod factor are sufficient to induce responses in root epidermis such as, membrane depolarization of epidermal cells (Ehrhardt *et al.*, 1992; Felle *et al.*, 1995; Kurkdjian, 1995), calcium spiking in root hairs (Ehrhardt *et al.*, 1996) and deformation of root hairs (Journet *et al.*, 1994; Cook *et al.*, 1995). Together, these factors suggest a receptor mediated activity for Nod factor. If rhizopine has a modulating influence on the Nod factor, it could also be acting at a very low concentration.

Where could rhizopine be exerting its effect on nodulation? An initial delay in nodulation might be explained by rhizopine influencing infection thread development. Alternatively, rhizopine may be affecting nodule meristematic development. It is known that the initiation of cortical cell division is autoregulated by the plant, which produces a feedback signal to control the number of nodules that develop (Caetano-Anolles and Gresshoff, 1991). Mutants with a defective autoregulatory mechanism continued to nodulate in the presence of nitrate and formed 5 to 10 times the number of nodules than the wild type strain (Carroll *et al.*, 1985). In this situation, fewer meristematic regions are arrested (Mathews *et al.*, 1989). Development of meristematic activity is the critical point in inducing the shoot-mediated systemic autoregulatory response (Caetano-Anolles and Gresshoff, 1990). Rhizopine may cause a transient arrest of cell division centres followed by a rapid development of nodules, as the accumulation of nodulation signals dislodges the action of rhizopine and permit cortical cell division to proceed. How this would occur is open to speculation.

Future work may involve the use of chemically synthesised rhizopine to ascertain if any other component in the semi-pure extract interferes with the effect of rhizopine

on nodulation enhancement and to study the dose-response relationship of added rhizopine.

7.5 Rhizopine - A Nitrogen Sink?

During the course of this study, a likely pathway for the synthesis of rhizopine was proposed. In *Streptomyces griseus*, the α -amino group from L-glutamine is transferred to *scyllo*-inosose to yield *scyllo*-inosamine (a rhizopine) and α -ketoglutarate, in a transamination reaction catalysed by L-glutamine:*scyllo*-inosose aminotransaminase [Fig. 7.1] (Ahlert *et al.*, 1997). L-glutamine:*scyllo*-inosose aminotransaminase (StsC) shares extensive homology and possibly a similar function, with MosB (see Chapter 3). Therefore, it is conceivable that a similar biochemical reaction occurs in the synthesis of rhizopine by the bacteroids. This pathway would utilise a considerable amount of glutamine, especially when the bacteroids are extensively producing rhizopine during active stages of nitrogen fixation.

Plant cells convert the ammonia fixed by bacteroids into L-glutamine, which can be utilised directly as an important nitrogen source for the host plant (see Fig. 1.7; Vance, 1997) or converted to amino acids, glutamate and asparagine (Givan, 1979). As well, the additional conversion of L-glutamine to rhizopine by bacteroids would further reduce the level of glutamine. The removal of glutamine may have several effects. It could mean less nitrogen is supplied to the plant. Alternatively, rhizopine may function as a storage compound removing excess glutamine from the system. A feedback system where the accumulation of glutamine prevents the conversion of ammonia to L-glutamine could lead to the accumulation of large amounts of ammonia which may interfere with host plant metabolism, thereby causing cell damage (Givan, 1979; Sauer *et al.*, 1987).

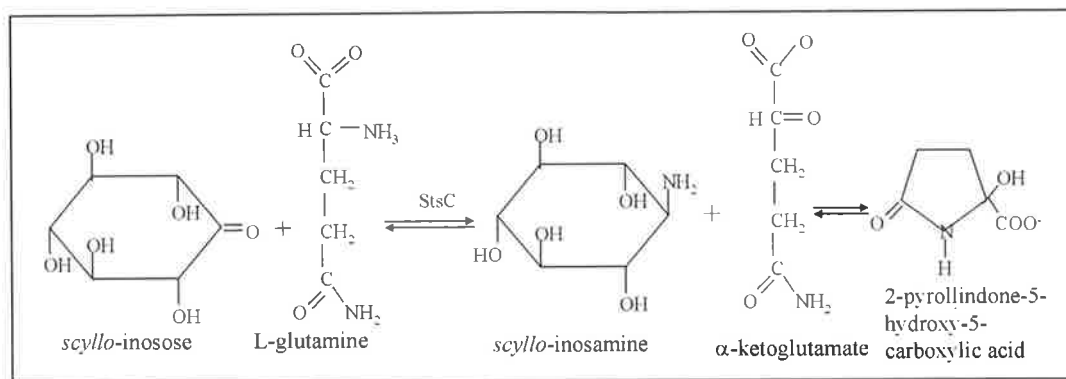


Fig. 7.1 Transamination reaction catalysed by L-glutamine:*scyllo*-inosose amino-transferase (StsC). α -ketoglutarate undergoes a spontaneous cyclization to 2-pyrrolidone-5-hydroxy-5-carboxylic acid (Ahlert *et al.*, 1997).

To further investigate this, it would be interesting to compare the nitrogen fixation of rhizopine strains and their mutants.

Chapter 8

Bibliography



THE UNIVERSITY OF ADELAIDE

CHAPTER 8

Bibliography

Abebe, H. M., Sadowsky, M. J., Kinkle, B. K., and Schmidt, E. L. (1992). Lysogeny in *Bradyrhizobium japonicum* and its effect on soybean nodulation. *Appl. Environ. Microbiol.* **58**:3360-3366.

Aguilar, J. M., Ashby, A. N., Richards, A. J. M., Loake, G. J., Watson, M. D., and Shaw, C. H. (1988). Chemotaxis of *Rhizobium leguminosarum* bv. *phaseoli* towards flavonoid inducers of the symbiotic nodulation genes. *J. Gen. Microbiol.* **134**:2741-2744.

Ahlert, J., Distler, J., Mansouri, K. and Piepersberg, W. (1997). Identification of *stsC*, the gene encoding the L-glutamine:scyllo-inosose aminotransferase from streptomycin-producing *Streptomyces*. *Arch Microbiol.* **168**: 102-113.

Allen, R. M., Chatterjee, R., Madden, M. S., Ludden, P. W., and Shah, V. K. (1994). Biosynthesis of the iron-molybdenum cofactor of nitrogenase. *Crit. Rev. Biotech.* **14**: 225-249.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**:403-410.

Arnold, W., Rump, A., Klipp, W., Priefer, U. B., and Puhler, A. (1988). Nucleotide sequence of a 24,206-base-pair-DNA fragment carrying the entire nitrogen fixation gene cluster of *Klebsiella pneumoniae*. *J. Mol. Biol.* **203**: 715-738.

→ Bagdasarjan et al., (1981) see p170a

Bahar, M. (1997). Molecular biology of rhizopine genes in *Rhizobium leguminosarum* bv. *viciae*. *Ph.D Thesis, University of Adelaide*.

Bahar, M., de Majnik, J. Wexler, M., Fry, J., Poole, P.S. and Murphy, P.J. (1998). A model for the catabolism of rhizopine in *Rhizobium leguminosarum* involves a

ferredoxin oxygenase complex and the inositol degradative pathway. *Mol. Plant-Microbe Interact.* **11**:1057-1068.

Banfalvi, Z., Nieuwkoop, A., Schell, M., Besl, L., and Stacey, G. (1988). Regulation of *nod* gene expression in *Bradyrhizobium japonicum*. *Mol. Gen. Genet.* **214**:420-424.

Barker, D. G., Bianchi, S., London, F., Dattee, Y., Duc, G., Essad, S., Flament, P., Gallusci, P., Genier, G., Guy, P., Muel, X., Tourneur, J., Dénarie, J., and Huguet, T. (1990). *Medicago truncatula*, a model plant for studying the molecular genetics of the *Rhizobium*-legume symbiosis. *Plant Mol. Biol.* **8**:40-49.

Batut, J., Daveran-Mingot, M. L., David, M., Jacob, J., Garnerone, A. M., and Kahn, D. (1989). *fixK*, a gene homologous with *fnr* and *crp* from *Escherichia coli* regulates nitrogen fixation genes both positively and negatively in *Rhizobium meliloti*. *EMBO J.* **8**:1279-1286.

Bauer, W. D., and Caetano-Anollés, G. (1990). Chemotaxis induced gene expression and competitiveness in the rhizosphere. *Plant and Soil* **129**:45-52.

Bergersen, F. J. (1961). The growth of *Rhizobium* in synthetic media. *Aust. J. Biol. Sci.* **14**:349-360.

Better, M., Lewis, B., Corbin, D., Ditta, G., and Helinski, D. R. (1983). Structural relationship among *Rhizobium meliloti* symbiotic promoters. *Cell* **35**:479-485.

Black, S., Wright, N.G. (1955) Aspartate β -semialdehyde dehydrogenase and aspartate β -semialdehyde. *J. Biol. Chem.* **213**:39-50.

Blattner F. R., Plunkett, G. III, Block, C. A., Perna, N. T., Burland, V., Riley, N., Collado-Vides, J., Glasner, F. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N.

W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., Shao, Y. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* **277**:1453-1474.

Blickling, S., Renner, C., Laber, B., Pohlenz, H., Holak, T. A., and Huber, R. (1997). Reaction mechanism of *Escherichia coli* dihydrodipicolinate synthase investigated by X-ray crystallography and NMR spectroscopy. *Biochem.* **36**:24-33.

Boivin, C., Barran, L. R., Malpica, C. A., and Rosenberg, C. (1991). Genetic analysis of a region of the *Rhizobium meliloti* pSym plasmid specifying catabolism of trigonelline, a secondary metabolite present in legumes. *J. Bacteriol.* **173**:2809-2817.

Boivin, C., Camut, S., Malpica, C. A., Truchet, G., and Rosenberg, C. (1990a). *Rhizobium meliloti* genes encoding catabolism of trigonelline are induced under symbiotic conditions. *The Plant Cell* **2**:1157-1170.

Boivin, C., Malpica, C., Rosenberg, C., Denaire, J., Goldmann, A., Fleury, V., Maille, M., Message, B., and Tepfer, D. (1989). Metabolic signals in the rhizosphere: catabolism of calystegins and trigonelline by *Rhizobium meliloti*, B. Lugtenberg, ed. Berlin: Springer-Verlag KG.

Title!

Boivin, C. C., Malpica, C. A., Rosenberg, C., Goldmann, A., Fleury, V., Maille, M., Message, B., Pamboukdjian, N., and Tepfer, D. (1990b). Catabolism of the plant secondary metabolites calystegins and trigonelline by *Rhizobium meliloti*. *Symbiosis* **9**:147-154.

Bonnin, I., Huguet, T., Gherardi, M., Prosperi, J. M., and Olivieri, I. (1996). High level of polymorphism and spatial structure in a selfing plant species *Medicago truncatula* Gaertn. using RAPDs markers. *Am. J. Bot* **83**:843-855.

Boyd, D.A., Cvitkovitch, D.G., and Hamilton, I.R. (1995). Sequence expression and function of the gene for the non-phosphorylating, NADP-dependent glyceraldehyde-3-phosphate dehydrogenase of *Streptococcus mutans*. *J. Bacteriol.* **177**:2622-2627.

Boyer, H. W., and Roulland-Dussoix, D. (1969). A complementation analysis of restriction and modification of DNA in *Escherichia coli*. *J. Mol. Biol.* **41**:459-472.

Brewin, N. J. (1991). Development of the legume root nodule. *Ann. Rev. Cell Biol.* **7**:191-236.

Brewin, N. J., Wood, E. A., and Young, J. P. W. (1993). Contribution of the symbiotic plasmid to the competitiveness of *Rhizobium leguminosarum*. *J. Gen. Microbiol.* **129**:2873-2977.

Brewin, N. J., Robertson, J. G., Wood, E. A., Wells, B., Larkins, A. P., Galfre, G., and Butcher, G. W. (1985). Monoclonal antibodies to antigens in the peribacteroid membrane from *Rhizobium*-induced root nodules of pea cross-react with plasma membranes and golgi bodies. *EMBO J.* **4**:605-611.

Brockwell, J., and Bottomley, P. J. (1995). Recent advances in inoculant technology and prospects for the future. *Soil Biol. Biochem.* **27**:683-697.

Brom, S., Gracia de los Santos, A., Stepkowsky, T., Flores, M., Davila, G., Romero, D., and Palacio, R. (1992). Different plasmids of *Rhizobium leguminosarum* bv. *phaseoli* are required for optimal symbiotic performance. *J. Bacteriol.* **174**:5183-5189.

Brockwell, J., and Bottomley, P. J. (1995). Recent advances in inoculant technology and prospects for the future. *Soil Biol. Biochem.* **27**:683-697.

Bueno, A., and Russell, P. (1992). Dual functions of CDC6: a yeast protein required for DNA replication also inhibits nuclear division. *EMBO J.* **11**:2167-2176.

Bukhari, A. I., and Taylor, A. L. (1971). Genetic analysis of diaminopimelic acid and lysine requiring mutants of *Escherichia coli*. *J. Bacteriol.* **105**:844-854.

Caetano-Anolles, 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**:3164-3167.

Caetano-Anolles, G., and Gresshoff, P. M. (1990). Early induction of feedback regulatory responses governing nodulation in soybean. *Plant Sci.* **71**:69-81.

Caetano-Anolles, G., and Gresshoff, P. M. (1991). Plant genetic control of nodulation. *Ann. Rev. Microbiol.* **45**:345-382.

Caetano-Anolles, G., Joshi, P. A., and Gresshoff, P. M. (1990). Spontaneous nodules induced feedback suppression of nodulation in alfalfa. *Planta* **183**:77-82.

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.

> Cannon, W., Charlton, W., and Buck, ^M~~B~~ (1991). Organisation and function of binding sites for the transcription activator NifA in *Klebsiella pneumoniae* *nifE* and *nifU* promoters. *J. Mol. Biol.* **220**:915-931.

Calvert, H. E., Pence, M. K., Pierce, M., Malik, N. S. A., and Bauer, W. D. (1984). Anatomical analysis of the development and distribution of *Rhizobium* infection in soybean roots. *Can. J. Bot.* **62**:2375-2384.

Carlson, R. W., Price, N. P. J., and Stacey, G. (1994). The biosynthesis of rhizobial lipo-oligosaccharide nodulation signal molecules. *Mol. Plant-Microbe Interact.* **7**:684-695.

Carlson, R. W., Sanjuan, J., Bhat, U. R., Glushka, J., Spaink, H. P., Wijffes, A. H. M., van Brussel, A. A. N., Stokkerman, T. J. W., Peters, N. K., and Stacy, G. (1993). The structures and biological activity of the lipo-oligosaccharide nodulation signals produced by type I and type II strains of *Bradyrhizobium japonicum*. *J. Biol. Chem.* **268**:18372-18381.

Cardenas, L., Vidali, L., Dominguez, J., Pirez, H., Sanchez, F., Helper, P.K. and Quinto, C. (1998). Rearrangement of actin microfilaments in plant root hairs responding to *Rhizobium elti* nodulation signals. *Plant Physiol.* **116**:871-877.

Carrol, B. J., McNeil, D. L., Gresshoff, P. M. (1985). Isolation and properties of soybean (*Glycine max* L. Merr.) mutants that nodulate in the presence of high nitrate concentration. *Proc. Natl. Acad. Sci. USA* **82**:4162-4166.

Cook, D., Dreyer, D., Bonnet, D., Howell, M., Nony, E., Van den Bosch, K. (1995). Transient induction of peroxidase gene in *Medicago trunculata* precedes infection by *Rhizobium meliloti*. *Plant Cell* **7**:43-55.

Cook, D. R., VandenBosch, K., de Bruijn, F. J., and Huguet, T. (1997). Model legumes get the nod. *The Plant Cell Meeting Report*, pp. 275-281.

Cooper, J. B., and Long, S. R. (1994). Morphogenetic rescue of *Rhizobium meliloti* nodulation mutants by trans-zeatin secretion. *Plant Cell* **6**:215-225.

Coronado, C., Zuanazzi, J. A. A., Sallaud, C., Quirion, J. C., Esnault, R., Husson, H. P., Kondorosi, A., and Ratet, P. (1995). Alfalfa root flavonoid production is nitrogen regulated. *Plant Physiol.* **108**:533-542.

Craig, N. L., and Nash, H. A. (1984). *E. coli* integration host factor binds to specific sites in DNA. *Cell* **39**:707-716.

Cullimore, J. V., and Bennett, M. J. (1992). Nitrogen assimilation in legume root nodule: current status of the molecular biology of the plant enzymes. *Can. J. Microbiol.* **38**:461-466.

Cullimore, J. V., Cock, J. M., Farnden, K. J. F., and Bennett, M. J. (1990). Regulation of expression of the GLN- τ gene of *Phaseolus vulgaris* L. in Nitrogen Fixation: Achievement and Objectives (eds. P.M. Gresshoff, L.E. Roth, Stacey and W.E. Newton), Chapman and Hall, New York, London, pp. 713-714.

Dakora, F. D. (1995). Plant flavonoids: biological molecules for useful exploitation. *Aust. J. Plant Physiol.* **22**:87-94.

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 encodes a ferredoxin-like protein. *Cell* **54**:671-683.

Day, D. A. (1992). Metabolite transport between endosymbiotic partners. In: Lectures from the International Symbiosis Congress. Jerusalem, Nov. 1991.

Dazzo, F. B., and Truchet, G. L. (1983). Interaction of lectins and their saccharide receptors in the *Rhizobium*-legume symbiosis. *J. Membr. Biol.* **73**:1-4.

→ Dean et al., 1999 see p 170a

de Bruijn, F. J. (1992). Use of repetitive (repetitive extragenic palindromic and enterobacterial repetitive intergenic consensus) sequences and the polymerase chain reaction to fingerprint the genome of *Rhizobium meliloti* isolates and other soil bacteria. *Appl. Environ. Microbiol.* **58**:2180-2187.

de Lajudie, P., Willems, A., Pot, B., Dewettinck, D., Maestrojuan, G., Neyra, M., Collins, M. D., Dreyfus, B., Kersters, K., and Gillis, M. (1994). Polyphasic taxonomy of rhizobia: emendation of the genus *Sinorhizobium* and description of *Sinorhizobium meliloti* comb. nov., *Sinorhizobium saheli* sp. nov., and *Sinorhizobium teranga* sp. nov. *Int. J. Syst. Bacteriol.* **44**:715-733.

de Philip, P., Batut, J., and Boistard, P. (1990). *Rhizobium meliloti* FixL is an oxygen sensor and regulates *R. meliloti nifA* and *fixK* genes differently in *Escherichia coli*. *J. Bacteriol.* **172**:4225-4262.

Dénarié, J., Debelle, F., and Prome, J. C. (1996). *Rhizobium* lipochitooligosaccharides modulation factors: signalling molecules mediating recognition and morphogenesis. *Annu. Rev. Biochem.* **65**:503-535.

Dénarié, J., Debelle, F., and Rosenberg, C. (1992). Signalling and host range variation in nodulation. *Annu. Rev. Microbiol.* **46**:497-531.

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

Donner, S. C., Jones, D. A., McClure, N. C., Rosewarne, G. M., Tate, M. E., Kerr, A., Fajardo, N. N., and Claire, B. G. (1993). Agrocin 434, a new plasmid encoded agrocin from the biocontrol *Agrobacterium* strains K84 and K1021, which inhibits biovar 2 agrobacteria. *Physiol. Mol. Plant Path.* **42**:185-194.

Dowling, D. N., and Broughton, W. J. (1986). Competition for nodulation of legumes. *Ann. Rev. Microbiol.* **40**:131-157.

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

Downie, J. A. (1994). Signalling strategies for nodulation of legumes by rhizobia. *Trends Microbiol.* **2**:318-324.

Drummond, M., Clements, J., Merrick, M., and Dixon, R. (1983). Positive control and autogenous regulation of the *nifLA* promoter in *Klebsiella pneumonia*. *Nature* **301**: 302-307.

Dudman, W.F. and Belbin, L. (1988). Numerical taxonomic analysis of some strains of *Rhizobium* spp. that uses a qualitative coding of immunodiffusion reactions. *Appl. Environ. Microbiol.* **54**:1825-1830.

Dunn, M. F., Pueppke, S. G., and Krishnan, H. B. (1992). The *nod* gene inducer genistein alters the composition and molecular mass distribution of extracellular polysaccharides produced by *Rhizobium fredii* USDA193. *FEMS Microbiol. Lett.* **97**:107-110.

Dusha, I., Kovalenko, S., Banfalvi, Z., and Kondorosi, A. (1987). *Rhizobium meliloti* insertion element IS*Rm2* and its use for identification of the *fixX* gene. *J. Bacteriol.* **169**:1403-1409.

Dylan, T., Ielpi, S., Stanfield, S., Kashyap, L., Douglas, C., Yanofsky, M., Nester, E. W., Helinski, D. R., and Ditta, G. (1986). *Rhizobium meliloti* genes required for nodule development are related to chromosomal virulence genes in *Agrobacterium tumefaciens*. *Proc. Natl. Acad. Sci. USA.* **83**:4403-4407.

Eckhardt, T. (1978). A rapid method for the identification of deoxyribonucleic acid in bacteria. *Plasmid* **1**:584-588.

Economou, A., Davis, A. E., Johnston, A. W. B., and Downie, J. A. (1994). The *Rhizobium leguminosarum* biovar *viciae* *nodO* gene can enable a *nodE* mutant of *Rhizobium leguminosarum* biovar *trifolii* nodulate vetch. *Microbiol.* **140**:2341-2345.

Edberg, S. C., Allen, M. J., Smith, D. B. and The National Collaborative Study (1988). National field evaluation of a defined substrate method for the simultaneous

enumeration of total coliforms and *E. coli* from drinking-water: Comparison with the standard multiple tube dilution method. *Appl. Environ. Microbiol.* **54**:1595-1601.

Egli, M. A., Larson, R. J., Hruschka, W. R., and Vance, C. P. (1991a). Synthesis of nodulins and nodule enhanced polypeptides by plant gene-controlled ineffective alfalfa. *J. Exp. Bot.* **43**:441-448.

Egli, M. A., Vance, C. P., and Larson, R. J. (1991b). Regulation of nodule gene expression in plant-controlled ineffective alfalfa, S. S. Patil, S. Ouchi, D. Mills and C. Vance, eds. (New York: Springer-Verlag).

↳ Title

Ehrhardt, D. W., Atkinson, E. M., and Long, S. R. (1992). Depolarization of alfalfa root hair membrane potential by *Rhizobium meliloti* Nod factors. *Science* **256**:998-1000.

Elkan, G. H. (1992). Taxonomy of rhizobia. *Can. J. Microbiol.* **38**:446-451.

Fed, E. B., Baldwin, I. R., and McCoy, E. (1932). Root nodule bacteria and leguminous plants (Wisconsin University of Wisconsin Press).

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.

Felle, H. H., Kondorosi, E., Kondorosi, A., Schultze, M. (1995). Nod signal-induced plasma membrane potential changes in alfalfa root hairs are differentially sensitive to structural modifications of the lipo-chitooligosaccharide. *Plant J.* **7**:939-947.

Finan, T., Kunkel, B., De Vos, G. F., and Signer, E. R. (1986). Second symbiotic plasmid in *Rhizobium meliloti* carrying exopolysaccharide and thiamine synthesis genes. *J. Bacteriol.* **167**:66-72.

Fischer, H. M. (1996). Environmental regulation of rhizobial symbiotic nitrogen fixation genes. *Trends Microbiol.* **4**:317-320.

Fischer, H. M. (1994). Genetic regulation of nitrogen fixation in rhizobia. *Microbiol. Rev.* **58**:352-386.

Fischer, R. F., and Long, S. R. (1992). *Rhizobium* signal exchange. *Nature* **367**:655-660.

Fischer, H. M., Babst, M., Kaspar, T., Acuña, G., Arigoni, F., Hennecke, H. (1993). One member of a *groESL*-like chaperonin multigene family in *Bradyrhizobium japonicum* is co-regulated with symbiotic nitrogen fixation genes. *EMBO J.* **12**:2901-2902.

Forster-Hartnett, D., Cullen, P. J., Gabbert, K. K. and Kranz, R. G (1993). Sequence genetic, and *lacZ* fusion analyses of a *nifR3-ntrB-ntrC* operon in *Rhodobacter capsulatus*. *Mol. Microbiol.* **8**:903-914.

Franssen, H. J., Vijn, I., Yang, W. C., and Bisseling, T. (1992). Developmental aspects of *Rhizobium*-legume symbiosis. *Plant Mol. Biol.* **19**:89-96.

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

Freiberg ^{Bairoch}
~~Fellay~~ G., Fellay, R., ~~Bairoch~~ A., Broughton, W.J., Rosenthal, A., and Perret, X. (1997). Molecular basis of symbiosis between *Rhizobium* and legumes. *Nature* **387**:394-401.

Friedman, D. I. (1988). Integration host factor: a protein for all seasons. *Cell* **55**:545-554.

Gallon, J. R., and Chaplin, A. E. (1987). An introduction to nitrogen fixation. *Cassel Education Limited, London*.

Gao, W. M., and Yang, S. S. (1995). A *Rhizobium* strain that nodulates and fixes nitrogen in association with alfalfa and soybean plants. *Microbiology* **141**:1957-1962.

Gault, R. R., and Schwinghamer, E. A. (1993). Direct isolation of *Bradyrhizobium japonicum* from soil. *Soil Biol. Biochem.* **25**:1161-1166.

Gepts, P. (1990). Biochemical evidence bearing on the domestication of Phaseolus (*Fabaceae*) beans. *Econ. Bot.* **44**:28-38.

Gibson, A. H. (1990). The Genetics of Nitrogen-Fixing Symbiotic Associations. In: Nitrogen in Higher Plants. (Y.P. Abrol Ed.), pp. 461-488. *Research Studies Press Ltd., Taunton, Somerset*.

Giladi, H., Gottesman, M., and Oppenheim, A. B. (1990). Integration host factor stimulates the phage Lambda pL promoter. *J. Mol. Biol.* **213**:109-121.

Giladi, H., Koby, S., Gottesman, M., and Oppenheim, A. B. (1992). Supercoiling, integration host factor and a dual promoter system, participate in the control of bacteriophage λ pL promoter. *J. Mol. Biol.* **224**:937-948.

Givan, C. V. (1979). Metabolic detoxification of ammonia in tissues of higher plants. *Phytochem.* **18**:375-382.

Gober, J. W., and Kashket, E. R. (1987). K^+ regulates bacteroid-associated functions of *Bradyrhizobium*. *Proc. Natl. Acad. Sci. USA.* **84**:4650-4654.

Goldman, A., Boivin, C., Fleury, V., Message, B., Lecoeur, L., Maille, M., and Tepler, D. (1991). Betaine use by rhizosphere bacteria: genes essential for

trigonelline, stachydrine, and carnitine catabolism in *Rhizobium meliloti* are located on pSym in the symbiotic region. *Mol. Plant-Microbe Interact.* **4**:571-578.

Gordon, D. M., Ryder, M., Heinrich, K., and Murphy, P. J. (1996). An experimental test of the rhizopine concept in *Rhizobium meliloti*. *Appl. Environ. Microbiol.* **62**:3991-3996.

Göttfert, M., Holzhauser, D., Bani, D., and Hennecke, H. (1992). Structural and functional analysis of two different *nodD* genes in *Bradyrhizobium japonicum* USDA110. *Mol. Plant-Microbe Interact.* **5**:257-260.

Göttfert, M., Lamb, J. W., Gasser, R., Semenza, J., and Hennecke, H. (1989). Mutational analysis of *Bradyrhizobium japonicum* common *nod* genes and further *nod* box-linked genomic DNA regions. *Mol. Gen. Genet.* **215**:407-410.

Grzemeski, W. (1994). Biochemical and molecular genetic studies of the *Rhizobium meliloti* *mos* locus. *Ph.D Thesis, The University of Adelaide*.

Gussin, G. N., Ronson, C. W., and Ausubel, F. M. (1986). Regulation of nitrogen fixation genes. *Annu. Rev. Genet.* **20**:567-591.

Haack, A. (1964). Über den Einfluß der Knöllchenbakterien auf die wurzelhaare von Leguminosen und Nichtleguminosen. *Zbl. Bakt. Abt. II* **117**: 343-347.

Handberg, K., and Stougaard, J. (1992). *Lotus japonicus*, an autogamous, diploid legume species for classical and molecular genetics. *Plant J.* **2**:487-496.

Hardison, R. C. (1996). A brief history of hemoglobins: plant, animal, protist and bacteria. *Proc. Natl. Acad. Sci. USA.* **93**:5675-5679.

Harlow, E., and Lane, D. (1988). Antibodies. A laboratory manual. Cold Spring Harbor Laboratory.

Hartwig, U. A., Joseph, C. M., and Phillips, D. A. (1991). Flavonoids released naturally from alfalfa seeds enhance growth rate of *Rhizobium meliloti*. *Plant Physiol.* **95**:797-803.

Hawkins, F. K. L., Kennedy, C., and Johnston, A. W. B. (1991). A *Rhizobium leguminosarum* gene required for symbiotic nitrogen fixation melanin synthesis and normal growth on certain media. *J. Gen. Microbiol.* **137**:1721-1728.

Hawkins, F. K. L. and Johnston, A. W. B. (1988). Transcription of a *Rhizobium leguminosarum* biovar *phaseoli* gene needed for melanin synthesis is activated by *nifA* of *Rhizobium* and *Klebsiella pneumoniae*. *Mol. Microbiol.* **2**:331-337.

Hayat, M. A. (1989). Principles and techniques of electron microscopy. CRS Press.

Heinrich, K. (1999). Ecological and molecular studies of rhizobial rhizopines. *Ph.D Thesis, University of Adelaide*.

Hill, S. (1992). Physiology of nitrogen fixation in free-living heterotrophs. In Biological nitrogen fixation. G. Stacey, Burris, R. H. and Evans, H. J. eds., pp. 87-134. Chapman and Hall, New York.

Hirsch, A. M. (1992). Developmental biology of legume nodulation. *New Phytol.* **122**:211-237.

Hirsch, A. M., Bhavaneswari, T. V., Torrey, J. G., and Bisseling, T. (1989). Early nodulin genes are induced in alfalfa root outgrowths elicited by auxin transport inhibitors. *Proc. Natl. Acad. Sci. USA* **86**:1244-1248.

Ho, S. C., Schindler, M., and Wang, J. L. (1990). Carbohydrate binding activities of *Bradyrhizobium japonicum*. II. Isolation and characterization of galactose specific lectin. *J. Cell Biol.* **111**:1639-1645.

Holgate, C. S., Jackson, P., Cowen, P. N., and Bird, C. C. (1983). Immunogold-silver staining; new method of immunostaining with enhanced sensitivity. *J. Histochem. Cytochem.* **31**:938-942.

Honeycutt, R. J., McClelland, M., and Sobral, B. W. S. (1993). Physical map of the genome of *Rhizobium meliloti* 1021. *J. Bacteriol.* **175**:6945-6948.

Honma, M. A., and Ausubel, F. M. (1987). *Rhizobium meliloti* has three functional copies of the *nodD* symbiotic regulatory gene. *Proc. Natl. Acad. Sci.* **84**:8558-8562.

Hoover, T. R., Santero, E., Porter, S., and Kustu, S. (1990). The integration host factor stimulates the interaction of RNA polymerase with NIFA, the transcriptional activation for nitrogen fixation operons. *Cell* **63**:11-22.

Hooykaas, P. J. J., den Dulk-Ras, H., Regensburg-Tuink, A. J. G., vanBrussel, A. A. N., and Schilperoort, R. A. (1985). Expression of a *Rhizobium phaseoli* sym plasmid in *R. trifolii* and *Agrobacterium tumefaciens* incompatibility with a *R. trifolii* sym plasmid. *Plasmid* **14**:47-54.

Hooykaas, P. J. J., Klapwijk, P. M., Nuti, M. P., Schilperoort, R. A., and Rorsch, A. (1977). Transfer of the *Agrobacterium tumefaciens* Ti plasmid to avirulent agrobacteria and to *Rhizobium ex planta*. *J. Gen. Microbiol.* **98**:477-484.

Hooykaas, P. J. J., Snijdwint, F. G. M., and Schilperoort, R. A. (1982). Identification of the sym plasmid of *Rhizobium leguminosarum* strain 1001 and its transfer to, and expression in other rhizobia and *Agrobacterium tumefaciens*. *Plasmid* **8**:73-82.

Hungria, M., and Stacey, G. (1997). Molecular signals exchange between host plants and rhizobia: basic aspects and potential application in agriculture. *Soil Biol. Biochem.* **29**:819-830.

Hynes, M. F., and McGregor, N. F. (1990). Two plasmids other than the nodulation plasmid are necessary for the formation of nitrogen-fixing nodules by *Rhizobium leguminosarum*. *Mol. Microbiol.* **4**:567-574.

Inoue, J., Shaw, J.P., Rekik, M., and Harayama, S. (1995). Overlapping substrate specificities of benzaldehyde dehydrogenase (the *xylC* gene product) and 2-hydroxymuconic semialdehyde dehydrogenase (the *xylG* gene product) encoded by TOL plasmid pww0 of *Pseudomonas putida*. *J. Bacteriol.* **177**:1196-1201.

Island, M. D., Wei, B. -Y., and Kadner, R. J. (1992). Structure and function of *uhp* genes for the sugar phosphate transport system in *Escherichia coli* and *Salmonella typhimurium*. *J. Bacteriol.* **174**:2754-2762.

Jarvis, B. D. W., and Tighe, S. W. (1994). Rapid identification of *Rhizobium* species based on cellular fatty acid analysis. *Plant Soil* **161**:31-34.

Jensen, E. S. (1987). Inoculation of pea by application of *Rhizobium* in the planting furrow. *Plant and Soil* **97**:63-70.

Jordan, D. C. (1984). Rhizobiaceae conn 1938, 321AL., Volume 1, N. R. Krieg and J. G. Holt, eds. (Baltimore: William and Wilkins).

Journet, E. P., Pichon, M., Dedieu, A., De Billy, F., Truchet, G., Barker, D. G. (1994). *Rhizobium meliloti* Nod factors elicit cell-specific transcription of the ENOD12 gene in transgenic alfalfa. *Plant J.* **6**:241-249.

Kaijalainen, S., and Lindström, K. (1989). Restriction fragment length polymorphism analysis of *Rhizobium galgae* strains. *J. Bacteriol.* **171**:5561-5566.

Kannenberg, E. J., and Brewin, N. J. (1989). Expression of a cell surface antigen is regulated by oxygen and pH for *Rhizobium leguminosarum* strain 3841. *J. Bacteriol.* **171**:4543-4548.

Kape, R., Parniske, M., and Werner, D. (1991). Chemotaxis and *nod* gene activity of *Bradyrhizobium japonicum* in response to hydroxycinnamid acids and isoflavonoids. *Appl. Environ. Microbiol.* **57**:316-323.

Kennedy, C. (1989). The genetics of nitrogen fixation., D. A. Hopwood and K. E. Chater, eds. (London: Academic Press).

Kerstens, K., and de Ley, J. (1984). *Agrobacterium*, 9 Edition, Volume 1, N. R. Kreig and J. C. Holt, eds. (Baltimore: Williams and Wilkins).

Keyser, H. H., Bohlool, B. B., Hu, T. S., and Weber, D. F. (1982). Fast growing rhizobia isolated from root nodules of soybean. *Science* **215**: 1631-1632.

Koes, R. E., Quattrocchio, F., and Mol, J. N. M. (1994). The flavonoid biosynthetic pathway in plants: function and evolution. *BioEssays* **16**:123-128.

Kosturko, L. D., Daub, E., and Murialdo, H. (1989). The interaction of *E. coli* integration host factor and lambda Cos DNA: multiple complex formation and protein-induced bending. *Nucl. Acids Res.* **17**:317-334.

Kowalski, M. (1967). Transduction in *Rhizobium meliloti*. *Acta Microbiol. Polonica.* **16**:7-12.

Kurkdjian, A. C. (1995). Role of differentiation of root epidermal cells in Nod factor (from *Rhizobium meliloti*)-induced root-hair depolarization of *Medicago sativa*. *Plant Physiol.* **107**:783-790.

Krishnan, H. B., and Pueppke, S. G. (1991). Repetitive sequences with homology to *Bradyrhizobium japonicum* DNA and the T-DNA of *Agrobacterium rhizogenes* are closely linked to *nodABC* of *Rhizobium fredii* USDA257. *Mol. Plant-Microbe Interact.* **4**:521-529.

Krishnan, H. B., and Pueppke, S. G. (1993). Flavonoid inducers of nodulation genes stimulate *Rhizobium fredii* USDA257 to export proteins into the environment. *Mol. Plant-Microbe Interact.* **6**:107-113.

Kyte, J., and Doolittle, R. F. (1982). A simple method for displaying the hydropathy character of a protein. *J. Mol. Biol.* **157**:105-132.

Laber, B., Gomis-Ruth, F., Romao, M. J., and Huber, R. (1992). *Escherichia coli* dihydroxydipicolinate synthase. Identification of the active site and crystallisation. *Biochem. J.* **288**:691-695.

Lamb, J. W., Hombrecher, G., and Johnston, A. W. B. (1982). Plasmid-determined nodulation and nitrogen-fixation abilities in *Rhizobium phaseoli*. *Mol. Gen. Genet.* **186**:449-452.

Languerre, G., Bardin, M., and Amarger, N. (1993). Isolation from soil of symbiotic and non-symbiotic *Rhizobium leguminosarum* by DNA-hybridisation. *Can. J. Microbiol.* **39**:1142-1145.

Leong, J. M., Nunes-Duby, S., Lesser, C. F., Youderian, P., Susskind, M. M., and Landy, A. (1985). The $\phi 80$ and P22 attachment sites: primary structure and interaction with *Escherichia coli* integration host factor. *J. Biol. Chem.* **260**:4468-4477.

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

Lie, T. A. (1978). Symbiotic specialisation in pea plants: The requirement for specific *Rhizobium* strains for peas from Afganistan. *Ann. Appl. Biol.* **88**:462-469.

Lie, T. A., Goktan, D., Engin, M., Pijnenborg, J., and Aularsal, E. (1987). Co-evolution of the legume-*Rhizobium* association. *Plant Soil* **100**:171-181.

Ligon, J. M. (1990). Molecular genetic of nitrogen fixation in plant-bacteria symbiosis, J. P. Nakas and C. Hahedorn, eds. New York: Macmillan Publishing Company.

Lindström, K., Paulin, L., Roos, C., and Souminen, L. (1995). Nodulation genes of *Rhizobium galegae*, in Nitrogen fixation: Fundamentals and application, Tikhonovich, I. A. *et al.*, eds. The Netherlands.

(Publisher!)

Lochner, H. H., Strijdom, B. W. and Law, I. J. (1989). Unaltered nodulation competitiveness of a strain of *Bradyrhizobium* sp. (Lotus) after a decade in soil. *Appl. Environ. Microbiol.* **55**:3000-3008.

Loh, J., Garcia, M., Yuen, J., and Stacey, G. (1996). *Nod* gene regulation in *Bradyrhizobium japonicum*. In: Proceedings of the 8th International symposium of Molecular Plant-Microbe Interaction. (Stacey, G., Mullin, B., and Gresshoff, P. M.). Knoxville, Tennessee.

Long, S. R. (1996). *Rhizobium* symbiosis: Nod factors in perspectives. *The Plant Cell* **8**:1885-1889.

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

López-Lare, I. M., van den Berg, J. D. J., Thomas-Oates, J. E., Glushka, J., Lugtenberg, B. J. J., and Spaink, H. P. (1995). Structural identification of lipo-

chitin oligosaccharide nodulation signals of *Rhizobium loti*. *Mol. Microbiol.* **15**:627-638.

Machado, H. R., Yates, M. G., Funayama, S., Rigo, L. U., Steffens, M. B. R., Souza, E. M., and Pedrosa, F. O. (1995). The *ntrBC* genes of *Azospirillum brasilense* are part of a *nifH3*-like-*ntrB*-*ntrC* operon and are negatively regulated. *Can. J. Microbiol.* **41**:674-684.

Madsen, S.M., Albrechtsen, B., Hansen, E.B., and Israelsen, H. (1996). Cloning and transcription analysis of two threonine biosynthetic genes from *Lactococcus lactis* MG1614. *J. Bacteriol.* **178**:3689-3694.

Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982). Molecular cloning. A laboratory manual. Cold Spring Harbor laboratory.

Mansouri, K., and Piepersberg, W. (1991). Genetics of streptomycin production in *Streptomyces griseus*: nucleotide sequence of five genes, *strFGHIK*, including a phosphatase gene. *Mol. Gen. Genet.* **228**:459-469.

Martínez, E., Palacios, R., and Sanchez, F. (1987). Nitrogen fixing nodules induced by *Agrobacterium tumefaciens* harbouring *Rhizobium phaseoli* plasmids. *J. Bacteriol.* **169**:2828-2834.

Martínez-Romero, E. (1994). Recent developments in *Rhizobium* taxonomy. *Plant Soil* **161**:11-20.

Martínez, E., Romero, D., and Palacios, R. (1990). The *Rhizobium* genome. *Crit. Rev. Plant Sci.* **9**:59-93.

Martínez-Romero, E., and Caballero-Mellado, J. (1996). *Rhizobium* phylogenies and bacterial genetic diversity. *Crit. Rev. Plant Sci.* **15**:113-140.

- Martínez-Romero, E., and Rosenblueth, M. (1990). Increased bean (*Phaseolus vulgaris* L.) nodulation competitiveness of genetically modified *Rhizobium* strains. *Appl. Environ. Microbiol.* **56**:2384-2388.
- Martínez-Romero, E., Segovia, L., Mercante, F. M., Franco, A. A., Graham, P., and Pardo, M. A. (1991). *Rhizobium tropici*, a novel species nodulating *Phaseolus vulgaris* L. beans and *Leucaena* sp. trees. *Int. J. Syst. Bacteriol.* **41**:417-426.
- Mathews, A., Carrol, B. J., Gresshoff, P. M. (1989). Development of *Bradyrhizobium* infection in a super-nodulating and non-nodulating mutant of soybean (*Glycine max* (L) Merr.). *Protoplasma* **150**:40-47.
- Mauro, V. P., and Verma, D. P. S. (1988). Transcriptional activation in nuclei from uninfected soybean of a set of genes involved in symbiosis with *Rhizobium*. *Mol. Plant-Microbe Interact.* **1**:46-51.
- McSpadden Gardener, B.B. and de Bruijn, F.J. (1998). Detection and isolation of novel rhizopine-catabolising bacteria from the environment. *Appl. Environ. Microbiol.* **64**:4944-4949.
- Meade, H. M., Long, S. R., Ruvkun, G. B., Brown, S. E., and Ausubel, F. M. (1982). Physical and genetic characterisation of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn5 mutagenesis. *J. Bacteriol.* **149**:114-122.
- Mellor, R. B., Christensen, T. M. I. E., and Werner, D. (1986). Choline kinase II is present in nodules that synthesize stable peribacteroid membranes. *Proc. Natl. Acad. Sci. USA* **83**:659-663.
- Mercado-Blanco, J., and Toro, N. (1996). Plasmids in rhizobia: the role of non-symbiotic plasmids. *Mol. Plant-Microbe Interact.* **9**:535-545.

Merrick, M. J. (1980). Analysis of the regulation of the *Klebsiella pneumonia* nitrogen fixation (*nif*) gene cluster with gene fusions. *Nature* **286**:128-132.

Merrick, M. J. (1993). In a class of its own - the RNA polymerase sigma factor σ^{54} (σ^N). *Mol. Microbiol.* **10**:903-909.

Merrick, M., Filser, M., Dixon, R., Elmerich, C., Sibold, J., and Houmard, J. (1980). Use of translocatable genetic elements to construct a fine structure map of the *Klebsiella pneumonia* nitrogen fixation (*nif*) gene cluster. *J. Gen. Microbiol.* **117**:509-520.

Miflin, B. J., and Cullimore, J. V. (1984). Nitrogen assimilation in the legume-rhizobium symbiosis: a joint endeavour, D. P. S. Verma and T. Hohn, eds. (New York: Springer-Verlag). (Title !)

Moreira, F., Gillis, M., Pot, B., Kersters, K., and Franco, A. A. (1993). Characterisation of rhizobia isolated from different divergent groups of tropical Leguminosae by comparative polyacrylamide gel electrophoresis of their total proteins. *Syst. Appl. Microbiol.* **16**:135-138.

Morret, E., and Buck, M. (1989). In vivo studies on the interaction of RNA polymerase- σ^{54} with the *Klebsiella pneumonia* and *Rhizobium meliloti nifH* promoters. The role of NifA in the formation of open promoter complex. *J. Mol. Biol.* **210**:65-77.

Morret, E., and Segovia, L. (1993). The sigma 54 bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains. *J. Bacteriol.* **175**:6067-6074.

Mulligan, M. E., Buikema, W. J., and Haselkorn, R. (1988). Bacterial-type ferredoxin genes in the nitrogen fixation regions of the *Cyanobacterium Anabaena* sp. strain PCC7120 and *Rhizobium meliloti*. *J. Bacteriol.* **170**:4406-4410.

Murphy, P. J., and Ryder, M. H. (1994). The use of rhizopines for artificial rhizosphere colonisation. In: Third international workshop on plant growth-promoting rhizobacteria, (M. H. Ryder, P. M. Stephens and G. D. Bowen eds.). Adelaide, South Australia.

Murphy, P. J., and Saint, C. P. (1992). Rhizopines in the Legume-*Rhizobium* Symbiosis, D. P. S. Verma, ed. (Boca Raton, Fla: CRC Press).

Murphy, P. J., Heycke, N., Banfalvi, Z., Tate, M. E., de Bruijn, F. J., Kondorosi, A., Tempe, J., and Schell, J. (1987). Genes for the catabolism and synthesis of an opine like compound in *Rhizobium meliloti* are closely linked and on the Sym plasmid. *Proc. Natl. Acad. Sci. USA*. **84**:493-497.

Murphy, P. J., Heycke, N., Trenz, S. P., Ratet, P., de Bruijn, F. J., and J., S. (1988). Synthesis of an opine-like compound, a rhizopine, in alfalfa nodules is symbiotically regulated. *Proc. Natl. Acad. Sci. USA*. **85**:9133-9137.

Murphy, P. J., Trenz, S. P., Grzenski, W., de Bruijn, F. J., and Schell, J. (1993). The *Rhizobium meliloti* rhizopine *mos* locus in a mosaic structure facilitating its symbiotic regulation. *J. Bacteriol.* **175**:5193-5204.

Murphy, P. J., Wexler, W., Grzenski, W., Rao, J. P., and Gordon, D. (1995). Rhizopines-their role in symbiosis and competition. *Soil Biol. Biochem.* **27**:525-529.

Mylona, P., Pawlowski, K., and Bisseling, T. (1995). Symbiotic nitrogen fixation. *The Plant Cell* **7**:869-885.

Nap, J. P., and Bisseling, T. (1990). Developmental biology of a plant-prokaryote symbiosis: the legume root nodule. *Science* **250**:948-954.

- Nash, H. A., and Robertson, C. A. (1981). Purification and properties of the *Escherichia coli* protein factor required for λ integrative recombination. *J. Biol. Chem.* **256**:9246-9253.
- Niegemann, E., Schulz, A., and Bartsch, K. (1993). Molecular organization of the *Escherichia coli* *gab* cluster: nucleotide sequence of the structural genes *gabD* and *gabP* and expression of the GABA permease gene. *Arch. Microbiol.* **160**:454-460.
- Niner, B.M. and Hirsch, A.M. (1998). Review: How many *Rhizobium* genes in addition to *nod*, *nif*fix and *exo* are needed for nodule development and function? *Symbiosis* **24**:51-102.
- North, A. K., Klose, K. E., Stedman, K. N., and Kustu, S. (1993). Prokaryotic enhancer binding proteins reflect eukaryotic-like modularity: the puzzle of nitrogen regulatory protein C. *J. Bacteriol.* **175**:661-668.
- Paau, A. S., and Brill, W. J. (1982). Comparison of the genetic arrangement and relative transcription of the nitrogenase gene in *Rhizobium meliloti*. *Microbiol.* **28**:1330-1339.
- Paau, A. S., Cowles, J. R., and Raveed, D. (1978). Development of bacteroids in alfalfa (*Medicago sativa*) nodules. *Plant Physiol.* **62**:526-530.
- Pagel, J. M., Winkelman, J. W., Adams, C. W., and Hatfield, G. W. (1992). DNA topography-mediated regulation of transcription initiation from the tandem promoters of the *ilvGMEDA* operon of *Escherichia coli*. *J. Mol. Biol.* **224**:919-935.
- Palmer, B.R. and Marinus, M.G. (1994). The *dam* and *dcm* strains of *Escherichia coli*. *Gene* **143**:7-8.

Pankhurst, C. E., and Craig, A. S. (1978). Effect of oxygen concentration, temperature and combined nitrogen on the morphology and nitrogenase activity of *Rhizobium* sp. strain 32H1 in agar culture. *J. Gen. Microbiol.* **106**:207-219.

Patriarca, E. J., Riccio, A., Tate, R., Colonna-Romano, S., Inccarino, M., and Defez, R. (1993). The *ntrBC* genes of *Rhizobium leguminosarum* are part of a complex operon subject to negative regulation. *Mol. Microbiol.* **9**:569-577.

Pawlowski, K., and Bisseling, T. (1996). Rhizobial and actinorhizal symbiosis: What are the shared features? *The Plant Cell* **8**:1899-1913.

Peréz-Martín, J., Rojo, F., and De Lorenzo, V. (1994). Promoters responsive to bending: a common theme in prokaryotic gene expression. *Microbiol. Rev.* **58**: 268-290.

Perret, X., Frelberg, C., Rosenthal, A., Broughton, J., and Fellay, R. (1999). High-resolution transcriptional analysis of the symbiotic plasmid of *Rhizobium* sp. NGR234. *Mol. Microbiol.* **32**:415-425.

Peters, N. K., and Verma, D. P. S. (1990). Phenolic compound^s as regulators of gene expression in plant-microbe interactions. *Mol. Plant Microbe Interact.* **3**:4-8.

Petes, T. D., and Hill, C. W. (1988). Recombination between repeated genes in microorganisms. *Annu. Rev. Genet.* **22**:147-156.

Phillips, D. A. (1992). Flavonoids: plant signals to soil microbe, H. A. Stafford and R. K. Ibrahim, eds. (New York: Plenum Press).

/Title!

Phillips, D. A., Dakora, F. D., Sande, E., Joseph, C. M., and Zon, J. (1994). Synthesis, release and transmission of alfalfa signals to rhizobial symbionts. *Plant Soil* **161**:69-80.

Postma, J., Hok-A-Hin, C. H., and van Veen, J. A. (1990). Role of microniches in protecting introduced *Rhizobium leguminosarum* biovar *trifolii* against competition and predation in soil. *Appl. Environ. Microbiol.* **56**:495-502.

Prinsen, E., Chauvaux, N., Schmidt, J., John, M., Wieneke, U., De Greet, J., Schell, J., and Van Onckelen, H. (1991). Stimulation of indole-3-acetic acid production in *Rhizobium* by flavonoids. *FEBS. Lett.* **282**:53-55.

Prosperi, J. M. (1993). Selection of annual medics for French Mediterranean regions. In: Workshop on introducing the Ley farming system in the Mediterranean basin, (S. Christensen, L. Materon, M. Falcinelli and P. Cocks eds.). Aleppo Syria : International Centre for Agricultural Research in the Dry Areas, 173-191.

Pueppke, S. G. (1996). The genetic and biochemical basis for nodulation of legumes by rhizobia. *Crit. Rev. Plant Sci.* **16**:1-51.

Quinto, C., de la Vega, H., Flores, M., Leemans, J., Cevallos, M. A., Pardo, M. A., Azpiroz, R., Girard, M. D. L., Calra, E., and Palcios, R. (1985). Nitrogenase reductase: a functional multigene family in *Rhizobium phaseoli*. *Proc. Natl. Acad. Sci. USA.* **82**:1170-1174.

Rae, A. L., Bonfante-Fasolo, P., and Brewin, N. J. (1992). Structure and growth of infection thread in the legume symbiosis with *Rhizobium leguminosarum*. *Plant J.* **2**:385-391.

Rao, J. P., Grzinski, W., and Murphy, P. J. (1995). *Rhizobium meliloti* lacking *mosA* synthesizes the rhizopine *scyllo*-inosamine in place of 3-*O*-methyl-*scyllo*-inosamine. *Microbiol.* **141**:1683-1690.

Renalier, M. H., Batut, J., Ghai, J., Terzaghi, B., Gherardi, M., David, M., Garnerone, A. M., Vasse, J., truchet, G., Huguet, T., and Boistard, P. (1987). A new

symbiotic cluster on the pSym megaplasmid of *Rhizobium meliloti* 2011 carries a functional *fix* gene repeat and a *nod* locus. *J. Bacteriol.* **169**:2231-2234.

Reuhs, B. L., Kin, M. J. S., Badgett, A., and Carlson, R. W. (1994). Production of cell associated polysaccharides of *Rhizobium fredii* USDA205 is modulated by apigenin and host root extract. *Mol. Plant-Microbe Interact.* **7**:240-247.

Rjtzsckke, O., Falk, K., Mack, J., Lau, J.M., Jung, G. and Strominger, J.L. (1999). Conformational variants of class II MHC/peptide complexes induced by N- and C-terminal extensions of minimal peptide epitopes. *Proc. Natl. Acad. Sci. USA.* **13**:7445-7450.

Roberts, G. P., and Bill, W. J. (1980). Gene-product relationships of the *nif* (nitrogen fixation) regulon of *Klebsiella pneumonia*. *J. Bacteriol.* **144**:210-216.

Robertson, C. A., and Nash, H. A. (1988). Bending of the bacteriophage Lambda attachment site by *Escherichia coli* integration host factor. *J. Mol. Chem.* **263**:3554-3557.

Robertson, J.G., Warburton, M.p., Lyttleton, P., Fordyce, A.M., and Bullivant, S. (1978). Membrane in lupin root nodules II Preparation and properties of peribacteroid membranes and bacteroid envelope inner membranes from developing lupin nodules. *J. cell Sci.* **30**:151-174.

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

Rodriguez-Quinones, F., Banfalvi, Z., Murphy, P., and Kondorosa, A. (1987). Interspecies homology of nodulation genes in *Rhizobium*. *Plant Mol. Biol.* **8**:61-75.

- Rome, S., Fernandez, M. P., Brunel, B., Normand, P., and Cleyet-Marel, J.-C. (1996). *Sinorhizobium medicae* sp. nov., isolated from annual *Medicago* spp. *Int. J. Syst. Bacteriol.* **46**:972-980.
- Romero, D., Brom, S., Martínez-Salazar, J., de Lourdes Girard, M., Palacios, R., and Davila, G. (1991). Amplification and ~~deletion~~^{deletion} of a *nod-nif* region in the symbiotic plasmid of *Rhizobium phaseoli*. *J. Bacteriol.* **173**:2435-2441.
- Rossbach, S., Kulpa, D. A., Rossbach, U., and de Bruijn, F. J. (1994). Molecular and genetic characterisation of rhizopine catabolism (*mocABRC*) genes of *Rhizobium meliloti* L5-30. *Mol. Gen. Genet.* **245**:11-24.
- Ruvkun, G. B., Sundaresan, V., and Ausubel, F. M. (1982). Directed transposon Tn5 mutagenesis and complementation analysis of *Rhizobium meliloti* symbiotic nitrogen fixation genes. *Cell* **29**:551-559.
- Saint, C. P., Wexler, M., Murphy, P. J., Tempe, J., Tate, M. E., and Murphy, P. J. (1993). Characterisation of genes for synthesis and catabolism of a new rhizopine induced in nodules of *Rhizobium meliloti* Rm220-3: extension of the rhizopine concept. *J. Bacteriol.* **175**:5205-5215.
- Sambrook, J., Maniatis, T., and Fritsch, E (1989). Molecular cloning: A Laboratory Manual. Cold Spring Harbour Laboratory, Cold Spring Harbour, N.Y.
- Sanchez, F., Padilla, J. E., Perez, H., and Lara, M. (1991). Control of nodulin genes in root-nodule development and metabolism. *Annu. Rev. Plant Physiol.* **42**:507-528.
- Sanger, F., Nicklen, S., and Coulson, R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* **74**:5463-5467.

Santero, E., Hoover, T. R., North, A. K., Berger, D. K., Porter, S. C., and Kustu, S. (1992). Role of integration host factor in stimulating transcription from the σ^{54} -dependent *nifH* promoter. *J. Mol. Biol.* **227**:602-620.

Sauer, H., Wild, A., and Rühle, W. (1987). The effect of phosphinothricin (glufosinate) on photosynthesis II. The causes of inhibition of photosynthesis. *Z. Naturforsch.* **42c**:270-278.

Sawada, H., Leki, H., Oyaizu, H., and Matsumoto, S. (1993). Proposal for rejection of *Agrobacterium tumefaciens* and ~~raise~~^{revised} description^s for the genus *Agrobacterium* and for *Agrobacterium radiobacter* and *Agrobacterium rhizogens*. *Int. J. Syst. Bacteriol.* **43**:694-702.

Scapin, G., Blanchard, J.S., and Sacchettini, J.C. (1995). Three ~~dimensional~~^s structure of *Escherichia coli* dihydrodipicolinate reductase. *Biochem.* **34**:3502-3512.

Scheres, B., van de Wiel, C., Zalensky, A., Horvath, B., Spaink, H. P., *et al.* (1990). The ENOD12 gene product is involved in the infection process during the pea-*Rhizobium* interaction. *Cell* **60**:281-294.

Schlaman, H. M., Horvath, B., Vijgenboom, E., Okker, R. J. H., and Lugtenberg, B. J. J. (1991). Suppression of nodulation gene expression in bacteroids of *Rhizobium leguminosarum* biovar *viciae*. *J. Bacteriol.* **173**:4277-4287.

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.

Schmidt, P. E., Broughton, W. J., and Werner, D. (1994). Nod factors of *Bradyrhizobium japonicum* and *Rhizobium* sp. NGR234 induce flavonoid accumulation in soybean root exudate. *Mol. Plant-Microbe Interact.* **7**:384-390.

- Schofield, P. R., Ridge, R. W., Rolfe, B. G., Shine, J., and Watson, J. M. (1984). Host specificity nodulation is encoded on a 14 kb DNA fragment *Rhizobium trifolii*. *Plant Mol. Biol.* **3**:3-6.
- Schultze, M., and Kondorosi, A. (1998). Regulation of symbiotic root nodule development. *Ann. Rev. Genet.* **32**:33-57.
- Segovia, L., Pinero, D., Palacios, R., and Martínez-Romero, E. (1991). Genetic structure of a soil population of non-symbiotic *Rhizobium leguminosarum*. *Appl. Environ. Microbiol.* **57**:426-433.
- Skøt, L., and Egsgaard, H. (1984). Identification of ononitol and *O*-methyl-scyllo inositol in pea root nodules. *Planta* **161**:32-36.
- Simon, R., Priefer, U., and Pühler, A. (1983). A broad host range mobilisation system for *in vivo* genetic engineering: Transposon mutagenesis in Gram negative bacteria. *Biotechnol.* **1**:784-791.
- Smit, G., Kijne, J. W., and Lugtenberg, B. J. J. (1987). Both cellulose fibril 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., Swart, S., Lugtenberg, B. J. J., and Kijne, J. W. (1992). Molecular mechanism of attachment of *Rhizobium* bacteria to plant roots. *Mol. Microbiol.* **6**:2897-2903.
- Soto, M. J., Zorzano, A., Mercadoblanco, J., Lepek, V., Olivares, J., and Toro, N. (1993). Nucleotide sequence and characterisation of *Rhizobium meliloti* nodulation competitiveness gene *nfe*. *J. Mol. Biol.* **229**:570-576.
- Spaink, H. P. (1995). The molecular basis of infection and nodulation in rhizobia: the ins and outs of symbiogenesis. *Annu. Rev. Phytopathol.* **33**:345-368.

Spaink, H. P., Sheeley, D. M., Van Russel, A. A. N., Glushka, J. N., York, W. S., Tak, T., Geiger, O., Kennedy, E. P., Reinhold, V. N., and Lugtenberg, B. J. J. (1991). A novel, highly unsaturated fatty acid moiety of lipo-oligosaccharide signals determines host specificity of *Rhizobium leguminosarum*. *Nature* **354**:125-130.

Spaink, H. P., Wijfjes, A. H. M., van der Drift, K. M. G. M., Hoverkamp, J., Thomas-Oates, J. E., and Lugtenberg, B. J. J. (1994). Structural identification of metabolites produced by the NodB and NodC proteins of *Rhizobium leguminosarum*. *Mol. Microbiol.* **13**:821-831.

Sprent, J. (1989). Which steps are essential for the formation of functional legume nodules? *New Phytol.* **111**:129-153.

Stephanou, M. and Manetas, Y. (1997). The effects of seasons, exposure, enhanced UV-B radiation, and water stress on leaf epicuticular and internal UV-B absorbing capacity of *Cistus creticus*: a Mediterranean field study. *J. Expt. Bot.* **48**:1977-1985.

Stokkermans, J. J., and Peters, N. K. (1994). *Bradyrhizobium elkanii* lipo-oligosaccharides signals induce complete nodules structure on *Glycine soja* Siebold et Zucc. *Planta* **193**: 413-420.

Sutton, J. M., Lea, E. J. A., and Downie, A. (1994). The nodulation-signal protein NodO from *Rhizobium leguminosarum* biovar *viciae* forms ion channel in membranes. *Proc. Natl. Acad. Sci. USA.* **91**:9990-9994.

Swart, S., Logman, T. J., Smit, G., Lugtenberg, B. J. J., and Kijne, J. W. (1994). Purification and partial characterisation of a glycoprotein from pea (*Pisium sativum*) with receptor activity for rhicadhesin, an attachment protein of *Rhizobiaceae*. *Plant Mol. Biol.* **24**:171-183.

Tempé, J., and Petit, A. (1983). La Piste des opines, [↑]A. Puhler, ed. (Berlin: Springer-Verlag). *Molecular genetics of the bacteria plant interaction.*

Tepfer, D., Goldmann, A., Fleury, V., Maille, M., Message, B., Pamboukdjian, N., Boivan, C., Denaire, A., Rosenberg, C., Lallemant, J. Y., Descoins, C., Chappin, I., and Amarger, N. (1988a). Calystegins, nutritional, mediators in plant-microbe interaction, [↑]R. Palacois and D. P. S. Verma, eds. (St Paul, Minn: American Phytopathological Society Press). *Molecular genetics of plant-microbe interactions*

Tepfer, D., Goldman, A., Pamboukdjian, N., Maille, M., Lepingle, A., Chevalier, D., Denarie, J., and Rosenberg, C. (1988b). A plasmid of *Rhizobium meliloti* 41 encodes catabolism of two compounds from root exudate of *Calystegia sepium*. *J. Bacteriol.* **170**:1153-1161.

Thompson, J., and Donkersloot, J. A. (1992). N-(carboxyalkyl) amino acids: occurrence, synthesis, and functions. *Ann. Rev. Biochem.* **61**:517-557.

Thompson, J. F., and Landy, A. (1988). Empirical estimation of protein induced DNA bending angles: application of lambda site-specific recombination complexes. *Nucl. Acids Res.* **16**:9687-9705.

Timmers, A.C.J., Auriac, M., and Truchet, G. (1999). Refined analysis of early symbiotic steps of *Rhizobium-Medicago* interaction in relationship with microtubular cytoskeleton rearrangements. *Development* **126**:3617-3628.

Triplett, E. W. (1990a). The molecular genetics of nodulation competitiveness in *Rhizobium* and *Bradyrhizobium*. *Mol. Plant -microbe Interact.* **3**:199-206.

Triplett, E. W. (1990b). Construction of a symbiotically effective strain of *Rhizobium leguminosarum* bv. *trifolii* with increased nodulation competitiveness. *Appl. Environ. Microbiol.* **56**:98-103.

Triplett, E. W., and Sadowsky, M. J. (1992). Genetics of competition for nodulation of legumes. *Ann. Rev. Microbiol.* **60**:4163-4166.

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

Unwin, N. (1998). The nicotinic acetylcholine receptor of the *Torpedo* electric ray. *J. Struct. Biol.* **121**:181-190.

van Egeraat, A. W. S. M. (1975). The possible role of homoserine in the development of *Rhizobium leguminosarum* in the rhizosphere of pea seedlings. *Plant and Soil* **42**:381-386.

van Spronsen, P., Bakhuizen, R., vanBrussel, A. A. N., and Kijne, J. N. (1994). Cell wall degradation during infection thread formation by the root nodule bacterium *Rhizobium leguminosarum* in two-step process. *Eur. J. Cell Biol.* **64**:88-94.

Van Etten, H. D., and Pueppke, S. G. (1976). Isoflavonoid phytoalexins, J. Friend and D. R. Threlfall, eds. (London: Academic Press). *Title!*

Van Rhijn, P. and Vanderleyden, J. (1995). The *Rhizobium*-plant symbiosis. *Microbiol. Rev.* **59**:124-142.

Vance, C. P. (1997). Nitrogen fixation capacity. In: Biotechnology and the improvement of forage legumes. B. D. Mckerside and D. C. D. Brown Des. University Press, Cambridge. pp. 357-407.

Vasse, J., de Billy, F., Camut, S., and Truchet, G. (1990). Correlation between ultra-structural differences of bacteroids and nitrogen fixation in alfalfa nodules. *J. Bacteriol.* **172**:4295-4306.

Verma, D. P. S., Miao G. H., Cheon, C. I., and Suzuki, H. (1991). Genesis of root nodules and function of nodulins, in *Advances in Molecular Genetics of Plant-Microbe Interactions*. H. ~~Hann~~^{Hennecke} and D. P. S. Verma, eds, vol. 1, Kluwer Academic, Dordrecht, pp. 291-299.

Vesper, S. J., and Bauer, W. D. (1986). Role of pili (fimbriae) in attachment of *Bradyrhizobium japonicum* to soybean roots. *Appl. Environ. Microbiol.* **52**:134-140.

Vincent, J. M. (1970). A manual for the practical study of the root-nodule bacteria. IBP handbook No. 15 Blackwell, Oxford, pp. 1-45.

Walker, J.B. and Walker, M.S. (1967). Enzymatic synthesis of streptidine from *Scyllo*-inosamine. *Biochem* **6**:3821-3829.

Walker, J.B. and Walker, M.S. (1969). Streptomyces biosynthesis. Transamination reactions involving inosamines and inosadialamines. *Biochem.* **8**:763-770.

Wang, Y., Kolb, A., Buck, M., Wen, J., O'Gara, F., and Buc, H. (1998). CRP interaction with promoter bound σ^{54} RNA polymerase and blocks transcriptional activation of the *dctA* promoter. *EMBO J.* **17**:786-796.

Wehrmann, A., Phillipp, B., Sahm, H. and Eggeling, L. (1998). Different modes of diaminopimelate synthase and their role in cell wall integrity: a study with *Corynebacterium glutamicum*. *J. Bacteriol.* **180**:3159-3165.

Werner, D. (1992). Symbiosis of plants and microbes (London: Chapman and Hall).

Wexler, M. (1994). Molecular biology and distribution of rhizopine genes in *Rhizobium* species. *Ph.D Thesis, The University of Adelaide*.

- Wexler, M., Gordon, D., and Murphy, P. J. (1995). The distribution of inositol rhizopine genes in *Rhizobium* populations. *Soil. Biol. Biochem.* **27**:531-537.
- Wexler, M., Gordon, D., and Murphy, P. J. (1996). Genetic relationships among rhizopine-producing *Rhizobium* strains. *Microbiology* **142**:1059-1066.
- Wheatcroft, R., and Laberge, S. (1991). Identification and nucleotide sequence of *Rhizobium meliloti* insertion sequence IS*Rm3*: similarity between the putative transposase encoded by IS*Rm3* and thise encoded by *Staphylococcus aureus* IS256 and *Thiobacillus ferrooxidans* IST2. *J Bacteriol.* **173**:2530-2538.
- Willems, A., and Collins, M. D. (1992). Evidence for a close genealogical relationship between Afipia, the casual organism of cat scratch disease, *Bradyrhizobium japonicum* and *Blastobacter denitrificans*. *FEMS Microbiol. Lett.* **96**:241-246.
- Willems, A., and Collins, M. D. (1993). Phylogenetic analysis of rhizobia and agrobacteria based on 16S rRNA gene sequences. *Int. J. Syst. Bacteriol.* **43**:305-313.
- Wilson, K. (1991). Preparation of genomic DNA from bacteria. In Current Protocols in Molecular Biology vol. 1 (F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J. G. Seidman, J. A. Smith and K. Struhl) pp. 2.4.1-2.4.5. Green Publishing Associates, Inc. and John Wiley and Sons Inc., Brooklyn, N.Y.
- Wong, E. (1970). Structural and biogenetic relationships of isoflavonoids, W. Hertz, H. Grisebach and A. I. Scott, Eds. Vienna, Austria: Springer Verlag.
- Wright, S. F., Foster, J. G. and Bennett, O. L. (1986). Production and use of monoclonal antibodies for identification of strains of *Rhizobium trifolii*. *Appl. Environ. Microbiol.* **52**:119-123.

Yanagi, M., and Yamsato, K. (1993). Phylogenetic analysis of the family Rhizobiaceae and related bacteria by sequencing of 16S rRNA gene using PCR and DNA sequencer. *FEMS Microbiol. Lett.* **107**:115-120.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC19 vector. *Gene* **33**:103-119.

Young, J. P. W. (1992). Phylogenetic classification of nitrogen fixing organisms, G. Stacey, R. H. Burris and H. J. Evans, eds. (New York: Chapman and Hall).

Young, J. P. W., Downer, H. L., and Eardly, B. D. (1991). Phylogeny of the phototrophic *Rhizobium* strain BTail by polymerase chain reaction-based sequencing of a 16S rRNA gene segment. *J. Bacteriol.* **173**:2271-2277.

Young, J. P. W., and Haukka, K. E. (1996). Diversity and phylogeny of rhizobia. *New Phytol.* **133**:87-94.

Yugari, Y., and Gilvarg, C. (1965). The condensation step in diaminopimelate synthesis *J. Biol. Chem.* **240**:4710-4716.

Zakin, M.M., Duchange, N., Ferrara, P., and Cohen, G.N. (1983). Nucleotide sequence of the *metL* gene of *Escherichia coli*, its product, the bifunctional aspartokinase II-homoserine dehydrogenase II, and the bifunctional product of *thrA* gene, aspartokinase-homoserine dehydrogenase I, derive from a common ancestor. *J. Biol. Chem.* **258**:3028-3031.

Zhang, L., and Kerr, A. (1993). Rapid purification of Ti plasmids from *Agrobacterium* by ethidium bromide treatment and phenol extraction. *Lett. Appl. Microbiol.* **16**:265-268.

Ahlert, J., Distler, J., Mansouri, K. and Piepersberg, W. (1997) Identification of *stsC*, the gene encoding the L-glutamine:*scyllo*-inosose aminotransferase from streptomycin-producing *Streptomyces*. *Arch. Microbiol.* 168: 102-113.

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

Dean, R.M., Rivers, R.L., Zeidel, M.L. and Roberts, D.M. (1999). Purification and functional reconstitution of soybean nodulin 26. An aquaporin with water and glycerol transport properties. *Biochemistry* 38:347-353.

APPENDICES

Appendix 1 - Media and antibiotics list

- 1.1 Hoagland's solution
- 1.2 M9 minimal medium
- 1.3 GTS minimal medium
- 1.4 Rhizopine catabolism medium
- 1.5 Luria Bertani (LB) broth
- 1.6 LB + X-gal +IPTG
- 1.7 SOC medium
- 1.8 TY ^rbroth
- 1.9 Antibiotic concentrations

Appendix 2 - Restriction maps of cloning vectors used

Appendix 3 - BLAST analysis of *R. leguminosarum* *mosB* and *mosC*

Appendix 4 - Tables on rhizopine effect on nodulation

Appendix 5 - Conference papers: oral presentations and poster

Appendix 1 Media and Antibiotics

1.1 Hoagland's medium (without nitrogen, for nodulation) [per litre]

1M $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	2ml
1M KH_2PO_4	1ml
1M KCl	5ml
1M $\text{CaCl}_2 \cdot \text{H}_2\text{O}$	5ml
Micronutrients	1ml
FeEDTA (40g/L MilliQ H_2O)	1ml

Micronutrients

H_3BO_4	2.86g/L
$\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$	1.81g/L
$\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$	0.22g/L
$\text{CuSO}_4 \cdot 7\text{H}_2\text{O}$	0.08g/L
$\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$	0.025g/L

1.2 M9 minimal medium (per litre)

NaHPO_4	6g
KH_2PO_4	3g
NH_4Cl	1g
NaCl	0.5g

Adjust pH to 7.4, add 1.5% agar, autoclave and cool to 55°C, then add

1M $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	1ml
0.1M $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$	1ml
1M Thiamine HCl	1ml
20% (w/v) glucose	10ml

Sterilise separately the above solutions by filtration (glucose and ~~lysine~~ ^{Thiamine}) or autoclaving.

1.3 GTS minimal medium (per litre)

Autoclave ~~15~~¹⁵ g agar in 767 ml water, cool to about 55°C and add the following sterile solutions (solution 6 was filter sterilised while others were autoclaved):

200 ml solution 1

1 ml solutions 2, 3 and 5

10 ml solutions 4, 6 and 7

Solution 1	K ₂ HP0 ₄	0.131 g/l
	NaCl	5.0 g/l
	Tris	15.0 g/l
	Sodium succinate	13.5 g/l
	(NH ₄) ₂ SO ₄	10.0 g/l ; (pH 7.5)
Solution 2	MgSO ₄ .7H ₂ O	246 g/l
Solution 3	CaCl ₂ .2H ₂ O	14.5 g/l
Solution 4	FeCl ₃ .6H ₂ O	27.0 mg/l
Solution 5	Na ₂ MoO ₄	242.0 mg/l
	H ₃ BO ₃	3.0 g/l
	MnSO ₄ .H ₂ O	1.83 g/l
	ZnSO ₄ .7H ₂ O	287.0 mg/l
	CuSO ₄ .5H ₂ O	125.0 mg/l
	CoCl ₂ .H ₂ O	119.0 mg/l
Solution 6	Biotin	200.0 mg/l
Solution 7	Glucose	200.0 g/l

1.4 Rhizopine catabolism medium (per litre)

Bergersen's salts solution (1×)	987 ml
Biotin (0.2 mg/ml)	1 ml
Thiamin (1 mg/ml)	1 ml
Calcium pantothenate (1 mg/ml)	1 ml
Ammonium sulphate (200 mg/ml)	10 ml

Bergersens's salts solution (5×) (per litre)

K_2HPO_4	1.1 g
$MgSO_4 \cdot 7H_2O$	0.5 g
$FeCl_2 (H_2O)_6$ in 0.1M HCl	0.1 g
$CaCl_2 \cdot (H_2O)$	0.2 g

Adjust to pH 6.8 with 0.1M NaOH and autoclave

1.5 Luria-Bertani (LB) broth (per litre)

Bacto-yeast extract	5 g
Bacto-tryptone	10 g
NaCl	5 g

Adjust to pH 7.2 with NaOH. For solid medium, add 1.5% agar and autoclave.

1.6 LB + X-gal + IPTG agar

LB agar (1% agar)	200 ml
20 mg/ml X-gal (5-bromo-4-chloro-3-indolyl- β-D-galactosidase [Sigma] in dimethylformamide)	1 ml
IPTG (isopropyl-thiogalactoside; Sigma) (23.8 mg/ml in H_2O)	1 ml

1.7 SOC medium

Bacto-tryptone	2%
Bacto-yeast extract	0.5%
NaCl.	10 mM
KCl	2.5 mM
$MgCl_2$	10 mM
$MgSO_4$	10 mM
Glucose	20 mM

1.8 TY broth (per litre)

CaCl₂.2H₂O 0.9 g

Bacto-tryptone 5.0 g

Bacto-yeast extract 3.0 g

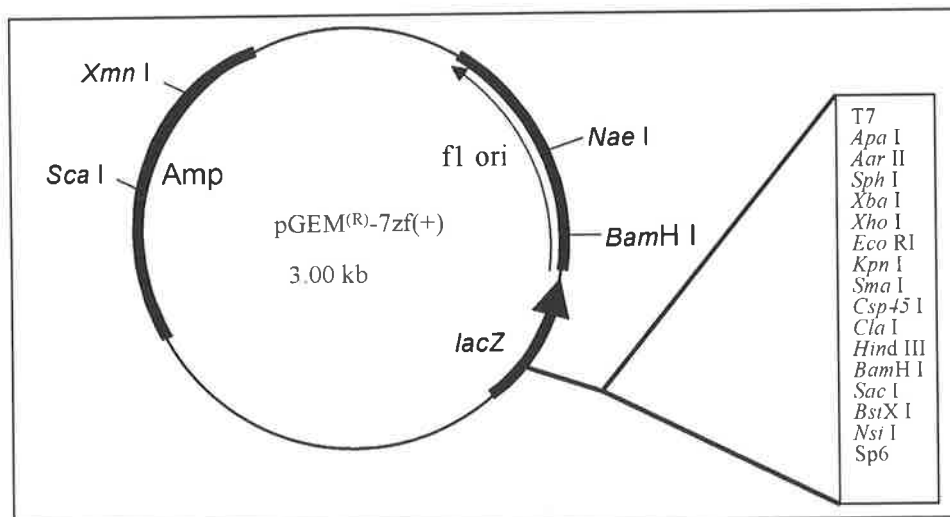
Adjust to pH 6.8. For solid medium add 1.5% agar and autoclave.

1.9 Antibiotic concentrations used in this work

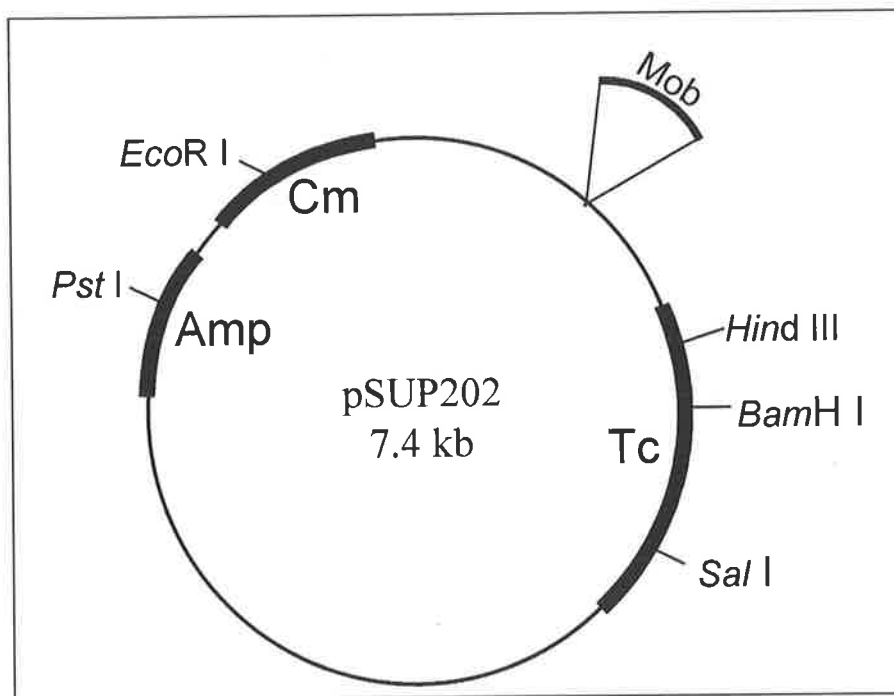
Antibiotic	Concentration for <i>Rhizobium</i> spp. (µl/ml)	Concentration for <i>E.coli</i> (µl/ml)	Stock solution (mg/ml)	Solvent
Ampicillin (sodium salt)	100	100	100	Water
Chloramphenicol	30	30	100	Ethanol
Kanamycin monosulphate	250	30	50	Water
Rifampicin	100	100	10	Methanol
Tetracycline hydrochloride	10	10	10	50% ethanol
Streptomycin sulphate	250	250	250	water

Appendix 2 Restriction maps of vectors

Appendix 2.1 Restriction map of pGEM®-7zf(+)



Reference: Promega corporation.

Appendix 2.2 Restriction map of pSUP202

Reference: Simon *et al.* 1983

Appendix 3 - BLAST analysis of *R. leguminosarum* *mosB* and *mosC*

Appendix 3.1 - BLAST analysis of *R. leguminosarum* *mosB*

Currently in Directory ~waite04/chike

Filter ...

Save sequence in GCG format for marked records to

~waite04/chike

Create a list file as when writing sequences.

		Smallest Sum		N
Sequences producing High-scoring Segment Pairs:		High Probability Score	P (N)	
☐	sp <u>MOSB RHIME</u> MOSB PROTEIN	1561	2.3e-214	1
☐	sp <u>SPSC BACSU</u> SPORE COAT POLYSACCHARIDE BIOSYNTHESIS PR...	333	2.7e-41	2
☐	sp <u>YA66 METJA</u> HYPOTHETICAL PROTEIN MJ1066	331	5.8e-39	2
☐	sp <u>YFBE ECOLI</u> HYPOTHETICAL 42.9 KD PROTEIN IN AIS-PMRD ...	256	1.4e-35	3
☐	sp <u>RFFA ECOLI</u> LIPOPOLYSACCHARIDE BIOSYNTHESIS PROTEIN RFFA	282	1.4e-31	3
☐	sp <u>DEGT BACST</u> PLEIOTROPIC REGULATORY PROTEIN	270	1.4e-29	2
☐	sp <u>ERBS SACER</u> ERYTHROMYCIN BIOSYNTHESIS SENSORY TRANSDU...	155	7.5e-28	4
☐	sp <u>DNRJ STRPE</u> DAUNORUBICIN BIOSYNTHESIS SENSORY TRANSDU...	241	1.4e-26	2
☐	sp <u>RFBH SALTY</u> LIPOPOLYSACCHARIDE BIOSYNTHESIS PROTEIN RFBH	216	1.0e-25	3
☐	sp <u>Y09R MYCTU</u> HYPOTHETICAL 44.9 KD PROTEIN CY78.26	76	5.8e-12	5
☐	sp <u>HIS8 ZYMMO</u> HISTIDINOL-PHOSPHATE AMINOTRANSFERASE (EC...	58	0.030	2
☐	sp <u>PATA BACSU</u> PUTATIVE AMINOTRANSFERASE A (EC 2.6.1.-)	74	0.15	2
☐	sp <u>DCLY BACSU</u> LYSINE DECARBOXYLASE (EC 4.1.1.18) (LDC)	72	0.25	2
☐	sp <u>DHBE BACSU</u> 2,3-DIHYDROXYBENZOATE-AMP LIGASE (EC 6.3....	69	0.44	1
☐	sp <u>YAAO BACSU</u> HYPOTHETICAL 53.2 KD PROTEIN IN XPAC-ABRB...	55	0.96	2
☐	sp <u>NPRA BACST</u> TRANSCRIPTIONAL ACTIVATOR NPRA	63	0.98	1
☐	sp <u>GLYA SYN3</u> SERINE HYDROXYMETHYLTRANSFERASE (EC 2.1.2...	61	0.9995	1
☐	sp <u>CYSD EMENI</u> O-ACETYLHOMOSERINE (THIOL)-LYASE (EC 4.2....	48	0.9996	3
☐	sp <u>AAT BACST</u> ASPARTATE AMINOTRANSFERASE (EC 2.6.1.1) (...)	50	0.9997	3

>sp|MOSB RHIME MOSB PROTEIN
Length = 507

Score = 1561 (723.5 bits), Expect = 2.3e-214, P = 2.3e-214
Identities = 292/401 (72%), Positives = 336/401 (83%)

Query: 20 LHRWPILTGANRKYVIACISSGELSGSGLEVINEFETTVENWIGGGHVSTNSGTTALMV 79
+H WP LT Y V ACISSGELSGSGL +IN FE +E WIGGG+VVS +SGT AL V
Sbjct: 103 MHPWPALTRTAYEDVAACISSGELSGSGLGIINAFERRMEEWIGGGYVVSASSGTAALTV 162

Query: 80 ALLTLGVGPEDVVLVPSYTWASATAFAATLIGAIPRFVDIDSDTYNLSTEAVAKAITPDVK 139
AL+ LG+ P DVVL+PSYTW+ATA A LIGAIPRFVDID ++YN+S A+A AITPDVK
Sbjct: 163 ALIALGIQPGDVVLLPSYTWAAATALAPLLIGAIPRFVDIDPNSYNISPTALAAAITPDVK 222

Querv: 140 AIIVVHMHGLSCEMDEITAIAREKGVPIIEDCAOHGALYKKOYVGTLSDIGCFSMOKSK 199

```

-
Sbjct: 223 AIIIVVHMHG+SC+MDEI ARE+G+ +IEDCAQAHGALYK Q+VG LSDIGCFMSMQSKS 282
Query: 200 HFSAGDGGFLVTRNAALAQKARDICNFGFLPTPKHNYHFEEGLREGYAVFRECEQIGGMFR 259
      H SAGDGGF+VTR+ LAQK RDICNFGFLPTPK NY F+E +R+GYAVFRECEQIGGMFR
Sbjct: 283 HLSAGDGGFMVTRDPTLAQKMRDICNFGFLPTPKANYRFEDEVVRDGYAVFRECEQIGGMFR 342
Query: 260 LNPLSAALVIDQLEHLQRIEQLQEQAMKPLVAEAAQVPFLKITQPGSHRTHVWHKIRVGI 319
      L P+SAALV+ QLEHL QRI WLQ AM+PLV E+A++PF KIT+ RTHVWHKIRVGI
Sbjct: 343 LQPMASALVMHQLEHLRQRIAWLQTAMEPLVEESAKIPFFKITRSHVDRTHVWHKIRVGI 402
Query: 320 DYGAVDHYGRSMADIRQGLRNTLARSIGPSTLWTAPILPLQKVFRPYAGTIDWTKTAGHL 379
      DY AVD++GRSMA+IR+GLRN+LA GI STLWTAPILPLQ FRPYAG ++WTK+AGHL
Sbjct: 403 DYAAVDYFGRSMAEIRRLRNSLAERGISSTLWTAPILPLQTAFRPYAGVVEWTKSAGHL 462
Query: 380 AIENSFIVFDEKYPLIAQDPSTMTALVSKLRQVWSDYFASL 420
      AIENSFIVFDE YPLIAQ+P+ M LV KL+Q W ++F SL
Sbjct: 463 AIENSFIVFDENYPLIAQEPAKMAELVVKLQQAWEHFTSL 503

```

>sp|SPSC_BACSU SPORE COAT POLYSACCHARIDE BIOSYNTHESIS PROTEIN SPSC
Length = 389

Score = 333 (154.3 bits), Expect = 2.7e-41, Sum P(2) = 2.7e-41
Identities = 69/190 (36%), Positives = 104/190 (54%)

```

Query: 51 INEFETTENVWIGGGHVSTNSGTTALMVALLTLGVGPEDVVLVPSYTWSATAFAATLIG 110
      + +FE ++G H V+ NS T AL +AL G+GP D V+ T+S+TA G
Sbjct: 39 VQQFEKEFAAFVGAHVAVNSCTAALFLALKAKGIGPGDEVITSPLTFSSTANTIHTG 98
Query: 111 AIPRFVDIDSDTYNLSTEAVAKAITPDVKAIIVVHMHGLSCMEITAAREKGVPIED 170
      A P F DID +T N+ + A+TP KA++ VH G SC+MD I A+A+ G+ ++ED
Sbjct: 99 ATPVFADIDENTLNPVKLEAAVTPRTKAVVPVHFGGQSCMDAILAVAQNHGLFVLED 158
Query: 171 CAQAHGALYKKQYVGTLSDIGCFMSMQSKHFSAGDGGFLVTRNAALAQKARDICNFGFLPT 230
      A A YK++ +G++ D FS +K+ + G+GG L T + LA K R + G+
Sbjct: 159 AAHAVYTTYKQRMIGSIGDATAFSFYATKNLATGEGGMLTTDDEELADKIRVLSLHGMSK 218
Query: 231 PKHNYHFEEG 240
      N + G
Sbjct: 219 AAWNRYSSNG 228

```

Score = 56 (26.0 bits), Expect = 2.7e-41, Sum P(2) = 2.7e-41
Identities = 19/73 (26%), Positives = 28/73 (38%)

```

Query: 248 FRECEQIGGMFRLNPLSAALVIDQLEHLQRIEQLQEQAMKPLVAEAAQVPFLKITQPGSH 307
      + E E G + L AAL + QL+ LD + +E Q+P L
Sbjct: 231 YVEVESPGYKMMFDLQAALGLHQLKRLDDMQKRREEIAGRYQTAFQQIPGLITPFVHDD 290
Query: 308 RTHVWHKIRVGID 320
      H WH + +D
Sbjct: 291 GRHAWHLYVLQVD 303

```

>sp|YA66_METJA HYPOTHETICAL PROTEIN MJ1066
Length = 386

Score = 331 (153.4 bits), Expect = 5.8e-39, Sum P(2) = 5.8e-39
Identities = 73/188 (38%), Positives = 105/188 (55%)

```

Query: 53 EFETTENVWIGGGHVSTNSGTTALMVALLTLGVGPEDVVLVPSYTWSATAFAATLIGAI 112
      +FE N+IG +S NS T L ++L+ L + +D V+ YT++AT A
Sbjct: 39 KFEELFRNYIGSKFAISLNSCTAGLHLSLVALNIKDKDEVITTPYTFAATGNVIVHQRK 98

```

Score = 37 (17.1 bits), Expect = 5.8e-12, Sum P(5) = 5.8e-12
Identities = 8/21 (38%), Positives = 12/21 (57%)

Query: 159 IAREKGVPIIEDCAQAHGALY 179
+A E +PI+ D A G+ Y
Sbjct: 186 LAIEWELPIVLDSAGFGSTY 206

Score = 35 (16.2 bits), Expect = 2.3e-09, Sum P(5) = 2.3e-09
Identities = 9/34 (26%), Positives = 16/34 (47%)

Query: 222 DICNFGLPKPHNYHFEEGLREGYAVFRECEQIG 255
DIC+ + P H++ + + A +E E G
Sbjct: 378 DICSRIVSLPVHDHMAPDDVARVVAQAEVRG 411

>sp|HIS8_ZYMMO HISTIDINOL-PHOSPHATE AMINOTRANSFERASE (EC 2.6.1.9) (IMIDAZOLE
ACETOL-
Length = 369

Score = 58 (26.9 bits), Expect = 0.031, Sum P(2) = 0.030
Identities = 12/27 (44%), Positives = 16/27 (59%)

Query: 87 GPEDVVLVPSYTSATAFAATLIGAIP 113
G +D VL P Y++S AA +GA P
Sbjct: 104 GQDDEVLYPRYSFSVYPLAARRVGATP 130

Score = 56 (26.0 bits), Expect = 0.031, Sum P(2) = 0.030
Identities = 9/30 (30%), Positives = 17/30 (56%)

Query: 116 VDIDSDTYNLSTEAVAKAITPDVKAIIVVH 145
V+ D Y S +A+ KA+TP + + + +
Sbjct: 131 VEAPDDDYRCSVDALLKAVTPRTRVVFIAN 160

>sp|PATA_BACSU PUTATIVE AMINOTRANSFERASE A (EC 2.6.1.-)
Length = 392

Score = 74 (34.3 bits), Expect = 0.16, Sum P(2) = 0.15
Identities = 16/58 (27%), Positives = 28/58 (48%)

Query: 86 VGPEDVVLVPSYTSATAFAATLIGAIPRFVDIDSDTYNLSTEAVAKAITPDVKAIIV 143
+ P D V++P + L GA P VD S + L+ + A+TP+ K +++
Sbjct: 109 LSPGDEVIMPGPIYPGYEPIINLCGAKPVIVDTTSHGFKLTARLIEDALTPNTKCVVL 166

Score = 35 (16.2 bits), Expect = 0.16, Sum P(2) = 0.15
Identities = 6/13 (46%), Positives = 9/13 (69%)

Query: 139 KAIIVVHMHGLSC 151
K I+ VH + +SC
Sbjct: 250 KHILKVHQYNVSC 262

Score = 34 (15.8 bits), Expect = 0.22, Sum P(2) = 0.19
Identities = 10/26 (38%), Positives = 15/26 (57%)

Query: 164 GVPIIEDCAQAHGALYKKQYVGTLSLSD 189
GV + E+ ++ AL K + V LSD
Sbjct: 174 GVTLSSEELKSIAALLKGRNVFVLSLSD 199

>sp|DCLY_BACSU LYSINE DECARBOXYLASE (EC 4.1.1.18) (LDC)
Length = 490

Score = 72 (33.4 bits), Expect = 0.29, Sum P(2) = 0.25

>sp|-|GLYA SYN3 SERINE HYDROXYMETHYLTRANSFERASE (EC 2.1.2.1) (SERINE
METHYLASE)
Length = 427

Score = 61 (28.3 bits), Expect = 7.6, P = 1.0
Identities = 18/68 (26%), Positives = 27/68 (39%)

Query: 152 EMDEITAIAREKGVPIIEDCAQAHGALYKKQYVGTLSDIGCFSMQKSKHFSAGDGGFLVT 211
E D+ AIA E G ++ D A G + + L + K GG ++T
Sbjct: 184 EFDKFRADAEVGAAYLMADIAHAGLVASGHHPSPPLPYCDVVTTTTHKTLRGPGRGLIMT 243

Query: 212 RNAALAQK 219
N L +K
Sbjct: 244 NNEELGKK 251

>sp|-|CYSD EMENI O-ACETYLHOMOSERINE (THIOL)-LYASE (EC 4.2.99.10)
(O-ACETYLHOMOSERINE
Length = 437

Score = 48 (22.2 bits), Expect = 7.7, Sum P(3) = 1.0
Identities = 7/19 (36%), Positives = 13/19 (68%)

Query: 152 EMDEITAIAREKGVPIIED 170
+ + I +A EKG+P++ D
Sbjct: 167 DFEAIAKVAHEKGIPLVVD 185

Score = 44 (20.4 bits), Expect = 7.7, Sum P(3) = 1.0
Identities = 11/42 (26%), Positives = 21/42 (50%)

Query: 51 INEFETTVENWIGGGHVSTNSGTTALMVALLTLGVGPEDVV 92
++ FE + GG V+ +SG A +A+ L +++V
Sbjct: 66 VDVFEKRIAALGGVAAVAASSGQAAQFMAISALAHAGDNIV 107

Score = 36 (16.7 bits), Expect = 7.7, Sum P(3) = 1.0
Identities = 10/21 (47%), Positives = 12/21 (57%)

Query: 260 LNPLSAALVIDQLEHLDQRIE 280
LNP +A +I LE L R E
Sbjct: 279 LNPFAAQQILGLLETLSLRAE 299

>sp|-|AAT BACST ASPARTATE AMINOTRANSFERASE (EC 2.6.1.1) (TRANSAMINASE A)
(ASPAT)
Length = 393

Score = 50 (23.2 bits), Expect = 8.1, Sum P(3) = 1.0
Identities = 8/26 (30%), Positives = 18/26 (69%)

Query: 118 IDSDTYNLSTEAVAKAITPDVKAIIV 143
++ + + ++ E + +AITP KA+I+
Sbjct: 144 LEQNHFKITPEQLKQAITPRTKAVII 169

Score = 43 (19.9 bits), Expect = 8.1, Sum P(3) = 1.0
Identities = 8/28 (28%), Positives = 15/28 (53%)

Query: 90 DVVLVPSYTSATAFAATLIGAIPRFVD 117
D V++P+ W + L G +P +V+
Sbjct: 115 DEVIIPTPYWVSYPEQVKLAGGVPVYVE 142

Score = 34 (15.8 bits), Expect = 8.1, Sum P(3) = 1.0
Identities = 10/23 (43%), Positives = 12/23 (52%)

BLAST

http://www.angis.org.au/TempFiles/BLAST_FASTAwaite0494705752672.628784179687518428.html

Query: 263 LSAALVIDQLEHLDQRIEWLQEA 285
LS A +D LE +RI EA
Sbjct: 369 LSYATSLDALETAVERIHRFMEA 391

Appendix 3.2 - BLAST analysis of *R. leguminosarum* mosC

```

0.0040    4    0
0.0025    4    0
0.0016    4    1

```

Currently in Directory ~waite04/chike

Filter ...

Save sequence in GCG format for marked records to

~waite04/chike

Create a list file as when writing sequences.

		Smallest Sum	High Probability Score P(N)	N
Sequences producing High-scoring Segment Pairs:				
☐	sp - MOSC_RHIME	MEMBRANE PROTEIN MOSC	406 1.0e-118	3
☐	sp - Y4WD_RHISN	HYPOTHETICAL TRANSPORT PROTEIN Y4WD	196 9.3e-45	3
☐	sp - YDIN_ECOLI	HYPOTHETICAL 46.3 KD PROTEIN IN LPP-AROD ...	83 0.00071	3
☐	sp - GUDT_BACSU	PROBABLE GLUCARATE TRANSPORTER	64 0.0012	2
☐	sp - YDIM_ECOLI	HYPOTHETICAL 44.6 KD PROTEIN IN LPP-AROD ...	73 0.019	3
☐	sp - YJHB_ECOLI	HYPOTHETICAL 44.1 KD PROTEIN IN LEUX-FECE...	76 0.024	2
☐	sp - UHPT_SALTY	HEXOSE PHOSPHATE TRANSPORT PROTEIN	60 0.035	3
☐	sp - FSR_ECOLI	FOSMIDOMYCIN RESISTANCE PROTEIN	61 0.084	2
☐	sp - YEGT_ECOLI	PUTATIVE NUCLEOSIDE TRANSPORTER YEGT	52 0.094	3
☐	sp - YEGB_ECOLI	HYPOTHETICAL 50.9 KD PROTEIN IN ALKA-BAES...	62 0.21	3
☐	sp - DGOT_ECOLI	D-GALACTONATE TRANSPORTER	57 0.38	2
☐	sp - GLPT_ECOLI	GLYCEROL-3-PHOSPHATE TRANSPORTER (G-3-P T...	53 0.60	4
☐	sp - GUDT_PSEPU	PROBABLE GLUCARATE TRANSPORTER	64 0.60	2
☐	sp - YHCA_BACSU	HYPOTHETICAL 58.3 KD PROTEIN IN GLPD-CSPB...	65 0.63	3
☐	sp - YCDG_ECOLI	HYPOTHETICAL 48.1 KD PROTEIN IN WRBA-PUTA...	68 0.72	2
☐	sp - UHPT_ECOLI	HEXOSE PHOSPHATE TRANSPORT PROTEIN	54 0.74	3
☐	sp - YJQ3_YEAST	HYPOTHETICAL 61.5 KD PROTEIN IN TPK1-CCW7...	65 0.77	1
☐	sp - YRDR_BACSU	HYPOTHETICAL 35.9 KD PROTEIN IN BLTD-TRKA...	64 0.85	1
☐	sp - NU4C_MARPO	NADH-PLASTOQUINONE OXIDOREDUCTASE CHAIN 4...	50 0.93	2
☐	sp - YCXA_BACSU	HYPOTHETICAL 44.9 KD PROTEIN IN SRFA4-SFP...	46 0.94	4
☐	sp - YDJK_ECOLI	HYPOTHETICAL METABOLITE TRANSPORT PROTEIN...	63 0.94	1
☐	sp - CCMF_RHOCA	CYTOCHROME C-TYPE BIOGENESIS PROTEIN CCL1	43 0.95	4
☐	sp - YHAU_ECOLI	HYPOTHETICAL 49.0 KD PROTEIN IN RNPB-SOHA...	59 0.95	3
☐	sp - YIDY_ECOLI	HYPOTHETICAL 41.5 KD PROTEIN IN TNAB-BGLB...	57 0.95	3
☐	sp - Y17M_MYCTU	HYPOTHETICAL 43.5 KD PROTEIN CY15C10.22C	60 0.95	2
☐	sp - YOU1_CAEL	HYPOTHETICAL 59.1 KD PROTEIN ZK637.1 IN C...	63 0.98	2
☐	sp - PATC_DROME	MEMBRANE PROTEIN PATCHED	62 0.98	1
☐	sp - YIEO_ECOLI	HYPOTHETICAL 51.5 KD PROTEIN IN RBSR-RRSC...	59 0.99	3
☐	sp - Y38K_THETE	HYPOTHETICAL 38 KD PROTEIN IN 23S RNA OPERON	58 0.990	2

☐	sp VGLB_PRVIF	GLYCOPROTEIN GII PRECURSOR	65	0.992	2
☐	sp EMRB_HAEIN	MULTIDRUG RESISTANCE PROTEIN B HOMOLOG	64	0.993	2
☐	sp Y651_METJA	PUTATIVE PROTEASE MJ0651 (EC 3.4.-.-)	61	0.994	1
☐	sp YDJE_ECOLI	HYPOTHETICAL METABOLITE TRANSPORT PROTEIN...	61	0.996	1
☐	sp Y0C2_MYCTU	HYPOTHETICAL 45.7 KD PROTEIN CY19G5.02	59	0.996	2
☐	sp YQGH_BACSU	PROBABLE ABC TRANSPORTER PERMEASE PROTEIN...	48	0.996	3
☐	sp RECG_STRPN	ATP-DEPENDENT DNA HELICASE RECG (EC 3.6.1.-)	61	0.996	1
☐	sp GUDT_ECOLI	PROBABLE GLUCARATE TRANSPORTER	55	0.997	2
☐	sp WZXC_ECOLI	LIPOPOLYSACCHARIDE BIOSYNTHESIS PROTEIN WZXC	51	0.998	2
☐	sp PRCB_ARCFU	PROTEASOME BETA SUBUNIT (EC 3.4.99.46) (M...	60	0.999	1
☐	sp NANT_ECOLI	PUTATIVE SIALIC ACID TRANSPORTER	50	0.999	4
☐	sp AC22_STRCO	PROBABLE ACTINORHODIN TRANSPORTER	50	0.999	3
☐	sp YXAM_BACSU	HYPOTHETICAL 44.7 KD PROTEIN IN GNTR-HTPG...	61	0.999	2
☐	sp POLK_RHOSH	ATP-BINDING TRANSPORT PROTEIN POLK	60	0.9993	1
☐	sp Y4YK_RHISN	PROBABLE TRANSLOCATION PROTEIN Y4YK	60	0.9993	1
☐	sp YIM0_YEAST	HYPOTHETICAL 61.8 KD PROTEIN IN KGD1-SIM1...	54	0.9994	2
☐	sp NU5M_ANOQU	NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5 (E...	47	0.9997	4
☐	sp GLCP_SYNY3	GLUCOSE TRANSPORT PROTEIN	49	0.9997	2
☐	sp YJDL_ECOLI	HYPOTHETICAL 53.1 KD PROTEIN IN LYSU-CADA...	49	0.9998	3
☐	sp YBHS_ECOLI	HYPOTHETICAL 42.1 KD PROTEIN IN MOAE-RHLE...	50	0.9999	3
☐	sp PUR8_STRLP	PUROMYCIN RESISTANCE PROTEIN PUR8	54	0.9999	3
☐	sp NME4_MOUSE	GLUTAMATE [NMDA] RECEPTOR SUBUNIT EPSILON...	51	0.99992	2

>sp|MOSC_RHIME MEMBRANE PROTEIN MOSC
Length = 407

Score = 406 (191.8 bits), Expect = 1.0e-118, Sum P(3) = 1.0e-118
Identities = 85/160 (53%), Positives = 110/160 (68%)

Query: 163 HPLVQAIALLVGTLMLSWRLFPPLDLDADRSRISLELGALGNRLILLGFLSFLTMMSDG 222
H + +++ LL G L+++ LFPPL D L +G L N +L++ G LSFLTM++DG
Sbjct: 181 HYVARSVTLTGLILIVARWLFPPLDLDVRSGEHRLAIGELRNCKLLMFGILSFLTMVTDG 240

Query: 223 AIADWSTIYLEDYWPTSPQTAVLGYLAFVFMIVGRISGDRIGSVIGDRAMIAISGGLMS 282
AIA+WS +YL + Q LGY+AF + MI GRISGDR+ IG RA+IAISG L S
Sbjct: 241 AIAEWSKLYLIRVEQVTDQVGSGLGYVAFTLLMIAGRISGDRVKDAIGCRALIAISGSLAS 300

Query: 283 AGMTIALSIPFFAAALAGFSLGIGMANLVPIIFRNASEV 322
AGMT AL +P FA LAGF+LLG+GMANLVPIIF A+ +
Sbjct: 301 AGMTTALFMPSPFAGKLAGFALLGLGMANLVPIIFSEAASM 340

Score = 353 (166.8 bits), Expect = 1.0e-118, Sum P(3) = 1.0e-118
Identities = 69/138 (50%), Positives = 92/138 (66%)

Query: 13 VMVAFQFFINGVVGWATNHLHRRSFGLSEFGLSIVFLTMGAAAVLIMSLTGVLVQI 72
+ AVAFQFFING+V WAT+I H++ ++ L ++ L M A A+L MSL GY +
Sbjct: 33 ITAVAFQFFINGLVLAAWATSIPHVKNAYSFNDAELGVLLLIMAAGALLFMSLAGYFHSV 92

Query: 73 FGSRRISILSCLFFPAALCFVFPVAPNHIALLVGVGLFGAANGSMDVSMNDQGTNLGKRGS 132
FGSRR+SI S L FP+AL +F APN + L + LFGAANG+MDV MN Q L + G
Sbjct: 93 FGSRRMSIQSALLFPFSAVLIFAAPNCMTFLCSIVLFGAANGAMDVLMNHQAKALEENG 152

Query: 133 DPVISLLHGFSSIGALAG 150
+++ LHG SS G LAG
Sbjct: 153 PRIMAFLHGCSSTGILAG 170

Score = 115 (54.3 bits), Expect = 1.0e-118, Sum P(3) = 1.0e-118
Identities = 24/45 (53%), Positives = 29/45 (64%)

Query: 327 LGTAFVSVCGYSGFLVGPPIIGGAAGAVGLNKALLIVSAVGATLA 371
+G FVSVCGYSGFLVGPPIIG + +G + AVG +A
Sbjct: 347 VGLTFVSVCGYSGFLVGPPIIGASRRPLGSGELCSSSFAVGVIVA 391

Score = 34 (16.1 bits), Expect = 2.3e-07, Sum P(2) = 2.3e-07
Identities = 9/34 (26%), Positives = 18/34 (52%)

Query: 280 LMSAGMTIALSIPFFAAALAGFSLLGIGMANLVP 313
+M+AG + +S+ + + + G + I A L P
Sbjct: 74 IMAAGALLFMSLAGYFSHVFGSRRMSIQSALLFP 107

Score = 33 (15.6 bits), Expect = 5.7e-55, Sum P(3) = 5.7e-55
Identities = 10/32 (31%), Positives = 17/32 (53%)

Query: 257 GRISGDRIGSVIGDRAMIAISGGLMSAGMTIA 288
G ++G VIGD +A S L++ + +A
Sbjct: 166 GILAGIMTFGVIGDGHYVARSVTLTGILIVA 197

>sp|-|Y4WD_RHISN HYPOTHETICAL TRANSPORT PROTEIN Y4WD
Length = 377

Score = 196 (92.6 bits), Expect = 9.3e-45, Sum P(3) = 9.3e-45
Identities = 41/140 (29%), Positives = 76/140 (54%)

Query: 16 VAFQFFINGVVYGTWATNILHIRRSFGLSEFGLSIVFLTMGAAAVLIMSLTGYLQVIFGS 75
VA F +NG++ GTWA+ + + F +++ ++ L +G A++ +TG L G+
Sbjct: 11 VAAGFLLNGILLGTWASRPVAVMGHFHAKASFVLLLLLLGLGALISFPITGRLSDSLGA 70

Query: 76 RRISILSCLFFPAALCFVFPVAPNHIALLVGVGLFGAANGSMDVSMNDQGTNLGKRGSDPV 135
R++ + + F ++ + +AP L + + LFG +GSMDV++N + + K PV
Sbjct: 71 VRVARMIAIPFLVSIAALGLAPTIPLLAIALFLFGMCHGSMDVAVNSWASEVEKHMGRP 130

Query: 136 ISLLHGFSSIGALAGSLLSY 155
+S H S+GA+ G+ Y
Sbjct: 131 MSSFHAMWSVGA VLGAAGGY 150

Score = 175 (82.7 bits), Expect = 9.3e-45, Sum P(3) = 9.3e-45
Identities = 40/133 (30%), Positives = 66/133 (49%)

Query: 202 LGNGRLILLGFLSFLTMMSDGAIDWSTIYLEDYWPTSPQTAVLGYLAFVFMIVGRISG 261
L N L L+G ++ + + +G DWS +YL D T A LGY+ F+ M++ R+
Sbjct: 198 LPNSGLFLVLGLIALASGLGEGTALDWSAVYLHDVVGTESDAALGYMGFSAAMVMMRLKA 257

Query: 262 DRIGSVIGDRAMIAISGGLMSAGMTIALSIPFFAAALAGFSLLGIGMANLVPIIFRNASE 321
D + + G + ISG G+ + + +AGF L+G+G A ++P+ F A+
Sbjct: 258 DSLVTRWGSATVARISGFSAVFGILLIVLGETLPLVVAGFVLMGVGYAAVLPLAFSRAAA 317

Query: 322 VVPDGLGTAFVSV 334
+ G A SV
Sbjct: 318 DLVVPAGKAIASV 330

Score = 63 (29.8 bits), Expect = 9.3e-45, Sum P(3) = 9.3e-45
Identities = 18/53 (33%), Positives = 23/53 (43%)

Query: 319 ASEVVPDGLGTAFVSVCGYSGFLVGPPIIGGAAGAVGLNKALLIVSAVGATLA 371
A VVP G A V++ Y +GP IG A A + IV A +A
Sbjct: 317 ADLVVPAGKAIASVAIFAYGAMTLGPFAIGLLAEAAATMRLCFFIVGLFAALVA 369

Score = 39 (18.4 bits), Expect = 0.73, Sum P(2) = 0.52
Identities = 12/31 (38%), Positives = 17/31 (54%)

Query: 273 MIAISGGLMSAGMTIALSIPFFAAALAGFSL 303
MIAI + A + +A +IP A AL F +
Sbjct: 76 MIAIPFLVSIAALGLAPTIPLLAIALFLFGM 106

Score = 37 (17.5 bits), Expect = 6.6e-21, Sum P(3) = 6.6e-21
Identities = 9/20 (45%), Positives = 13/20 (65%)

Query: 194 RISLELGALGNRLILLGFL 213
R+S LGA+ R+I + FL
Sbjct: 63 RLSDSLGAVRVARMIAIPFL 82

Score = 35 (16.5 bits), Expect = 2.5, Sum P(2) = 0.92
Identities = 10/41 (24%), Positives = 20/41 (48%)

Query: 42 GLSEFGLSIVFLTMGAAVLIMSLTGVLVQIFGSRRLSILS 82
G E ++ ++ AA V++ LV +GS ++ +S
Sbjct: 233 GTEESDAALGYMGFSAAMVMRLKADSLVTRWGSATVARIS 273

Score = 33 (15.6 bits), Expect = 6.9e-19, Sum P(2) = 6.9e-19
Identities = 11/38 (28%), Positives = 18/38 (47%)

Query: 326 GLGTA FVSVCYSGFLVGPPIIGGAAGAVGLNKALLIV 363
G GTA Y +VG A G +G + A++++
Sbjct: 216 GEGTALDWSAVYLHDVVGTEESDAALGYMGFSAAMVMM 253

>sp|-|YDIN_ECOLI HYPOTHETICAL 46.3 KD PROTEIN IN LPP-AROD INTERGENIC REGION
Length = 423

Score = 83 (39.2 bits), Expect = 0.00071, Sum P(3) = 0.00071
Identities = 24/136 (17%), Positives = 53/136 (38%)

Query: 13 VMAVAFQFFINGVVGTVATNLIHRRSFGLSEFGLSIVFLTMGAAVLIMSLTGVLVQI 72
V+ + F +F++G+ T A N+ + F G++ + +G ++ + G +
Sbjct: 17 VLCIYFSYFLHGIVITLAQNMSLAKEFSTDNAGIAYLISGIGLGRILSILFFGVISDK 76

Query: 73 FGSRRISILSCLFFPAALCFVFVAPNHIALLVGVGLFGAANGSMDVSMNDQGTNLGKRGS 132
FG R + +++ + + + PN G AN ++D + S
Sbjct: 77 FGRRVILMAVIMYLLFFFGIPACPNLTAYGLAVCVGIANSALDTGGYPALMECFPKAS 136

Query: 133 DPVISLLHGFSSIGAL 148
+ L+ S G +
Sbjct: 137 GSAVILVKAMVSFGQM 152

Score = 40 (18.9 bits); Expect = 0.00071, Sum P(3) = 0.00071
Identities = 11/30 (36%), Positives = 16/30 (53%)

Query: 161 GYHPLVQAIALLVGTLMLSWRLFPHLDLDA 190
GY ++ I ++ TLML FP +DA
Sbjct: 169 GYGLIIPGILFVLITLMLLKSKFPSQLVDA 198

Score = 33 (15.6 bits), Expect = 0.00071, Sum P(3) = 0.00071
Identities = 5/21 (23%), Positives = 10/21 (47%)

Query: 236 WPTSPQTAVLGYLAFAVFMIV 256
W + + G AF+ F ++
Sbjct: 214 WLEGVSSVLFGVAAAFSTFYVI 234

>sp|-|GUDT_BACSU PROBABLE GLUCARATE TRANSPORTER
Length = 455

Score = 64 (30.2 bits), Expect = 0.0012, Sum P(2) = 0.0012

Identities = 16/50 (32%), Positives = 24/50 (48%)

Query: 37 IRRSFGLSSEFGLSIVFLTMGAAVLIMSLTGYLVQIFGSRISILSCLFF 86
++ GL + VF G A V+ G+L+ FGS+ I LS F+
Sbjct: 46 VQHDGLGLDSVAMGYVFSAFGWAYVIGQLPGGWLLDRFGSKTIIALSIFFW 95

Score = 59 (27.9 bits), Expect = 0.0012, Sum P(2) = 0.0012
Identities = 11/48 (22%), Positives = 23/48 (47%)

Query: 326 GLGTAFVSVCGYSGFLVGPPIIGGAAGAVGLNKALLIVSAVGATLAYV 373
G +AF + Y ++ PP++G + G + +++ G LA +
Sbjct: 149 GTASAFFNSAQYFAIVIFPPLMGWLTHSFGWHSVFVVMGIAGILLAVI 196

Score = 48 (22.7 bits), Expect = 0.19, Sum P(2) = 0.17
Identities = 14/53 (26%), Positives = 26/53 (49%)

Query: 247 YLAFAVFMIVGRISGDRIGSVIGDRAMIAISGGLMSAGMTIALSIPFFAAALA 299
+ AF ++G++ G + G + +IA+S S + +I FF+A A
Sbjct: 61 FSAFGWAYVIGQLPGGWLLDRFGSKTIIALSIFFWSFFTLLQGAIGFFSAGTA 113

Score = 39 (18.4 bits), Expect = 0.16, Sum P(3) = 0.15
Identities = 8/25 (32%), Positives = 12/25 (48%)

Query: 140 HGFSSIGALAGSLLSYWWIGRGYHP 164
H + +AG LL+ W+ Y P
Sbjct: 180 HSVFVVMGIAGILLAVIWLKTVYEP 204

Score = 38 (18.0 bits), Expect = 4.3, Sum P(2) = 0.99
Identities = 7/27 (25%), Positives = 14/27 (51%)

Query: 237 PTSPQTAVLGYLAFVFMIVGRISGDR 263
P +T+V ++ F +F++ DR
Sbjct: 10 PAGKKTSVRWFIVFMLFLVTSINYADR 36

Score = 36 (17.0 bits), Expect = 0.16, Sum P(3) = 0.15
Identities = 6/32 (18%), Positives = 17/32 (53%)

Query: 312 VPIIFRNASEVVPDGLGTAFVSVCGYSGFLVG 343
V ++ ++ G + ++CG++G ++G
Sbjct: 275 VYLVQARGMSILEAGFVASLPALCGFAGGVLG 306

Score = 36 (17.0 bits), Expect = 0.16, Sum P(3) = 0.15
Identities = 16/79 (20%), Positives = 24/79 (30%)

Query: 293 FFAAALAGFSLLGIGMANLVPIIFRNASEVVPDGLGTAFVSVCGYSGFLVGPPIIGGAAG 352
+ + + G G L + + S GL + G + P IIG
Sbjct: 348 WLVVVIMSLAFFGKGFGALGWAVVSDTSPKECAGLSGGLFNTFGNIASITTPPIIIGYIVN 407

Query: 353 AVGLNKALLIVSAVGATLA 371
A G L+ A A
Sbjct: 408 ATGSFNGALVFVGANAIAA 426

Score = 34 (16.1 bits), Expect = 0.29, Sum P(3) = 0.25
Identities = 7/17 (41%), Positives = 10/17 (58%)

Query: 338 SGFLVGPPIIGGAAGAV 354
+GF+ P + G AG V
Sbjct: 288 AGFVASLPALCGFAGGV 304

>sp|-|YDIM_ECOLI HYPOTHETICAL 44.6 KD PROTEIN IN LPP-AROD INTERGENIC REGION
Length = 404

Score = 73 (34.5 bits), Expect = 0.019, Sum P(3) = 0.019
Identities = 23/104 (22%), Positives = 44/104 (42%)

>sp|-|UHPT_SALTY HEXOSE PHOSPHATE TRANSPORT PROTEIN
Length = 463

Score = 60 (28.3 bits), Expect = 0.036, Sum P(3) = 0.035
Identities = 17/58 (29%), Positives = 26/58 (44%)

Query: 202 LGNGRLILLGFLSFLTMMSDGAIADWSTIYLEDYWPTSPQTAVLGYLAFVFMIVGRI 259
L N + LL F + + I WST+Y S + A+ G+ F V +VG +
Sbjct: 256 LKNKVIWLLCFSNIFLYVVRIGIDQWSTVYAFQELKLSKEVAIQGFTLFEVGALVGTL 313

Score = 47 (22.2 bits), Expect = 0.036, Sum P(3) = 0.035
Identities = 11/30 (36%), Positives = 16/30 (53%)

Query: 346 IIGGAAGAVGLNKALLIVSAVGATLAYVIG 375
+IG AA KA+ + + T AY+IG
Sbjct: 370 LIGVAAVGFVPKKAIADGKGTFAYLIG 399

Score = 37 (17.5 bits), Expect = 0.036, Sum P(3) = 0.035
Identities = 7/13 (53%), Positives = 11/13 (84%)

Query: 93 VFVAPNHIALLVG 105
+F+ P+ IAL+VG
Sbjct: 194 MFIFPSIIALIVG 206

Score = 36 (17.0 bits), Expect = 0.049, Sum P(3) = 0.048
Identities = 10/40 (25%), Positives = 19/40 (47%)

Query: 141 GFSSIGALAGSLLSYWWIGRGYHPLVQAIALLVGTMLMLSW 180
GFS + +L+SY+ G+ + + +L ML +
Sbjct: 73 GFSITYGVGKTLVSYYADGKNTKQFLPFMLILSAICMLGF 112

Score = 34 (16.1 bits), Expect = 0.090, Sum P(3) = 0.086
Identities = 11/33 (33%), Positives = 14/33 (42%)

Query: 100 IALLVGVGLFGAANGSMDVSMNDQGTNLGKRGS 132
IA G F + GS S + T KRG+
Sbjct: 127 IAFYALSGFFQSTGGSCSYSTITKWTPRKRGT 159

>sp|-|FSR_ECOLI FOSMIDOMYCIN RESISTANCE PROTEIN
Length = 406

Score = 61 (28.8 bits), Expect = 0.088, Sum P(2) = 0.084
Identities = 19/66 (28%), Positives = 29/66 (43%)

Query: 209 LLGFLSFLTMMSDGAIADWSTIYLEDYWPTSPQTAVLGYLAFVFMIVGRISGDRIGSVI 268
+L- L F +I+ + T YL + S Q A L AF + G + G +G I
Sbjct: 231 ILLILIFSXYFYMASISSYYTFYLMQKFLSIQNAQLHLFAFLFAVAAGTVIGGPVGDKI 290

Query: 269 GDRAM I 274
G + +I
Sbjct: 291 GRKYVI 296

Score = 48 (22.7 bits), Expect = 0.088, Sum P(2) = 0.084
Identities = 17/99 (17%), Positives = 40/99 (40%)

Query: 15 AVAFQFFINGVVYGTWATNILHIRRSFGLSEFGLSIVFLTMGAAVLIMSLTGVLVQIFG 74
A++ +N ++ ++ F L+ + ++ LT A+ L+ + GY +
Sbjct: 29 AISLSHLLNDMIQSLILAIYPLLQSEFSLTFMQIGMITLTFQLASSLLQPVVGYWTDKYP 88

Query: 75 SRRISILSCLFFPAALCFVVFAPNHIALLVGVGLFGAAN 113
+ F + L + +A + A+L+ L G +
Sbjct: 89 MPWSLPIGMCFTLSGLVLLALAGSFGAVLLAAALVGTGS 127

Score = 33 (15.6 bits), Expect = 2.7, Sum P(4) = 0.94
Identities = 6/19 (31%), Positives = 10/19 (52%)

Query: 235 YWPTSPQTAVLGYLAFV 253
+W Q+ ++ L F VF
Sbjct: 208 FWNVMKQSRILIHILYFGVF 226

>sp|-|YDJK_ECOLI HYPOTHETICAL METABOLITE TRANSPORT PROTEIN IN ANSA-GAPA
INTERGENIC
Length = 459

Score = 63 (29.8 bits), Expect = 2.8, P = 0.94
Identities = 17/71 (23%), Positives = 31/71 (43%)

Query: 254 MIVGRISGDRIGSVIGDRAMIAISGGLMSAGMTIALSIPFFAAALAGFSLGIGMANLVP 313
M +G + G IG G R + + A M + P +A ++G+G+ L+
Sbjct: 71 MFLGALVGGIIGDKTGRRNAFILYEAIHIASMVVGAFSPNMDFLIACRFVMGVGLGALLV 130

Query: 314 IIFRNASEVVP 324
+F +E +P
Sbjct: 131 TLFAGFTEYMP 141

>sp|-|CCMF_RHOCA CYTOCHROME C-TYPE BIOGENESIS PROTEIN CCL1
Length = 653

Score = 43 (20.3 bits), Expect = 2.9, Sum P(4) = 0.95
Identities = 15/63 (23%), Positives = 26/63 (41%)

Query: 27 YGTWATNILHIRSFGLSEFGLSIVFLTMGAAVLMISLTGYLVQIFGSRIRISILSCLFF 86
+G A + +F S+F L +V+ ++ +TG GS + +L F
Sbjct: 47 FGLIAIAFAALTYAFVTSDFSLKLVYENSHTDKPMLYKVTGVWGNHEGSMLLWVLILAMF 106

Query: 87 PAA 89
AA
Sbjct: 107 GAA 109

Score = 42 (19.8 bits), Expect = 2.9, Sum P(4) = 0.95
Identities = 8/23 (34%), Positives = 13/23 (56%)

Query: 158 IGRGYHPLVQAIALLVGTLMLSW 180
+G+ PL+ A+ L + L L W
Sbjct: 418 LGKLVRLPALVLTAVLALVW 440

Score = 40 (18.9 bits), Expect = 2.9, Sum P(4) = 0.95
Identities = 11/21 (52%), Positives = 12/21 (57%)

Query: 194 RISLELGALGNRLILLGFLS 214
R L LGA G G IL G L+
Sbjct: 446 RPMLALGAAGLGAWILFGALA 466

Score = 36 (17.0 bits), Expect = 2.9, Sum P(4) = 0.95
Identities = 8/20 (40%), Positives = 12/20 (60%)

Query: 142 FSSIGALAGSLLSYWWIGRG 161
F +IG GS +Y+ +G G
Sbjct: 217 FLTIGIALGSWWAYYELGWG 236

>sp|-|YHAU_ECOLI HYPOTHETICAL 49.0 KD PROTEIN IN RNPB-SOHA INTERGENIC REGION
Length = 444

Appendix 4 - Tables on rhizopine effect on nodulation

Table 7.2 Nodules formed on lucerne roots inoculated with L5-30 and treated with or without 3-O-MSI

Days	Six replicates of nodules from roots treated with L5-30/3-O-MSI						Stdev	Sterr
	1	2	3	4	5	6		
0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0
6	2	0	0	2	0	0	1.032796	0.421637
8	2	0	2	2	0	1	0.983192	0.401386
10	3	0	3	0	1	2	1.378405	0.562731
12	4	0	3	0	1	3	1.722401	0.703167
14	4	1	5	1	4	4	1.722401	0.703167
16	4	1	6	5	5	5	1.75119	0.71492
18	7	3	7	10	14	8	3.656045	1.492574
20	7	3	8	12	15	8	4.167333	1.701307
22	11	4	9	14	17	11	4.427189	1.807392
24	12	4	9	15	17	12	4.593474	1.875278
26	12	7	9	15	17	12	3.687818	1.505545
28	13	12	9	15	17	16	2.94392	1.20185
30	13	12	9	15	17	16	2.94392	1.20185
32	13	13	9	15	18	17	3.250641	1.327069
34	13	13	11	16	19	17	2.994439	1.222475
36	15	15	11	17	19	18	2.857738	1.166667
38	15	15	11	17	19	18	2.857738	1.166667

Days	Six replicates of nodules from roots inoculated only with L5-30						Stdev	Sterr
	1	2	3	4	5	6		
0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0
6	0	1	0	0	0	0	0.408248	0.166667
8	0	2	0	1	0	5	1.966384	0.802773
10	2	2	3	1	1	6	1.870829	0.763763
12	4	2	3	1	1	7	2.280351	0.930949
14	5	2	4	8	2	8	2.71416	1.108051
16	5	2	6	8	3	9	2.738613	1.118034
18	5	4	6	9	5	10	2.428992	0.991632
20	5	4	8	10	5	11	2.926887	1.194897
22	6	8	11	12	5	11	2.926887	1.194897
24	6	8	11	13	6	11	2.926887	1.194897
26	6	8	11	13	6	12	3.076795	1.256096
28	6	9	14	14	6	13	3.829708	1.563472
30	6	9	17	16	7	13	4.676181	1.909043
32	6	10	20	18	8	15	5.671567	2.315407
34	6	10	24	19	10	15	6.663332	2.720294
36	6	10	25	20	11	15	7.007139	2.860653
38	6	10	25	20	11	15	7.007139	2.860653

Table 7.3 Nodules formed on lucerne roots inoculated with Rm1021 treated with or without 3-O-MSI

Days	Six replicates of nodules from Rm1021/3-O-MSI treated roots						Stdev	Sterr
	1	2	3	4	5	6		
0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0
8	3	0	0	1	0	3	1.47196	0.600925
10	3	0	0	1	0	3	1.47196	0.600925
12	3	1	1	0	2	3	1.21106	0.494413
14	6	1	4	1	2	4	2	0.816497
16	7	2	5	3	2	4	1.94079	0.792324
18	9	2	6	6	2	5	2.683282	1.095445
20	10	2	6	8	2	5	3.209361	1.310216
22	11	3	6	9	4	5	3.076795	1.256096
24	14	4	7	11	5	6	3.868678	1.579381
26	14	5	7	11	9	6	3.386247	1.382429
28	15	5	7	12	9	7	3.710346	1.514742
30	18	5	8	15	9	9	4.844241	1.977653
32	20	6	8	15	9	9	5.269409	2.151227
34	20	8	10	15	11	9	4.535049	1.851426
36	20	9	11	15	11	9	4.27785	1.746425
38	20	11	11	15	11	9	4.020779	1.641476

Days	Six replicates of nodules formed on Rm1021 inoculated roots						STDEV	STERR
	1	2	3	4	5	6		
0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0
6	2	2	3	3	0	2	1.095445	0.447214
8	4	2	4	3	3	2	0.894427	0.365148
10	2	3	3	4	2	4	0.894427	0.365148
12	2	3	3	4	2	5	1.169045	0.477261
14	2	4	3	4	3	6	1.36626	0.557773
16	3	4	3	4	4	6	1.095445	0.447214
18	4	3	3	4	5	6	1.169045	0.477261
20	8	3	3	4	5	8	2.316607	0.945751
22	10	4	4	5	7	10	2.804758	1.145038
24	11	5	4	6	8	10	2.804758	1.145038
26	11	5	4	6	8	11	3.016621	1.23153
28	16	5	5	6	8	11	4.32435	1.765408
30	16	6	7	6	10	11	3.88158	1.584649
32	18	8	9	7	8	12	4.131182	1.686548
34	19	10	9	9	12	12	3.763863	1.536591
36	19	11	9	10	12	12	3.544949	1.44722
38	19	11	9	10	12	12	3.544949	1.44722

Table 7.4 Nodules formed on lucerne roots inoculated with Moc- and treated with or without 3-O-MSI

Time	Five replicates of nodules from Moc-/3-O-MSI					STDEV	Sterr
	1	2	3	4	5		
0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0
12	0	0	0	1	0	0.447214	0.2
14	1	0	1	1	0	0.547723	0.244949
16	5	0	1	3	2	1.923538	0.860233
18	7	3	2	7	4	2.302173	1.029563
20	8	4	3	7	7	2.167948	0.969536
22	10	6	4	8	9	2.408319	1.077033
24	15	6	7	8	9	3.535534	1.581139
26	15	6	8	8	9	3.420526	1.529706
28	21	7	8	8	9	5.85662	2.61916
30	21	7	10	8	9	5.700877	2.54951
32	26	9	12	9	9	7.382412	3.301515
34	28	9	14	9	9	8.228001	3.679674
36	30	10	14	9	9	8.961027	4.007493
38	31	10	14	9	10	9.257429	4.140048

Days	Five replicates of nodules from Moc- treatment					STDEV	Sterr
	1	2	3	4	5		
0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0
6	1	3	3	0	0	1.516575	0.678233
8	1	3	3	0	0	1.516575	0.678233
10	1	3	3	0	0	1.516575	0.678233
12	1	3	3	0	0	1.516575	0.678233
14	3	3	3	0	1	1.414214	0.632456
16	4	4	3	0	1	1.81659	0.812404
18	6	4	3	0	2	2.236068	1
20	9	5	3	0	4	3.271085	1.462874
22	12	5	3	3	6	3.701351	1.655295
24	14	7	3	8	8	3.937004	1.760682
26	14	7	4	8	8	3.63318	1.624808
28	15	7	4	8	10	4.086563	1.827567
30	17	7	4	8	10	4.868265	2.177154
32	17	7	4	9	10	4.827007	2.158703
34	17	8	4	10	12	4.816638	2.154066
36	17	8	4	10	14	5.07937	2.271563
38	19	10	4	10	17	6.041523	2.701851

Appendix 5 Conference presentations and poster*

- 1.***Owuama, C.I.**, Bahar, M., Grzinski, W. & Murphy, P.J. (1998) Genes involved in the biosynthesis of the nodulation competition enhancing factor, rhizopine are not similar in two species of rhizobia. Australian Society for Biochemistry and Molecular Biology, Adelaide, Australia, September 1998.
2. **Owuama, C.I.**, Bahar, M., Grzinski, W. & Murphy, P.J. (1999). Genetic variation in rhizopine genes from *Sinorhizobium meliloti* and *Rhizobium leguminosarum*. International Union of Microbiological Societies. IXth International Congress of Bacteriology and Applied Microbiology, Sydney, Australia, August, 1999.
3. Murphy, P.J., Heinrich, K., **Owuama, C.**, Bahar, and M., de Majnik, J. (1999). Rhizobial rhizopines in nodulation competition and infection. Proceedings of the 43rd Annual conference of the Australian Society for Biochemistry and Molecular Biology, Gold Coast, Australia, 1999.
4. Murphy, P.J., Heinrich, K., **Owuama, C.**, Bahar, M., and de Majnik, J. (1999). Competition and infection studies in *Rhizobium*. The twelfth Australian nitrogen fixation conference. Wagga Wagga, Australia, October, 1999.
5. **Owuama, C.**, Bahar, M., Grzinski, W., and Murphy, P.J. (1999). Variation in rhizopine biosynthetic genes in rhizobia. The twelfth Australian nitrogen fixation conference. Wagga Wagga, October, 1999.