myo-Inositol metabolism

in *Rhizobium leguminosarum*

JUDITH FRY PH.D. THESIS 2000
myo-Inositol is a sugar/polyol that is present in soil (Sulochana, 1962, Wood and Stanway, 2000), plant tissues and nodules (Skøt and Egsgaard, 1984, Streeter, 1987, Streeter and Salminen, 1985). *Rhizobium leguminosarum* biovar *viciae* has an inducible pathway for myo-inositol utilisation as the sole carbon source but bacteroids do not utilise myo-inositol (Poole *et al*., 1994). In this study, the pathway of myo-inositol catabolism was further investigated.

Three discrete loci involved in myo-inositol utilisation in *R. leguminosarum* bv. *viciae* were mutated, cloned and characterised. Genes with homology to *iolA* and *iolD*, which encode proteins involved in the final steps of myo-inositol degradation in *Bacillus subtilis* (Yoshida *et al*., 1997) and a putative transport system for myo-inositol were identified. The catabolic mutants RU360 (*iolD*) and RU361 (*iolA*) were not able to induce the first two enzymes in the proposed catabolic pathway for myo-inositol, which suggests end-product induction of the pathway. The pathway is subject to catabolite repression by a range of carbon compounds.

The mutants nodulated peas (*Pisum sativum*) and vetch (*Vicia sativa*) at the same rate as the wild type. Acetylene reduction rates and overall plant dry weights were also the same. When the wild type, 3841, and the catabolic mutants, RU360 and RU361, were co-inoculated onto plants, nodules were predominantly infected with the wild type. However, a transport mutant (RU307) was not at a competitive disadvantage. Therefore, the ability to
utilise myo-inositol is essential for competitive ability of the wild type, but the advantage of the wild type is unlikely to be due to myo-inositol being used purely as a carbon source. Instead, myo-inositol might be involved in signalling between host plants and rhizobia or myo-inositol catabolism might be important in the early steps of the developing infection process.
ACKNOWLEDGEMENTS

I would like to thank my supervisors Dr Philip Poole and Dr Martin Wood for their guidance and assistance throughout my PhD.

I am indebted to Mrs Mary Leonard and Mrs Kim Carter for their invaluable support, both in and out of the laboratory. I should also like to thank Dr David Allaway and the people of Lab 160 for their help. Thank you also to the people of the Central Services Section and the electrical workshop in the School of Animal and Microbial Sciences for technical assistance.

This work would not have been possible without the University of Reading studentship awarded to me, for which I am very grateful.

Finally, a big thank you to Mazda, my parents and friends for their unwavering support and encouragement.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-O-MSI</td>
<td>L-3-O-methyl-scyllo-inosamine</td>
</tr>
<tr>
<td>ADP</td>
<td>adenosine diphosphate</td>
</tr>
<tr>
<td>AMA/AMS</td>
<td>acid minimal agar/acid minimal salts</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>bp</td>
<td>base pairs</td>
</tr>
<tr>
<td>cfu</td>
<td>colony forming units</td>
</tr>
<tr>
<td>CoA</td>
<td>coenzyme A</td>
</tr>
<tr>
<td>CO₂</td>
<td>carbon dioxide</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescent protein</td>
</tr>
<tr>
<td>HPLC</td>
<td>high performance liquid chromatography</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase pairs</td>
</tr>
<tr>
<td>LCO</td>
<td>lipo-chito-oligosaccharide</td>
</tr>
<tr>
<td>NAD</td>
<td>nicotinamide adenine dinucleotide</td>
</tr>
<tr>
<td>N/T</td>
<td>not tested</td>
</tr>
<tr>
<td>ORF</td>
<td>open reading frame</td>
</tr>
<tr>
<td>PBM</td>
<td>peribacteroid membrane</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PHB</td>
<td>polyhydroxybutyrate</td>
</tr>
<tr>
<td>SD</td>
<td>Shine-Dalgarno</td>
</tr>
<tr>
<td>SI</td>
<td>scyllo-inosamine</td>
</tr>
<tr>
<td>TY</td>
<td>tryptone yeast medium</td>
</tr>
</tbody>
</table>
Contents

CHAPTER 1 - INTRODUCTION.......................................................................................... 1

1.1 Introduction........................................................................................................... 2

1.2 Rhizobium and Symbiosis.................................................................................. 2
  1.2.1 Nodulation .......................................................................................................4
  1.2.2 Nitrogen Fixation by Bacteroids .................................................................... 9
  1.2.3 Carbon Metabolism by the Bacteroid ...........................................................11

1.2 Soil Carbon......................................................................................................... 14

1.3 myo-Inositol...................................................................................................... 15
  1.3.1 Synthesis of myo-Inositol .............................................................................16
  1.3.2 myo-Inositol Utilisation by Bacteria ..............................................................17
    1.3.2.1 myo-Inositol Uptake by Bacteria............................................................21
  1.3.3 myo-Inositol in the Environment ...................................................................22
    1.3.3.1 Plant myo-Inositol..............................................................................22
    1.3.3.2 Soil myo-Inositol..............................................................................25
  1.3.4 myo-Inositol Derivatives ..............................................................................27
    1.3.4.1 Auxins....................................................................................................28
    1.3.4.2 Methyl Esters of myo-Inositol ...............................................................28
    1.3.4.3 Di-my-o-inositol 1,1'-phosphate............................................................30
    1.3.4.4 Phosphatidylinositol ...........................................................................30
    1.3.4.5 myo-Inositol Hexaphosphate .................................................................31
    1.3.4.5.1 Phytase ............................................................................................33
  1.3.6 Rhizopines ....................................................................................................34
1.3.6.1 Rhizopine Synthetic and Catabolic Genes of S. meliloti .........................36
1.3.6.2 Agrobacterium Opines .................................................................38

1.4 Rhizosphere Growth and Competition for Nodulation .................. 40
  1.4.1 Competition for Rhizosphere Colonisation ..................................40
  1.4.2 Competition for Nodulation .........................................................42
    1.4.2.1 Utilisation of Secondary Metabolites ........................................45
  1.4.3 Improving Nodulation Competitiveness .....................................48

1.5 Research Objectives ........................................................................50

CHAPTER 2 – MATERIALS AND METHODS ...........................................52

2.1 Bacterial Strains and Plasmids .........................................................53

2.2 Culture Conditions ...........................................................................56
  2.2.1 Calculation of Mean Generation Time ..........................................56

2.3 DNA and Genetic Manipulations ....................................................57
  2.3.1 PCR Amplification .........................................................................58
  2.3.2 DNA Sequencing ...........................................................................59

2.4 Analysis of Sequence Data .............................................................61

2.5 Statistical Analysis of Data .............................................................61

2.6 Transport Assay ...............................................................................62

2.7 Enzyme Assays .................................................................................62
  2.7.1 myo-Inositol Enzyme Assay ..............................................................62
  2.7.2 β-Galactosidase Assay .................................................................63
  2.7.3 Alkaline Phosphatase Assays ..........................................................64
CHAPTER 3 - MYO-INOSITOL CATABOLIC MUTANTS OF RHIZOBIUM LEGUMINOSARUM BIOVAR VICIAE .............................................................. 70

3.1 Introduction ............................................................................................ 71

3.2.1 Growth Characteristics of the myo-Inositol Mutants ......................... 72
3.2.2 Cosmids Carrying myo-Inositol Catabolic Genes ................................. 76
3.2.3 Southern Hybridisation ....................................................................... 79
3.2.4 Transduction of RU307 ...................................................................... 85
3.2.5 Sequencing of myo-Inositol Catabolic Genes ..................................... 86
   3.2.5.1 RU360 ......................................................................................... 86
   3.2.5.1.1 Complementation of the iol region of RU360 ......................... 102
   3.2.5.2 RU361 ....................................................................................... 104
   3.2.3.3 RU307 ....................................................................................... 119
3.2.5 Uptake of Glucose and myo-Inositol ................................................. 125

3.3 Discussion.............................................................................................. 133
CHAPTER 4 - NODULATION AND RHIZOSPHERE GROWTH

4.1 Introduction

4.2 Results

4.2.1 Rate of Nodulation

4.2.2 Acetylene Reduction Assay on Pea Nodules

4.2.3 Plant Dry Weights for Pea

4.2.4 Nodule Number and Mass for Pea

4.2.5 Competition for Nodulation

4.2.5.1 Competitive Ability of the Complemented Mutants

4.2.6 Measurement of Nodule Co-Residence

4.2.7 Average Number of Nodules per Vetch Plant

4.2.9 Rhizosphere Growth Assay

4.2.10 Analysis of a nodC-phoA Fusion

4.3 Discussion

CHAPTER 5 - REGULATION OF MYO-INOSITOL CATABOLISM
5.2.2 Sequence Analysis of the myo-Inositol Inducible Promoter Error! Bookmark not defined.

5.2.3 Induction of the myo-Inositol Inducible Promoter Error! Bookmark not defined.

5.2.4 Catabolite Repression Error! Bookmark not defined.

5.2.5 Relationship to myo-Inositol Mutants Error! Bookmark not defined.

5.2.6 Expression in the Rhizosphere and Nodules Error! Bookmark not defined.

5.2.7 Stability of pRU601 in the Rhizosphere Error! Bookmark not defined.

5.2.8 Construction of a Specific myo-Inositol Inducible GFP-UV Promoter Probe Error! Bookmark not defined.

5.2.9 RU360 β-Galactosidase Assay Error! Bookmark not defined.

5.2.10 myo-Inositol Catabolic Enzyme Assays Error! Bookmark not defined.

5.3 Discussion Error! Bookmark not defined.

CHAPTER 6 - FINAL DISCUSSION Error! Bookmark not defined.

6.1 Conclusions Error! Bookmark not defined.

6.2 Future Work Error! Bookmark not defined.

BIBLIOGRAPHY Error! Bookmark not defined.

APPENDIX Error! Bookmark not defined.
Chapter 1 - Introduction
1.1 Introduction

There has been extensive study of the *Rhizobium*-legume symbiosis and many of the rhizobial genes required for nodulation and nitrogen fixation have been elucidated. However, the circumstances enabling growth and survival of free-living *Rhizobium* in soil and the rhizosphere are still something of an enigma. Identifying molecules that have effects on bacterial growth in the rhizosphere and elucidating the genes that are involved in responding to these factors is essential to understanding how communities of microorganisms develop in the rhizosphere. This could also lead to a greater understanding of the mechanisms of the *Rhizobium*-plant symbiosis and the relationship between the two distinct growth states. Ultimately, this may lead to the construction of strains that increase yields of agricultural crops.

1.2 *Rhizobium* and Symbiosis

*Rhizobia* are Gram negative, motile, aerobic bacteria that are capable of fixing atmospheric nitrogen. They are ubiquitous in the soil where they live saprotrophically, but they can also live in association with leguminous plants (family *Leguminosae*) (Jordan, 1984) and some members of the genus *Parasponia* (family *Ulmaceae*) (Trinick, 1988). Several of the specific rhizobial-plant symbioses are listed in Table 1.1.
Table 1.1 Some *Rhizobium* Species and Their Plant Hosts

<table>
<thead>
<tr>
<th><em>Rhizobium</em> species</th>
<th>plant host</th>
<th>species name</th>
<th>common name</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Bradyrhizobium japonicum</em></td>
<td><em>Glycine max</em></td>
<td><em>Phaseolus vulgaris</em></td>
<td>soybean</td>
</tr>
<tr>
<td><em>R. leguminosarum</em> bv. <em>phaseoli</em></td>
<td><em>Phaseolus vulgaris</em></td>
<td><em>Phaseolus vulgaris</em></td>
<td>common bean</td>
</tr>
<tr>
<td><em>R. etli</em></td>
<td><em>Phaseolus vulgaris</em></td>
<td><em>Phaseolus vulgaris</em></td>
<td>common bean</td>
</tr>
<tr>
<td><em>R. leguminosarum</em> bv. <em>trifolii</em></td>
<td><em>Trifolium species</em></td>
<td><em>Pisum sativum</em></td>
<td>pea</td>
</tr>
<tr>
<td><em>R. leguminosarum</em> bv. <em>viciae</em></td>
<td><em>Vicia sativa</em></td>
<td><em>Vicia sativa</em></td>
<td>vetch</td>
</tr>
<tr>
<td><em>Mesorhizobium loti</em></td>
<td><em>Lotus species</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Sinorhizobium mellilotil</em></td>
<td><em>Medicago sativa</em></td>
<td><em>Medicago sativa</em></td>
<td>alfalfa</td>
</tr>
<tr>
<td><em>Rhizobium NGR234</em></td>
<td>most legume genera</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>and <em>Parasponia</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Biological processes account for approximately 60% of the biosphere’s fixed nitrogen. It is difficult to quantify the amount of nitrogen fixed globally by rhizobia, but legume symbioses are estimated to contribute at least 70 million tonnes of fixed nitrogen *per annum* (Brockwell *et al.*, 1995). This represents approximately 50% of the total nitrogen that enters terrestrial ecosystems from the process of biological nitrogen fixation (Tate, 1995). The association is of major agronomic importance as many food crops can be grown without requiring nitrogenous fertiliser to give economically viable yields (reviewed in Zahran, 1999).
1.2.1 Nodulation

As stated in Section 1.1, nodulation and nitrogen fixation have been extensively studied. These topics are not the subject of this thesis, so instead of an exhaustive review, an overview follows (detailed reviews include Poole and Allaway, 2000, van Rhijn and Vanderleyden, 1995, Rosendahl et al., 1991, Schultze and Kondorosi, 1998, Whitehead and Day, 1997).

*Rhizobium*-legume interactions begin with the exchange of signals between the plant and microbe. Plants release signalling molecules such as flavonoids and chalcone into the rhizosphere. There are several classes of flavonoids, including flavones, flavonols, flavanones and isoflavones, which are produced from the phenylpropanoid pathway of plants (van der Meer et al., 1993). The compounds secreted differ according to the plant species. Some are even released by non-legumes. Many flavonoids are constantly released by plants, but when inoculated with a rhizobial symbiont, the composition and quantities often change (Dakora et al., 1993, Lawson et al., 1996, Recourt et al., 1991, Schmidt et al., 1994, Schmidt et al., 1991).

Flavonoids act as chemoattractants (Aguilar et al., 1988, Caetano-Anollés et al., 1992, Kape et al., 1991) and in conjunction with the transcriptional activator protein NodD, they induce expression of nodulation (nod) genes in *Rhizobium* (Schlaman et al., 1992). The effects of flavonoids are both species and strain specific. For example, the flavone apigenin is able to induce the NodD proteins of many species, whereas the isoflavone daidzein is only a

Transcriptional activation of *nod* genes occurs when NodD binds to *nod* boxes. These are conserved DNA sequences upstream of the *nod* operon (Rostas *et al*., 1986). The NodD proteins belong to the LysR family of transcriptional activators, which means that they share common features such as a characteristic helix-turn-helix motif (Goethals *et al*., 1990, Henikoff *et al*., 1988). They are highly conserved between rhizobial species, although the numbers of copies of the *nodD* gene varies between species. NodD is one of the factors responsible for host specificity, as mutations in *nodD* cannot usually be complemented by a *nodD* gene from another species (Horvath *et al*., 1987, Spaink *et al*., 1987). The NodD proteins of narrow host range rhizobia such as *S. meliloti*, *R. leguminosarum* and *R. trifolii* only respond to a few flavonoids, whereas that of the broad host range *Rhizobium* NGR234 responds to most flavonoids (Györgypal *et al*., 1991).

The products of *nod* genes, Nod-factors, are lipo-chito-oligosaccharides (LCOs), which stimulate root hair curling and deformation, through which infection occurs in most legumes (Bhuvaneswari and Solheim, 1985, van Brussel *et al*., 1986, Yao and Vincent, 1969). In addition to nodulation genes, there is another class of rhizobial genes involved in infection and nodulation. These encode compounds such as exopolysaccharides (EPS), which are
necessary for proper formation of infection threads and root hair curling on plants that form indeterminate nodules, such as vetch (V. sativa) (Leigh et al., 1985, Rolfe et al., 1996, van Workum et al., 1998). It is thought that EPS either inhibit plant defence mechanisms that prevent damage of plant root hair cells or that they accelerate plant hair curling and infection processes so that root penetration by rhizobia precedes the plant defence response (van Workum et al., 1998).

After root hair deformation has occurred, plant root hair walls are degraded by hydrolysis, rhizobia enter the roots (Newcombe et al., 1979, Turgeon and Bauer, 1982) and an infection thread is formed. The thread is comprised of plant cell wall material, which surrounds the rhizobia within. (Callaham and Torrey, 1981, Turgeon and Bauer, 1985). Plant cortical cells divide to form nodule primordia, towards which the infection threads grow (Libbenga et al., 1973, Newcombe, 1981, Vasse and Truchet, 1984, Wood and Newcombe, 1989). Specialised structures called nodules are formed, whereupon rhizobia are released from the infection threads and differentiate into bacteroids. The peribacteroid membrane (PBM), which is derived from plant material, encloses the bacteroids (Brewin et al., 1985, Mellor and Werner, 1987 Robertson et al., 1978, Verma et al., 1978). The PBM is thought to have a vital role in regulating nutrient flow between the plant and the bacteroids and may also help maintain low oxygen conditions (reviewed in Rosendahl et al., 1991, Whitehead and Day, 1997).
Depending on the plant species, there are two major types of nodule development. Temperate legumes such as alfalfa, pea and vetch form indeterminate nodules. These nodules maintain an apical meristem and are elongated in shape. Most tropical legumes such as soybean form determinate nodules, which have a transient meristem and are round in shape (Newcombe, 1986, Newcombe, 1981, Newcombe et al., 1979, Turgeon and Bauer, 1985).

The nod and nitrogen fixation (nif, fix) genes are mostly clustered on symbiotic (Sym) plasmids in Rhizobium and Sinorhizobium (Martinez et al., 1990) but are chromosomal in Azorhizobium (Goethals et al., 1989), Bradyrhizobium, and M. loti (Chua et al., 1985, Pankhurst et al., 1983, Sullivan et al., 1995). The nodulation of legumes is species and strain specific according to the Nod-factors produced. The nodABC genes are found in all rhizobia and are known as the common nod genes as they are interchangeable between species (Kondorosi et al., 1984). The three proteins NodA, NodB and NodC are responsible for synthesis of the LCO backbone of Nod factors. NodC protein functions as a N-acetylglucosaminytransferase, causing polymerisation of N-acetylglucosamine units to form the backbone (Geremia et al., 1994, Spaink et al., 1994). NodB is a chitoooligosaccharide deacetylase that specifically removes the acetyl group at the non-reducing end of N-acetylglucosamine (John et al., 1993) and NodA transfers the acyl chain to the non-reducing end (Atkinson et al., 1994, Röhrig et al., 1994).
The *nodIJ* genes are also considered to be common *nod* genes as they are found in many rhizobial species, including *R. leguminosarum* bv. *viciae*, bv. *trifolii*, *R. etli* and *S. meliloti*. NodI and NodJ are involved in transport of Nod factors and are thought to be an ATP-binding cassette (ABC) transporter (Cardenas *et al*., 1996, Evans and Downie, 1986, Fernandez-Lopez *et al*., 1996, McKay and Djordjevic, 1993, Spaink *et al*., 1995, Vàzquez *et al*., 1993).

The organisation of the *nod* genes of *R. leguminosarum* bv. *viciae* is shown in Figure 1.1. The organisation of nod genes differs between different species, although *nodDABCIII* are often clustered into one organisational unit.

**Figure 1.1 The nod Genes of R. leguminosarum bv. viciae**

Adapted from van Rhijn and Vanderleyden, 1995.

The arrows indicate the direction in which the *nod* genes are transcribed. The common *nod* genes are dark blue and host-specific genes are purple. The asterisks indicate the position of *nod* boxes, which NodD binds to.

Host-specific *nod*-factors (*hsn*) are subtly different in each strain and are also responsible for host specificity. For example, the main factor that determines host specificity in *R. leguminosarum* bv. *viciae* and bv. *trifolii* is NodE (Spaink *et al*., 1991, Spaink *et al*., 1989). In *S. meliloti*, *nodH* and *nodPQ* are responsible for specifying nodulation of alfalfa (Cervantes *et al*., 1989,
Faucher et al., 1989, Roche et al., 1991, Schwedock and Long, 1989). In B. japonicum, nodVW, which are unique to this species, are required for nodulation of Macroptilium and Vigna species (Göttfert et al., 1990).

1.2.2 Nitrogen Fixation by Bacteroids

Inside bacteroids most nod genes are no longer expressed (Schlaman et al., 1991), possibly because large quantities of Nod factors can elicit plant defence reactions (Savouré et al., 1997) and bacteroids express nitrogen fixing genes (nif and fix). These genes are responsible for the reduction of atmospheric nitrogen to ammonium. The nif genes are structurally homologous to those of Klebsiella pneumoniae, a free-living bacterium in which nitrogen fixation has been extensively studied (Arnold et al., 1988). The fix genes are essential for nitrogen fixation, but do not have homologues in K. pneumoniae. These genes are usually organised into distinct clusters in rhizobia and have been studied most in S. meliloti and B. japonicum.

The two-component enzyme complex responsible for the process of nitrogen fixation is nitrogenase, which is encoded for by nifHDK. The α and β subunits of Component I (MoFe protein) are encoded for by nifD and nifK respectively. Component II (Fe protein) is encoded for by nifH. This enzyme is large and slow acting and may account for up to 30% of bacteroid cell protein (Haaker and Klugkist, 1987). The nifB, nifE and nifN genes are involved in the biosynthesis of a co-factor of component I (Dean et al., 1993). The nifH, nifM,
*nifQ* and *nifV* genes are also required for synthesis and maturation of the active enzyme complex (Filler et al., 1986, Howard et al., 1986, Imperial et al., 1984, Kennedy et al., 1986, Robinson et al., 1987).

High oxygen levels inactivate nitrogenase (Adams and Chelm, 1988, Hill, 1988) and expression of *nifA* is also dependent on low oxygen concentration (Ditta et al., 1987). NifA is a positive regulator of *nif* and *fix* genes and also activates other genes that are not involved in nitrogen fixation, but the functions are not known (Nienaber et al., 2000). The *fixL* and *fixJ* gene products sense and transduce the low oxygen signal. FixJ then activates transcription of *nifA* and *fixK* (Cebolla and Palomares, 1994). FixK is a positive regulator of *fixNOPQ*, which encode the high affinity membrane-bound cytochrome oxidase that supports bacteroid respiration in the microaerobic conditions of the nodule (Preisig et al., 1993). Several other fix genes, *fixABCX* and *fixGHIS* are also required for nitrogen fixation to occur in rhizobia, as mutants in these genes abolish the process (reviewed in Fisher, 1994).

There are also plant genes that are expressed specifically in root tissue as a result of the interaction with rhizobia. The plant-derived products are called nodulins. One nodulin is leghaemoglobin, which is transcribed in nodules (Fuller et al., 1983). The pigment in leghaemoglobin is responsible for nodules appearing pink. Leghaemoglobin maintains a high oxygen flux, but at a concentration about $10^4$ to $10^5$ times lower than in aerobic cultures, so that
nitrogen fixation is not inactivated and bacteroid respiration can still occur (Appleby, 1984, Hunt and Layzell, 1993).

The process of nitrogen fixation is highly energetic. The minimum stoichiometric requirements are eight electrons and 16 molecules of ATP hydrolysed per N₂ molecule fixed. Therefore, in addition to providing an environment free from oxidative stress, the plant also provides carbon and water so that the bacteroids can generate ATP and reductant for nitrogen fixation. In return, the bacteroids channel fixed nitrogen to the plant. Until recently, it was generally accepted that ammonium was the only product channelled to the plant. However, evidence has emerged that bacteroids provide alanine, in addition to ammonium. Ammonium is converted to alanine by the enzyme alanine dehyrogenase (AldA) (Allaway et al., 2000, Kretovich et al., 1986, Rosendahl et al., 1992, Waters et al., 1998, reviewed in Poole and Allaway, 2000). As the fixed nitrogen is supplied to the host plant, rather than being incorporated into compounds in the bacteroid, nitrogen fixation in the nodule is not generally accompanied by growth of bacteroids.

1.2.3 Carbon Metabolism by the Bacteroid

Free-living rhizobia are capable of utilising a wide range of carbon compounds for growth. However, the principal carbon compounds utilised by mature bacteroids are C₄-dicarboxylic acids, such as malate, succinate and fumarate, which are provided by the host plant. Mutants defective in C₄-dicarboxylic
acid uptake do not fix nitrogen, although they are able to form nodules (Arwas et al., 1986, Arwas et al., 1985, El-Din, 1992, Engelke et al., 1987, Finan et al., 1988, Finan et al., 1983, Humbeck and Werner, 1989, Lafontaine et al., 1989b, Ronson et al., 1981). The C₄-dicarboxylic acids are metabolised via the tricarboxylic acid cycle (TCA) and subsequently via malic enzyme and pyruvate dehydrogenase (Copeland et al., 1989, McKay et al., 1988).

Sucrose is the major carbohydrate translocated in legumes from shoots to nodules (Antoniw and Sprent, 1978, Kouchi and Yoneyama, 1984, Reibach and Streeter, 1983). Other carbon compounds present in nodules in addition to organic acids include glucose, fructose, maltose, trehalose and cyclitols (Kouchi and Yoneyama, 1984, Lafontaine et al., 1989a, Streeter, 1987, see Section 1.3.3.1). However, sugars are catabolised at very low levels in soybean and pea bacteroids (Glenn and Dilworth, 1981, Salminen and Streeter, 1987). Furthermore, pea and clover bacteroids with mutations in sugar catabolic enzymes are unaltered in symbiotic performance (Glenn et al., 1984, Ronson and Primrose, 1979).

As free-living cultures of rhizobia, including R. leguminosarum bv. viciae, utilise sugars, this indicates that there are fundamental differences in transport and metabolism in bacteroids. One reason why C₄-dicarboxylic acids are the principal energy source of bacteroids may be because they are more readily transported across the PBM than other carbon compounds such as sugars (McKay et al., 1985, Saroso et al., 1986, Udvardi et al., 1988).
Carbon compounds in nodules that are not catabolised may perform other roles. Such compounds include D-pinitol and proline (Ford, 1984, Keller and Ludlow, 1993, Straub et al., 1996), which might function as osmoprotectants in plants in times of drought stress (Section 1.3.4.2).
1.2 Soil Carbon

In order for nodulation to occur, rhizobia must first be able to grow and survive in the rhizosphere. Organic carbon, in the form of carbohydrates and organic acids, provides energy for growth and survival for non-photosynthesising micro-organisms. Carbohydrates comprise approximately 10% of soluble soil organic carbon content (Cheshire, 1985). Most is in the form of polysaccharides, with small amounts of monosaccharides and oligosaccharides also present (Cheshire, 1979). The origin of these compounds is dead plant tissue, root exudates and micro-organisms (Cheshire, 1985). Composition of root exudates varies with biotic factors such as species, age and nutrient status and abiotic factors such as temperature, soil structure, pesticides and water (Rovira, 1965).

Leakage of carbon compounds from plant roots varies with both the species and growth stage. Smith (1970) found that mature trees release different amounts and types of compounds respective to seedlings of the same species. It is worth noting that seedlings are the focus of most studies. The exudates may also differ in composition and patterns of availability along a root according to the age of the roots. For example, Jaeger et al. (1999) found that tryptophan was released from older sections and sucrose from younger sections of grass (*Avena barbata*).

The amount of exudate, root cap and mucilage available to micro-organisms has been estimated at between 3 to 15% of the dry weight of the root.
(Campbell and Greaves, 1990). Most reviewers conclude that around 20% of fixed photosynthate is lost to the rhizosphere, a process known as rhizodeposition (Campbell and Greaves, 1990, Metting, 1985, Sparling, 1985). Much of this loss is thought to be through passive leakage and from damaged regions of the plant (Klein et al., 1990, Whipps, 1990). The actual amount available to micro-organisms is unknown as plants may reclaim some of the exuded carbon by active uptake against a concentration gradient (Jones and Darrah, 1996).

1.3 myo-Inositol

One of the many soluble sugar compounds that Rhizobium can utilise as a sole carbon source is the cyclitol myo-inositol. This compound is abundant in soil and the rhizosphere (c.f. Section 1.5). It has been shown that in soil solution, Rhizobium numbers increase concomitantly with the disappearance of myo-inositol (Wood and Stanway, 2000). Raggio et al. (1959) found that supplementation of the growth medium of common bean (P. vulgaris) with mesoinositol (myo-inositol) resulted in an increase in both the percentage of nodulated roots and nodules per root containing R. leguminosarum bv. phaseoli. The growth medium of the plants contained sucrose, so the increase may not have simply been due to myo-inositol being utilised as a carbon source. These data imply that myo-inositol may be involved in the growth and survival of Rhizobium in soil and the rhizosphere.
myo-Inositol is a sugar alcohol with a six carbon ring structure. There are nine stereoisomeric forms of inositol (L-chiro, D-chiro, myo, neo, scylo, muco, cis, epi, allo) but the myo- form appears to be by far the most common (Figure 1.2). The other forms most commonly encountered in nature are D-chiro, L-chiro and scyllo-inositol (Loewus, 1990).

Figure 1.2 The Structural Formula of myo-Inositol

1.3.1 Synthesis of myo-Inositol

myo-Inositol de novo synthesis occurs by conversion from glucose-6-phosphate in two steps, which are catalysed by L-myoinositol 1-phosphate synthase and inositol monophosphatase (Loewus, 1990). Control of synthesis is through feedback inhibition of the synthase (Nelson et al., 1998) (Figure 1.3). The pathway was originally elucidated in yeast, but is believed to be the same for all organisms (Loewus, 1990).
1.3.2 myo-Inositol Utilisation by Bacteria

A pathway detailing the degradation of myo-inositol via 2-keto-myoinositol and D-2,3-diketo-4-deoxy epi-inositol, has been elucidated in the soil-dwelling bacterium, *Klebsiella aerogenes* (*Aerobacter aerogenes*) (Anderson and Magasanik, 1971). The final products were proposed to be dihydroxyacetone phosphate, acetyl coenzyme A (CoA) and carbon dioxide (CO₂).

It was postulated that the catabolic pathway is similar for *R. leguminosarum* bv. *viciae*, as the first two enzymes in the pathway, myo-inositol dehydrogenase and 2-keto-myoinositol dehydratase have been shown to be induced in the presence of myo-inositol (Poole et al., 1994). As a result of findings in this project, additional steps have been added to the pathway. In these steps, dihydroxyacetone phosphate forms pyruvate, which then combines with pyruvate to form acetolactate (Figure 1.4).
Figure 1.4  Proposed Pathway of myo-Inositol Degradation in *R. leguminosarum* bv. *viciae*

The proposed pathway of myo-inositol catabolism is adapted from that of *K. aerogenes*, *B. subtilis* and work in this thesis. Putative functions of Iol proteins from *B. subtilis* are shown.

* The pathway of conversion of D-2,3-diketo-4-deoxy-epi-inositol to 2-deoxy-5-keto-D-gluconic acid has not been characterised. Conversion may occur spontaneously, or through the action of a hydratase.
*Pseudomonas* species can also utilise myo-inositol as the sole carbon source (Deshusses and Reber, 1972), as can *Bacillus subtilis*. An operon of catabolic genes for the degradation of myo-inositol has been identified in *B. subtilis* (Fujita and Fujita, 1983, Yoshida *et al.*, 1994) (Figure 1.5). The proposed pathway of degradation corresponds to that of *K. aerogenes* (Yoshida *et al.*, 1997).

**Figure 1.5 iol Operon of B. subtilis**

![Diagram of iol Operon](image)

Adapted from Yoshida *et al.*, 1997.

The *iol* operon of *B. subtilis* is an 11.5kb region that contains 10 genes *iolA*-*J* which are believed to be transcribed from the *iol* promoter. All 10 *iol* genes are involved in myo-inositol catabolism, as disruption of each one individually eliminated the ability to degrade myo-inositol. The *iolR* and *iolS* genes are thought to be co-transcribed from the *iolRS* promoter (Yoshida *et al.*, 1997). The proposed functions of the *iol* genes are presented in Table 1.2.
Table 1.2 The iol Genes of *B. subtilis*

<table>
<thead>
<tr>
<th>gene</th>
<th>homologous proteins from Genbank database</th>
<th>expected function of gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>iolA</td>
<td>methylmalonic semialdehyde dehydrogenase</td>
<td>methylmalonic semialdehyde oxidative decarboxylase</td>
</tr>
<tr>
<td>iolB</td>
<td>no significant homology to known proteins</td>
<td>no known function</td>
</tr>
<tr>
<td>iolC</td>
<td>gluconokinase</td>
<td>gluconokinase</td>
</tr>
<tr>
<td>iolD</td>
<td>acetolactate synthase</td>
<td>no known function</td>
</tr>
<tr>
<td>iolE</td>
<td>MocC protein of <em>S. meliloti</em></td>
<td>no known function</td>
</tr>
<tr>
<td>iolF</td>
<td>proline/betaine or dicarboxylic acid transporter</td>
<td><em>myo</em>-inositol transporter</td>
</tr>
<tr>
<td>iolG</td>
<td>MocA protein of <em>S. meliloti</em></td>
<td><em>myo</em>-inositol 2-dehydrogenase</td>
</tr>
<tr>
<td>iolH</td>
<td>no significant homology to known proteins</td>
<td>no known function</td>
</tr>
<tr>
<td>iolI</td>
<td>no significant homology to known proteins</td>
<td>no known function</td>
</tr>
<tr>
<td>iolJ</td>
<td>aldolase</td>
<td>aldolase</td>
</tr>
<tr>
<td>iolR</td>
<td>glucitol operon repressor</td>
<td>operon repressor</td>
</tr>
<tr>
<td>iolS</td>
<td>MocA protein of <em>Agrobacterium tumefaciens</em></td>
<td>no known function</td>
</tr>
</tbody>
</table>

Compiled from the results of Yoshida *et al.*, (1997).

*B. subtilis* is the only bacterium to date that has been shown to possess an operon of *myo*-inositol degrading genes. The genes responsible for the degradation of *myo*-inositol may not be organised into an operon in rhizobial species and these genes may be found both in the chromosome and on plasmids. In *S. meliloti*, a gene encoding *myo*-inositol dehydrogenase was identified that was not part of an operon of *myo*-inositol utilising genes.
(Galbraith et al., 1998). Curing *R. leguminosarum* bv. *trifolii* of plasmids resulted in the loss of the ability to utilise *myo*-inositol (Baldani et al., 1992).

1.3.2.1 *myo*-Inositol Uptake by Bacteria

*myo*-Inositol enters the cell through dedicated transport systems in *K. aerogenes*. The system is bi-directional and can also transport *scyllo*-inositol, which is not metabolised (Deshusses and Reber, 1977b). Attempts to identify an active binding protein for *myo*- or *scyllo*-inositol were unsuccessful (Deshussses and Reber, 1977a). In contrast, Reber et al. (1977) reported that there are two uptake systems in *Pseudomonas putida*, a high affinity system involving a periplasmic binding protein and a different, low affinity system.

Mutation of a *myo*-inositol transport transport system in *Pseudomonas* sp. JD34 caused very slow growth when *myo*-inositol was the sole carbon source (Frey et al., 1983). The transport genes were mapped to a chromosomal 11.5kb region, but were not sequenced (Gauchat-Feiss et al., 1985). In another *Pseudomonas* species *myo*-inositol was transported by a high affinity transport system involving a 30kD *myo*-inositol-binding protein (Deshusses and Belet, 1984). The protein was purified and sequencing of the N-terminal segment revealed that it highly resembles the galactose binding protein of *E. coli* (Mahoney et al., 1981).
1.3.3 myo-Inositol in the Environment

1.3.3.1 Plant myo-Inositol

myo-Inositol is ubiquitous in the environment. It has been shown to accumulate in plant leaves exposed to air pollution, possibly due to membrane disintegration (Bücker and Guderian, 1994). Analysis of several foodstuffs such as oats (*Avena sativa* L.), soybean and barley (*Hordeum vulgare* L.) revealed free inositol (Norris and Darbre, 1956).

Sulochana (1962) reported the presence of inositol in soil and root exudates of cotton, along with many other compounds. Wood and Stanway, (2000) identified inositol in soybean exudates, although it was not the most abundant compound. However, a study of sugars in rice straw (*Oryza sativa* L.) and straw compost by high resolution gas chromatography revealed many sugars, predominantly xylose, arabinose and glucose, but no myo-inositol (Sugahara *et al.*, 1992). Sánchez-Mata *et al.* (1998) carried out high performance liquid chromatography (HPLC) analysis of legume carbon compounds and found that ciceritol, an inositol digalactoside was the principal sugar in chickpeas (*Cicer arietinum* L.). It was also present in lentils (*Lens sculenta* L.).

Sucrose is a major constituent of nodules, along with other sugars including glucose and fructose (Antoniw and Sprent, 1978, Kouchi and Yoneyama, 1984, Reibach and Streeter, 1983). However, there are several other carbon
compounds found in nodules and plant roots, including myo-inositol and other cyclitols. myo-Inositol was the most abundant compound found in bacteroids of *B. japonicum* on soybean (Skøt and Egsgaard, 1984, Streeter, 1987).

Streeter and Salminen (1985) reported that myo-inositol had been found in nodules formed on several different legumes including soybean, clover, pea, alfalfa and common bean. D-pinitol (1-D-4-O-methyl-myoinositol, D-chiro-inositol, oonitol and O-methyl-scyllo-inositol were also present in some nodules, with the largest amounts of cyclitols found in nodules on roots of soybean and clover. Only trace amounts of myo-inositol were found in nodules formed on pea and in contrast to other legumes, common bean only contains trace amounts of myo-inositol and no other cyclitols (Lafontaine et al., 1989a, Streeter and Salminen, 1985).

Davis and Nordin (1983) reported that D-pinitol was the most abundant compound in clover. D-pinitol, oonitol and chiro-inositol were also reported in soybean cytosol and bacteroids of *B. japonicum* (Kouchi and Yoneyama, 1986). Kouchi and Yoneyama (1984) found D-pinitol, chiro-inositol, sequoyitol and myo-inositol in the organs of plants, with high turnover in leaves, but low turnover in nodules. D-pinitol was again the major cyclitol. Streeter and Salminen (1985) reported that D-pinitol was present at a greater concentration than sucrose in nodules formed on lupin (*Lupinus angustifolius*). Aminated derivatives of myo-inositol called rhizopines have also been reported in nodules containing *S. meliloti* (Murphy et al., 1987, see Section 1.3.6).
Despite the prevalence of cyclitols in plants and nodules, it is believed that cyclitols are metabolically inert. When labelled $^{13}$CO$_2$ was supplied to plants, the labelled carbon was not readily incorporated into the cyclitols, suggesting a low turnover, particularly in nodules (Davis and Nordin 1983, Kouchi and Yoneyama, 1986). Studies on the first two enzymes in the catabolic pathway of myo-inositol showed that they were not expressed in the bacteroid (Poole et al., 1994, this work Section 5.2.10).

Clearly, myo-Inositol and derivatives are abundant in plant tissues and nodules, but as they are unlikely to be used as carbon sources, they must have other roles that have not yet been elucidated. Galbraith et al. (1998) hypothesised that bacteroids may synthesise cyclitols in nodules to act as osmoprotectants, as the bacteroids are subjected to high osmotic pressure. The role of myo-inositol derivatives as osmoprotectants is discussed in Section 1.3.4.2.
1.3.3.2 Soil myo-Inositol

Soil studies have primarily focused on inositol phosphates rather than free inositol, due to the importance of these compounds in phosphorus supply to plants. The results of the studies that have been carried out are conflicting, partly due to differences in methodology (e.g., extraction and analysis) and to the different composition of soils. Soil is an extremely complex and dynamic environment, where compounds may be complexed into unavailable forms. A variety of factors such as pH, presence or absence of plants, farming strategies and microbial activity all contribute to soil composition (Wood, 1995).

Yoshida (1940) isolated crystals of pure inositol, as well as myo-inositol hexaphosphate and other inositol phosphates, from three different samples of Hawaiian soil. The inositol was obtained by acid hydrolysis of soil organic phosphorus, extraction of which from soil was by solubility in acid and alkali. The author postulated that the inositol was probably not present in the free form, but as inositol monophosphate and that the separation technique may have been responsible for releasing free inositol.

Lynch et al. (1958) studied the composition of four soils with different organic matter content and discovered several carbon compounds, including inositol, also by using acid hydrolysis to extract compounds. McKercher and Anderson (1968) found small quantities of inositol in addition to other unidentified carbohydrates following acid hydrolysis of three Canadian soils.
However, others employing similar techniques to isolate organic phosphorus reported that they were unable to isolate pure inositol from soil, only myo-inositol hexaphosphate (Bower 1945, Wrenshall and Dyer, 1940). Similarly, no myo-inositol was found in soils derived from volcanic ash (andosols), where the most abundant sugars were mannose, fucose and ribose (Yoshida and Kumada, 1979).

Inositol was found in both the soils and exudates of cotton strains Neurospora crassa and N. sitophila (Sulochana, 1962). HPLC analysis of soil solutions extracted from two soils, showed that myo-inositol is the most common soluble carbon compound present (Wood and Stanway, 2000). This method does not involve any chemical changes to the soil and hence, it could be argued that it is a more accurate measure of soil composition.

These studies indicate that myo-inositol is likely to be present in soils, but not necessarily in the free form and when it is found in the free form, its presence may be overlooked, due to the techniques used.
1.3.4 myo-Inositol Derivatives

myo-Inositol is not only used as a carbon source by bacteria. It is an essential component of several biochemical pathways found in all organisms. myo-Inositol acts as a co-factor in the formation of galactinol from UDP-galactose (Frydman and Neufeld, 1963). This is the substrate for raffinose sugars, which are believed to be important in plants for stress tolerance and carbohydrate transport (Loewus and Dickinson, 1982).

A myo-inositol oxidation pathway has been identified in plants. myo-Inositol forms D-glucuronate (although D-glucuronate can also be made from UDP-glucose) (Loewus and Loewus, 1983), which then undergoes metabolism to be incorporated into cell wall polysaccharides, pentose phosphate cycle intermediates and hexonic and hexaric acids (Loewus, 1990). One product of D-glucuronate, D-galacturonate forms oligosaccharides that act as signals to activate plant defence mechanisms (Ryan, 1987). Therefore, myo-inositol is an essential carbon source for cell wall synthesis in growing plant tissues (Sasaki and Nagahashi, 1990).

In the following sections, the roles of various derivatives of myo-inositol are discussed.
1.3.4.1 Auxins

Auxins are plant hormones that are required for lateral root formation (Wightman et al., 1980). By conjugating myo-inositol to auxins, they are temporarily inactivated, which allows long-distance transport of these compounds from shoots to roots (Cohen and Bandurski, 1982). Auxin accumulates in the inner cortical cells of alfalfa after inoculation by S. meliloti (Hirsch et al., 1997). These are the cells that eventually become nodule primordia (Yang et al., 1994). The high auxin levels are thought to be necessary to stimulate the cell divisions that lead to nodule primordia development. Nod-factors (lipo-chitin-oligosaccharides) produced by rhizobia cause inhibition of auxin transport, which leads to an accumulation of auxins at the site of nodule initiation (Mathesius et al., 1998). This is mediated by flavonoids, which are also known to inhibit auxin transport (Jacobs and Rubery, 1988).

1.3.4.2 Methyl Esters of myo-Inositol

Another important pathway involves monomethylation or dimethylation of myo-inositol. It has been proposed that methyl esters of myo-inositol may play a role as a preformed store of myo-inositol that can be utilised when metabolic demands arise (Loewus, 1990). myo-Inositol and methylated derivatives are also implicated in high salinity tolerance by some animals and higher plants (Keller and Ludlow, 1994, Nelson et al., 1998).
For example, D-pinitol (1-D-4-O-methyl-myo-inositol) is a derivative of myo-inositol that may protect proteins and membranes from dehydration and denaturation in the tropical legume, pigeonpea (*Cajanus cajan*) and other legumes (Keller and Ludlow, 1994). D-pinitol and the amino acid proline are the main solutes to accumulate in drought-stressed pigeonpea leaves (Ford, 1984). Carbon flux is diverted from starch and sucrose accumulation into polyols in response to low water potential caused by drought and salt stress (Keller and Ludlow, 1994). D-pinitol and its precursor ononitol are thought to lower the cytoplasmic osmotic potential and balance sodium accumulation in the plant vacuole, in a similar mechanism to the role of myo-inositol and other solutes in the renal cells of vertebrates (Burg, 1994).

1.3.4.3 Di-	extit{myo}-inositol 1,1'-phosphate

Di-	extit{myo}-inositol 1,1'-phosphate (DIP) has been shown in the thermophilic bacteria 	extit{Pyrococcus woesei} (Scholz et al., 1992) and 	extit{Methanococcus igneus} to exert a thermo-stabilising effect on proteins within the cell. The cellular concentration of this small molecular mass compound increases comcomitantly with temperature increase (Ciulla et al., 1994). However, DIP is synthesised directly from glucose-6-phosphate rather than from 	extit{myo}-inositol (Scholz et al., 1998).

1.3.4.4 Phosphatidylinositol

	extit{myo}-Inositol is a constituent of membrane lipids, such as phosphatidylinositol (cardiolipin), which comprises 20% of the lipids of the inner mitochondrial membrane in animals, as well as a large proportion of bacterial membrane lipids. This compound is also important for signalling in eukaryotes.

Phosphatidylinositol is modified to produce the membrane lipid phosphatidylinositol-4,5-bisphosphate. This is hydrolysed by the enzyme phospholipase C to release inositol-1,4,5-triphosphate. Inositol-1,4,5-triphosphate then combines with a receptor that activates a calcium pump or transporter. Calcium then activates a variety of enzymes. Inositol-1,4,5-triphosphate is hydrolysed and then converted back to phosphatidylinositol. Similar pathways occur in plants and mammals (Salisbury and Ross, 1992).
Another role of phosphatidylinositol is in ice nucleation. Ice nucleating agents limit supercooling and are found in lichens, higher plants, insects, molluscs and bacteria (Duman et al., 1991). Kozloff et al., (1991) showed that *P. syringae* is able to degrade *myo*-inositol hexyaphosphate when grown under inorganic phosphorus limitation. This provides *myo*-inositol for production of a membrane protein containing phosphatidylinositol that enables the bacterium to cause ice nucleation. However, the bacterium was unable to utilise pure free *myo*-inositol as a component of the protein (Blondeaux and Cochet, 1994). *P. syringae*, *Erwinia herbicola* and *E. coli* all contain phosphatidylinositol and phosphatidylinositol synthase. The latter increases in amount in the presence of an ice nucleation gene (Duman et al., 1991).

1.3.4.5 *myo*-Inositol Hexaphosphate

*myo*-Inositol is abundant in animals and higher plants as *myo*-inositol hexaphosphate (phytate, or as the calcium-magnesium salt, phytin) (Lehninger, 1975, Loewus et al., 1990) (Figure 1.6). Phosphorus is often one of the major limiting factors for plant growth, so both plants and microorganisms have evolved systems for acquiring and safeguarding it. Conjugating phosphorus to *myo*-inositol acts as a phosphorus store (Cosgrove, 1980). The majority of inositol phosphates are the *myo*- isomer, then the *scyllo*- isomer and very small quantities of *dl*- and *neo*- isomers (Cosgrove, 1963, McKercher and Anderson, 1968, Omotoso and Wild, 1970).

Phytate is also abundant in soil, presumably due to exudation from plants. Studies have shown that phytate accounts for approximately 33% of organic phosphorus in soil (Anderson, 1956, Yoshida, 1940, Bower, 1945) and up to 50% of inorganic phosphate (Anderson, 1980, Dalal, 1978). Phytate has also been reported in both freshwater and saltwater sediments (De Groot and Golterman, 1993, Suzumura and Kamatani, 1995).

Phytate can chelate various metals and bind proteins, diminishing the bioavailability of proteins and nutritionally important minerals (Liu et al., 1998). Phytate may be particularly important as a siderophore, due to its high affinity
for iron. It is believed to be universal in eukaryotic tissues, where it may bind to iron without forming harmful hydroxyl radicals (Hawkins et al., 1993). A similar role has also been identified in prokaryotes as phytate can allow *P. aeruginosa* to grow in iron-limiting conditions (Smith et al., 1994). Other inositol phosphates may also have similar roles, as *myo*-inositol 1,2,3-trisphosphate was shown to be able to act as a siderophore in *P. aeruginosa*. The activity of iron uptake was similar to that of phytate (Spiers et al., 1996).

### 1.3.4.5.1 Phytase

The enzyme phytase catalyses the release of phosphate from phytate, yielding yield inositol trisphosphate, inositol monophosphate, phosphoric acid and inositol (Anderson, 1915). Phytases have been found in plants, bacteria and fungi. In the former two, they are usually intracellular, but fungal and yeast phytases are generally extracellular (Howson and Davis, 1983).

Phytase activity in seeds increases during germination, releasing phosphate (Reddy et al., 1989). Phytase may also be active in soil, as phytate derivatives have been found alongside phytate (Bower 1945, Yoshida, 1940). Soluble organic phosphorus that could be assimilated by maize was obtained in the root medium of maize plants supplemented with phytate and phytase (Findenegg and Nelemans, 1993). *Pseudomonas* species can synthesise phytase and utilise phytate as a source of inorganic phosphate when it is lacking in the culture medium (Cosgrove, 1970). Phytase production was enhanced by inorganic phosphorus-limiting conditions (Shieh and Ware,
Recently, several ruminal anaerobic bacteria have been found to have high phytase activity, which may release phosphorus for utilisation by their host animals (Yanke et al., 1998).

Of 200 randomly selected soil bacterial isolates, none could utilise phytate as a sole carbon and phosphorus source (Richardson and Hadobas, 1997), although they could utilise phytate as a phosphorus source when additional carbon was supplied. Of a further 238 isolates obtained, two fluorescent *Pseudomonas* strains CCAR53 and CCAR59, were able to utilise phytate as a sole carbon source, depending on the pH of the medium. Therefore, phytate release from plants and its subsequent breakdown by phytase may account for the presence of free myo-inositol in soil.

1.3.6 Rhizopines

Rhizopines are compounds produced exclusively in legume root nodules, which can be utilised by free-living rhizobia as both a carbon and nitrogen source. They are similar in structure to opines, of which many different forms have been identified in *Agrobacterium* species, which are Gram-negative, soil-dwelling bacteria (Tempé and Schell, 1977). The first rhizopine was discovered in alfalfa nodules containing *S. meliloti* strain L5-30 by Tempé and Petit (1983) and is called L-3-O-methyl-scyllo-inosamine (3-O-MSI) (Figure 1.5). Subsequently, a derivative called scyllo-inosamine (SI) was discovered in nodules containing *S. meliloti* strain Rm220-3 (Saint et al., 1993) (Figure
1.7). Both are aminated derivatives of myo-inositol, although by definition, rhizopines do not need to be inositol derived (Murphy and Saint, 1992).

**Figure 1.7 Structure of Rhizopines**

![Structure of Rhizopines](image)

The ability to synthesise and catabolise rhizopines appears to be an infrequent trait amongst species of *Rhizobium*, although such species are widespread in the environment (Wexler and Murphy, 1995). One study found that of over 300 strains of *Rhizobium*, approximately 10% of *S. meliloti* strains and 14% of *R. leguminosarum* bv. *viciae* strains possessed the genes to both synthesise and catabolise rhizopines. No *R. leguminosarum* bv *phaseoli* or *R. leguminosarum* bv. *trifolii* tested had the genes (Murphy *et al.*, 1995). Rossbach *et al.* (1995) found that strains of *Agrobacterium, Bradyrhizobium, Azorhizobium, Azospirillum, Escherichia* and *Pseudomonas* species were incapable of utilising rhizopines and did not contain known rhizopine catabolic genes (see below).
1.3.6.1 Rhizopine Synthetic and Catabolic Genes of *S. meliloti*

*S. meliloti* strains possess genes encoding both the synthesis of rhizopines (*mos*) and catabolism (*moc*). These genes are closely linked and located on the Sym plasmid in *S. meliloti* (Murphy *et al*., 1987). The expression of *mos* genes in *S. meliloti* is directly controlled by the symbiotic nitrogen fixation regulatory system (Murphy *et al*., 1988), so *mos* genes are expressed only when *Rhizobium* has differentiated into a bacteroid and is expressing *nif* genes. The *moc* genes are not regulated in the same way and are only expressed in free-living rhizobia (Saint *et al*., 1993). Despite the link to symbiosis, rhizopines are not essential for symbiosis, as mutants are still able to nodulate and fix nitrogen (Murphy *et al*., 1988).

The *mos* locus contains four genes in *S. meliloti* strain L5-30, *mosABC* and *orf1*. MosA is involved in a methylation step to form 3-O-MSI from SI (Rao *et al*., 1995). The *mosA* gene is not present in *S. meliloti* strain Rm220-3, which produces SI (Murphy *et al*., 1993). MosB may have a regulatory role and MosC may be involved in export of rhizopines or their precursors. A fourth gene, *orf1*, does not code for a known protein and is not required for rhizopine synthesis (Murphy *et al*., 1993).

The *moc* locus contains six open reading frames (ORFs), of which four genes are essential for rhizopine catabolism in *S. meliloti* (Rossbach *et al*., 1994). MocB is believed to be a periplasmic binding protein, probably involved in the transport of rhizopines into the cell. MocR is thought to be a regulatory
protein with a DNA-binding motif. MocA is thought to be a dehydrogenase and MocC has no known function but is believed to be involved in inositol catabolism (Rossbach et al., 1994). The latter two genes resemble genes in the \textit{B. subtilis ino} operon (Yoshida et al., 1997). The moc operon of \textit{R. leguminosarum} bv. \textit{viciae} contains nine ORFs, of which six are essential for degradation of 3-O-MSI (mocCABRDE) and two (mocCA) for degradation of SI (Bahar et al., 1998).

Additional genes involved in myo-inositol catabolism are also required for degradation of rhizopines. An inositol dehydrogenase mutant of \textit{S. meliloti} which could not grow on myo-inositol as the sole carbon source, was also unable to catabolise the rhizopine L-3-O-MSI (Galbraith et al., 1998). The same results were reported through a collaboration that used the myo-inositol catabolic mutants of \textit{R. leguminosarum} biovar \textit{viciae} used in this project (RU360 and RU361). The wild type strain 3841 does not possess the genes to synthesise or catabolise rhizopines but can utilise myo-inositol as the sole carbon source. After a plasmid containing rhizopine catabolic genes was introduced into 3841 and the mutants RU360 and RU361, 3841 gained the ability to utilise rhizopines as the sole carbon and nitrogen source, but the mutants did not (Bahar et al., 1998).

McSpadden Gardener and deBruijn (1998) isolated 21 soil isolates that could utilise SI as the sole carbon and nitrogen source. Hybridisation tests carried out using rhizopine catabolic genes as probes were negative, indicating that these strains do not use the known catabolic genes. Although 16S ribosomal
DNA analysis revealed that there were five distinct genera of bacteria, only four *Arthrobacter* and two *Sinorhizobium* isolates passed a standard rhizopine catabolism test and the *Sinorhizobium* isolates were unable to nodulate alfalfa. These data indicate that there may be alternative genes for rhizopine catabolism that are as yet unidentified.

### 1.3.6.2 Agrobacterium Opines

The *Agrobacterium* Opine Concept outlines a novel role for these compounds. It was postulated that pathogenic agrobacteria growing in tumours on host plants induce the production of selective growth substrates for other bacteria capable of inciting infection that they do not utilise themselves (Guyon *et al.*, 1980). This process of aiding genetic siblings is known as kin selection. A concept analogous to the opine concept, the Rhizopine Concept, was postulated by Murphy and Saint, (1992). It was suggested that bacteroids specifically produce and release nutrients into the rhizosphere to selectively aid the growth and survival of free-living rhizobia in infection threads and the rhizosphere. Strains that could utilise such compounds may thus have a competitive advantage over non-utilising strains. However, the selective advantage proposed by the concept would only apply if the environment were limited in carbon, as there are many other carbon compounds that rhizobia can utilise.

Another role for opines in *A. tumefaciens* involves a hierarchical cascade where opine availability regulates quorum sensing, which in turn regulates
transfer of the Ti plasmid between bacteria and plants (Piper et al., 1999). This plasmid contains the genes for synthesising and catabolising opines. Quorum sensing (auto induction) is a mechanism by which bacteria regulate the expression of certain genes in response to cell density. Bacteria gauge the size of the population by sensing small, diffusible signal molecules known as auto inducers that are produced by the bacteria. At a threshold concentration that corresponds with cell density, these molecules regulate transcription of certain genes, the products of which are assumed to be of benefit to the bacteria in the particular conditions. The autoinducers are usually N-acylated derivatives of L-homoserine lactone (acyl-HSL). It is postulated that regulation of quorum sensing by opines informs the Ti plasmid that nutritional conditions are conducive to conjugation Piper et al., 1999). This regulation may also serve to increase the copy number of the plasmid so that there is an increase in components of the opine catabolic system, which may be advantageous when nutritional resources are limited (Li and Farrand, 2000).

The evolutionary factors involved in rhizopine synthesis and catabolism, especially in relation to distribution of the genes, are not yet fully understood (Wexler et al., 1996). Rhizosphere colonisation studies using rhizopine positive and negative strains have found that rhizopine positive strains have a competitive advantage, but the exact role of rhizopines is not known. (Gordon et al., 1996, Heinrich et al., 1999) (c.f. Section 1.4).
1.4 Rhizosphere Growth and Competition for Nodulation

There has been extensive study of the genes involved in nodulation and nitrogen fixation by *Rhizobium*. However, relatively little is known about the traits that enable *Rhizobium* to grow and colonise the rhizosphere and successfully nodulate plants in the presence of other strains. The ability for one strain to dominate nodulation is termed competitiveness. Rhizobia chosen for their improved nodulation of legumes *in vitro* have frequently been out-competed by indigenous bacteria *in situ* due to unknown factors (Graham, 1981, Robleto et al., 1998).

There are two aspects of competition that are considered here, the ability to grow and survive in the rhizosphere and the ability to occupy nodules. To identify genes that are important for competition, many studies have compared a wild type with a defined mutant by co-inoculating an equal ratio into the rhizosphere and observing differences in phenotypic behaviour. Genes of particular interest are ones where mutants appear to grow in the rhizosphere and nodulate plants as successfully as the wild type when inoculated alone, but which are disadvantaged when co-inoculated.

1.4.1 Competition for Rhizosphere Colonisation

Studies on *Pseudomonas* have revealed several factors important for bacterial rhizosphere colonisation. These include motility (de Weger et al.,
1987), production of the O-antigen of lipopolysaccharide (de Weger et al., 1989), amino acids (Simons et al., 1997) and vitamin B1 (Simons et al., 1996).

Recently, a study identified a *P. fluorescens* mutant in a two-component regulatory system which, when inoculated singly onto potato cuttings, colonised as effectively as the wild type. However, when co-inoculated in equal numbers of $10^7$ colony forming units, the mutant was impaired by 300-1000-fold. The mutant was not distinguishable from the wild type with respect to *in vitro* growth rate and factors known to be important for colonisation, so the authors concluded that the two-component system must respond to some unknown factor that stimulates a trait crucial for colonisation (Dekkers et al., 1998).

Roberts et al. (1999) postulated that the quantities of carbon sources found in seed exudates and the ability of micro-organisms to utilise them are important in spermosphere colonisation. They studied a mutant of *Enterobacter cloacae* in *pfkA*, which encodes a phosphofructokinase. The mutant was severely deficient in its ability to colonise the spermosphere of several plant species. However, the mutant was only limited in the spermosphere of plants that released limited amounts of carbon.

Water-soluble vitamins may be important in competition for colonisation of the rhizosphere. Colonisation of alfalfa plant roots by *S. meliloti* can be limited by availability of biotin, thiamine and riboflavin (Streit et al., 1996). Alfalfa roots release biotin (Rovira and Harris, 1961), which is a co-factor for carboxylases.
Biotin auxotrophs created by transposon mutagenesis competed poorly for root colonisation when co-inoculated with the wild type into the rhizosphere, although they grew when inoculated alone in very low numbers (fewer than 100 cells). The results indicated that both synthesis and uptake of biotin from the rhizosphere were important for successful competition in the rhizosphere (Streit et al., 1996).

The ability to utilise carbon compounds in the rhizosphere and the role in competition has also been studied, particularly in relation to plasmid encoded catabolic genes. Plasmids are important in determining competitive ability in *Rhizobium* species as it has been shown that when cured of plasmids, competitive ability was severely impaired. (Brom et al., 1992, Hynes, 1990, Moënne-Loccoz and Weaver, 1995). However, the mechanisms are unknown as the plasmids contain many genes.

### 1.4.2 Competition for Nodulation

Oresnik *et al.*, (1998) showed that some catabolic loci present on plasmids are important for competition for nodulation by *R. leguminosarum* bv. *trifolii*. A mutant unable to utilise rhamnose occupied less than 10% of nodules when co-inoculated with the wild type onto clover. However, sorbitol and adonitol mutants were not impaired. The mutant was able to colonise the rhizosphere, nodulate clover, fix nitrogen at rates similar to the wild type and neither *in vitro* nor *in vivo* growth was slowed by the mutation. Rhamnose catabolic genes
were inducible by clover root extracts, indicating that this compound is present in the rhizosphere and may be important in nodulation.

Rhamnose is a component of plant cell walls in the polysaccharide rhamnogalacturonan (McNeil et al., 1984). Oresnik et al. (1998) postulated that as plant cell walls are continuous with infection threads, rhamnogalacturonan subunits or degradation products might be present in infection threads. The rhamnose mutant strain may then grow more slowly in the infection thread. Depending on when the mechanism that results in infection threads being aborted occurs, proportionately more infection threads formed by the mutant may be aborted, resulting in the wild type occupying the majority of nodules.

A mutant of R. etli in the rosR gene grew in the rhizosphere and nodulated bean as normal when inoculated alone. When co-inoculated with the wild type, a 17,000-fold excess of the mutant was required to achieve equal nodule occupancy with the wild type, although a high initial inoculum of $10^8$ cells was used. The mutant had altered cell surface characteristics (Araujo et al., 1994). There is also evidence that overexpression of rosR might enhance competitiveness (Bittinger et al., 1997). RosR is a negative transcriptional regulator of at least 42 loci and a positive regulator of at least one. The negatively regulated genes have homology with genes involved in polysaccharide and carbohydrate metabolism, synthesis and transport of opines and genes that may be involved in survival in the rhizosphere. Mutations in these genes did not affect nodulation competitiveness. However,
when two *rosR* mutants were mutated in loci that restored the normal cell surface characteristics, they became more competitive than the *rosR* mutant. Therefore, regulation by *rosR* is important for competition in some cases (Bittinger and Handelsman, 2000).

Milcamps *et al.* (1998) reported that three strains of *S. meliloti* mutated in genes induced by carbon deprivation were not at a competitive disadvantage for nodule occupancy. One strain that was mutated in a putative sugar transport system occupied more nodules than the wild type when co-inoculated onto alfalfa. As in other studies, an equal ratio of wildtype to carbon-deprivation induced mutants of *S. meliloti* were inoculated onto alfalfa. The mutants were able to nodulate and had equal nitrogenase activity to the wild type.

A region on the large plasmid of *S. meliloti* was identified as containing genes that are responsible for nodulation efficiency and competitive ability. These genes were designated nodulation formation efficiency genes (*nfe*) (Sanjuan and Olivares, 1989). The expression of these genes was found to be dependent on the NifA/NtrA regulatory system (Sanjuan and Olivares, 1991). Mutation in this region caused a delay in nodule formation and a reduction in nodulation competitiveness when co-inoculated with the wild type, although again, the mechanism for this is unknown.

Chun and Stacey (1994) identified a gene from *B. japonicum* termed *nfeC* that caused a significant delay in nodulation of soybean and was at a competitive
disadvantage for nodulation when co-inoculated with the wild type. The mutant exhibited a 6-day delay in nodule formation and occupied less than 12% of nodules when co-inoculated with the wild type. The gene was linked to the *nod* gene cluster and had two promoters, one that was expressed in bacteroids and the other in free-living rhizobia. The mutant was able to fix nitrogen effectively. However, the function of the gene was unknown.

1.4.2.1 Utilisation of Secondary Metabolites

Rhizobia may gain an advantage over other micro-organisms because of their ability to utilise novel nutritional compounds such as rhizopines and plant secondary metabolites, which include calystegins, betaines and homoserine. Plants may deliberately release these compounds in order to promote beneficial associations such as nodulation. For example, homoserine which is produced by peas, has been shown to be a selective substrate for *R. leguminosarum* bv. *viciae* (van Egeraat, 1975).

Calystegins are derivatives of tropane that can be utilised by *S. meliloti* as the sole carbon and nitrogen source, although they are only released by non-legumes (Tepfer *et al*., 1988). Catabolic genes (*cac*) are encoded on a non-symbiotic plasmid and when inoculated alone, both *cac*+ and *cac*− strains of *S. meliloti* were able to colonise the rhizosphere of calystegin positive and negative plants. When co-inoculated in the presence of calystegin-producing plants, the *cac*+ strain reached much higher population levels. There was no advantage in the rhizosphere of calystegin-negative plants (Guntli *et al*., 1988).
However, this study was carried out with strains cured of the plasmid containing the \textit{cac} genes and so it is possible that there were other factors on the plasmid that were responsible for the competitive advantage.

An opine producing strain of \textit{P. putida} was more competitive in the rhizosphere of mannopine producing tobacco plants (Wilson \textit{et al.}, 1995). Similarly, Gordon \textit{et al.} (1996) showed that the rhizopine-producing strain \textit{S. meliloti} L5-30 had a competitive advantage for nodulation of lucerne (\textit{Medicago sativa}) in soil when co-inoculated with a mutant strain, even though when inoculated alone, the mutant had a similar rate of growth and nodulation to the wild type. The mutant occupied less than 30% of nodules. This competitive advantage remained four years after inoculation, even though there had been turnover of nodules in that time (Heinrich \textit{et al.}, 1999). In Chapter 4 of this work, evidence is presented that the ability to utilise \textit{myo}-inositol as a carbon source is also important for competition for nodulation.

In contrast, Bosworth \textit{et al.} (1994) and Scupham \textit{et al.} (1996) found that disruption of the ability to utilise \textit{myo}-inositol was beneficial to host plants, although the mechanism for this is unknown. Bosworth \textit{et al.} (1994) created recombinant strains of \textit{S. meliloti} by insertion of combinations of an interposon, a modified \textit{nifA} regulatory gene and the \textit{dctABD} \textit{C}_{\text{4}}-dicarboxylic acid transport genes into an inositol locus. The strains could no longer grow when \textit{myo}-inositol was the sole carbon source. The mutations were in \textit{myo}-inositol dehydrogenase (Galbraith \textit{et al.}, 1998). Interruption of the inositol site by the interposon alone or by the interposon plus \textit{nifA/dctABD} sometimes
caused increased grain yields of alfalfa in field trials measured over four years but this was dependent on conditions in the fields. There was either no increase or a decrease in yield when the inositol locus was interrupted by an interposon plus \textit{nifA}, indicating that the inositol locus might have a role in nodulation competition (Bosworth et al., 1994, Scupham et al., 1996).

Catabolism of the betaine stachydrine (aka N,N-dimethylproline or proline betaine) may contribute to root colonisation. Stachydrine is a quaternary amine found in \textit{Medicago} species that is a catabolite, an osmoprotectant in times of osmotic stress (when catabolism is reduced) and an inducer of \textit{nod} genes. Although stachydrine catabolic genes are found on the Sym plasmid, they are not under the NifA/NtrA regulatory system and are therefore not essential for nitrogen fixation. Stachydrine mutants took three to four days longer to form nodules than the wild type on alfalfa, although the mutants colonised roots and formed nodules as effectively as the wild type when inoculated alone (Goldmann et al., 1994).

Two other \textit{S. meliloti} mutants impaired in the ability to utilise stachydrine as the sole carbon and nitrogen source were able to colonise roots as effectively as the wild type when inoculated separately. However, when co-inoculated with the wild type, the mutants were at a serious competitive disadvantage. Root colonisation was severely reduced and the final population count of the mutant was half that of the wild type (Phillips et al., 1998). One of the mutations was in the gene encoding proline dehydrogenase (\textit{putA}). Stachydrine is catabolised via proline, which is why \textit{putA} mutants cannot
utilise this compound. Trigonelline and stachydrine released from alfalfa exudates induce expression of NodD2 protein in S. meliloti (Phillips et al. 1992).

Jimenez-Surdo, et al., (1995) previously reported that a putA mutant of S. meliloti was unable to utilise ornithine or proline as the sole carbon source. The mutant was impaired in nodulation efficiency and competitiveness on alfalfa roots, although nitrogen fixation occurred as normal. These results were confirmed by Jimenez-Surdo, et al. (1997) who found that putA gene expression was induced by root exudates from alfalfa, during root invasion and nodule formation, but expression was not induced in differentiated bacteroids. Therefore, secondary metabolites do not seem to be used by bacteroids in nodules as carbon sources. However, the competitive disadvantage of the mutant may be related to the role of proline as an osmoprotectant (Straub et al., 1994), or stachydrine as an osmoprotectant or inducer of nod gene expression, rather than simply because it cannot be utilised as a carbon source.

1.4.3 Improving Nodulation Competitiveness

A potential method for improving nodulation competitiveness involves utilisation of antirhizobials such as trifolitoxin, produced by R. leguminosarum bv. trifolii T24. The tfx genes, which encode the production and resistance of trifolitoxin are present in several rhizosphere bacteria. When the tfx genes were inserted into R. leguminosarum bv. trifolii and co-inoculated onto clover plants
with the wild type, the trifolitoxin-producing strain occupied over 91% of nodules. The non-producing wild type occupied 41% (the numbers add up to more than 100% because many nodules contained both strains). When inoculated alone, both strains formed the same number of nodules and the rate of nitrogen fixation was the same (Triplett, 1990).

A strain of *R. etli* which had the *tfx* genes introduced into it was also more competitive than a non-producing strain in non-sterile soil (Robleto *et al.*, 1998), despite the fact that trifolitoxin is rapidly biodegraded in such conditions (Bosworth *et al.*, 1993). The trifolitoxin-producing strain occupied at least 20% more nodules, even when it constituted only 5% of the inoculum. There was no difference in the grain yield of common bean, indicating that there were no adverse effects on plant productivity (Robleto *et al.*, 1998). Similarly, Goel *et al.* (1999) found that a bacteriocin producing strain *Rhizobium* VF10 occupied significantly more nodules on green gram (*Vigna radiata*) than a bacteriocin sensitive strain when co-inoculated.

Clearly, there are many genes that are important for successful competition for rhizosphere colonisation and nodulation to occur, many of which have yet to be identified. It is extremely interesting that knocking out just one gene can cause loss of competitive ability, particularly when the function of the gene is unknown. The work of Scupham *et al.* (1996) highlights the fact that when engineering strains to improve competitive ability, it is important that sites of gene insertion are not themselves important for competitive ability.
1.5 Research Objectives

myo-Inositol and its derivatives are ubiquitous throughout nature and several micro-organisms are able to utilise myo-inositol as the sole carbon source. Therefore, myo-inositol may be important in the growth and survival of Rhizobium, either directly as a carbon source, in combination with other compounds such as phosphorus (inositol phosphates) and nitrogen (rhizopines), or through another as yet unidentified function.

The objectives of this project are as follows.

1. Identify and characterise the genes involved in myo-inositol utilisation in R. leguminosarum bv. viciae. Ascertain whether the genes are arranged in a regulon, an operon, or are located separately throughout the genome or plasmids. Identify promoters that are induced by myo-inositol to discover additional genes that are regulated by myo-inositol. Determine whether there is a specific transport system for myo-inositol.

2. Investigate catabolic regulation in free-living rhizobia and in bacteroids. Determine if myo-inositol is the preferred carbon source and whether myo-inositol inhibits catabolism of other sugars or vice versa. Investigate whether regulation of catabolism of myo-inositol is the same in free-living rhizobia and in bacteroids.
3. Ascertain the importance of the ability to utilise myo-inositol in the rhizosphere by studying growth in the rhizosphere of the mutants compared with the wild type. Compare the relative growth rates of the strains on different carbon sources in vitro, to determine whether the mutations are specific to myo-inositol catabolic ability.

4. Nodulation ability and competition for nodulation. Discover whether the myo-inositol mutants nodulate plants and if they do so at the same rate as the wild type. Compare the rate at which the mutants and wild type fix nitrogen. Determine whether the plants benefit from inoculation with the mutants to the same extent as when inoculated with the wild type. Investigate the mutants’ competitive ability to occupy nodules when co-inoculated with the wild type.
Chapter 2 – Materials and Methods
### 2.1 Bacterial Strains and Plasmids

<table>
<thead>
<tr>
<th>bacterial strain, relevant characteristics</th>
<th>source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>plasmid or bacteriophage</td>
<td></td>
</tr>
</tbody>
</table>

**R. leguminosarum bv. viciae**

| 3841 | Streptomycin and trimethoprim resistant derivative of strain 300 bv. viciae | Johnston and Beringer (1975) |
| RU360 | Strain 3841::Tn5-lacZ unable to grow when *myo*-inositol is the sole carbon source. | Poole *et al.* (1994) |
| RU361 | Strain 3841::Tn5-lacZ unable to grow when *myo*-inositol is the sole carbon source. | Poole *et al.* (1994) |
| RU307 | Strain 3841::Tn5-lacZ impaired in growth when *myo*-inositol is the sole carbon source. | P.S. Poole |

**S. meliloti**

| RMB7101 | Strain RCR2011::Ω unable to grow when *myo*-inositol is the sole carbon source. Streptomycin and spectinomycin resistant. | E.W. Triplett |

**E. coli**

| DH5α | Strain made competent for transformation, supE44, ΔlacU169 (φ80 lacZΔM15), hsdR17, recA1, endA1, gyrA96, thi-1, relA1. | Hanahan (1983) |
| XL1- and XL2-Blue | Super and ultracompetent Epicurian coli® cells for transformation, recA1, endA1. | Stratagene Ltd |

**bacteriophage**

| RL38 | General transducing phage of *R. leguminosarum* | Buchanan-Wollaston (1979) |

**plasmid**

<p>| Vector Name | Description | Source \n|-------------|-------------|------------------|
| pBluescript Phagemid; f1(^\d) origin of replication; ColE1 replicon; SK polylinker; 2.96kb; standard cloning vector. Ampicillin resistant. | Stratagene Ltd |
| pCR2.1 TOPO | TA vector for PCR product cloning. Kanamycin and ampicillin resistant. | Invitrogen |
| pRU426 | 12.4kb EcoRI fragment containing Tn5 from RU307 in pACYC184. | P.S. Poole |
| pRU3078, pLAFR1 (Friedman et al., 1982) cosmid in E. coli strain 803 containing strain 3841 genomic DNA that complements RU360. | Poole et al. (1994) |
| pRU3111, pLAFR1 (Friedman et al., 1982) cosmid in E. coli strain 803 containing strain 3841 genomic DNA that complements RU361. | this project |
| pRU438 | 11.5kb SalI fragment containing Tn5-lacZ from RU361 in pBluescript SK(^\d). | this project |
| pRU439 | 0.24kb BamHI fragment containing part of IS50L of Tn5-lacZ from pRU438 in pBluescript SK(^\d). | this project |
| pRU445 | 3.2kb NotI fragment from pRU438 in pBluescript SK(^\d). | this project |
| pRU472 | 9kb SstI fragment containing 6kb of Tn5-lacZ from RU360 in pBluescript SK(^\d). | this project |
| pRU476 | 1.2kb NotI fragment containing IS50R of Tn5-lacZ from pRU438 in pBluescript SK(^\d). | this project |
| pRU482 | Adjacent 7.5kb + 2.5kb PstI fragments from pRU3111 in pBluescript SK(^\d). | this project |
| pRU542 | 7.5kb SalI fragment from pRU3111 in pBluescript SK(^\d). | this project |</p>
<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>pRU544</td>
<td>0.2kb <em>EcoRI</em> fragment from pRU3111 in pBluescript SK&lt;sup&gt;-&lt;/sup&gt;.</td>
<td>this project</td>
</tr>
<tr>
<td>pRU555</td>
<td>6kb <em>SalI</em> fragment from pRU482 in pBluescript SK&lt;sup&gt;-&lt;/sup&gt;.</td>
<td>this project</td>
</tr>
<tr>
<td>pRU556</td>
<td>7.5kb <em>PstI</em> fragment, different to pRU482 from pRU3111 in pBluescript SK&lt;sup&gt;-&lt;/sup&gt;.</td>
<td>this project</td>
</tr>
<tr>
<td>pRU560</td>
<td>5kb <em>NotI</em> fragment from pRU482 in pBluescript SK&lt;sup&gt;-&lt;/sup&gt;.</td>
<td>this project</td>
</tr>
<tr>
<td>pRU683</td>
<td>0.7kb <em>PstI</em>-Spe&lt;sup&gt;I&lt;/sup&gt; PCR product from pRU3078 in pCR2.1 TOPO.</td>
<td>this project</td>
</tr>
<tr>
<td>pRU702</td>
<td>2.5kb <em>PstI</em>-Spe&lt;sup&gt;I&lt;/sup&gt; PCR product from pRU3078 in pCR2.1 TOPO.</td>
<td>this project</td>
</tr>
<tr>
<td>pRU703</td>
<td>0.7kb <em>PstI</em>-Spe&lt;sup&gt;I&lt;/sup&gt; fragment from pRU683 in pOT1.</td>
<td>this project</td>
</tr>
<tr>
<td>pRU705</td>
<td>2.5kb <em>XbaI</em>-Spe&lt;sup&gt;I&lt;/sup&gt; fragment from pRU702 in pOT1.</td>
<td>this project</td>
</tr>
<tr>
<td>pRU706</td>
<td>As pRU705, but opposite orientation.</td>
<td>this project</td>
</tr>
<tr>
<td>pRU707</td>
<td>4kb <em>Pmel</em>-Spe&lt;sup&gt;I&lt;/sup&gt; PCR product from pRU3078 in pCR2.1 TOPO.</td>
<td>this project</td>
</tr>
<tr>
<td>pRU713</td>
<td>4kb <em>Pmel</em>-Spe&lt;sup&gt;I&lt;/sup&gt; fragment from pRU707 in pOT1.</td>
<td>this project</td>
</tr>
</tbody>
</table>
2.2 Culture Conditions

Rhizobia were grown at 26-28°C either on tryptone yeast medium (TY) (Beringer, 1974) or on acid minimal salts (AMS) medium derived from that of Brown and Dilworth (1975). The changes to AMS were; phosphate (0.5mM), MgSO₄ (2mM) and buffering by MOPS (20mM) pH 7.0. All carbon and nitrogen sources added to AMS were used at 10mM, unless otherwise stated. Antibiotic concentrations (in µg ml⁻¹) were gentamycin 20, kanamycin 40, spectinomycin 100, streptomycin 500, tetracycline 2 (in AMS) and 5 (in TY). Liquid cultures were incubated in a rotary shaker at 200-250 rpm.

All *Escherichia coli* strains DH5α, 803, XL1-blue and XL2-blue were grown in Luria-Bertani (LB) liquid or solid medium at 37°C (Sambrook *et al.*, 1989). Antibiotic concentrations (in µg ml⁻¹) were ampicillin 50, gentamycin 10, kanamycin 20 and tetracycline 10.

2.2.1 Calculation of Mean Generation Time

Growth curves were plotted using the logarithm of cell density at OD₆₀₀nm. The mean generation time (g) was calculated using the equation \( g = \frac{0.693}{k} \). The growth rate constant (k) is equal to the slope of the graph x 2.303 (Stanier *et al.*, 1988).
2.3 DNA and Genetic Manipulations

Plasmid and cosmid DNA isolation was performed using the Flexiprep method (Pharmacia) according to the manufacturer's instructions. *Rhizobium* chromosomal DNA was isolated using DNA Isolator (Genosys) according to the manufacturer's instructions. Extracted DNA was dissolved in TE buffer (Sambrook *et al.*, 1989). Restriction enzymes digests, ligations and sub-clonings were performed according to Sambrook *et al.*, (1989). Transductions were performed with the bacteriophage RL38 as described by Buchanan-Wollaston (1979). Transductants were selected for on TY agar containing kanamycin (80µg ml⁻¹).

DNA was analysed by gel electrophoresis in 0.8% agarose with TAE buffer (Sambrook *et al.*, 1989). Visualisation of DNA was by staining of gels with ethidium bromide at a concentration of 10µg ml⁻¹ (10mg ml⁻¹ stock). Loading buffer allowed the movement of the DNA to be observed on the gel (Sambrook *et al.*, 1989). DNA fragments cut directly out of agarose gels for subcloning of plasmids were extracted using Qiaex II (Qiagen) according to the manufacturer’s instructions.

Southern analysis was carried out on DNA digested with different restriction enzymes and separated on 0.8% agarose gels. DNA was transferred to a positively charged, Hybond N+, nylon membrane (Amersham). The DNA was hybridised with a fluorescent labelled probe and the signal detected with
the CDP-star detection kit (Amersham) according to the manufacturer's instructions.

Triparental matings were carried out using the transfer functions of the helper plasmid pK2013 (Figurski and Helinski, 1979) in *E. coli* strain 803. Transformations were carried out using heat shock at 42°C, according to Sambrook *et al.*, (1989) with *E. coli* strain DH5α, or with Stratagene *E. coli* strains XL1-blue and XL2-blue according to the manufacturer's instructions.

### 2.3.1 PCR Amplification

Polymerase chain reaction (PCR) amplification of cosmid pRU3078 DNA was carried out using the primers detailed in table 2.1. All the primers were synthesised by Genosys. There was an initial denaturation for 5 minutes at 94°C, 1.5 minutes at 58°C, then 30 cycles of 1.5 minutes at 72°C and a final 10 minutes at 72°C. This final stage ensured an overhang A was added, for cloning into pCR2.1 TOPO (Invitrogen) according to the manufacturer's instructions. The reaction was carried out in a volume of 50µl, which contained 5µl 10x polymerase buffer, 1.5mM MgCl₂, 0.8mM dNTPs, 50pmol primers, 1% DMSO and 3 units of Bio-X-act DNA polymerase (Bioline).
2.3.2 DNA Sequencing

Most DNA sequencing was carried out with a Pharmacia ALF automated DNA sequencer. The enzyme used was ThermoSequenase. Universal M13 forward and reverse primers were used for sequencing of DNA in the plasmid Bluescript SK\(^+\), LacZB20 or IS50R primers were used for sequencing DNA adjacent to Tn5-lacZ. Custom Cy5-labelled primers were made by Pharmacia (Table 2.1). Other sequencing was carried out by MWG-Biotech Ltd, using a Li-Cor machine. Primers were designed by MWG-Biotech Ltd from a 100-200bp region supplied to the company. The company did not give details of the actual primers used.
Table 2.1  PCR Amplification and Custom DNA Sequencing Primers

<table>
<thead>
<tr>
<th>primer</th>
<th>Sequence</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>p164</td>
<td>ggaagcctatacccggttcc</td>
<td>Downstream of iolA, 5’-3’ direction.</td>
</tr>
<tr>
<td>p173</td>
<td>ccttgagatttgccattgttg</td>
<td>Upstream of iolB, 3’-5’ direction.</td>
</tr>
<tr>
<td>p174</td>
<td>gcacaagctctggccggc</td>
<td>Downstream of iolD, 5’-3’ direction.</td>
</tr>
<tr>
<td>p175</td>
<td>gctcgacatctcttatcccc</td>
<td>Sequencing further downstream from p164, 5’-3’ direction.</td>
</tr>
<tr>
<td>p176</td>
<td>gccgaccaggagacatacg</td>
<td>Upstream of iolA, 3’-5’ direction.</td>
</tr>
<tr>
<td>p182</td>
<td>gggcggtctgtttcctcg</td>
<td>Sequencing further upstream from p176, 3’-5’ direction.</td>
</tr>
<tr>
<td>p184</td>
<td>gcgcaagctcaacatctcg</td>
<td>Sequencing further downstream from p174, 5’-3’ direction.</td>
</tr>
<tr>
<td>p213</td>
<td>ttttttttctgcagcgagctgatttcctgcttcg</td>
<td>Upstream of iolD, 5’-3’ direction with PstI added.</td>
</tr>
<tr>
<td>p214</td>
<td>aaaaaaaaaactagtcaggaagagaaacgggata</td>
<td>In iolD, 3’-5’ direction with SpeI added.</td>
</tr>
<tr>
<td>p228</td>
<td>tttttttactagtgcggaggagcgggtgcgg</td>
<td>At the end of iolD, 3’-5’ direction with SpeI added.</td>
</tr>
<tr>
<td>p216</td>
<td>aaaaaaaaaactagttttgcgctgtcagatttt</td>
<td>At the end of iolB, 3’-5’ direction with SpeI added.</td>
</tr>
<tr>
<td>p230</td>
<td>ttttttttttaacccgagcttctctgcttcg</td>
<td>Upstream of iolD, 5’-3’ direction with Pmel added.</td>
</tr>
</tbody>
</table>
2.4 Analysis of Sequence Data

DNA sequence data was analysed using the Gapped BlastX algorithm with default parameters to search the SwissProt and Genbank/EMBL sequence databases. This was carried out via the National Centre for Biotechnology Information (NCBI) molecular biology server home page (www.ncbi.nlm.nih.gov). BlastX translated the DNA sequence into all six possible reading frames and a list was compiled of proteins most closely related to the query sequence. The statistical significance of the alignment score for each pairwise comparison was evaluated by comparing it to the mean score obtained from comparison of each sequence to random permutations of the other (Altschul et al., 1997). Other sequence analysis was carried out using software accessed via the EXPASY molecular biology server (www.expasy.ch) or using Genetics Computer Group (gcg) software, through the Human Genome Mapping Project (HGMP) website (www.hgmp.mrc.ac.uk).

2.5 Statistical Analysis of Data

Data were subjected to Analysis of Variance using Genstat 3.2 software.
2.6 Transport Assay

For *R. leguminosarum* bv. *viciae* strains, cells were prepared and transport assays performed as previously described (Poole *et al.*, 1985). The total substrate concentration was 25µM. Samples of 0.1ml were taken at intervals of 1 minute for up to 5 minutes, Millipore filtered and counted. For competition assays, a 5-fold excess (125µM) of a non-labelled substrate was added 5 seconds prior to addition of the substrate to be assayed. The specific activities of labelled substrates in the assays were; D-[U-14C]glucose (310 MBq mmol⁻¹) and D-[U-14C]myo-inositol (310 MBq mmol⁻¹).

2.7 Enzyme Assays

2.7.1 myo-Inositol Enzyme Assay

Supernatant from 400ml cultures of *R. leguminosarum* bv. *viciae* were obtained by centrifugation at 4000rpm for 30 minutes, then washed in 40mM N-2-hydroxyethylpiperazine-N'-ethanesulfonic acid (HEPES). pH7.0 and re-centrifuged (Poole *et al.*, 1994). Cells were then resuspended in 10ml 40mM HEPES, pH7.0, containing 2mM dithriothreitol (DTT). Bacteroids of 3841 were isolated under argon within an AtmosBag (Aldrich) with argon-purged
isolation buffer (Bergersen and Turner, 1990) and by using a Percoll density
gradient method (Reibach et al., 1981) by D. Allaway in the laboratory.

myo-Inositol dehydrogenase and 2-keto-myoinositol dehydratase were
assayed by procedures modified from Berman and Magasanik (1966). The
myo-inositol dehydrogenase assay contained in 1ml; NH$_4$Cl, 50mM; NAD$^+$,
0.4mM; sodium carbonate, pH10, 50mM; and myo-inositol, 100mM. The 2-
keto-myoinositol dehydratase assay contained in 1ml; Tris-HCl, pH8.5,
50mM; and 2-keto-myoinositol, 1mM.

The myo-inositol dehydrogenase assay was followed at 340nm. Activity was
calculated using the molar extinction coefficient of NADH at 340nm, which is
6220. The 2-keto-myoinositol dehydratase assay was followed at 260nm.
Activity was calculated using the molar extinction coefficient at 260nm, which
is 6000. The protein concentration was determined according to the method
of Lowry et al. (1951) using bovine serum albumin as standard.

2.7.2 β-Galactosidase Assay

β-galactosidase assays were conducted according to the method of Miller
(1972) except that chloroform permeabilisation was replaced by 5 minutes
incubation of cells in 0.5mg ml$^{-1}$ lysozyme, followed by 15 minutes incubation
in 1mM EDTA. Cells were lysed by addition of 0.001% sodium dodecyl
sulfate (SDS) (Poole et al., 1994).
2.7.3 Alkaline Phosphatase Assays

Cultures of *R. leguminosarum* bv. *viciae* containing a *nodC-phoA* reporter plasmid pIJ1687 (Economou, 1990) were grown in 10ml AMS with appropriate carbon sources, in the presence of hesperetin 1µM to induce expression of the *nod* gene. Controls were grown in the absence of hesperetin. Cells were harvested and the assay carried out according to the method of Brickman and Beckwith (1975). Cells were assayed for alkaline phosphatase activity with *p*-nitrophenylphosphate (Sigma). The reaction was stopped with 1M K$_2$HPO$_4$. The samples were cleared by centrifugation for two minutes and the optical density read at 420nm.

2.8 GFP-UV Assay

GFP-UV has an excitation maximum of 395nm and an emission maximum of 509nm (Clontech). GFP-UV expression was measured in cultures grown overnight in AMS plus appropriate nutritional factors. 200µl was aliquoted into a 96-well microtitre plate (Iwaki) and the assay performed using a Biolumin960 plate reader (Molecular Dynamics). Cell density was measured at OD$_{630\text{nm}}$ and fluorescence of GFP-UV at an excitation of 405/10nm and emission of 505/10nm. Specific fluorescence was calculated using the equation A-X/B-Y. A is the fluorescence of the sample, X is fluorescence of the blank, B is the OD$_{630\text{nm}}$ of the sample and Y is the OD$_{630\text{nm}}$ of the blank. Colonies grown on AMA plus appropriate nutritional factors were observed on
an UV transilluminator (UVP model TL33E), fitted with 420nm bulbs and a long wave emission filter.

Bacteria harvested from the rhizosphere and bacteroids in nodules were assessed for GFP-UV expression under a Nikon Optiphot epifluorescence microscope, magnification 1000x using a GFP filter set 11003 (Chroma Technology). The filter consisted of a band pass exciter of 425nm with a 40nm width, a long pass emitter of 475nm and a dichroic long pass of 460nm (Figure 2.1). The filter set is non-specific which means that only material expressing GFP-UV will appear green and all auto-fluorescent material will appear yellow.

Figure 2.1  GFP-UV Filter Set
2.9 Plant Assays

2.9.1 Seed Sterilisation

Seeds of *Pisum sativum* cv. Feltham First (common pea) (Sutton seeds) and
*Vicia sativa* (common vetch) (Chiltern seeds) were surface sterilised by
immersion in 100% ethanol for 5 minutes, followed by immersion in 3%
sodium hypochlorite for 10 minutes, rinsing after each treatment with sterile
distilled water three times. Sterilised seeds were germinated on wetted sterile
filter paper in petri dishes for up to seven days in the dark at 26°C (M. Wood,
pers. comm.). The germination rate was approximately 90%. Seeds
contaminated by fungi after surface sterilisation numbered fewer than 1%,
whereas more than 50% of unsterilised seeds were contaminated after
germination.

2.9.2 Nodulation Competition

*P. sativum* seedlings were added to sterile 250ml glass conical flasks
containing 250ml sterile vermiculite. *V. sativa* seedlings were added to glass
boiling tubes containing 50ml sterile vermiculite or 0.2% water agar containing
nitrogen-free rooting solution. Only seedlings with an established root and
shoot that were not visibly contaminated were planted. The vermiculite was
then wetted with sterile nitrogen-free rooting solution. This contained 1mM
CaCl$_2$.H$_2$O, 100µM KCl, 800µM MgSO$_4$.7H$_2$O, 10µM FeEDTA, 35µM H$_3$BO$_3$, 9µM MnCl$_2$.4H$_2$O, 0.8µM ZnCl$_2$, 0.5µM NaMoO$_4$.2H$_2$O, 0.3µM CuSO$_4$.5H$_2$O. 7.2mM KH$_2$PO$_4$ and 7.2mM Na$_2$HPO$_4$ were autoclaved separately and added just prior to use, to prevent precipitation. There was no exogenous carbon or nitrogen in the system (Poole et al., 1994).

At three to five days after planting, the seeds were inoculated with 1ml of a $10^3$, $10^4$, $10^5$, or $10^6$ cfu ml$^{-1}$ bacterial culture of 3841, RU360, RU361 or RU307. The bacterial number was calculated by cell density at OD$_{600nm}$. An OD$_{600}$ of 1 was taken to be $10^9$ cfu ml$^{-1}$. The cultures were washed twice in AMS to ensure removal of all nutritional compounds and antibiotics. When a mutant and wildtype were co-inoculated, the mutant was inoculated first to avoid absorption effects. Controls were inoculated with sterile AMS. Aliquots of the bacterial cultures were plated onto TY agar with appropriate antibiotics to confirm that an OD$_{600}$ of 1 was actually $10^9$ cfu ml$^{-1}$.

*P. sativum* plants were incubated in a growth cabinet with a 16 hour photoperiod at 25°C. The temperature was lowered to 16°C for the dark period. Relative humidity was 94%. *V. sativa* seeds were incubated at 22°C in a growth room illuminated by a Philips Sont-Agro grow light, with a 16 hour photoperiod. Plants were harvested four to six weeks post-inoculation.
2.9.3 Nodule Harvesting

Nodules were removed from roots and surfaced sterilised by immersion in sodium hypochlorite for 10 minutes, followed by three rinses in sterile distilled water (Poole et al., 1994). Nodules were crushed and plated onto TY agar, with appropriate combinations of streptomycin, kanamycin and tetracycline to distinguish between the wildtype and mutant strains. Sample colonies were subsequently streaked onto AMA with appropriate nutritional factors, to confirm the phenotype of the colonies.

2.9.4 Acetylene Reduction

Acetylene reduction was carried out on whole P. sativum plants four weeks post-inoculation according to the method described by Trinick et al., (1976).

2.9.5 Plant Dry Weight

P. sativum and V. sativa plants with the root systems removed were dried for 48 hours in an oven at 55°C. The plants were then weighed.
2.9.6 Nodule Mass

All nodules were removed from each *P. sativum* plant and weighed to give total wet nodule mass.

2.9.7 Rhizosphere Growth

*P. sativum* and *V. sativa* seedlings were planted in 20ml glass Universal bottles containing sterile vermiculite and 12ml nitrogen-free rooting solution (c.f. section 2.9.2). The seeds were inoculated on the same day as planting with 1ml of a $10^3$ or $10^6$ cfu ml$^{-1}$ bacterial culture of 3841, RU360, RU361 or RU307. Harvesting occurred every one to two days post-inoculation, for up to two weeks. The contents of the Universal, minus the plant shoot, were macerated with a pestle and mortar with 10ml nitrogen-free rooting solution added. The mixture was filtered through sterile muslin and centrifuged for 2 minutes at 4000rpm, to remove debris. The supernatant was then plated onto TY agar with appropriate antibiotics to determine colony numbers for each strain (Schofield, 1999).

CHAPTER 3 - MYO-INSITOL CATABOLIC MUTANTS OF *RHIZOBIUM LEGUMINOSARUM* BIOVAR *VICIAE* ........................................................... 70

3.1 Introduction ............................................................................................ 71

3.2.1 Growth Characteristics of the *myo*-Inositol Mutants ....................... 72

3.2.2 Cosmids Carrying *myo*-Inositol Catabolic Genes ............................ 76
3.2.3 Southern Hybridisation................................................................. 79
3.2.4 Transduction of RU307 ............................................................... 85
3.2.5 Sequencing of myo-Inositol Catabolic Genes ............................. 86
  3.2.5.1 RU360 ...................................................................................... 86
  3.2.5.1.1 Complementation of the iol region of RU360 ..................... 102
  3.2.5.2 RU361 ...................................................................................... 104
  3.2.3.3 RU307 ...................................................................................... 119
3.2.5 Uptake of Glucose and myo-Inositol .......................................... 125

3.3 Discussion...................................................................................... 133
Chapter 3 - *myo*-Inositol Catabolic Mutants of

*Rhizobium leguminosarum* biovar *viciae*
3.1 Introduction

In order to identify the genes involved in myo-inositol utilisation by *R. leguminosarum* bv. *viciae*, three mutant strains of *R. leguminosarum* bv. *viciae* strain 3841 were characterised. Strain 3841 is capable of growth when myo-inositol is the sole carbon source. Following introduction of a plasmid containing rhizopine degradation genes, 3841 grew on rhizopines as the sole carbon and nitrogen source (Bahar et al., 1998).

Strains RU360 and RU361 are Tn5-lacZ mutants that cannot grow on myo-inositol as the sole carbon source (Poole et al., 1994). They also cannot grow when rhizopines are the sole carbon source following introduction of a plasmid containing rhizopine degradation genes (Bahar et al., 1998). Strain RU360 was previously shown to be unable to induce expression of the first two enzymes in the proposed myo-inositol degradation pathway. However, the strain retains the ability to nodulate *V. sativa* and reduces acetylene at the same rate as the wildtype (Poole et al., 1994).

Strain RU307 is a Tn5 mutant that was isolated by its ability to grow in the presence of glutamic acid gamma hydrazide, a compound toxic to the wild type strain 3841 (P.S. Poole, pers. comm.). During this project, this strain was found to be severely impaired in the ability to utilise myo-inositol as the sole carbon source. Sequencing data indicated that RU307 might be mutated in an uptake system for myo-inositol, so the ability of the mutants and RU307 to transport myo-inositol into the cell was studied.
3.2 Results

3.2.1 Growth Characteristics of the myo-Inositol Mutants

The mutants RU360 and RU361 were identified by their inability to grow when myo-inositol was the sole carbon source (Poole et al., 1994). Strain RU307 was identified as being severely impaired in the ability to grow when myo-inositol was the sole carbon source. In order to characterise the catabolic pathway for myo-inositol in R. leguminosarum bv. viciae, the mutants and 3841 were tested for the ability to grow in AMS on different carbon sources. These were 10mM myo-inositol, 10mM glucose, 10mM glucose plus 10mM myo-inositol, 20mM pyruvate, 20mM pyruvate plus 10mM myo-inositol.

Growth of the strains on each carbon source was measured at OD600nm. The mean generation times for each strain in each carbon source were calculated over a period of at least six hours during linear growth. The mean generation times were similar for the mutants and 3841 in all combinations of carbon sources tested other than myo-inositol (Table 3.1). Strains RU360 and RU361 could not grow when myo-inositol was the sole carbon source. Strain RU307 grew on myo-inositol, but the mean generation time was almost four times higher than that for 3841 grown on myo-inositol. Growth by all strains was slightly slower on pyruvate than on glucose. The mean generation times were also similar when the strains were grown in glucose or pyruvate plus myo-inositol. This indicates that the presence of myo-inositol in the medium did not inhibit growth on other carbon compounds.
Two cosmids (pRU3078 and pRU3079) were previously isolated that complemented RU360, restoring the ability to utilise *myo*-inositol (*Poole et al.*, 1994). A cosmid (pRU3111) was isolated during this project that complemented RU361 (c.f. Section 3.2.2). These complemented strains, RU360/pRU3078, RU360/pRU3079 and RU361/pRU3111 were also tested for growth on 10mM glucose, 10mM glucose plus 10mM *myo*-inositol and 10mM *myo*-inositol (Table 3.1).

The mean generation times for the complemented strains were similar to the mutants and 3841 on glucose and glucose plus *myo*-inositol. When grown on *myo*-inositol as the sole carbon source, the cosmids restored the ability to grow as well as 3841.
## Table 3.1  Mean Generation Times for the Mutants and 3841

<table>
<thead>
<tr>
<th>strain</th>
<th>myo-inositol (hrs)</th>
<th>glucose (hrs)</th>
<th>glucose + myo-inositol (hrs)</th>
<th>pyruvate (hrs)</th>
<th>pyruvate + myo-inositol (hrs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3841</td>
<td>3.41</td>
<td>3.52</td>
<td>3.57</td>
<td>4.39</td>
<td>4.12</td>
</tr>
<tr>
<td>RU307</td>
<td>13.36</td>
<td>3.67</td>
<td>3.85</td>
<td>4.05</td>
<td>4.1</td>
</tr>
<tr>
<td>RU360</td>
<td>*</td>
<td>3.41</td>
<td>3.37</td>
<td>4.45</td>
<td>3.91</td>
</tr>
<tr>
<td>RU361</td>
<td>*</td>
<td>3.58</td>
<td>3.6</td>
<td>4.16</td>
<td>3.93</td>
</tr>
<tr>
<td>RU360/pRU3078</td>
<td>3.78</td>
<td>4.15</td>
<td>3.79</td>
<td>N/T</td>
<td>N/T</td>
</tr>
<tr>
<td>RU360/pRU3079</td>
<td>3.4</td>
<td>3.85</td>
<td>3.56</td>
<td>N/T</td>
<td>N/T</td>
</tr>
<tr>
<td>RU361/pRU3111</td>
<td>3.28</td>
<td>3.58</td>
<td>3.54</td>
<td>N/T</td>
<td>N/T</td>
</tr>
</tbody>
</table>

Each value represents the average of three cultures, which was then used to determine the mean generation time.

* = too low to be measured.

N/T = not tested.
Strains RU360, RU361 and RU307 were also able to grow on AMS and AMA containing 10mM of the sugar polyols mannitol, sorbitol and glycerol, with no apparent deficiency compared with 3841. The mutant strains and 3841 were also tested for their ability to utilise myo-inositol hexaphosphate (phytate) as the sole carbon source. Phytate is abundant in the rhizosphere and some micro-organisms produce enzymes that can cleave phytic acid, to release the phosphorus stored in the compound (c.f. Section 1.6.3). It was postulated that Rhizobium might be able to utilise the myo-inositol in phytate for growth, especially if phosphorus were limiting. However, there was no growth by any of the strains in the presence or absence of phosphate when phytate was the sole carbon source. It was not tested whether phytate could be used as a phosphorus source in the presence of other carbon sources. Therefore, phytate is unlikely to be a source of energy in the rhizosphere for R. leguminosarum bv. viciae.

The data presented above indicate that there is a specific pathway for myo-inositol utilisation in R. leguminosarum bv. viciae, mutation of which does not affect growth on other carbon sources. There was no detectable effect on growth on pyruvate or glucose caused by the presence of the transposon insertion or by the presence of complementing cosmids. The presence of myo-inositol did not inhibit growth of the mutants on pyruvate or glucose.
3.2.2 Cosmids Carrying myo-Inositol Catabolic Genes

In order to characterise the region surrounding the myo-inositol mutations and to try to identify other genes involved in myo-inositol utilisation, a *R. leguminosarum* bv. *viciae* cosmid library was used to isolate cosmids that could complement the mutants, restoring the ability to utilise myo-inositol. The cosmids pRU3078 and pRU3079 were previously isolated from this library for their ability to restore to RU360 the ability to utilise myo-inositol (Poole *et al.*, 1994).

The cosmid library was conjugated into RU361 and colonies spread on AMA containing 10mM myo-inositol as the sole carbon source. Two colonies were obtained that grew on myo-inositol and were tetracycline resistant, indicating the presence of a cosmid. Restriction enzyme analysis showed that both colonies contained the same cosmid, which was designated pRU3111. Attempts to isolate a cosmid that complemented RU307 for the ability to utilise myo-inositol were unsuccessful. The same library was also used to try to complement a *S. meliloti* strain RMB7101, a myo-inositol dehydrogenase (*idh*) mutant. Again, this was unsuccessful.

Each cosmid was conjugated into each of the mutants to determine whether they could cross-complement. The cosmids pRU3078 and pRU3079 were unable to complement RU361 and pRU3111 was unable to complement RU360. None of the cosmids complemented RU307. The profiles of the
three cosmids digested with five restriction enzymes, *EcoRI, NotI, PstI, SalI* and *SstI* are shown in Figure 3.1.

Cosmids pRU3078 and pRU3079 contain five *EcoRI* bands of the same size, two *NotI* bands, two *PstI* bands, two *SalI* bands and three *SstI* bands. Cosmid pRU3111 has two *EcoRI* bands of the same size as pRU3078 and pRU3079, two *NotI* bands, one *SalI* band, one *SstI* band and no *PstI* bands.

The results suggest that pRU3078 and pRU3079 contain fragments of the same DNA although they are not identical, as each cosmid contains additional fragments. They probably contain overlapping parts of the chromosome. Cosmid pRU3111 does not appear to contain the same DNA as there are only a few bands of the same size as those in pRU3078 and pRU3079.
Figure 3.1 Restriction Enzyme Digests of Cosmids pRU3078, pRU3079 and pRU3111.

0.8% agarose gel showing DNA bands stained with ethidium bromide visualised on a UV transilluminator.

1 = pRU3078 EcoRI, 2 = pRU3079 EcoRI, 3 = pRU3111 EcoRI, 4 = pRU3078 NotI, 5 = pRU3079 NotI, 6 = pRU3111 NotI, 7 = 1kb ladder 8 = pRU3078 PstI, 9 = pRU3079 PstI, 10 = pRU3111 PstI, 11 = pRU3078 SalI, 12 = pRU3079 SalI, 13 = pRU3111 SalI, 14 = pRU3078 SstI, 15 = pRU3079 SstI, 16 = pRU3111 SstI.
3.2.3 Southern Hybridisation

In order to determine whether each complementing cosmid contained the DNA interrupted by the transposon, or whether they contained genes that suppressed the mutation, Southern hybridisation studies were carried out. The cosmids pRU3078, pRU3079 and pRU3111 were probed with DNA cloned from RU360 (pRU472, c.f. Section 3.2.5.1) and RU361 (pRU476, c.f. Section 3.2.5.2). These clones both contain part of Tn5-lacZ and flanking downstream DNA. A 1kb DNA ladder was used as a positive control in the Southern blot, as both probes were cloned into Bluescript SK\(^-\), which contains DNA that is homologous to DNA in the mass ladder. The cosmids were each digested with five enzymes, \(EcoRI\), \(NotI\), \(PstI\), \(SalI\) and \(SstI\).

The results of the cosmids probed with pRU472 are shown in Figure 3.2. Plasmid pRU472 DNA bound to cosmids pRU3078 and pRU3079 but not to pRU3111. There are two \(EcoRI\) sites within pRU472. The probe bound to two fragments in both pRU3078 and pRU3079, one of approximately 900bp and one of approximately 3kb. There was also faint binding to a fragment approximately 2.7kb in pRU3079. This corresponds with sequence data that there should be binding to fragments of 902bp, at least 2115bp and at least 2593bp (c.f. Section 3.2.5.1). There should have been binding to another fragment in pRU3078, but this may not have been observed if it was a similar size to the 3kb fragment. There are no known \(NotI\) sites within the sequence and the probe only bound to a large fragment at the top of the gel in both pRU3078 and pRU3079.
There is one *PstI* site in the probe DNA sequence. The probe bound to a fragment of approximately 4.5kb in both cosmids and to a large fragment at the top of the gel in pRU3078. The sequence data indicate that there should have been binding to fragments of at least 1938bp and 3668bp. There should have been binding to another fragment in pRU3079, but this may not have been observed if it was a similar size to the 4.5kb fragment (c.f. Section 3.5.2.1). There is no *SalI* site in pRU472 but there are two *SalI* sites within the region sequenced, which encompass the probe DNA, giving a fragment of 4694bp, to which the probe bound to in both pRU3078 and pRU3079 (c.f. Section 3.2.5.1). There is one *SstI* site at one end of pRU472 and one in the region sequenced. The probe bound to one fragment of 3117bp in both pRU3078 and pRU3079, which corresponds with the sequence data (c.f. Section 3.2.5.1).

The data indicate that cosmids pRU3078 and pRU3079 contained the same region of DNA that restores the ability to utilise *myo*-inositol to the mutants. Cosmid pRU3079 contained additional DNA to pRU3078, but this was not required for complementation of RU360.
Figure 3.2 Hybridisation of pRU3078, pRU3079 and pRU3111 to pRU472

Southern blot with each lane containing digested DNA as detailed below.

1 = 1kb ladder, 2 = pRU3078 EcoRI, 3 = pRU3079 EcoRI, 4 = pRU3111 EcoRI, 5 = pRU3078 NotI, 6 = pRU3079 NotI, 7 = pRU3111 NotI, 8 = pRU3078 PstI, 9 = pRU3079 PstI, 10 = pRU3111 PstI, 11 = pRU3078 Sall, 12 = pRU3079 Sall, 13 = pRU3111 Sall, 14 = pRU3078 SstI, 15 = pRU3079 SstI, 16 = pRU3111 SstI.
The results of the cosmids probed with pRU476 are shown in Figure 3.3. Plasmid pRU476 DNA bound to pRU3111 DNA but not to pRU3078 or pRU3079. Two EcoRI fragments of approximately 1.1kb and 2.2kb hybridised. There was one EcoRI site within pRU476 and two within the entire region sequenced. The binding corresponds with a fragment of 1143bp and one larger than 1535bp (c.f. Section 3.2.5.2). A NotI fragment of approximately 7.5kb hybridised. There is a NotI site at the end of pRU476. The binding corresponds with a fragment larger than 771bp. There was also binding to a fragment of approximately 3kb. This does not correspond with known data and it is not known why this fragment bound (c.f. Section 3.2.5.2). A PstI fragment of approximately 7.5kb hybridised. There were no PstI sites within pRU476 or the entire region sequenced. Therefore, this fragment should encompass the entire iolA gene and surrounding region.

A SalI fragment of approximately 7.5kb hybridised. There were no SalI sites within pRU476 and two sites in the region sequenced, giving a fragment of approximately 7.5kb (Figure 3.8). Based on the size of pRU438, which was isolated as a SalI clone, binding was expected to a fragment that was approximately 4kb in size. It was subsequently discovered by restriction mapping that the upstream SalI site did not exist in the complementing cosmid pRU3111 (c.f. Section 3.2.5.2). It is postulated that when RU361 was digested with SalI, star activity occurred, followed by a forced ligation. This could have caused the creation of a SalI site, which gave the fragment that was then cloned into pBluescript SK+, resulting in pRU438. Alternatively, there could have been a deletion between SalI sites. Only one colony was obtained
from the RU361 subcloning. This is probably because the true SalI fragment containing Tn5-lacZ would have been too large to clone into pBluescript SK−.

One SstI fragment of approximately 4kb hybridised. There were no SstI sites within pRU476 and one site downstream of iolA. This corresponds with binding to a fragment encompassing the entire iolA gene at least 3167bp in size (c.f Section 3.2.5.2).

The results indicate that cosmid pRU3111 contains the DNA that was interrupted by Tn5-lacZ in RU361. The data also confirmed the results that the two mutants RU360 and RU361 are mutated in different loci.
Figure 3.3 Hybridisation of pRU3078, pRU3079 and pRU3111 to pRU476

Southern blot with each lane containing digested DNA as detailed below.

1 = pRU3078 EcoRI, 2 = pRU3079 EcoRI, 3 = pRU3111 EcoRI, 4 = pRU3078 NotI, 5 = pRU3079 NotI, 6 = pRU3111 NotI, 7 = 1kb ladder 8 = pRU3078 PstI, 9 = pRU3079 PstI, 10 = pRU3111 PstI, 11 = pRU3078 SalI, 12 = pRU3079 SalI, 13 = pRU3111 SalI, 14 = pRU3078 SstI, 15 = pRU3079 SstI, 16 = pRU3111 SstI.
3.2.4 Transduction of RU307

Poole et al. (1994) carried out transduction of Tn5-lacZ from RU360 with the bacteriophage RL38 back into 3841 and showed that the insertion of Tn5-lacZ was tightly linked to the inability to grow on myo-inositol. Transduction of Tn5 from RU307 with RL38 had previously shown that the ability to grow on glutamic acid gamma hydrazide was tightly linked to the transposon insertion (P.S. Poole, pers. comm.). To ascertain whether the reduced growth capability on myo-inositol was also due to the insertion of Tn5, the transduction of Tn5 from RU307 into 3841 was repeated.

Colonies were selected for kanamycin resistance. Eleven transductants were tested for growth on AMA with 10mM myo-inositol as the sole carbon source. No growth was visible for any of the transductants until at least eight days after they had been streaked out. All the transductants grew as normal on glucose. This corresponds with the phenotype of RU307 and confirms that the reduced ability to grow on myo-inositol is tightly linked to the transposon insertion.
3.2.5 Sequencing of myo-Inositol Catabolic Genes

3.2.5.1 RU360

Part of Tn5-lacZ was cloned from RU360 as a 12378bp SstI fragment into pBluescript SK− (pRU472), with 4030bp chromosomal DNA flanking the IS50R (Figure 3.4). This enzyme digests in the lacZ part of Tn5-lacZ, so only DNA downstream of the transposon was obtained. Attempts to obtain an intact Tn5-lacZ using the enzyme SalI, which does not digest Tn5-lacZ were unsuccessful, probably because the resulting fragment that would be obtained would be too large to clone into pBluescript SK− efficiently.
Figure 3.4  Plasmid pRU472

SstI fragment from RU360 containing part of Tn5-lacZ and adjacent chromosomal DNA cloned into pBluescript SK−.
The plasmid pRU472 was completely sequenced on both strands and three open reading frames (ORFs) were identified. BlastX analysis of the deduced amino acid sequences revealed homology with proteins encoded by genes in the \textit{myo}-inositol degradation operon (\textit{iol}) of \textit{Bacillus subtilis} (Figure 3.5). The amino acid sequence of the ORF interrupted by the transposon had identity to several bacterial acetolactate synthases. A 9bp repeat was created in RU360 by insertion of Tn\textit{5-lacZ} (1775-1784bp). The sequence that was repeated is highlighted in Figure 3.7. Highest identity was 41\% (p-value of $10^{-127}$) with \textit{iolD} found in the \textit{myo}-inositol catabolic \textit{iol} operon of \textit{B. subtilis}. The ORF was designated \textit{iolD}. The beginning of the gene was not contained on pRU472.

Downstream of \textit{iolD} was an ORF whose deduced amino acid sequence had homology with three proteins in the databases, MocC of \textit{R. leguminosarum} bv. \textit{viciae}, MocC of \textit{S. meliloti} and \textit{iolE} of \textit{B. subtilis}. These proteins have no known function. The highest identity was 42\% (p-value of $3 \times 10^{-62}$) over the entire amino acid sequence of MocC of \textit{R. leguminosarum} bv. \textit{viciae}. The ORF was designated \textit{iolE} and began 46bp downstream of the end of \textit{iolD}.

Downstream of the \textit{iolE} gene was an ORF with 62\% identity (p-value of $5 \times 10^{-58}$) with \textit{iolB} of \textit{B. subtilis}. This protein has no other homologues in the databases and has no known function. The ORF was designated \textit{iolB} and began 246bp downstream of the end of \textit{iolE}. The end of the gene was not contained on pRU472. All three genes are predicted to be transcribed in the same direction as the transposon but they are not organised in the same order as in \textit{B. subtilis}.
Figure 3.5 RU360 iol operon

The diagram represents the entire region of DNA sequenced from pRU3078 and encompasses the DNA from pRU472. The entire region sequenced was 5.606kb.
In order to obtain the full DNA sequence of the interrupted genes and to identify any other genes in the region surrounding the myo-inositol locus, attempts were made to subclone pRU3078 and pRU3079. No fragment large enough to encompass the iol region and surrounding DNA was identified by Southern hybridisation. Therefore, sequencing was carried out on cosmid pRU3078 directly with primers designed to sequence upstream and downstream of the region encompassed by pRU472. A region of 5606bp was sequenced on both strands. This contained the entire sequence of the three putative iol genes (EMBL Accession Number AJ276296) (Figure 3.5). Putative Shine-Dalgarno sequences indicating ribosome binding sites (Shine and Dalgarno, 1974) were identified for each putative gene (Figure 3.7).

A TTG start site for iolD corresponds with a TTG start of the closest homologue, iolD of B. subtilis. A putative Shine-Dalgarno sequence was identified upstream of the TTG start. However, a TTG start is less common than ATG and it is not known whether the start of the putative iolD gene is the same as the closest homologue. If the entire gene were considered as the largest ORF starting at ATG, which also had a putative Shine-Dalgarno sequence upstream, then the putative iolD gene would contain an additional 201bp upstream of the start of its closest homologue.

Other homologues in the Genbank and EMBL databases are larger than the B. subtilis iolD and are of a similar size to the iolD gene identified here, with ATG start sites. Testcode analysis was carried out on the sequence to determine whether the sequence is likely to be coding (Figure 3.6). The
results indicate that the 201bp upstream of the TTG start site are likely to be coding.

Further analysis is required to confirm the correct sequence, but in order to ensure that the start of iolD is definitely present, the iolD gene was considered to be the largest ORF, which starts with ATG. The iolD gene is therefore 1881bp, encoding a protein of 627 amino acids with a predicted molecular weight of 66841.9 and an Isoelectric Point (PI) of 6.1. The smaller ORF would be 1680bp, encoding a protein of 560 amino acids, with a predicted molecular weight 59363.5 and a PI of 5.8.

The iolE gene is 939bp, encoding a predicted protein of 313 amino acids, with a molecular weight of 344440.3 and a PI of 5.6. The iolB gene is 795bp, encoding a predicted protein of 295 amino acids, with a molecular weight of 29157 and a PI of 5.8.

Further sequence was obtained for 1kb either side of the iol genes. A putative ORF was identified that began 407bp downstream of iolB. Sequencing did not extend far enough to give the end of the putative ORF. BlastX analysis of the deduced amino acid sequence of the region surrounding the iol genes revealed no homology with known sequences. The ORF was designated orf1 but testcode analysis indicates that this region is unlikely to be coding. Sequencing of the region upstream of iolD did not reveal any homology with sequences in the databases and no putative ORFs. The entire sequence,
with the deduced amino acid sequence for the four ORFs is presented in Figure 3.7.

The sequence data obtained suggest that the putative genes *iolD*, *iolE* and *iolB* comprise an operon that is a distinct locus to other genes involved in *myo*-inositol utilisation.
Figure 3.6 Testcode Analysis of RU360

Testcode analysis of the iol region from pRU472 and pRU3078. Sequence that is in the top portion of the graph is likely to be coding. Sequence in the middle portion is ambiguous and sequence in the bottom portion is unlikely to be coding.

The ATG and TTG start sites of iolD are marked with arrows.

T = testcode.
Figure 3.7 Entire Sequence of pRU472 and Surrounding DNA

```
1    CGACTTGAAG ACGGAGCAGC AGGAGAAGCT GCGGACGCTG TTCGAGGCCG
GCTGAACCTTTC TGCCTCTGCT TCCTCTTCGA CGCCTGCGAC AAGCTCCGGC
51   CCCGCAAGGT CGGCCGCGAG CTGCTGGTGG AGATCATCGC CGGCAAGAAC
GGGCGTTCCA GCCGGCGCTC GACGACCACC TCTAGTAGCG GCCGTTCTTG
101  GGCCACTCAC CGACGACACG ATCGCGACCG CGCTGGAGGA ACTCTATGCG
CCGGTGAGTG GCTGCTGTGC TAGCGCTGGC GCGACCTCCT TGAGATACGC
151  CTCGGCATCA AGGCGGATTG GACGCTGG则是 GACGTCTAGA TCTAGCTTCT
GGCCACTCAC CGACGACACG ATCGCGACCG CGCTGGAGGA ACTCTATGCG
201  CTGCTGCTCC CTGAGCGTGG GCTGCTGTGC TAGCGCTGGC GCGACCTCCT TGAGATACGC
GACGACCGCT CGCTGCTGCT GCCTGCTGCT GCCTGCTGCT GCCTGCTGCT GCCTGCTGCT
251  CTGCTCAGCG GCGACGCTGG CAGCTCCCTC CGTGAAGGGT TTCGCCGTCG
GACGAAGTCG CGCTGCTGCT GCCTGCTGCT GCCTGCTGCT GCCTGCTGCT GCCTGCTGCT
301  GACGCTCCTGC CTTAATGCTG GCTGCTGTGC TAGCGCTGGC GCGACCTCCT TGAGATACGC
GCCGGACGAT CTTTGGCGAT CGGCGCGCGC CCCGGCGCGC CCCGGCGCGC CCCGGCGCGC
351  GGCGCTGGGT GAGGGGGGTC GCTGCTGTGC TAGCGCTGGC GCGACCTCCT TGAGATACGC
GCCGGACGAT CTTTGGCGAT CGGCGCGCGC CCCGGCGCGC CCCGGCGCGC CCCGGCGCGC
401  AAGGCGACAG AAGGCGACAG AAGGCGACAG AAGGCGACAG AAGGCGACAG AAGGCGACAG
GGCGCTGCTG GCTGCTGCT GCCTGCTGCT GCCTGCTGCT GCCTGCTGCT GCCTGCTGCT
451  GACGCTCCTGC CTTAATGCTG GCTGCTGTGC TAGCGCTGGC GCGACCTCCT TGAGATACGC
GCCGGACGAT CTTTGGCGAT CGGCGCGCGC CCCGGCGCGC CCCGGCGCGC CCCGGCGCGC
501  GTCGGGTATT GCTGCTGTGC TAGCGCTGGC GCGACCTCCT TGAGATACGC
GCCGGACGAT CTTTGGCGAT CGGCGCGCGC CCCGGCGCGC CCCGGCGCGC CCCGGCGCGC
551  GACGCTCCTGC CTTAATGCTG GCTGCTGTGC TAGCGCTGGC GCGACCTCCT TGAGATACGC
GCCGGACGAT CTTTGGCGAT CGGCGCGCGC CCCGGCGCGC CCCGGCGCGC CCCGGCGCGC
601  GACGCTCCTGC CTTAATGCTG GCTGCTGTGC TAGCGCTGGC GCGACCTCCT TGAGATACGC
GCCGGACGAT CTTTGGCGAT CGGCGCGCGC CCCGGCGCGC CCCGGCGCGC CCCGGCGCGC
651  AAAATGGCTG AAGAAAGATT GCGGGAGGAC CCGATGGCAGA AGACGATACG
TTTACGCTT GCTGCTGTGC TAGCGCTGGC GCGACCTCCT TGAGATACGC
701  GTGAGCTGATG GCGCAGCCGC TTCACATGCT CTCGGACGCT CTCGGACGCT CTCGGACGCT
GCCGGACGAT CTTTGGCGAT CGGCGCGCGC CCCGGCGCGC CCCGGCGCGC CCCGGCGCGC
751  AAGGCGACAG AAGGCGACAG AAGGCGACAG AAGGCGACAG AAGGCGACAG AAGGCGACAG
GGCGCTGCTG GCTGCTGCT GCCTGCTGCT GCCTGCTGCT GCCTGCTGCT GCCTGCTGCT
801  GACGCTCCTGC CTTAATGCTG GCTGCTGTGC TAGCGCTGGC GCGACCTCCT TGAGATACGC
GCCGGACGAT CTTTGGCGAT CGGCGCGCGC CCCGGCGCGC CCCGGCGCGC CCCGGCGCGC
851  GACGCTCCTGC CTTAATGCTG GCTGCTGTGC TAGCGCTGGC GCGACCTCCT TGAGATACGC
GCCGGACGAT CTTTGGCGAT CGGCGCGCGC CCCGGCGCGC CCCGGCGCGC CCCGGCGCGC
901  GACGCTCCTGC CTTAATGCTG GCTGCTGTGC TAGCGCTGGC GCGACCTCCT TGAGATACGC
GCCGGACGAT CTTTGGCGAT CGGCGCGCGC CCCGGCGCGC CCCGGCGCGC CCCGGCGCGC
951  GACGCTCCTGC CTTAATGCTG GCTGCTGTGC TAGCGCTGGC GCGACCTCCT TGAGATACGC
GCCGGACGAT CTTTGGCGAT CGGCGCGCGC CCCGGCGCGC CCCGGCGCGC CCCGGCGCGC
1001 GCCGGACGAT CTTTGGCGAT CGGCGCGCGC CCCGGCGCGC CCCGGCGCGC CCCGGCGCGC
```
751  TCGTGCACGG CAAGAAAGTG CCGATCTTCG GCGGTGTCTG GGCGATCTTC
AGCAGCTGCC GTTCTTTCAC GGCTAGAAGC CGCCACAGAC CCGCTAGAAG
+3  G H G N V A G I G E A L Y Q V R E
801  GGTACGCCGCA ACCTACCGCG TGATGGCGAG GCGCTCTATC AGGTGGCAGA
CCAGTGCCGT TGGATAGCGC GGGTGTTGCT TGTCCCGTAC CGCGTACGGC
putative ttg start
+3  E L T Y R A N Q M A H A
851  GGAATTTCAG ACCTACTCCG CCCACACACG AAACGCATCG GGCACGTCGG
CCTAGCTGAC TGGATAGCGC GGGTGTTGCT TGTCCCGTAC CGCGTACGGC
+3  A A Y A K A N F R T R F M A C T
901  CGATGCCCTTA TGCCAGCGGC ATAGCCGCTC CGCGAGATAG TCCACGCGCT
putative ttg start
+3  H V N R I P V L F L P G D V F A
1001 GCATGTCAAT CGTATCCCGG TTCTCTCTCT GCCGGCGGAC GTCTTCCGCA
CGTACAGTGA GCATAGGGGC AAGAGAAGGA GGCGGCGCTA TAAAGCTGGC
+3  N R A P D P V L Q Q I E D S A T A
1051 ACCCGCGCGCC GAATCCGGTG CTGCAACAGA TCGAGGATTC GGCGACGGCA
TGGCGCGCGG CCTAGGCCAC GACGTTGTCT AGCTCTTAAG CCGCTACGGC
+3  S V S A N D A F R S V S R Y F D R
1101 TCGATCCCGAC CCAATGGATT GCATGGCTTCG GTTGGCCTGGCT ATTTCCCGCG
AGCCAAAGCC GGTACTACGC GAAGGGCAGA CAAAGCGCAG TAAAGCTGGC
+3  I T R P E Q I I T A L K R A M Q
1151 CATCACCAGG CCCGAGCAGA TCATACCGCG GCCTAGGCGG GCACATCGGA
GTAGTGCCCG GCCGCGCCCG GCCGCGCCCG GCCGCGCCCG GCCGCGCCCG
+3  V L T D P L D C G P V T L S L Q
1201 TCGGTACGCC CCGCGCTTCCG TGGCGCGCGG TGATCCGCACT GCCTTCCGCA
AGGAGTTGCGT GGGCGGGCC ACCGCGCGCG CCCTGAGCGAC CCAGGCGCGC
+3  D V Q A E A Y D Y P E S L F A E K
1251 GACGTTTCCAG CAAAGGCGTA TGATTTTCCG GAAGAACCTGT TGGCGGAAAA
CTGCAAGTGC GTCCTCCCGCT ACTAATAGGC CTTTCCGACA AAGCCTCCTT
+3  V W T T R P Q P D A D E L A N
1301 AGCTGGCAGC ACCGCGCGCG CGCGCGCGCG GCCTAGGCGA TGCCCGCGATG
TGGACGATCG TGGCGCGGCG GCCGCGCGCG GCCGCGCGCG GCCGCGCGCG
+3  A I A L I K A S Q K P V I V A G G
1351 CCGTGCGAC GAGCAAGCAG TCGAGGATTC GTCCGGCGCG GCCTTGCGCG
GGTACGGCGA CTTAGGCTCG GCCGCTAAGA ACAGGCGCGC
+3  G V L Y S Q A T K E L A A A F A E A
1401 GCGTGCGCGG ATTCCGCGCG GACGAGAGAA CTGGCGCGGT TTGGCGGAGG
CGCGAGAAAT TAAGGTCCGT CTGCTTCTCT GAAGGCGGGA AACGGCTCCG
9bp overlap by Tn5-lacZ in RU360
GGATGTACTA CTACTGAGG CTCGAACGCT GGAGCCAGTA CGAGCCGGAG
2151
+3
K L N I I VL DNR G YGC N R
AAAGCTCAAACA TCATCGTGCT CGATAATCGG CTTGAGCGGC CACTAAACGG
TTGAGTTTGT AGTAGCAGGA GCTATTAGCC CCAAAGCGCA CTTAGTGGGC
2201
+3
LQ MGT GANFN NNLK D
ATGGCAAATG GGAACCCCGCC GGGCCAACTT CAAACATCGT CTCAGGACT
TAAGTTTACT CTGGCGCCGC CGCGTTGAAA TGGTTAGAC GAGTTCTGGA
2251
+3
SY HEVMPEDFRAHES
CCTATCAGCA GGTAGTGCCG GAGATCGATT TCCGGGCGCA TGCCGAAGGC
GGATAGTGGCT CCACTACGGG CTCTAGCTAA AGGGCGCCGT ACGGCTTTCG
2301
+3
MGAIAVKSIAIEQLA
ATGGGCGCCA TCCGCGCCAA AGTCCGCTTG ATCGCCGAGC TGGAACAGGC
TACCAGGGGT AGGGCCAGTT CAGGCGGAGC TACCGGCTCG ACCTTGTCGG
2351
+3
LADSRKNDRTSVID
GCTGCGCGAT TCCAGGAAGA AGCAGCGCAG TCCGCGCTTC GCAGTGCCAC
CGACGGCGTA AGGTCTCTCT TGTCGAGCTG CAGCCAGAAG CATGAGCTG
2401
+3
TDPLITTETAGEHWGWDA
CCGATCCGCT GTACCCCAACA GAGACGCCGGC GCACTGTGGT GGATGTCCGG
GGCTAGGGGC CTATGGTGGT CTGCGGCGCC CGGTGACCAC CCTAGCAGCC
2451
+3
VPEVSRSRSEVNRDHAY
GTGCTTGAGG TGACGCGCAG CAGTGGATTG AACAGCGGACC ATGAGGCTTA
CAGCGACTCC AGTCCGCAGC GCAGTCCCCG TGGTTGCGGC TACTGCGGAT
2501
+3
VKARAARQVRG
TGTCGAAAGCA CGTGCTCGCC ACGCGCTCGG CCGTCCGGTGT TGGTCGCTTT
ACAGTTTCGT GCACGCAGGC TCGCGAGCTG ACAGTGCCTA GCACAGTAAA
2551
io1E
+3
SDMLLLRRMLL
TTCTGGAGGA GCCGCCCGGC TCCTCGCCAA TGGCTTTTAA GGAAAATGGT GCAGACCCTC TCCGAGCTAC
2601
+3
MKAKLGMSPIAWWNDDL
ATGAGGAGCA AACTCGGGCT GTGGGCCATC GCTGCGCTTG AACAGGCAGTT
TACTGCGGCT TGAGCCCGTA CAGCGGTTAG CGAACACCTC TGCTGCTGGA
2651
+3
PELSVDSLESVSRQ
TCCTGAACTC AGCGCGAGCG TGCTGACGAG GGAATGCGGT CGCGATAGCC
AGGACGTGAG TGCTGCTGC ACAGACGTG CTTGCGGCG CCGGTCAGGG
2701
+3
RSAGFTGMEQGRRFPSN
GAGTGCTGAG GCTTTACCGGC AGGGAGGAGG GCCAGGCATT CCCAGCAACT
CTTCAGGCAC GAAATGCGGG TACCCGAGAG CGGCGTCAGG
2751
+3
PEEMLPILRALADVTLCG
CCCGAGGAGA TGCGCTGGCC TGGCGCGGGG CGGCGCAGTG CGCTGGCGG
GGCTCCCTCT AGCAAGGTGA GAGAGCGCAG GGGCTGCACT GCAGACGCGC
2801
+3
GWFSGTLVNEELAANK
2801  TGGCTGGTTC TCCGGCACGC TGGTCAATGA GGAGCTTGCC GCCAACAAGG
      ACCGACCAAG AGGCCGGTTCG ACCAGTTACT CCTCGAACGG CGGTTGCTTC
+3 D R I A P M I A L F K A V N A P C
2851  ACCGACATCGG GCCGGATGATC GCGCTGTTCA AGGCGGTCAA TGCACCTTGC
      TAGCAGGATCG CCTGGCCGCTGG AAAGCTCTAC GCTCTACTTC CGGATGACGG
+3 I V Y G E V G R S I Q G D R S K P
2901  ACCGCATCGC GCCGATGATC GCGCTGTTCA AGGCGGTCAA TGCACCTTGC
      TGGCGTAGCG CGGCTACTAG CGCGACAAGT TCCGCCAGTT ACGTGGAACG
+3  I V Y G E V G R S I Q G D R S K P
2951  ACCGCCAGTG TTCGCGCCGCG AAATGGTGGCC CGAAGCCAGG CATGCCGCTT
      CGGCGCAGTG CCTTAAGCCC CTTACCACGC GGCTCGTTCC GTACGGCGAA
+3   L A T K P R L S D D E M K A Y A
3001  GCCTCAGCGACC AAGCCCGGCC GCCGATGATC GCGCTGTTCA AGGCGGTCAA
      ACCGCATCGC GCCGATGATC GCGCTGTTCA AGGCGGTCAA TGCACCTTGC
+3    R R V T E F G E W C A E Q G M P L
3051  TCCTATCACC ACCACATGGC GGCCGTGGTC GAGACCGAGC CGGAACTCGA
      AGGATAGTGG TGTTGATCCG CCCTGCTCCG GCTTTGAGCT GCTTTGAGCT
+3   S Y H H H M A A V V E T E P E L D
3101  CGCCCTTTCAT CGTCAATTCCG GCAGCTGGGTC GCGAAGGGCAT CCGCTGTGCT
      GCGGAAGTAC GCAGTAAGCC CGCTTCCGTA GGGCGACGAC GAGCTAGCG
+3    H A R I N H V H V K D I R K P V V
3151  GCCATCTCGC CTTTGCCGGC GGCGACGTGC TGCGCGCCAT CGGAAACCAC
      CGGTAGAAGCC GAAACGCCGG CCCTGCCGCTG AGCTCGAGCG CCCTGCGTGC
+3  H A R I N H V H V K D I R K P V V
3201  CACGCCCCCA GACACCATGGC CTACGTCAAG CACATCAGCC AGGCTGTGCT
      GTGCGGGCTCG CTGATGCTGC CGGGCGTCTT TTACCGGCTC TAGCCGGTGC
+3   D G L D R S R Q S F L D A V A L
3251  GGATGGGCTG GACCGCAGCC GGCAGTCCTT CCTCGAATCT GCCTGGGCTC
      CCTCGGAGG CGAGTGATCC CCTGCTGGTG CCCTGCTGGG GCGGTAGCAG
+3  Q R L A D H Y E G W F V E A E
3301  GCGCCTTCAC GGTGCCGGGC GACGGCTCGC TCGATTTCGG CGCCATCGTC
      CGCGGAAGTG CCACGGCAGA AGGCTGACCC CCGCGGATCC CGCGGTAGCAG
+3   Q R L A D H Y E G W F V E A E
3351  ACAGGATCCG CGCAAGGCCC CGCCGCAGAA AATGGCCGGA ATCCGCCGAC
      TGTCTCATGC CCGTCTCCGC CGCTTGACAA GCAGTACGCG ACGGCCGCCA
+3  Q D R K A P P Q K M A E I G H
3401  ACAGGATCCG CGCAAGGCCC CGCCGCAGAA AATGGCCGGA ATCCGCCGAC
      TGTCTCATGC CCGTCTCCGC CGCTTGACAA GCAGTACGCG ACGGCCGCCA
+3    A E L M R V M T A A G Y T V E T E
3451  CGGAATTCATG GCCGCGCTATG ACGCGCCCGCC GCCTATCGAGT GGAACCGCAG
      GGCTTAAGCTC CGCGGACGAG CGATATGAGT CTTGGCATCG GCCTGGCAGA
+3    G F P K G
<table>
<thead>
<tr>
<th>Position</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>3501</td>
<td>GGCTTCCCCA AGGAG<strong>TAG</strong>G</td>
</tr>
<tr>
<td></td>
<td>AGCCACAAGAC CCCCTCCCAC AAGGGGGAGA</td>
</tr>
<tr>
<td></td>
<td>CCGAAGGGGT TCCCTATTCG CTGTTGTTCTG GGGAGGGGTG TTCCCCCTCT</td>
</tr>
<tr>
<td>3551</td>
<td>GGCTGATCAG GACAGCCGGGA GCAAGACGCC CCCCTATCCG CCCTCCGGGAC</td>
</tr>
<tr>
<td></td>
<td>CCGACTAGTC CTGTCGGGTCCT CTTTCTGGCG GGGAGTAGGC GGGAGGCCTG</td>
</tr>
<tr>
<td>3601</td>
<td>ACCTTTGCGG TGGTTAGGCC ACCCGACTCA ACCCGATCTCTT CCAGACCCCG</td>
</tr>
<tr>
<td></td>
<td>TGGAAACGGG CCCAATTCGCT TGGCCTGAGT TGGGCAGGAA GCTCTGGGGGC</td>
</tr>
<tr>
<td>3651</td>
<td>CGCGGAAAGA GGGGAAAGAG GGGCAGCCGGC ACATCCCTTC TCCCCTCCGG</td>
</tr>
<tr>
<td></td>
<td>GACCCCCCTT CCCCCCGCGC CTTGAGGAGA AGGGAGGCCC</td>
</tr>
<tr>
<td>3701</td>
<td>GAGAAGGTGG CGCGACCCGG ATGAGGGGGG GCCCGACAGC CCAGATTTACG</td>
</tr>
<tr>
<td></td>
<td>CTCTTCACC ACCGTCGGGC TACTCCCCCC GCACGCTGGT GGTTAAATAC</td>
</tr>
<tr>
<td>3751</td>
<td>+2 SD M P N L K V K P S G T H</td>
</tr>
<tr>
<td>3801</td>
<td>AGGAGAAACA CCA<strong>ATG</strong>CCAA ATCTCAAGGT GAAACCATCG GGCACGCAGT</td>
</tr>
<tr>
<td></td>
<td>GCTCTTTTGT GGGTTACGGGT TAGAGTTCCA CTTTGGTAGC CCGTGCGTAC</td>
</tr>
<tr>
<td>3851</td>
<td>+2 G R V T H V T P E N A G W T Y V G</td>
</tr>
<tr>
<td>3901</td>
<td>CGATCGCGAA GTCTGCTTCC ACCTGGGTACACCCTGGGATG GAGAGACGGG AAGCTAGACG TAGCCTACTT CGGCCCGCTC TGGCAAAGCC CTCTCTGCCC</td>
</tr>
<tr>
<td>3951</td>
<td>+2 D R E V C L V W V T G K G K A S</td>
</tr>
<tr>
<td>4001</td>
<td>GCCGCGTACAC CCATGTCACC CCGGAAAACG CCGGCTGGAC CTATGTCGGC CGGCGCAGTG GGTACAGTGG GGCCTTTTGC GGCCGACCTG GATACAGCCG</td>
</tr>
<tr>
<td>4051</td>
<td>+2 A E T D L E L A V C S A P G G G</td>
</tr>
<tr>
<td>4101</td>
<td>GGGCAGGACC GATCTGGAGC TCGCGGTCTG CTCGGCACCC GGCGGCGGCA CAGCTCGGCG CATTCGCTGC TGGTCGTCGA GGTGATCACG CCGGGTGGAC GTCGAGCCGC GTAAGCGACG ACCAGCAGCT CCACTAGTGC GGCCCACCTG</td>
</tr>
<tr>
<td>4151</td>
<td>+2 T Y Q A K E I P P G T H P Q V T R</td>
</tr>
<tr>
<td>4201</td>
<td>GGGCAGGAGCCA CAGCTCGGCG TGGTTCGCTG CCGCCTCTGG CTTAACACGC GATAGCTGGG GCTGCGCGGC TGTAGGGCTAC CTTAGCTGTA CCAGCCACGT</td>
</tr>
<tr>
<td>4301</td>
<td>+2 G K G T N V R Y V N N I M P E D D</td>
</tr>
<tr>
<td>4351</td>
<td>GGGCAGGAGCCA CAGCTCGGCG TGGTTCGCTG CCGCCTCTGG CTTAACACGC GATAGCTGGG GCTGCGCGGC TGTAGGGCTAC CTTAGCTGTA CCAGCCACGT</td>
</tr>
<tr>
<td>4401</td>
<td>+2 S S A H S L L V V E V I T P G G</td>
</tr>
<tr>
<td>4451</td>
<td>CAGCTCGGCG CATTCGCTGC TGGTTCGCTG CCGCCTCTGG CTTAACACGC GATAGCTGGG GCTGCGCGGC TGTAGGGCTAC CTTAGCTGTA CCAGCCACGT</td>
</tr>
</tbody>
</table>
ACGGCAGTAT TCTCGCTGCT CCTTACGCGG GTGTCCTGCT CGGAGCAGAT
+3  P Y R G V F M R H V G R N D T V
5051  TCCCTATCCG GGCCTCTTCA TGCGGCATGT CGGCGCGCAAT GATACGGTG
AGGGATAGCG CCACGAGAAT ACTCGGTACA CGGCGGTTTA CTATGCGACC
+3  A E A I R S C S S I P A R D T G S
5101  CCGAACCAAT CAGGTCTTGT TCTTCAATAC TCGGCTGCGA TACCGGATCA
GCTTGCCTTA GTCAGTTATG AGCGTGGTCC ATGGCCCTAG
+3  A T L S R V A T P A S I W R S T T
5151  GCCACCCATT CGAGGGTGTCG GACGGCTGCA TCGATCTGCG GATTCAGACG
CGTTGGGATA GCTCCACCG GTTCGGAGCT AGCTAGACC CTAGCTGCTG
+3  P C S K S Y A E G A G A R P A
5201  TCCATGCTCG AAGAGTTACG CCGAAGGAGC AGGCGCAGCC CGGCCCGCCC
AGGTACGAGC TTCTCAATGC GGCTTCCGTC TCCGCGTCGG GCCGGGCGGG
+3  V A L L R H G L S R N V A E T L E
5251  GTTCGCTCCT TCGCCATGG TCTCAGCCGC AACGTCGCCG AAACACTGGA
CAACCGGAGG CCGCGGTCCG GCGCGCTAG CTAGGGCGCG CGCGTCCGAC
+3  A E I L A T L V R R S L G R A
5301  GCGCGGAATA TTGGCGCTGA CGCTCGTGCG CCGTTCGCTC GGGCGAGCGC
CCGCCTTTAT AACCACGACT GCGAGCACGC GCGAAGCGAG CCCGCTCGCG
+3  H I G T A G G A P A P A G R N S S
5351  ACATCGGCAC GGCAGGCGGG GCCCGAGCGC CGCCGGCGAG AAACTCCTCG
TGATGCGCTG TCTCCGCCGC CGGCCTGTCG GCCGGCCGTC TTTGAGCAGC
+3  T R G K S W L L F P P I L A R A L
5401  ACCCGGGGGA AAAGCTGGTT GCTTTTCCCT CCGATTCTCG CGCGGGCGCT
TGGGGCCTGG TTTGACCAA CGAAAGGGA GGCTAAGACG CGCGCGCGGA
+3  G T T F R E N R H P K W G V S A
5451  TGGGACCACT TTCCGGAAA ATCGGCCAACG GAAGTGGGCG GCTCTCGGC
ACCCTGGTGA AAGCCCTTTG TAGCGGTTGG CCTTACCCCG CAGACCGGCG
+3  V L S H P G S F Q Q V R R R T P L
5501  TTTATATCTCA CCCAGGTCTT TTCACGAGG TCCGAAGGGC GACGCCACCT
AAAATAGAGT GGGTCCCAGA AAGTGCCTCC AGGCCTTCGG CTGGGCGTAA
+3  F T
5551  TTTACC AAATGG

Letters in red indicate amino acids. For amino acid code, see Appendix

3.1. Start and stop codons for each ORF are bold and underlined.

Putative Shine-Dalgarno sequences (SD) are bold.
3.2.5.1.1 Complementation of the ioI region of RU360

In order to determine whether all three genes identified in RU360 are necessary for myo-inositol catabolism, PCR amplification was carried out on the ioI region of pRU3078 (Figure 3.8). The fragments obtained were cloned into pCR2.1 TOPO and then subcloned into pOT1, which is a broad host range vector. Primers p213 and p228 were designed to amplify a 2308bp fragment containing the entire ioID gene. Primer p228 was designed with a SpeI site. This fragment was then cloned into the plasmid pOT1, as an XbaI-SpeI fragment. The resulting plasmid was named pRU706.

Primers p230 and p228 were designed to give a 4343bp fragment containing the entire three ioI genes identified, ending 38bp beyond the end of ioIB. The restriction enzyme sites Pmel and SpeI were included in the primers p230 and p228 respectively, to enable directional cloning of the fragment. The resulting plasmid was named pRU713.

The plasmids pRU706 and pRU713 were conjugated into RU360. Six colonies from each conjugation that were streptomycin, kanamycin and gentamycin resistant were streaked onto AMA containing 10mM myo-inositol as the sole carbon source. None of the colonies containing strain RU360/pRU706 were able to grow on myo-inositol, indicating that the presence of the entire ioID gene was not sufficient to restore the ability to utilise myo-inositol to RU360. All six colonies containing RU360/pRU713 grew on myo-inositol as the sole carbon source, indicating that the ioID, ioIE
and iolB genes enable complementation. This suggests that the transposon insertion in iolD has a polar effect on the downstream genes and therefore one or both of these genes is essential for myo-inositol catabolism. It is not known whether the putative gene orf1 is required for myo-inositol utilisation, as it may have its own promoter, but as the sequence is unlikely to be coding, this region was not investigated further.

**Figure 3.8 Subclones of RU360**

The 5.606kb region sequenced from pRU3078 with arrows detailing the sequence encompassed by the subclones.
3.2.5.2 RU361

An intact Tn5-lacZ with flanking chromosomal DNA was cloned from RU361 as an approximately 12400bp SalI fragment into pBluescript SK− (pRU438) (Figure 3.9). The clone contained 230bp of chromosomal DNA upstream of the transposon and approximately 3800bp downstream. A subclone of pRU438 was subsequently constructed in pBluescript SK− (pRU476) in order to facilitate sequencing and for use in Southern hybridisation studies. Plasmid pRU476 was a 1370bp NotI clone, consisting of 599bp of Tn5-lacZ DNA with 771bp adjacent chromosomal DNA.

Sequencing was carried out on both strands of 230bp of chromosomal DNA upstream of Tn5-lacZ and on the 771bp downstream of Tn5-lacZ contained in pRU476. A 9bp repeat was created in RU361 by insertion of Tn5-lacZ (1525-1533bp). The sequence that was repeated is highlighted in Figure 3.12. BlastX analysis of the deduced amino acid sequence of the DNA revealed an ORF with homology to several methylmalonate semialdehyde dehydrogenases, including IolA of the B. subtilis iol operon. Highest identity was 68% with MmsA of Mycobacterium tuberculosis (p-value of $1 \times 10^{-145}$). The ORF was designated iolA and is predicted to be transcribed in the opposite direction to the transposon. The end of the gene was not contained on pRU438.
Sequencing was also carried out on 500bp of DNA from the SalI site downstream of the transposon. The deduced amino acid sequence had no homology with any known sequences.

Figure 3.9 Plasmid pRU438

SalI fragment from RU361 containing intact Tn5-lacZ and flanking chromosomal DNA cloned into pBluescript SK-. The size of the plasmid is approximate, as the entire chromosomal insert DNA was not sequenced.
The transposon clone pRU438 did not contain the entire \textit{iolA} gene. As the \textit{SalI} fragment in pRU438 was smaller than the \textit{SalI} fragment that bound in the Southern hybridisation, it was decided to utilise the complementing cosmid to obtain the entire gene sequence and to identify any adjacent genes. This was in case pRU438 contained deletions of crucial sequence. The cosmid was known to contain the correct sequence as it complemented RU361 (Figure 3.10). Attempts were made to clone a \textit{PstI} fragment approximately 7.5kb in size, as this was the size of the \textit{PstI} fragment that bound to the probe in the Southern blot and was expected to encompass the entire \textit{iolA} gene.

A 10.2kb fragment was obtained, containing three \textit{PstI} fragments, one of approximately 7.5kb fragment, one of approximately 2.5kb and one of 213bp. The 7.5kb fragment was cloned into pBluescript SK\(^{-}\) and the resulting plasmid designated pRU482. Attempts to clone the 7.5kb fragment alone by cutting the fragment out of an agarose gel and cloning into pBluescript SK\(^{-}\) failed. It is postulated that a lethal gene product would be produced if the end of the 7.5kb fragment was not covered by a piece of DNA that would prevent transcription from the cloning vector.

Sequencing using custom primers to the \textit{iolA} gene revealed that pRU482 contained DNA identical to that interrupted by Tn5-\textit{lacZ} in RU361. The entire \textit{iolA} gene was sequenced on both strands (EMBL Accession Number AJ276297) and a Shine-Dalgarno sequence identified upstream of the start site (Shine and Dalgarno, 1974). The \textit{iolA} gene was 1494bp, encoding a
predicted protein of 498 amino acids, which has a molecular weight of 53528 and a PI of 5.8.

Downstream of *iolA* an ORF was identified whose deduced amino acid sequence had identity to Aau3 of *S. meliloti*. The Aau3 protein is required for growth on polyhydroxybutyrate (PHB) cycle intermediates. There was 74\% identity (p-value of $6 \times 10^{-54}$) over the amino acid sequence. Other homologues included Yhde of *B. subtilis* and YjeB of *E. coli*. None of the other homologues had any known function. The ORF was designated *aau3* and began 336bp downstream of the putative end of *iolA*. A putative Shine-Dalgarno sequence was identified (Shine and Dalgarno, 1974). The *aau3* gene is predicted to be transcribed in the same direction as *iolA*. Sequencing was only carried out on one strand.

Sequencing was continued for 1206bp upstream of the *iolA* gene, using custom primers. An ORF was identified beginning 405bp upstream of the start of *iolA*. BlastX analysis of the deduced amino acid sequence showed that the ORF had 35\% identity (p-value of $9 \times 10^{-12}$) over the first 173 amino acids of 282 with a hypothetical gene from *A. tumefaciens*. The gene does not have any known function. The next highest homology was 31\% (p-value of $3 \times 10^{-9}$) with a hypothetical transcriptional activator gene, *act*, from *P. aeruginosa*. The ORF was designated *orf1* and is predicted to be transcribed in the opposite direction to *iolA*. Sequencing did not extend far enough to give the end of *orf1*. This was because when the sequencing was carried out,
there were no homologous genes registered in the databases, so it was not considered necessary to sequence further.

Testcode analysis was carried out on the region sequenced in pRU482 (Figure 3.11). The results show that *iolA* and *aau3* are likely to be coding, but *orf1* is unlikely to be coding. Based on sequencing results, it is postulated that the *iolA* gene is not part of an operon of *myo*-inositol utilising genes. The complete DNA and amino acid sequence of *iolA* along with flanking DNA is presented in Figure 3.12.
Restriction map of part of approximately 17.5kb of pRU3111, based on restriction mapping and sequence data. Putative genes are marked according to sequencing results. The three PstI fragments of pRU482 are arranged in the proposed correct orientation, with each fragment indicated in a different colour.

The orientation as cloned into pRU482 is shown below. The two smallest fragments are flipped round with respect to the correct orientation.
Figure 3.11 Testcode Analysis of RU361

Testcode analysis of the iol region from pRU438 and pRU482. Sequence that is in the top portion of the graph is likely to be coding. Sequence in the middle portion is ambiguous and sequence in the bottom portion is unlikely to be coding.

T = testcode.
Figure 3.12  Sequence of ioL and Surrounding DNA

1  AGAACGTTCAT  CCTGCTTTTT  TGATCAGGCG  GCGGGGCAAC  GATCGCGGGT  TCTTCAAGTA  GACACAGAAA  ACTAGTCCGC  CGCCCCGTTG  CTAGCCGCCA  
-3  G P R K I L R R P L S R P

51  TTGCGATAGGG  CGGCTGCAGCG  GTGATTCCGA  GAAGGAAGTT  GATCTGCGGA  AACGCTATCC  GCCGAGCGGC  CACTAAGGCT  CTTCCTTCAA  CTAGACGCCT  
-3  A I P P E G T I G L L F N I Q P

101  CGGCTTTTGA  CGAGGTGTCG  GATGGAATAT  CGAAGTACCG  CAAAGAAAGC  GCGCGAAACT  GCTCCAGCAG  CTACCTTATA  GCTTCATGGC  GTTTCTTTCG  
-3  R A K V L D D I S Y R L V A F F A

151  GTTGAGCGGC  TTGCGCAGCG  CCGAATTCAA  ACCGCAGCTG  TCGACCAGGG  CAACTCGGCC  GGCTTAAGTT  TGGCGTGCAG  AGCTCTGCCC  
-3  N L A K R L A S N L G C S D V L

201  GGCATTCCGC  CTGCGCGTCC  TCCTCGAAGC  ATTCGGCGAT  GCCGAAGCTG  CCGTATTTCG  TAAGCCCGTA  CGGCTTTGAC  
-3  C E V E G D F E C E A M A F S

251  CTCTCGTGCA  CCCGAACGAC  GTCGAAAAAG  CTAATATCCG  CAGGCAGGTG  ACACTCAGCA  GTCGAAAGTT  GATCTGCGGA  AACGCTATCC  GCCGAGCGGC  
-3  D E T V R V D F L S I D A A P K

301  GCCCCAAGCGC  ACGCCACCCTG  TGCGACCGCG  CACGTTTTCG  ACCAGACCG  CGGGTTCGCC  TGCCAAAGGC  TGGCTCGGCC  AGCTTGCGAC  
-3  G L R V G G N R G R V T E V L G

351  CCTGTTAGCC  CGGCTGAAGG  ATCTTGAAAA  GAAAGAAGCT  GCCGAGCGGC  CCGTATTTCG  TAAGCCCGTA  CGGCTTTGAC  
-3  K N L P Q L I K F L F L E S V G

401  TAGGCCCTGG  CGATTTCCGG  AATCCGGCTC  AAGTGCCCGT  CGTTGGACAG  ATCCGGGACC  GCTAAAGGCC  TTAGGCCGAG  TTCACGGGCA  GCAACCTGTC  
-3  Y A R A I E P R S L H G D N S L

451  CACAGTACAC  AAGTCGCGGC  CGCATAGTCC  GCTCTGCTTC  GTCAACCGCA  GTGTTATGTA  TTGACGCTKG  GCGTATCAAG  CGAGACGAAG  CAGTGGCGGT  
-3  V I C L Q S G C L E S Q K T L R M  aau3

501  TGCCCCCTCG  CTGCACTCGT  GTCTTGCGAC  CBVTVTACGG  GGTGTCSTT  ACCTGAAAGGC  GCAAGTCTCG  GVBABATCCG  CCAACGCGAT  
-3  Y A R A I E P R S L H G D N S L

551  GTCGGCGGCA  GTCCATCTCC  GCTCCAGGCT  GCTGATKTAS  YCGAGGATGG  AAACGGCCTG  CCCAGTTGGG  CGAGCTCTGC  GAATAMATS  RGCTCTTACC  
-3  V I C L Q S G C L E S Q K T L R M  aau3
651  CTCGCCTGGG CTCGGGCATG CCTTGCACAT CAGTTTGACC TCGGCACTGG
GAGCGGACCC GAGCCCCCAG GGAACGTGTA GTCAAACTTG AGCCGCTGACC
701  GGGAAATAGGA GATGTCGAGC AGACCGCGCC CTTGTTTCAG GACATCCTCG
CCCTATCTCT CTACAGCTCG TCTGGCGCGG GAACAAAGTC CTGTAGGAGCC
751  GGGAATAGGA GATGTCGAGC AGACCGCGCC CTTGTTTCAG GACATCCTCG
CCCTTATCCT CTACAGCTCG TCTGGCGCGG GAACAAAGTC CTGTAGGAGCC
801  GCCAAAAGGCC CCTCCCAATTA CAGATTTCAG ATTTGGAATGT CGTGAAGATC
CGGTATTCCG AGGGTTATAAT GTCTAAAAGTC TAACCTTACA GCCACTTCTAG
851  ACTTCATCGT CGGATCGAGC AATTCAGCGC CGCTCTTGAT GCCCGAGGGC
TGAAGTAGCA GCCGTACTGC TTAAGTCGCG GCGAGAACTA CGGGCTCCCG
-1  K M T P M V F E A G S K I G S P
901  CAGCGGGCGCG TACAGGTCTT GTGTTCTCTCG CAGAAGTGCTA TCGAATCGCT
GTCGGCCCTGG ACTGCCAGAA CCAGAAGCAG GTCTTGAACT AGCTTAGGCA
-1  W R A T V T K T K T W F K I S D T
951  GCCGTGCTGG TTGAGGTCGC CGAAGCTCGA GGCCTTCCAG CCGCCGAAGG
CGGACGAGCC AACTCCGCGG CTGTCAGACT CGGGAGCTTC GCCGGCTTCC
-1  G H Q N L D G F S S A K W G G F
1001 AGTGCGAAGC GACCGGAAACC GGGATCGGGAG CTTTGATGCA ATCATGTGCC
TCACCAATCGT CTCGCTTCGG CCTAGCCCTGC GCAACTACCG CTAGTACGCC
-1  H Y A L P V I P V N I G M I
1051 ATATTTGATGC GTGAGGCAAAG ATCGCGGGCG GCATCGCCGT CACGCGTGAA
TATAAACTAG CACTCGGTTT TAGCGCCCGG CGTACCGCGA GTGCGCAGTT
-1  I N I R S A F D R A D G D R T F
1101 GATGGCGACG CCGTTGCCAT ATTCGAGCAGC ATGCGCCGCT CAGCGCCGCT
CTACCGCTTCG GGAAGCGGATG TAGGCAGCTG CTGTCATGG GTGCGGCTGG
-1  I A V G N G Y E H K M P L S L A
1151 CTCGCTTAGT CGTGCCGAGA CGAGAGGGAGA GGACAGGCTCC CAGGATCTCG
GGAGATCAGA GACCCGCTCC TGCTGCTCTC CTGTCAGCGG CCGTCAGGCC
-1  E Y N Q A R V V S L V P G F I E
1201 GTCTTATAGA TATCCATCTG AGGGCGTACAG TGATCGCAACA GCCAAACCGCC
CAGAATATCT ATAGGTACAG TCCGACACTG ACTAGCTCTG CCGTGGCGGG
-1  T K Y I D M D P T V H D F L C G G
1251 GACGAACTAGG CCGTCCTTGCA AGCCCTGGAG TTTGAATCGC CGGCCGCTGA
CTGCTCTTCAT CGCGAGAGTG TCGGAGCTTC AAAGTCTAGG CCGCGCACGT
-1  V F Y G D E Y G Q L K F D R G D
1301 CGACGAGCTTG GGCACGTCGCC TCGATCGCGC GGTGATCGCG GCCGCGCGAGC
GCTGCTCCGA CCAGGGAGGG AGCTAGCCGC CCGACTGTGC CGGGCGCTGT
-1  V L K A G E I G R D I L G R V
1351 CGGGTATAGG CTTCCCTTGGT AAGGCAGGGG CCCATGTCGG CCGTGCAGCT
GCCCATATCC GAAGGAAACCA TTGCTGCGCC GGTACAGG GGAACAGGAG
-1  R T Y A E K T V L P G M D A K D D
Letters in red indicate amino acids. For amino acid code, see Appendix 3.1. Start and stop codons for each ORF are bold and underlined. Putative SD sequences are bold.
Sequencing of the end of the pRU482 insert DNA using a primer from the vector revealed that approximately 4kb upstream of the start of iolA, there were two partial ORFs. BlastX analysis of the deduced amino acid sequence of one ORF revealed homology to gluconokinases. Highest identity was with the \textit{E. coli} thermosensitive gluconokinase 2 (GTNK). There was 54\% identity (p-value of $6.2 \times 10^{-13}$) from the start of the protein to amino acid 63 of 187. The ORF was designated \textit{gtnK}.

Ending 20bp upstream of the start of \textit{gtnK}, the other partial ORF had identity with the ATP binding protein component of dipeptide or oligosaccharide ABC transporter systems (Dpp). Highest identity was 67\% with DppF of \textit{E. coli} (p-value of $1.0 \times 10^{-20}$). The identity was from amino acid 281 to 233, the end of the protein.

When the other end of the 10kb fragment of pRU482 was sequenced, the deduced amino acid sequence of the 213bp \textit{PstI} fragment had identity to the middle part of DppF of \textit{E. coli}. There was 66\% identity (p-value of $1.0 \times 10^{-19}$). The sequence identity was from amino acid 209 to 279 of 334. Based on the sequence homology, it was postulated that the true position of this fragment should be at the other end of the 10kb fragment, as the amino acid sequence followed on immediately to the above sequence. The 213bp fragment was not naturally linked to the adjacent 2.5kb \textit{PstI} fragment because the homology to DppF ended abruptly at the next \textit{PstI} site. The deduced amino acid sequence of the DNA from the next \textit{PstI} site had 75\% identity (p-value of $4 \times 10^{-3}$) homology over the first 133 amino acids of 334 to DppB of \textit{E.coli}. Sequencing
of another PstI fragment detailed below contained sequence that was identical to this sequence, indicating that the 2.5kb fragment is also not naturally in that position. When subcloning non-specific fragments, it is possible that fragments will re-ligate to each other in the wrong order, as the ends are identical. Based on restriction mapping of pRU3111 and sequencing of another subclone, pRU544, it is postulated that the 2.5kb fragment was also created unnaturally and that this PstI fragment should have been much bigger.

Two further fragments that were subcloned from pRU3111 yielded partial ORFs whose deduced amino acid sequence had homology with other DPP proteins. A second 7.5kb PstI fragment was cloned into pBluescript SK−. The resulting plasmid was designated pRU556. Sequencing of 500bp of each end of the insert in pRU556 revealed that this was not the same fragment as that in pRU482. At one end of the fragment the deduced amino acid sequence revealed a gene with 66% identity (6 x 10^{-24}) from amino acid 115 to 270 of 339 to DppB of E. coli. At the other end of the insert there was no homology with known sequences.

A 2kb EcoRI fragment was also cloned from pRU3111 into pBluescript SK−. The resulting fragment was designated pRU544. Sequencing of approximately 500bp of each end of the insert revealed DNA whose deduced amino acid sequence had homology to DppB and at the other end with DppC. There was 51% identity (p-value of 3.5 x 10^{-28}) over the first 110 amino acids of DppB of E. coli. There was 67% identity (p-value of 1.2 x 10^{-48}) over the
final 149 amino acids (152 to 300) of DppC of *E. coli*. Restriction mapping and sequence homologies to DppB indicate that this *Eco*RI fragment overlaps the *Pst*I fragment of pRU556 (Figure 3.8).

Restriction mapping of pRU556 and pRU544, along with homology of the deduced amino acid sequences with known genes meant that it was possible to postulate the relative positions of pRU556, pRU544 and pRU482. A restriction map of part of pRU3111 with putative gene positions marked is shown in Figure 3.10.

Apart from *iola*, none of the putative genes identified had any homology with ones involved in *myo*-inositol utilisation. For this reason, these subclones were not investigated further during this project.
3.2.3.3 RU307

An intact Tn5 was cloned from RU307 as an approximately 7900bp EcoRI fragment, with approximately 2100bp of chromosomal DNA flanking both ends of the transposon into the cloning vector pACYC184 by co-workers in the laboratory (Figure 3.13). The resulting plasmid was designated pRU426.

The upstream region of Tn5 consisted of 700bp of chromosomal DNA, which was sequenced on both strands. Downstream of Tn5 were approximately 1400bp of chromosomal DNA. Sequencing was carried out for approximately 500bp from either end of this region of chromosomal DNA. A region of approximately 400bp in the middle was not sequenced. A 9bp overlap was created in RU307 by insertion of Tn5 (691-699bp). The sequence that was repeated is highlighted in Figure 3.14.

BlastX analysis of the deduced amino acid sequence revealed that the transposon interrupted an ORF with identity to the ATP binding component of ABC transport systems involved in the uptake of D-galactose and methyl-galactoside (Mgl). The highest homology was 54% identity over the entire sequence of MglA of E. coli (p-value of $2 \times 10^{-71}$), with a gap of approximately 400bp in the middle of the gene as the entire region was not sequenced. The gene encoded by the ORF was designated intA (myo-inositol transport). A putative Shine-Dalgarno sequence was identified (Shine and Dalgarno, 1974). There was no stop codon present on pRU426, so it is postulated that the
putative gene is at least 200bp longer than its homologues. The gene is predicted to be transcribed in the same direction as Tn5.

Ending 5bp upstream of the putative start of the intA gene was an ORF whose deduced amino acid sequence had 75% identity (p-value of $1 \times 10^{-38}$) over the final 100 amino acids of MocB of S. meliloti and R. leguminosarum bv. viciae. This gene encodes a periplasmic binding protein that binds to rhizopines and is thought to be part of a rhizopine transport system. The gene encoded by the ORF was designated intB and is predicted to be transcribed in the same direction as intA. There was also high homology with D-galactose binding proteins. This was also the case for the myo-inositol binding protein of a Pseudomonas species (Deshusses and Belet, 1984). Plasmid pRU426 did not contain the beginning of the gene. The sequence of pRU426 is presented in Figure 3.14.

On the basis of the sequence data, it was postulated that RU307 is mutated in an operon of genes encoding an ABC transport system that might be responsible for uptake of myo-inositol or a derivative. This led to the discovery during this project that RU307 was severely impaired in the ability to utilise myo-inositol as the sole carbon source (c.f. Section 3.2.1).
Approximately 7900bp EcoRI fragment from RU307 containing intact Tn5 and flanking chromosomal DNA cloned into pACYC184.

The size of the plasmid is approximate, as the entire chromosomal insert DNA was not sequenced.
Figure 3.14 Sequence of pRU426

**intB**

+1
1 E F D A V I S N N D E M A I G A

+1
1 GAATTCGACG CCGTGATCTC CAACAAACGAC GAAATGGCGA TCGGCGCCAT

+1
Q A L K A A G K D M T K V V V G

51 CCAAGCGCTG AAGGCCGCCC GCAAAGATAT GACCAAGGTGC GTGCGTCCGGC

+1
G V D A T Q D A L A A M Q A G D L

101 GTGTCGATTGC GACGAGAACGC CGGCCTGCCC CGCAAGGATAT GGGCGATCTC

+1
D V T V F Q D A A G Q G K G S L

151 GAGCAGCTAGG TGTGCCAGCA TGGCCGCCCG CAGGGCAAGG GCTCTCTCGA

+1
A A L K L A K G E K I E K V Y

201 TCAGCCGGCTA AAGGAGGGAC GCTGGGGCGG GCCATGGGCT ACCAGCTGGC

+1
I P F Q L V T P A N V K D F V T K

251 TTCCCTTCCA GCTGCTACGG CCTGCCAACG TCAAGGACTT CGTCACCAAG

+1
M P H R M R A L P I L F R

301 AAC TAAGGGCC CAGT CACAC AAGGATCGGG GCTTGGCCC ATCCCTCCC

+1
L P V S G G D D M A V S P T T M A

351 TTACCGGCTA CCGGAGGAGG AGATATGTGC GCAGGCGGG GCTATACCGG

+1
A V R A S G A V P N A E Y L S

401 CGCCGCACGG CGAGGCACGG CAGTCCCCGA AGCAGGGATT CGTTGGACGG

+1
A E G V R K E F P G V V A L D D V

451 CGCAAGGCTT CCGAGAGGA A TTCTGTGTC GTGCTGAGTC GCAGGATGG

+1
Q F R L K R A S V H A L M G E N G

501 CAGTGGGCCT TAAAGGCGCC CTTGGGCTAC GCAGAGATGC GAGGAAACCG

+1
A G K S T L M K I L A G I Y T P

551 CGCCGCCCAA TCGACATTTG TGAAGATCCT CCGCCGCTAC TATAGCCGGG

intA

+3
N SD

301 AAC TAAGGGCC CAGT CACAC AAGGATCGGG GCTTGGCCC ATCCCTCCC

+3
L P V S G G D D M A V S P T T M A

351 TTACCGGCTA CCGGAGGAGG AGATATGTGC GCAGGCGGG GCTATACCGG

+3
A V R A S G A V P N A E Y L S

401 CGCCGCACGG CGAGGCACGG CAGTCCCCGA AGCAGGGATT CGTTGGACGG

+3
A E G V R K E F P G V V A L D D V

451 CGCAAGGCTT CCGAGAGGA A TTCTGTGTC GTGCTGAGTC GCAGGATGG

+3
Q F R L K R A S V H A L M G E N G

501 CAGTGGGCCT TAAAGGCGCC CTTGGGCTAC GCAGAGATGC GAGGAAACCG

+3
A G K S T L M K I L A G I Y T P

551 CGCCGCCCAA TCGACATTTG TGAAGATCCT CCGCCGCTAC TATAGCCGGG
+3 D K G D I R L K G I E I Q L K S P
601 ACAAGGGCGA TATTCGCCCTG AAAGGGATCG AGATCCAGCT GAAATCTCCG
TGTTCCCCGCT ATAAAGCGGAC TTTCCCTAGC TCTAGGTCGA CTTTAGAGGC

9bp overlap by Tn in RU307
+3 L D A L E N G I A M I H Q E L N L
651 CTCGACGCAC TCGAAGATCGG GATTGCCATTG ATCCATCAGG AGCTGAAACCT
GAGCTGGCTG AGCTTCCAGT CTAACGGTAC TAGTGTCCG TCGACTTTGGGA

+3 M P F M T V A E N I W I R R E P
701 GATGCCGTTC ATGACGCGTTT CCGAAATTAT CTGGATTCGC CGCGAAGCGA
CTACGCAGAG TACTGCAATC GACCTAAGCG GCCGACTGGCT

+3 K N R L G F I D H G V M H R M T E
751 AGAACCAGCT CCGTTTCTAC GATACGGCGG TGATCGACCA CATGACCCAGAG
TCTTGGCGGA GCCAAAGTAG CTAGTGCCCG ACTACGTGGC GTACGAGGGCT

+3 F L S V A R Q M V E I A K A V
801 ATTCCTCTCG GCCTGCCAACCA GGCAGATGTTG CGAGATGCCC AACGGGCTTTT
TAAGAGGAGC CAGGCGGTTG CGCTCTACCA GCTCTAGCGG TTACGCCAAA

+3 S Y N S D V L I M D E P T S A L T
851 CCTATAATTC CGATGTGCTG TACATGGACG AGCCACCTTG CAGCGCAGAC
GATGCCGTGC ATCAGCAGGC ATAGCCATAG CAGGCGGCTG AGCTCGGGGT

+3 E R E V E H L F R I I R D L K A Q
901 GACGCACCGA TCGACAGATT TTTGCTATCG ACGCCACGGA TCAACCCCGA
CTGCAGCTGC AGTGCCGATGA ATAGCCATAG CAGGCGGCTG AGCTCGGGGT

+3 G I G I V Y I T H K M N E L F E
951 GGGATCGCGG ATCGCTACCC TGCGGTCCGCG ATCCGCCATGA TGGTCGGGCG
CGGAGCTGGC TGCAGTGGGC GCTGCTGTAG TAGGCGTACT ACCAGCCCGG

+3 I A D E F S V F R D G R Y I G T H
1001 GCCCTCGACG GTCTCGCGTAC TCCAGCCGAGG GCAGCAGCAT ACGGCGTCCGCT
CAAGAGGCGG AAGGCACTGC GCCGTATATA GCCGTCGAGT

+3 A S T D V T R D D I I R M M V G R
1051 GCCCTGACG CACGCTAGGC TGCGGTCTAC GCTGCTGACTA AGCCGAGAGG
CGACGAGGGC GCCCTCGCGC TGGTCCGGCG TGGTCCGGCG TGCTGAGGAC

+3 E I T Q M F P K E E G A D R R G
1101 CGARATCCAC CARATGTTTCC CAGAAGGAGATG ACGGCGGTCGCT GCTYTAGTGG GTYTACAAGA GCTGCTTCTT TACCGCAGCC ACGGCGGTCGCT

+3 H A F T Q G F L S Q R R L Q K V S
1151 ATGCTTTCCG TCAAAGATTT TTTGTCCTACAC GGCGCCCTAC AAAAGTYYCC
TACGAAAGGC AGTCTCTAAA AAGCAGGTTG CGCCAGAAGT TTTTCCARAGG
Unsequenced region of approximately 400bp.

+3     G P P E A P A K K G K K S A R E N
1551  GGGCCCGTGC AGGCACCTGC GAAGAATGGC AAAAAAGTG CGCGTGAATAA
       CCGCGGCGAG TCCGTGGAGCG CTTCTTCACGG TTTTTTPCAG GCCGACTTTT

+3     A N L Y E R V E N L S G G N Q Q
1601  CGGCAAATCTT TACGACCGGG TGGAAATCTT TTCCGGCGCG AAAATCGACRA
       GGGTTAGAAT ATGCTTCGCC ACCCTTNTAG AAGCCCGCGG TTAGTCTGTYT

+3     K V L I G R W L L T N P R I L L
1651  AGGTGCTGAT CGGCCGCGTA CTGCTCACCA ATCCGCGCAT CCTCATCCTC
       ACGGCGTCTA GCCCGCGACC GACGAGTGGT TAGGCGCGTA GGAGTAGGAG

+3     D E P T R G I D V G A K A E I H R
1701  GACGAGCCGA CGGCCGCAT CGATGTCGCC GCCAAAGGCAG AATCCACCG
       CTGCTTCGCT GCCCGCGTA CTCTGCGCACA CGGCACCACC TTAGTGGCGC

+3     L V T E M A R D G V A V V M I S
1751  GTTGGTCACG GAGATGGCGC GAGACGGTGT GGCGGTGGTG ATGATCTCGT
       CAACCAGTGC CTCTACCGGC GTTGGTCACG GAGATGGCGC GAGACGGTGT

+3     S E M P E V L G M S D R I M V M H
1801  CGGAGATGCC TGAGGTTCTC GGGATGAGCG ACCGCATCAT GGTCATGCAC
       GGCTCTACGG ACTCCAAGAG CCCTACTCGC TGGCGTAGTA CCAGTACGTG

+3     E R V T G F L N R D E A T Q I K
1851  GAGGGACGCG TGACCGGTTT CCTCAATCGC GACGAAGCAA CGCAGATCAA
       CTCCCTGCGC ACTGGCCAAA GAGGTTAGCG CTGCTTCGTT GCGTCTAGTT

+3     V M E L V R S D A L T A I A Q G
1901  GGTGATGGAG CTGGTGCGCA GTGATGCGTT GACGGCGATC GCCCAAGGGA
       CCACTACCTC GACCACCGGT CACTACGGAA CTGCGCCGCTG CGGGTGCCCT

+3     R F D M T T K A A E G A A P L A T
1951  GGT7TGACAT GACTACCAAG GCACGAGAG GCGCGCGCTC CTGCGGCAACC
       CCAAACTGTG ATGATGCTGT CGTCTGCCTC CGCGCGAGG CAGGTTGGTG

+3     R Q R R R I P T E L S I F L V L
2001  CGGCAAAGGC GGGCCGCGCT ACCGACCGAG CTCAGCATTT TCCTCGTGCT
       GCCGTTCGCGG CCGCGGCGTT TGGCTGCCTC GAGTGCTAAA AGGAGCAGCA

+3     V G I A L I Y E V L G W M F I G
2051  CGTCCGCCATG GCCGTTCATCT AGCAAGTGCT TGGCTGGATG TTCACTCGGCC
       GCACCGGCTG TAGCCGAGTTA TGCTCTCAAG ACCGACCTAC AAGTAGCGCG

+3     Q S F L M N
2101  AAAGCTTCCT CATGAATTC TTTCGAAGGA GTACTTAAG

Letters in red indicate amino acids. For amino acid code, see Appendix

3.1. Start and stop codons for each ORF are bold and underlined.

Putative SD sequences are bold.
3.2.5 Uptake of Glucose and myo-Inositol

Results presented above show that RU307 was severely impaired in the ability to utilise myo-inositol as the sole carbon source and that this phenotype was linked to the transposon insertion. The DNA sequence of the region interrupted by the transposon had identity with components of ABC sugar transport systems. The deduced amino acid sequence of a partial ORF immediately upstream of the mutation resembled MocB of the rhizopine catabolic system and D-galactose binding proteins. A myo-inositol uptake mutant of *Pseudomonas* sp. JD34 also only grew very slowly on myo-inositol (Frey *et al.*, 1983). The myo-inositol binding protein of another *Pseudomonas* species also resembled D-galactose binding proteins (Deshusses and Belet, 1984). Therefore, it was hypothesised that RU307 is mutated in an uptake system for myo-inositol.

Previously, Poole *et al.* (1994), measured uptake of myo-inositol in 3841, RU360 and RU361, using myo-[2-^3^H]-inositol. The authors reported that there was constitutive uptake of radio-labelled myo-inositol by the strains but uptake was at a lower rate than glucose uptake. It was reported that in *K. aerogenes*, active transport of myo-[2-^3^H]-inositol is complicated by the presence of a dehydrogenase that attacks the substrate at C-2, causing loss of the label (Deshusses and Reber, 1972).

Therefore, ^14^C-labelled myo-inositol was used to measure uptake in 3841 and in the mutants RU360, RU361 and RU307 (Figure 3.15). The strains were
grown on 20mM pyruvate or 20mM pyruvate plus 10mM myo-inositol as the sole carbon sources. Uptake was also measured for 3841 and RU307 grown on 10mM myo-inositol alone. All uptake experiments were also carried out with $^{14}$C-labelled glucose, to determine whether the mutants were affected in their ability to transport glucose (Figure 3.16).

The uptake data were subjected to Analysis of Variance (Appendix 3.2). There was no significant difference in the rates of uptake of myo-inositol by each strain grown on 20mM pyruvate, RU360 and RU307 grown on pyruvate plus myo-inositol and RU307 grown on myo-inositol (P>0.1) (Figure 3.15). When grown on 20mM pyruvate and 10mM myo-inositol, there was a 2.5-fold increase in the rate of uptake of myo-inositol by 3841 and a 1.7-fold increase by RU361 compared with the rate of uptake when grown on pyruvate.

Although the increase in uptake by RU361 was significantly higher than RU361 grown on pyruvate alone, (P<0.05), it was still significantly lower than the rate of uptake by 3841 (P<0.05). There was a 7-fold increase in uptake by 3841 when grown on 10mM myo-inositol as the sole carbon source. These data indicate that there is an inducible system for the transport of myo-inositol in *R. leguminosarum biovar viciae* that is expressed in 3841 but not in the mutants.

Uptake of glucose occurred in all strains when grown either on 20mM pyruvate, 20mM pyruvate with 10mM myo-inositol, or 10mM myo-inositol as the sole carbon source (Figure 3.15). The data were subjected to Analysis of
Variance (Appendix 3.3). There was no significant difference in the rate of uptake of glucose by the strains on different carbon sources ($P>0.1$). These data indicate that the mutants were not impaired in the ability to transport glucose into the cell.
**Figure 3.15 Uptake of $^{14}$C-labelled myo-inositol**

Cells were grown in AMS containing either 20mM pyruvate (P), 10mM myo-inositol (I) or 20mM pyruvate plus 10mM myo-inositol (PI).

Mean uptake of myo-inositol per minute over four minutes. Mean results of three experiments with standard deviation values.
Figure 3.16  Uptake of Glucose

Cells were grown in AMS containing either 20mM pyruvate (P), 10mM myo-inositol (I) or 20mM pyruvate plus 10mM myo-inositol (PI).

P = 20mM pyruvate, I = 10mM myo-inositol.

Mean uptake of glucose over four minutes. Mean results of three experiments with standard deviation values.
The deduced amino acid sequence of the putative \textit{intA} gene of RU307 has high homology to MglA, a component of the galactose and methyl-galactoside uptake system. Therefore, the uptake system might not be specific to one compound. To determine the specificity of the transport system, the uptake of \textit{myo}-inositol by 3841 was measured after addition of a 5-fold excess of different carbon compounds (Figure 3.17). The uptake data were subjected to Analysis of Variance (Appendix 3.4). Addition of a 5-fold excess of unlabelled \textit{myo}-inositol significantly inhibited uptake of radio-labelled \textit{myo}-inositol \((P<0.05)\) but none of the other carbon compounds inhibited \textit{myo}-inositol uptake. There was no significant difference in the rate of uptake of \textit{myo}-inositol when different carbon compounds were added \((P>0.1)\).

The inhibition experiment was also carried out using radio-labelled glucose. The uptake system for glucose is well defined in bacteria, so no inhibition was expected. Unlabelled carbon compounds were added in 5-fold excess prior to addition of the radio-labelled carbon (Figure 3.18). The uptake data were subjected to Analysis of Variance (Appendix 3.5). Addition of a 5-fold excess of unlabelled \textit{myo}-inositol significantly inhibited uptake of radio-labelled \textit{myo}-inositol \((P<0.05)\) but none of the other carbon compounds inhibited \textit{myo}-inositol uptake. There was no significant difference in the rate of uptake of \textit{myo}-inositol when different carbon compounds were added \((P>0.1)\).

The results indicate that the inducible transport system for \textit{myo}-inositol uptake in \textit{R. leguminosarum} bv. \textit{viciae} is \textit{myo}-inositol-specific.
Figure 3.17 Uptake of myo-Inositol When Excess Carbon Compounds are Added

Cells were grown in AMS containing 10mM myo-inositol.

Mean uptake by 3841 of radio-labelled myo-inositol over four minutes when 5-fold excess carbon compound added. Mean results of three experiments with standard deviation values.
Figure 3.18 Uptake of Glucose When Excess Carbon Compounds are Added

Cells were grown in AMS containing 10mM myo-inositol.

Mean uptake of glucose over four minutes when 5-fold excess of carbon compound added. Mean results of three experiments with standard deviation values.
3.3 Discussion

There is a specific pathway of myo-inositol catabolism in \textit{R. leguminosarum} bv. \textit{viciae}, as the mutants RU360, RU361 and RU307 were all able to grow at similar rates to 3841 on several difference carbon sources, but RU360 and RU361 were unable to grow \textit{in vitro} when myo-inositol was the sole carbon source. Strain RU307 was able to grow on myo-inositol but was severely limited, as growth was at a quarter of the rate of 3841 on myo-inositol. The data also show that the presence of a transposon did not affect growth on other carbon sources. myo-Inositol did not inhibit growth of the mutants when present in the growth medium in addition to pyruvate or glucose, indicating that there are no toxic products produced by partial breakdown of myo-inositol by the mutants.

When complemented by cosmids, RU360/pRU3078, RU360/pRU3079 and RU361/pRU3111 grew at the same rates on glucose or myo-inositol as 3841. These data indicate that the ability to utilise myo-inositol was completely restored and also that the presence of a cosmid did not inhibit growth. The inability of the mutants and 3841 to grow when phytic acid was the sole carbon source, either in the presence or absence of phosphate, indicates that this compound is not a likely energy source in the rhizosphere for \textit{R. leguminosarum} bv. \textit{viciae}.

Sequencing of the genes interrupted by Tn5-\textit{lacZ} in the mutants RU360 and RU361 showed that the pathway for myo-inositol utilisation in \textit{R. leguminosarum} bv. \textit{viciae}. 
*Rhizobium leguminosarum* bv. *viciae* is not organised in a single operon, as occurs in *B. subtilis* (Yoshida *et al.*, 1997). There must be at least one other locus of *myo*-inositol degradation genes, as only putative genes involved in the final stages of *myo*-inositol catabolism were discovered. These genes may not be encoded on the chromosome, but on plasmids as occurs in *R. leguminosarum* bv. *trifolii* (Baldani *et al.*, 1992) and with rhizopine catabolism genes of *S. meliloti* (Murphy *et al.*, 1987) (c.f. Section 1). This could be ascertained by curing 3841 of plasmids and determining whether the ability to grow on *myo*-inositol as the sole carbon source is retained.

The discovery that the interrupted gene in RU360 may encode a protein that has high homology to acetolactate synthases, including *iolD* of *B. subtilis*, led to the hypothesis that the catabolic pathway as elucidated for *K. aerogenes* is not complete. Acetolactate synthase is required for the synthesis of acetolactate from pyruvate and acetyl CoA. This compound is the precursor of the amino acids valine, leucine and isoleucine. In *B. subtilis*, the product of the *iolD* gene is also essential for *myo*-inositol breakdown (Yoshida *et al.*, 1997). Therefore, it is proposed that the catabolic pathway requires an extra step (Figure 3.17).
Figure 3.17 Proposed Pathway of myo-Inositol Catabolism in *R. leguminosarum* bv. *viciae*

Putative functions of Iol proteins from *B. subtilis* are shown.

* The pathway of conversion of D-2,3-diketo-4-deoxy-epi-inositol to 2-deoxy-5-keto-D-gluconic acid has not been characterised. Conversion may occur spontaneously, or through the action of a hydratase.
*E. coli* and *S. typhimurium* have three acetolactate synthases that are all under different modes of regulation, but the results of one enzyme activity study suggested that *R. leguminosarum* bv. *viciae* has only one copy (Royuela *et al.*, 1998). If *iolD* encodes an acetolactate synthase that is essential for *myo*-inositol catabolism, then there must be more than one gene encoding acetolactate synthase in *R. leguminosarum* bv. *viciae*, as RU360 would not be able to grow on any carbon source without supplementation in the medium of valine and isoleucine.

Other acetolactate synthase genes must be under regulatory systems that are not induced by *myo*-inositol, otherwise they would be expected to compensate for the mutation in RU360. To test whether *iolD* is essential for *myo*-inositol catabolism in *R. leguminosarum* bv. *viciae*, just the *iolE* and *iolB* genes could be cloned and tested for complementation, or point mutation could be carried out on *iolD*, to eliminate any polar effect on downstream genes. It was shown that *iolD* of *B. subtilis* is essential (Yoshida *et al.*, 1997) and due to time constraints this was not considered a priority in this project. To test whether *iolD* is a 1680bp or 1881bp gene, PCR amplification of the two ORFs could be carried out in order to produce two different size clones, which could then be assayed for acetolactate synthase activity. The N-terminus of the purified protein could also be sequenced.

The lack of complementation by pRU706, which contains just the *iolD* gene and the complementation by pRU713, which contains all three genes, indicates that the mutation in *iolD* is polar on the downstream genes, *iolE* and
iolB. Therefore, at least two of these three genes probably share the same promoter and one or both of these genes is essential for myo-inositol catabolism. This concurs with data obtained from *B. subtilis* that *iolE* and *iolB* are essential for myo-inositol catabolism, although the specific functions of the genes are not known (Yoshida *et al.*, 1997). It is not currently possible to predict their role in myo-inositol utilisation, as other homologues in the databases have no known function. The requirement for *iolE* and *iolB* could also be tested by point mutation of the individual genes. It is not known whether *orf1* is required for myo-inositol utilisation, although testcode analysis indicated that this gene is unlikely to be coding again. This could be tested by mutation of the gene.

The deduced amino acid sequence of the DNA interrupted by Tn5- lacZ in RU361 had homology with methylmalonate semialdehyde dehydrogenases, including IolA of *B. subtilis*. Based on sequence homology, the putative function of the protein is malonic semialdehyde oxidative decarboxylase, an enzyme active in the final stages of the myo-inositol breakdown pathway. To determine whether just *iolA* would be required to restore the ability to utilise myo-inositol to RU361, the putative gene could be obtained by PCR amplification and conjugated into RU361 to see whether it would restore the ability to utilise myo-inositol. Due to time constraints, this was not attempted.

Further sequencing of the complementing cosmid pRU3111 revealed the presence of several putative genes. Upstream of *iolA* was a putative gene whose deduced amino acid sequence had homology to gluconokinases.
However, there was no significant homology with IolC of the *B. subtilis iol* operon, which is postulated to be a gluconokinase (Yoshida *et al*., 1997). Whether this putative gene is involved in *myo*-inositol breakdown could be ascertained by mutation of the gene in 3841 to see if *myo*-inositol catabolism is knocked out.

Downstream of the *iolA* gene was a putative gene whose deduced amino acid sequence had homology to Aau3 of *S. meliloti*. This is required for utilisation of PHB cycle intermediates. PHB is one of the principal carbon storage compounds in nodules. Other proteins in the databases that have identity with the putative gene have no known function. It is not thought that this gene has any role in *myo*-inositol utilisation due to its lack of homology with known *iol* genes, but this could also be investigated by mutation of the gene in 3841. Sequencing of other subclones of pRU3111 revealed ORFs whose deduced amino acid sequence had strong homology with components of the DPP operon. These proteins are not expected to be involved in *myo*-inositol utilisation. Hence, they were not investigated further.

A complementing cosmid was not obtained for RU307, so no information was obtained for genes surrounding the Tn5 insertion. Transduction of RU307 showed that both the *myo*-inositol and the toxic escape phenotypes were tightly linked to the transposon insertion. Without further sequence information and a complementing cosmid, it was not possible to elucidate the molecular basis of the glutamic acid gamma hydrazide phenotype.
It was hypothesised that RU307 might be impaired in the ability to utilise myo-inositol because the deduced amino acid sequence of the putative gene upstream of the Tn5 mutation had homology with putative rhizopine and D-galactose binding proteins. *A Pseudomonas* species myo-inositol binding protein also has homology to D-galactose binding proteins (Deshusses and Belet, 1984) and a *Pseudomonas* sp. JD34 myo-inositol uptake mutant was impaired in growth on myo-inositol, but not completely unable to grow. This evidence strongly suggests that RU307 is mutated in an uptake system for myo-inositol.

Uptake experiments using radio-labelled myo-inositol showed that 3841 has an inducible system for the uptake of myo-inositol that the mutants RU360, RU361 and RU307 do not express. This transport system in *R. leguminosarum* bv. *viciae* is myo-inositol specific, as there was no inhibition in 3841 by an excess of a variety of different carbon compounds, including D-galactose. The ATP binding component of ABC transport systems are highly conserved, so it was not unexpected that the transport system was not actually a D-galactose transporter, despite the high homology. If the transport system were responsible for transporting glutamic acid gamma hydrazide into the cell, then an excess of this compound would have been expected to block transport of myo-inositol in 3841, but it did not. Therefore, the uptake system mutated in RU307 is not a high affinity glutamic acid gamma hydrazide transport system. It is not possible from results gained during this project to explain the ability of RU307 to grow in the presence of glutamic acid gamma hydrazide.
The inducible system for uptake of *myo*-inositol by 3841 was repressed by the presence of pyruvate in the growth medium. This suggests that *myo*-inositol is not preferentially utilised when other carbon sources are present. However, this might be specific to pyruvate, as uptake was not tested in the presence of other sugars. The mutants were all able to transport glucose into the cell at the same rate as 3841, indicating that there was no deficiency in uptake of this sugar.

There was a low rate of uptake of *myo*-inositol into the cell in all the strains when grown on pyruvate. This suggests the presence of a constitutive system for *myo*-inositol transport which may not necessarily be *myo*-inositol specific. There are two uptake systems in *P. putida*, one high affinity system involving a periplasmic binding protein and a low affinity system not associated with a binding protein (Reber *et al.*, 1977). The presence of this second system may explain the slow growth by RU307 on *myo*-inositol. As the rate of uptake was so much lower, the growth rate was also greatly reduced. Therefore, it is likely that RU307 is mutated in a *myo*-inositol uptake system. If one or more of the final breakdown products of *myo*-inositol induce uptake, this would explain why there was no induction of uptake in RU360 and RU361 as they were blocked in the breakdown of *myo*-inositol. To locate the second uptake system would require a second mutation to be introduced into RU307.

Bahar *et al.* (1998) reported that 3841 can catabolise rhizopines when a plasmid containing *moc* genes is introduced. Although there is a rhizopine binding gene (*mocB*) amongst the *moc* genes, there is no known transport
system. Therefore, it was hypothesised that rhizopines entered the cell using a chromosomally encoded transport system. This could be the system mutated in RU307. The moc gene plasmid could be introduced into RU307 and the strain tested for the ability to grow on rhizopines as the sole carbon and nitrogen source. It is postulated that RU307 will grow very slowly, or not at all.

In order to fully elucidate the myo-inositol catabolic pathway in *R. leguminosarum* bv. *viciae*, several more genes need to be identified. An attempt was made to complement a *S. meliloti* myo-inositol dehydrogenase (*idh*) mutant using the 3841 cosmid library in the hope of isolating a region of DNA containing a *R. leguminosarum* bv. *viciae* myo-inositol dehydrogenase gene. However, this was unsuccessful. Another strategy would be PCR amplification using primers designed from the *S. meliloti idh* gene. This might also be possible using primers from the *B. subtilis* myo-inositol dehydrogenase gene *iolG*, as based on results gained in this project, this gene is likely to have high homology with *R. leguminosarum* bv. *viciae* genes.

**CHAPTER 4 - NODULATION AND RHIZOSPHERE GROWTH ............. 142**

4.1 Introduction ........................................................................................................... 143

4.2 Results ..................................................................................................................... 145

4.2.1 Rate of Nodulation ............................................................................................... 145

4.2.1.1 Dry Weight of Vetch Plants ........................................................................ 149

4.2.2 Acetylene Reduction Assay on Pea Nodules .................................................... 151
4.2.3 Plant Dry Weights for Pea................................................................. 155
4.2.4 Nodule Number and Mass for Pea.................................................. 157
4.2.5 Competition for Nodulation................................................................. 161
  4.2.5.1 Competitive Ability of the Complemented Mutants...................... 170
4.2.6 Measurement of Nodule Co-Residence .......................................... 173
4.2.7 Average Number of Nodules per Vetch Plant .................................. 176
  4.2.9 Rhizosphere Growth Assay............................................................... 186
4.2.10 Analysis of a $\text{nodC-phoA}$ Fusion...................................................... 192

4.3 Discussion........................................................................................... 195
Chapter 4 - Nodulation and Rhizosphere Growth
4.1 Introduction

myo-Inositol is present in soil and in the rhizosphere (Lynch et al., 1958, McKercher and Anderson, 1968, Sulochana, 1962, Yoshida, 1940, Wood and Stanway, 2000). It is also found in plant tissues, including legume root nodules (Kouchi and Yoneyama, 1984, 1986, Skøt and Egsgaard, 1984, Streeter, 1987; Streeter and Salminen, 1985). There are many micro-organisms competing in the rhizosphere for resources. Therefore, there may be a competitive advantage for bacteria that are able to utilise myo-inositol or derivatives such as rhizopines.

The results of research carried out to date are mixed. Heinrich et al. (1999) reported that a rhizopine producing strain of S. meliloti occupied more nodules than a non-producing strain when co-inoculated into the rhizosphere of alfalfa. The competitive advantage remained over a period of four years. However, Bosworth et al (1994) and Scupham et al. (1996) found that disrupting a locus of inositol genes sometimes enhanced competitive ability. Raggio et al. (1959) found that supplementation of the growth medium of bean with mesoinositol (myo-inositol) resulted in an increase in the percentage of nodulated roots and nodules per root of R. leguminosarum bv. phaseoli. The reasons for these effects are unknown.

In Chapter 3, it was shown that the myo-inositol mutants are able to grow in vitro on different carbon sources at the same rate as 3841 and that the presence of myo-inositol in the culture media does not inhibit growth. RU360
was previously shown to nodulate vetch and reduce acetylene at the same rate as 3841 (Poole et al., 1994). It was decided to test whether the myo-inositol mutants RU360, RU361 and RU307 would be at a competitive disadvantage when co-inoculated into a plant rhizosphere with 3841 in order to determine whether the ability to utilise myo-inositol is important for rhizosphere growth and survival.

Experiments were carried out in the rhizosphere of pea (P. sativum) and vetch (V. sativa) plants, both of which R. leguminosarum bv. viciae nodulates. Vetch was used primarily as the plants are smaller and so are easier to cultivate in large numbers. It would also have been impractical to analyse every nodule on pea plants. The plants were grown in sterile vermiculite or water agar with a defined nutrient solution, to ensure that no exogenous factors could affect results and to ensure standardisation between experiments.

Several aspects of growth in the rhizosphere were considered for the mutants and 3841. These were the ability to grow in the rhizosphere, nodulate plants, nitrogen fixation rates, average plant mass, average number of nodules and nodule mass per plant and the time taken for nodules to form. The ability of the mutants to successfully form nodules when co-inoculated into the rhizosphere with 3841 was also investigated. The ability to respond to plant signals was measured using the nodC gene fused to the phoA gene.
4.2 Results

4.2.1. Rate of Nodulation

It was shown previously that RU360 can nodulate vetch plants (Poole et al., 1994). This needed to be established for the mutants RU361 and RU307 and whether they were affected in the rate or number of nodule formation compared with 3841. Vetch seedlings were grown in boiling tubes containing 0.2% water agar supplemented with nitrogen-free rooting solution. This enabled nodule formation to be observed without disturbing the plants. Representative plants and roots inoculated with each strain are shown in Figure 4.1 and Figure 4.2.

Seven day-old seedlings were inoculated with $10^3$ cfu of RU360, RU361, RU307 or 3841. Each strain was inoculated onto eight plants. Eight plants were uninoculated. Nodules were first observed on at least one plant per replicate four days post-inoculation (Figure 4.3). Nodule numbers increased rapidly until 11 days post-inoculation after which there was just a slight increase in nodules of RU361 and 3841 17 days post-inoculation. The increase in nodule numbers was similar for the mutants and wildtype throughout the experiment. No nodules were observed on any of the uninoculated plants. The plants were harvested 21 days post-inoculation. The data were subjected to Analysis of Variance (Appendix 4.1). There was no significant difference in the overall average number of nodules per plant.
(P>0.1). All nodules appeared pink in colour, which signifies the presence of leghaemoglobin and suggests that the nodules were actively fixing nitrogen.

**Figure 4.1 Vetch Plants Grown in Water Agar**

A representative vetch plant inoculated with each strain as indicated. The plants were harvested three weeks post-inoculation.
Figure 4.2 Roots of Vetch Plants

Roots of a representative vetch plant inoculated with each strain as indicated. Nodules are indicated by arrows.
Figure 4.3  Average Number of Nodules per Plant Formed on Vetch

Each point represents the mean of eight plants.
4.2.1.1 Dry Weight of Vetch Plants

Twenty-one days post-inoculation, the shoots of the plants were harvested, dried and weighed to see if there were any differences between plants inoculated with different strains (Figure 4.4). The data were subjected to Analysis of Variance (Appendix 4.2). The uninoculated plants weighed significantly less than plants that were inoculated (P<0.05). There was no significant difference in average dry weight between plants inoculated with the mutant strains or 3841 (P>0.1).
Figure 4.4  Average Plant Dry Weight for Vetch

Each result represents the mean of eight vetch plants with standard deviation.
The results presented above indicate that the mutants are able to form pink nodules on plants. In order to confirm that the nodules were effective, it was necessary to determine whether nitrogen fixation was taking place. The acetylene reduction assay was used to indirectly measure nitrogenase activity of bacteroids in nodules. The nitrogenase enzyme is able to reduce acetylene to ethylene at a rate that is proportional to the amount of nitrogen that can be fixed, giving results in moles of acetylene reduced. This is not a direct measurement and there are many factors which may cause erroneous results (reviewed in Giller, 1987). However, it is useful for establishing that the nitrogenase enzyme is active and is easier and cheaper than direct measurement techniques.

Fifteen pea plants were inoculated with $10^6$ cfu of 3841, RU361 or RU307. Peas were chosen for this experiment as vetch are so small that precise measurements per plant would be more difficult to obtain. Representative plants are shown in Figure 4.5. Three uninoculated plants were randomly placed amongst the inoculated plants to test that cross-contamination did not occur. The plants were harvested four weeks post-inoculation. Each plant had pink nodules on the roots, confirming the ability of the mutants to nodulate peas. The uninoculated plants had no nodules and were not analysed further. The acetylene reduction assay was then performed. The plants were assayed in five jars, each containing three plants. The average amount of acetylene reduced per plant was then calculated (Figure 4.6).
The data were subjected to Analysis of Variance (Appendix 4.3). There was no significant difference in the rates of acetylene reduction between RU361 and 3841 (P>0.1). Strain RU307 reduced more acetylene than RU361 and 3841 but this was not statistically significant (0.1>P>0.05). It was shown previously that RU360 reduced acetylene at the same rate as 3841, so this was not measured again (Poole et al., 1994).
Figure 4.5  Pea Plants Used for the Acetylene Reduction Assay

Representative pea plants inoculated with each strain as indicated.
Figure 4.6 Acetylene Reduction Rates

Each series represents the average result of 15 pea plants assayed in duplicate or triplicate with standard deviation.
4.2.3 Plant Dry Weights for Pea

Plant dry weights were obtained for the pea plants that were used in the acetylene reduction assay (Figure 4.7). The data were subjected to Analysis of Variance (Appendix 4.4). There was no significant difference in plant dry weight between 3841, RU361 and RU307 (P>0.1).
Figure 4.7  Average Pea Dry Weight

Each series represents the average of 15 pea plants with standard deviation.
4.2.4 Nodule Number and Mass for Pea

The data for average total nodule number per plant and average total wet nodule mass per plant were subjected to Analysis of Variance (Appendix 4.5, Appendix 4.6). The mutants RU361 and RU307 formed significantly more nodules per plant than 3841 (P<0.05) (Figure 4.8). However, there was no significant difference in the average total nodule mass per plant between the mutants and 3841 (P>0.1) (Figure 4.9). These data suggest that nodules formed by the mutants are smaller and less efficient at nitrogen fixation but that the plants compensated by permitting more nodules to be formed. This is in contrast to vetch, where nodule numbers were the same (Figure 4.3).
Figure 4.8  Total Average Nodule Number per Pea Plant

Each series represents the average of 15 pea plants with standard deviation.
Figure 4.9  Total Average Nodule Mass per Pea Plant

Each series represents the average result of 15 pea plants with standard deviation.
4.2.5 Competition for Nodulation

The results presented above indicate that the mutants nodulate vetch and pea plants and fix nitrogen in nodules on pea as effectively as 3841. It was decided to examine whether the mutants would be just as competent at initiating nodulation when co-inoculated with 3841 into the rhizosphere of host plants, or whether 3841 would have a competitive advantage, due to its ability to utilise myo-inositol.

The mutants RU360 and RU361 were co-inoculated with 3841 into the rhizosphere of pea in equal numbers. Each strain was also inoculated alone. The inocula were each applied to four or five plants. An inoculum size of $10^3$ cfu was chosen to enable growth in the rhizosphere prior to nodulation. Four plants were uninoculated, to test whether cross-contamination occurred. Twelve nodules were harvested from each plant six weeks post-inoculation and stabbed onto TY agar containing different combinations of antibiotics, to distinguish between nodules containing the mutants and 3841 (Figure 4.10). None of the uninoculated plants had any nodules. Representative colonies were then streaked onto AMA containing 10mM myo-inositol to confirm the phenotype of each strain.

When inoculated alone, all of the nodules contained the strain inoculated, as defined by the ability to grow in the presence of the appropriate antibiotics. These results signify that there was no curing of antibiotic resistance genes from the strains. When co-inoculated in equal numbers with 3841, RU360
occupied 12.5% (6 of 48) of the nodules. At least one nodule containing RU360 was obtained from each plant harvested. RU361 occupied none of the nodules when co-inoculated with 3841. These data indicate that 3841 has a substantial advantage over the mutants in forming nodules when co-inoculated.
Figure 4.10 Pea Nodule Occupancy

1 = uninoculated, 2 = 3841 $10^3$ cfu, 3 = RU360 $10^3$ cfu, 4 = RU361 $10^3$ cfu, 5 = 3841 $10^3$ cfu + RU360 $10^3$ cfu, 6 = 3841 $10^3$ cfu + RU361 $10^3$ cfu.
The competition experiment was carried out on vetch plants, with equal numbers of 3841 with RU360, RU361 and RU307 (10^3 cfu) co-inoculated into the rhizosphere. Based on the results obtained on peas, it was anticipated that 3841 would have a competitive advantage. Therefore, plants were also co-inoculated with 10-fold and 100-fold more of the mutants than 3841 (10^4 cfu and 10^5 cfu) to investigate whether the mutants would occupy proportionately more nodules. Each inoculum was applied to at least six plants. All the nodules from each plant were harvested five weeks post-inoculation.

As expected, when inoculated alone, all nodules contained the appropriate strain (Figure 4.11). When co-inoculated in equal numbers with RU360, 3841 dominated the nodules, occupying 97.6% (82 of 84) of nodules. The nodules containing RU360 were obtained from two plants. Strain 3841 maintained the dominance when higher numbers of RU360 were inoculated as it occupied 99.4% (175 of 176) of nodules when RU360 was in 10-fold excess. When RU360 was in 100-fold excess, 95.9% (186 of 194) of nodules were occupied by 3841. Nodules containing RU360 were obtained from three plants.

The same trend was observed with RU361 co-inoculated with 3841 (Figure 4.12). No nodules contained RU361 when co-inoculated in equal numbers (0 of 116). The dominance was maintained by 3841 when higher numbers of RU361 were inoculated. Strain 3841 occupied 89.5% (171 of 191) of nodules when RU361 was in 10-fold excess. At least one nodule containing RU361 was obtained from each plant harvested. When RU361 was in 100-fold
excess, 98.1% (156 of 159) of nodules were occupied by 3841. Nodules containing RU361 were obtained from two plants.
Figure 4.11  Nodule Occupancy of Vetch by RU360 and 3841

1 = uninoculated, 2 = 3841 $10^3$ cfu, 3 = RU360 $10^3$ cfu, 4 = 3841 $10^3$ cfu + RU360 $10^3$ cfu, 5 = 3841 $10^3$ cfu + RU360 $10^4$ cfu, 6 = 3841 $10^3$ cfu + RU360 $10^5$ cfu.
Figure 4.12  Nodule Occupancy of Vetch by RU361 and 3841

1 = uninoculated, 2 = 3841 10³ cfu, 3 = RU361 10³ cfu, 4 = 3841 10³ cfu + RU361 10³ cfu, 5 = 3841 10³ cfu + RU361 10⁴ cfu, 6 = 3841 10³ cfu + RU361 10⁵ cfu.
The dominance exhibited by 3841 when co-inoculated with RU360 and RU361 was not maintained when co-inoculated with RU307 (Figure 4.13). At least nine plants were treated with each inoculum. When co-inoculated in equal numbers, RU307 occupied 47.1% (81 of 172) of the nodules. When inoculated at 10-fold and 100-fold higher than 3841, RU307 occupied 69.3% (131 of 189) and 80.9% (148 of 183) of the nodules respectively. At least one nodule containing RU307 was obtained from each plant harvested.

The data presented above indicate that the myo-inositol catabolic mutants RU360 and RU361 suffer a severe disadvantage in competition for nodulation when co-inoculated with 3841. However, RU307, which is not a complete myo-inositol mutant, is competitive when co-inoculated with 3841. This suggests that there is a requirement for the ability to utilise myo-inositol in order to be competitive for nodulation.
Figure 4.13 Nodule Occupancy of Vetch by RU307 and 3841

1 = uninoculated, 2 = 3841 $10^3$ cfu, 3 = RU307 $10^3$ cfu, 4 = 3841 $10^3$ cfu + RU37 $10^3$ cfu, 5 = 3841 $10^3$ cfu + RU307 $10^4$ cfu, 6 = 3841 $10^3$ cfu + RU307 $10^5$ cfu.
4.2.5.1 Competitive Ability of the Complemented Mutants

It was shown in Chapter 3 that the mutants RU360 and RU361 are complemented by cosmids containing the region of DNA interrupted by Tn5-lacZ, so that the ability to grow on myo-inositol is restored. The complemented mutants grew at similar rates to 3841 on glucose and myo-inositol \textit{in vitro}, indicating that the presence of a cosmid did not have any adverse effects on growth. Therefore, it was postulated that the complemented mutants would not be at a competitive disadvantage when co-inoculated with 3841.

The complemented mutants RU360/pRU3078, RU360/pRU3079 and RU361/pRU3111 were co-inoculated in equal numbers with 3841 (10^3 cfu) onto five or six vetch plants. All nodules were harvested from the plants five weeks post-inoculation. After stabbing of the nodules onto TY agar plus appropriate antibiotics, it became apparent that the cosmids were not always retained by the mutants (Table 4.1). When inoculated alone, 24.5% of nodules contained RU360/pRU3078, 74.2% contained RU360/pRU3079 and 60% contained RU361/pRU3111. The remainder of the nodules contained the appropriate mutant without a cosmid. Plasmids based on pLAFR1 are known to be unstable in the absence of selection pressure (Long \textit{et al.}, 1982).

When the nodules were harvested from plants that had been co-inoculated with a complemented mutant and 3841, none of the nodules contained RU360/pRU3078 or RU360/pRU3079 and 6.6% contained RU361/pRU3111.
However, 20.8% of nodules from plants inoculated with RU360/pRU3078 and 3841 contained RU360, 31.6% of nodules from plants inoculated with RU360/pRU3079 and 3841 contained RU360 and 8.9% of nodules from plants inoculated with RU361/pRU3111 and 3841 contained RU361 (Table 4.1).

The above numbers are considerably higher than those obtained for the mutants without a complementing cosmid when co-inoculated with 3841. Presumably, the complemented mutants were able to compete with 3841 more successfully and were able to form some nodules before the cosmid was lost. However, there was variability in the results. Strain RU360 was considerably more successful at occupying nodules when complemented than RU361. Although not conclusive, these data further suggest that the ability to utilise myo-inositol is important for the ability to compete successfully for nodulation.
Table 4.1  Cosmid Stability in the Rhizosphere

<table>
<thead>
<tr>
<th>inoculum</th>
<th>no. nodules containing mutant (%)</th>
<th>no. nodules containing mutant + cosmid (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RU360/pRU3078</td>
<td>53 (100)</td>
<td>13 of 53 (24.5)</td>
</tr>
<tr>
<td>RU360/pRU3078 + 3841</td>
<td>10 of 48 (20.8)</td>
<td>0</td>
</tr>
<tr>
<td>RU360/pRU3079</td>
<td>62 (100)</td>
<td>46 of 62 (74.2)</td>
</tr>
<tr>
<td>RU360/pRU3079 + 3841</td>
<td>18 of 57 (31.6)</td>
<td>0</td>
</tr>
<tr>
<td>RU361/pRU3111</td>
<td>70 (100)</td>
<td>42 of 70 (60)</td>
</tr>
<tr>
<td>RU361/pRU3111 + 3841</td>
<td>4 of 45 (8.9)</td>
<td>3 of 45 (6.7)</td>
</tr>
</tbody>
</table>

Figures represent nodules harvested from five or six vetch plants.
4.2.6 Measurement of Nodule Co-Residence

To determine the phenotype of bacteria in nodules from plants that were inoculated with mixtures of 3841 and mutants, whole nodules were stabbed successively onto TY agar, TY agar plus streptomycin, TY agar plus streptomycin and kanamycin and (if appropriate) TY agar plus streptomycin, kanamycin and tetracycline. Nodules may theoretically contain a mixture of strains. Nodules whose contents grew on streptomycin but not on kanamycin or tetracycline clearly only contained 3841. However nodules whose contents grew on kanamycin could also contain 3841 in addition to a mutant strain. Therefore, the dominance of 3841 may actually be higher than the results of the above experiments have already shown.

To test for co-residence of different strains, vetch nodule squashes that grew on kanamycin were streaked from the original master plate without antibiotics to determine the phenotype of individual colonies. Six individual colonies were then sub-streaked onto plates containing appropriate combinations of antibiotics. The results from experiments with different RU360 and RU361 inoculum sizes were combined, due to the limited number of nodules obtained that contained kanamycin resistant bacteria. From vetch inoculated with RU307, nodules were only tested from the plants that had received equal inocula of RU307 and 3841.

The same procedure was also applied to two nodules from plants inoculated with each strain alone. The results show that when the mutant was inoculated
alone, all colonies were kanamycin resistance and that when 3841 was inoculated alone, it did not acquire kanamycin resistance. Therefore, curing of the kanamycin gene was not responsible for colonies being obtained from mixed nodule squashes that were not kanamycin resistant. A representative number of colonies were also streaked onto myo-inositol, to confirm the phenotype.

The results indicate that 3841 was found in most of the nodules that contained the mutants, although the mutants were predominant (Table 4.2). Therefore, the advantage that 3841 may be even greater than suggested by data presented earlier. In addition, RU307 does not appear to be equally as competitive as 3841. However, only twelve nodules containing RU307, which might not have been a large enough sample to give a true representation of nodule occupancy.
Table 4.2 Individual Nodule Occupancy

<table>
<thead>
<tr>
<th>inoculum details</th>
<th>strain</th>
<th>no. colonies</th>
<th>% of colonies</th>
<th>frequency of nodules occupied</th>
</tr>
</thead>
<tbody>
<tr>
<td>3841</td>
<td>3841</td>
<td>12</td>
<td>100</td>
<td>2/2</td>
</tr>
<tr>
<td>RU360</td>
<td>RU360</td>
<td>12</td>
<td>100</td>
<td>2/2</td>
</tr>
<tr>
<td>RU361</td>
<td>RU361</td>
<td>12</td>
<td>100</td>
<td>2/2</td>
</tr>
<tr>
<td>RU307</td>
<td>RU307</td>
<td>12</td>
<td>100</td>
<td>2/2</td>
</tr>
<tr>
<td>RU360 + 3841</td>
<td>3841</td>
<td>16</td>
<td>22</td>
<td>7/12</td>
</tr>
<tr>
<td></td>
<td>RU360</td>
<td>56</td>
<td>78</td>
<td>12/12</td>
</tr>
<tr>
<td>RU361 + 3841</td>
<td>3841</td>
<td>7</td>
<td>9</td>
<td>7/12</td>
</tr>
<tr>
<td></td>
<td>RU361</td>
<td>65</td>
<td>91</td>
<td>12/12</td>
</tr>
<tr>
<td>RU307 + 3841</td>
<td>3841</td>
<td>14</td>
<td>19</td>
<td>9/12</td>
</tr>
<tr>
<td></td>
<td>RU307</td>
<td>58</td>
<td>81</td>
<td>12</td>
</tr>
</tbody>
</table>

The results represent six colonies streaked from each nodule.

Only nodules that contained kanamycin resistant bacteria were investigated. It should be noted that such nodules were rare, due to the dominance of 3841.
4.2.7 Average Number of Nodules per Vetch Plant

It was shown in Section 4.2.1 that vetch inoculated with the mutants alone formed the same number of nodules as plants inoculated with 3841. The average number of nodules per plant when mixed inocula were applied was also noted to see if the higher overall inoculum resulted in more nodules being formed (Figure 4.14, Figure 4.15, Figure 4.16). The plants were harvested five weeks post-inoculation.

The data were subjected to Analysis of Variance (Appendix 4.7). There was no significant difference in nodule numbers per plant between RU360, RU360/pRU3078, RU360/pRU3079 and 3841 inoculated alone and co-inoculated in equal numbers (P>0.1). There were however, significantly more nodules when RU360 was co-inoculated in 10-fold and 100-fold excess (P<0.05), although there was no significant difference in nodule numbers between the 10-fold and 100-fold experiments (Figure 4.14).

There was no significant difference (P>0.1) in the number of nodules when RU361, RU361/pRU3111 and 3841 were inoculated singly or co-inoculated in equal numbers (Appendix 4.8). When RU361 was co-inoculated in 10-fold excess, there were significantly more nodules formed than when RU361, RU361/pRU3111 and 3841 were inoculated alone and when RU361/pRU3111 was co-inoculated with 3841 (P<0.05). There was no significant difference between RU361 co-inoculated in equal numbers, 10-fold or 100-fold excess with 3841 (Figure 4.15).
For plants co-inoculated with RU307 in equal numbers, 10-fold and 100-fold excess with 3841, there were significantly more nodules formed than when either strain was inoculated alone (P<0.05) (Appendix 4.9). There was also a significant increase in nodule numbers between RU307 co-inoculated in equal numbers and 10-fold excess (P<0.05) but there was no significant difference between RU307 co-inoculated in equal numbers and 100-fold excess (P>0.1).

The results indicate that the plants were not disadvantaged in forming nodules when inoculated with the mutants. More nodules were formed when all the mutants were co-inoculated with 3841 in 10-fold and 100-fold excess. This was probably due to the higher inoculum size and indicates that the mutants are able to elicit nodule formation when co-inoculated with 3841. With RU307, this led to a higher proportion of nodules containing the mutant. However, with RU360 and RU361, the additional nodules formed did not contain the mutants but 3841. This suggests that even though the myo-inositol catabolic mutants can elicit nodule formation, they either do not gain access to the nodules, or competition from 3841 causes them be lost from the nodule.
Figure 4.14  Average Number of Nodules per Vetch Plant Inoculated with RU360 and 3841

Each result represents the average of at least six plants with standard deviation. 10x and 100x signify that RU360 was inoculated in 10-fold and 100-fold excess respectively.
Figure 4.15  Average Number of Nodules per Vetch Plant Inoculated with RU361 and 3841

Each result represents the average of at least six plants with standard deviation. 10x and 100x signify that RU361 was inoculated in 10-fold and 100-fold excess respectively.
Figure 4.16  Average Number of Nodules per Vetch Plant Inoculated with RU307 and 3841.

Each result represents the average of at least six plants with standard deviation. 10x and 100x signify that RU307 was inoculated in 10-fold and 100-fold excess respectively.
4.2.8 Average Plant Dry Weights for Vetch

The average dry weights were also obtained for vetch plants from each inoculum five weeks post-inoculation. The results of plants inoculated with RU360, RU360/pRU3078, RU360/pRU3079 and 3841 are shown in Figure 4.17. The data were subjected to Analysis of Variance (Appendix 4.10). There was a significant difference in the weight of the uninoculated plants and the inoculated plants (P<0.05). There was no significant difference in the average dry weight of plants inoculated with RU360, RU360/pRU3078, RU360/pRU3079 and 3841 either alone or when co-inoculated (P>0.1).

The average dry weight of plants inoculated with RU361, RU361/pRU3111 and 3841 are shown in Figure 4.18. The data were subjected to Analysis of Variance. There was a significant difference in the weight of the uninoculated plants and the inoculated plants (P<0.05). There was no significant difference in the average dry weight of plants inoculated with RU361, RU361/pRU3111, RU361 and 3841 co-inoculated in equal numbers or with RU361 in 100-fold excess (P>0.1). However, the average dry weight of plants inoculated with 3841 and RU361 in 10-fold excess were significantly higher than the other plants (P<0.05).

The average plant dry weight of plants inoculated with 3841 and RU307 are shown in Figure 4.19. The data were subjected to Analysis of Variance. There was a significant difference in the weight of the uninoculated plants and the inoculated plants (P<0.05). There was no significant difference in the
average dry weight of plants inoculated with RU307 and 3841 either alone or when co-inoculated in equal numbers or with RU307 in 10-fold excess (P>0.1). However, the average dry weight of plants inoculated with 3841 co-inoculated with RU307 in 100-fold excess was significantly higher than plants inoculated with either RU361 or 3841 alone (P<0.05).

These results show that even though more nodules were formed when higher mutant numbers were applied to plants, there was no obvious correlation with plant dry weight. All plants treated with all inocula weighed the same for RU360, but plants treated with a 10-fold excess of RU361 weighed more, although plants inoculated with 100-fold excess did not. With RU307, the plants inoculated with 100-fold excess weighed more, but plants inoculated with 10-fold excess did not. Therefore, the data are inconclusive as to whether more nodules result in greater plant dry weight. However, it can be concluded that plants inoculated with the mutants are not disadvantaged for growth compared to plants inoculated with 3841.
Figure 4.17  Average Plant Dry Weights for Vetch inoculated with RU360 and 3841.

Each series represents the average of at least six plants with standard deviation. 10x and 100x signify that RU360 was inoculated in 10-fold and 100-fold excess.
Figure 4.18  Average Plant Dry Weights for Vetch Inoculated with RU361 and 3841

Each series represents the mean of at least six plants, with standard deviation. 10x and 100x signify that RU361 was inoculated in 10-fold and 100-fold excess.
Figure 4.19  Average Plant Dry Weights for Vetch Inoculated with RU307 and 3841

Each series represents the mean of at least six plants, with standard deviation. 10x and 100x signify that RU307 was inoculated in 10-fold and 100-fold excess.
4.2.9 Rhizosphere Growth Assay

It was shown in Section 4.2.1 that nodules are formed on plants at the same time, whether inoculated with the mutants or 3841. In Section 3.2.1 it was shown that the mutants and 3841 have the same growth rates \textit{in vitro}. However, conditions in the rhizosphere may be completely different to those \textit{in vitro} and if 3841 has a growth advantage at the early stages of rhizosphere colonisation, this might be responsible for domination of the nodules.

This hypothesis was tested by measuring the growth of RU360, RU361, RU307 and 3841 in the rhizosphere of vetch in the eight days immediately following inoculation (Figure 4.20, Figure 4.21, Figure 4.22). This assay gives an indication of colony numbers for the entire plant system, so bacteria associated with plant roots as well as those in the rhizosphere were harvested. The data were subjected to Analysis of Variance (Appendix 4.13). When inoculated singly, all strains grew over eight days. Strain RU360 attained the highest colony numbers and RU307 the lowest, but the differences were not significant (0.1>P>0.05).

The results for RU360 and 3841 are shown in Figure 4.20. Numbers of both strains increased steadily over six days. The data after eight days were subjected to Analysis of Variance (Appendix 4.14). The numbers of both RU360 and 3841 when co-inoculated were not significantly different, but RU360 numbers were significantly lower when co-inoculated than when inoculated alone (P<0.05).
The results for RU361 and 3841 are shown in Figure 4.21. Both strains grew well over six days. The data after eight days were subjected to Analysis of Variance (Appendix 4.15). There was no significant difference in numbers of RU361 and 3841 inoculated alone, or for 3841 when co-inoculated, but numbers of RU361 when co-inoculated were significantly lower (P<0.05). Numbers of RU361 were consistently lower throughout the study. As RU361 grew well over the time period, there may have been an inoculation error.

The results for RU307 and 3841 are shown in Figure 4.22. Both strains grew over six days when inoculated alone and when co-inoculated. The data after eight days were subjected to Analysis of Variance (Appendix 4.16). When co-inoculated, numbers of 3841 were significantly higher than RU307 inoculated alone or co-inoculated (P<0.05). There was no significant difference between 3841 inoculated alone or co-inoculated (P>0.1), or for RU307 inoculated alone or co-inoculated (P>0.1).

The data indicate that the mutants were all able to grow in the rhizosphere in the eight days following inoculation. When inoculated into the rhizosphere alone, numbers of RU360 were highest, but this was not significant (P>0.05<0.1). However, when co-inoculated, numbers of 3841 were significantly higher (P>0.05) than the mutants. Overall, all strains grew well in the rhizosphere and based on these observations, if 3841 has a competitive advantage purely due to a higher growth rate, then RU307 would also be expected to be at a competitive disadvantage for nodulation. In addition, if
successful nodule occupancy were purely due a growth advantage, when the mutants were co-inoculated in 10- and 100-fold excess, they would have been expected to dominate the nodules.
Figure 4.20 Growth of RU360 and 3841 Recovered from the Rhizosphere

Each point represents the mean of at least four replicates.
Figure 4.21  Growth of RU361 and 3841 Recovered from the Rhizosphere

Each point represents the mean of at least four replicates.
Figure 4.22 Growth of RU307 and 3841 Recovered from the Rhizosphere

Each point represents the mean of at least four replicates.
4.2.10 Analysis of a nodC-phoA Fusion

Strains 3841 has a clear competitive advantage for nodulation over RU360 and RU361. The ability to utilise myo-inositol as a carbon source does not appear to be responsible for the competitive advantage, as the mutants were all able to grow well in the rhizosphere in the absence of exogenous carbon. However, RU360 and RU361 may not able to respond to signals given out by plants as effectively as 3841 because a breakdown product of myo-inositol is required, so that nodules are formed slightly more slowly than by 3841.

Nodulation (nod) genes encode proteins that are required for the formation of nodules on legumes. Their expression is induced by components of plant cell exudates. The nodC gene encodes an N-acetyl-glucosaminyl transferase involved in the formation of lipo-chito-oligosaccharide Nod factors that initiate nodule morphogenesis in legumes (Geremia et al., 1994, Spaink et al., 1994, see Section 1.2.1). If the mutants are unable to respond as effectively to plant signals as 3841, this may be reflected by reduced expression of nodC. Therefore, the ability of 3841 and the mutants to express nodC in response to a plant product was assessed in vitro.

A nodC-phoA fusion (pIJ1687) was conjugated into the mutants and 3841. Expression of NodC was assessed by measuring the activity of a promoter-less alkaline phosphatase (PhoA), encoded by the phoA gene fused to the nodC gene. Hesperetin was the plant product used to induce expression of nodC. The mutants and 3841 were assayed for alkaline phosphatase activity
in the presence and absence of hesperetin when grown on pyruvate and pyruvate plus myo-inositol. Strain 3841 was also tested when grown on myo-inositol only (Table 4.3).

Without induction by hesperetin, the activity of alkaline phosphatase was very low by all the strains tested, on all carbon sources. In the presence of hesperetin, activity of alkaline phosphatase was greatly increased for all strains on all carbon sources. The data were subjected to Analysis of Variance (Appendix 4.17). Highest activity was for RU360 grown on pyruvate plus myo-inositol and 3841 grown on myo-inositol. The activity of RU361 grown on pyruvate plus myo-inositol was lower than that of the other strains, but this was not significant (0.01>P>0.05). These data indicate that the mutants are able to express nodC in response to plant products in vitro as effectively as 3841.
Table 4.3 Induction of the *nodC* Gene

<table>
<thead>
<tr>
<th>strain</th>
<th>pyruvate</th>
<th>pyruvate + <em>myo</em>-inositol</th>
<th>myo-inositol</th>
</tr>
</thead>
<tbody>
<tr>
<td>RU360</td>
<td>258.84 ± 98.5</td>
<td>112.16 ± 8.16</td>
<td>N/A</td>
</tr>
<tr>
<td>uninduced</td>
<td>4911.58 ± 245.37</td>
<td>5531.9 ± 291.46</td>
<td>N/A</td>
</tr>
<tr>
<td>induced</td>
<td>3560.47 ± 154.26</td>
<td>3142.91 ± 53.41</td>
<td>N/A</td>
</tr>
<tr>
<td>RU361</td>
<td>100.37 ± 7.08</td>
<td>152.37 ± 8.52</td>
<td>N/A</td>
</tr>
<tr>
<td>uninduced</td>
<td>4510.98 ± 63.04</td>
<td>4591.21 ± 602.29</td>
<td>N/A</td>
</tr>
<tr>
<td>induced</td>
<td>4138.48 ± 507.28</td>
<td>4220.79 ± 370.75</td>
<td>5370.43 ± 1707.11</td>
</tr>
<tr>
<td>RU307</td>
<td>122.79 ± 10.43</td>
<td>157.66 ± 3.59</td>
<td>N/A</td>
</tr>
<tr>
<td>3841</td>
<td>59.48 ± 6.35</td>
<td>151.18 ± 39.14</td>
<td>227.24 ± 16.28</td>
</tr>
<tr>
<td>induced</td>
<td>4138.48 ± 507.28</td>
<td>4220.79 ± 370.75</td>
<td>5370.43 ± 1707.11</td>
</tr>
</tbody>
</table>

Each result represents the mean of at least two cultures assayed in triplicate.

N/A = not assayed.
4.3 Discussion

The mutants RU360, RU361 and RU307 were all able to form nodules on pea and vetch plants. There was no significant difference between the total numbers of nodules formed on vetch plants inoculated with the mutants and 3841 alone, or the time taken for them to form. There was also no difference in plant dry weight between the mutants and 3841, indicating that the mutants provided as much fixed nitrogen to the host plants as 3841. It is possible that 3841 formed nodule initials more quickly than the mutants. These are not visible to the naked eye and would have required the use of staining techniques. Due to time constraints, this was not attempted.

All the nodules formed on plants inoculated with the mutants and 3841 appeared pink. This indicates the presence of leghaemoglobin and suggests that the nodules were fixing nitrogen. The acetylene reduction assay confirmed this for pea nodules containing RU361, RU307 and 3841 and showed that nitrogenase activity was similar for all three strains. This was previously shown to be the case for RU360 (Poole et al., 1994).

Higher numbers of nodules were formed on peas inoculated with RU361 and RU307 than 3841, although total nodule mass and plant dry weights were the same. These data suggest that peas inoculated with the mutants formed nodules that were not quite as effective at nitrogen fixation as those inoculated with 3841. Presumably, the plants compensated for this by directing the formation of additional nodules. Vetch are considerably smaller
than peas so they require less fixed nitrogen and hence fewer nodules. Therefore, a small reduction in effectiveness per nodule may not affect the plant as much, so that there would be no noticeable difference in nodule number per plant. However, it is not possible to conclude from this study that the mutants are less effective at nodulating vetch than 3841 when inoculated alone.

As vetch inoculated with the mutants and 3841 formed nodules at the same rate, bacteroids from pea nodules fixed nitrogen at the same rate and plant dry weights were the same for both vetch and pea, it was expected that the mutants would be equally as competitive as 3841 for nodule formation when co-inoculated with 3841. However, co-inoculation of peas and vetch with RU360 or RU361 and 3841 resulted in massive domination of nodules by 3841, even when the mutants were inoculated in 10-fold and 100-fold excess. These data indicate a role for myo-inositol in nodulation competitiveness, which has not been shown before.

Bosworth et al. (1994) and Scupham et al. (1996) showed that interruption of the inositol locus of *S. meliloti* was sometimes beneficial to host plants. However, the effects of disruption of the *myo*-inositol locus were not consistent, suggesting that there is a role for *myo*-inositol in *S. meliloti*. There also appears to be a role for rhizopines, as Gordon et al. (1996) showed that the rhizopine-producing strain *S. meliloti* L5-30 had a competitive advantage for nodulation of lucerne (*Medicago sativa*) in soil when co-inoculated with a mutant strain, even though when inoculated alone, the mutant had a similar
rate of growth and nodulation to the wild type. The mutant occupied less than 30\% of nodules. This competitive advantage remained four years after inoculation, even though there had been turnover of nodules in that time (Heinrich et al., 1999).

The data from the competition experiments using RU360/pRU3078, RU360/pRU3079 and RU361/pRU3111 indicate that had the cosmids been stable in the environment, 3841 might not have had a competitive advantage. The transposon inserted into two completely different loci in RU360 and RU361, which suggests that it was the restoration of the ability to utilise myo-inositol, rather than the presence of other beneficial regions on the cosmid that was responsible for restoring the competitive advantage. However, to eliminate the possibility that independent factors on the cosmids were responsible, a more stable vector should be used. In Section 5.2.7, the pOT1 plasmid was shown to be stable in pea root nodules containing 3841, so this vector should be suitable. In Section 3.2.5.1.1, the cloning of iol genes from pRU3078 into pOT1 was described. The region of DNA required for complementation of RU361 could also be cloned into pOT1 and the competition experiment repeated. It is postulated that the complemented mutants would be as competitive as 3841.

Strain RU307 was able to compete successfully for nodulation when co-inoculated onto vetch with 3841. The nodule occupancy tests imply that 3841 may be slightly more competitive as the majority of nodules from the mixed inocula experiments contained 3841. However, this experiment needs to be
repeated with a much larger sample size, to confirm the results. Strain RU307 differs from RU360 and RU361 in that it is not a complete myo-inositol mutant, but is mutated in an uptake system for myo-inositol (c.f. Section 3.2.5). There is no apparent deficiency in the myo-inositol catabolic pathway as the first two enzymes in the catabolic pathway were induced when RU307 was grown on myo-inositol as the sole carbon source (c.f. Section 4 or 5). The requirement for myo-inositol was probably not as a carbon source because RU307 was competitive, despite the fact that its growth rate on myo-inositol is only a quarter of that of 3841 in vitro.

It is clear that there were other carbon sources present in the system, since RU360, RU361, RU307 and 3841 all grew in the rhizosphere. Substrate exhaustion was also unlikely to be a factor, as the strains all grew well over eight days, by which time nodules had begun to form. myo-Inositol did not have an inhibitory effect on the mutants in vitro. However, there was catabolite repression by other carbon compounds of some myo-inositol catabolic enzymes and the myo-inositol uptake system in vitro (c.f. Section 3 and Section 5). This suggests that even if myo-inositol were present in the rhizosphere, catabolite repression might occur, so other sugars would be used preferentially and 3841 would not gain any competitive advantage by being able to utilise myo-inositol.

In addition, the mutants and 3841 all grew well in the rhizosphere over eight days, although 3841 grew slightly better when co-inoculated. If the dominance of 3841 is purely because of a growth advantage in the
rhizosphere, then RU307 should also have been at a competitive
disadvantage. In addition, the competitive disadvantage should have been
overcome when RU360 and RU361 were co-inoculated in 10- and 100-fold
excess as the mutants growth should swamp that of 3841.

These results are in contrast to those of Wood and Stanway (2000), who
showed that in soil solution, levels of myo-inositol decreased concomitantly
with the increase of four different strains of rhizobia, which suggests that myo-
inositol was used preferentially. However, the soils that were used contained
myo-inositol as the dominant sugar. This is not the case in other soils that
have been studied (c.f. Section 1.3.3.2) and so myo-inositol might only be
used preferentially when it is present in such a large quantity that catabolite
repression by other sugars is not significant. However, all experiments were
carried out in sterile vermiculite in this project, which may give very different
results to experiments carried out in soil.

Higher numbers of nodules per vetch plant were observed when RU360 and
RU361 were co-inoculated in 10-fold and 100-fold excess with 3841. However, the majority of bacteria recovered from these nodules were still
3841 and there was no increase in plant dry weight, indicating that there was
not more nitrogen fixation occurring. These data indicate that the plants were
able to sense the additional rhizobia and produce signals that the rhizobia
could respond to. However, the mutants were still not able to get into the
nodule. This confirms that the dominance of the wild type is not because it
grows faster in the rhizosphere. Therefore, 3841 must be gaining an
advantage over the mutants, either at the commencement of infection or during early infection events.

Many more infection threads are formed than go onto make nodules (Bauer, 1981) and the mechanisms through which the plant controls which infection threads become nodules have not yet been identified. Oresnik et al. (1998) studied a rhamnose catabolic mutant. The authors postulated that as plant cell walls are continuous with infection threads, rhamnogalacturonan subunits or degradation products may be present in infection threads. Depending on when the mechanism that results in infection threads being aborted occurs, proportionately more infection threads formed by the mutant may be aborted if the mutant strain grew more slowly because of its inability to utilise rhamnose, resulting in a majority of nodules containing the wild type.

The myo-inositol oxidation pathway is the major pathway providing sugar residues for cell wall synthesis during germination of bean (Sasaki and Nagahashi, 1990). It is also present in other complex plant lipids and in plant cell membranes as phosphoinositides (Loewus, 1990). In addition, the presence of myo-inositol in nodules, including those of pea, has been widely reported (Kouchi and Yoneyama, 1984, 1986, Skøt and Egsgaard, 1984, Streeter, 1987; Streeter and Salminen, 1985).

Therefore, the theory postulated for rhamnose utilisation might also be applicable to myo-inositol. Although there mutants were able to grow well in the rhizosphere, the composition of available nutrients in the infection thread
might be very different to the rhizosphere. Strain RU307 would be expected to be disadvantaged because of its poor growth rate on myo-inositol, but if other nutrients are limited, it might be able to grow well enough to ensure its infection thread are not disproportionately aborted.

Another possibility is that the presence of myo-inositol or one of its derivatives acts as a signal to Rhizobium, encouraging infection thread development or the formation of nodules. myo-Inositol is likely to be present in areas that nodules are developing because of its role in cell wall synthesis. myo-Inositol might also be present near developing nodules as it used to transport auxins, which are also found at the site of developing nodules. Perhaps myo-inositol acts as a signal, highlighting the areas where nodules are developing. If 3841 can respond to plant signals and produce infection threads more quickly than RU360 or RU361, then the host plant may favour the formation of full nodules containing 3841 when there is a mixed inoculum.

It has been shown that when two sides of a root are spatially separated and inoculated at different time intervals, nodulation is prevented on the side inoculated 24 hours later (Kossak et al., 1984, Sargent et al., 1987). The suppression of nodulation increases as the time interval between the two inoculated sides increases (Kossak et al., 1983), indicating that if two strains were co-inoculated and one could initiate nodulation more quickly, then that strain would have a clear competitive advantage.
When inoculated alone, there was no competition from other strains, so nodules always contained the mutants. However, it is possible that the nodules were actually formed more slowly than on plants inoculated with 3841. If this difference were measured in hours then this would not have been observed in the experiments conducted in this project as the plants were assessed once a day. In a future experiment, 3841 could be inoculated into the rhizosphere of plants at different time intervals after the mutants to see if this would allow the mutants to overcome the disadvantage.

It was shown that when bean roots were inoculated with *R. phaseoli* and cultured in a growth medium supplemented with *myo*-inositol, the percentage of roots nodulated and the number of nodules per root increased (Raggio *et al.*, 1959). Another future experiment could investigate whether this occurs with vetch or peas and if there are any differences between plants inoculated with the mutants and 3841.

To determine the longevity of the mutants in the nodules, nodules could be harvested at different times post-inoculation. The nodules studied in the competition experiments were all harvested approximately five weeks post-inoculation. The few nodules that contained RU360 or RU361 from plants co-inoculated with two strains, usually also contained small numbers of 3841. This was not due to curing of the kanamycin resistance gene, but may have been due to intra-nodule competition between the mutants and 3841.
If nodules were harvested at different intervals after formation, there might be a difference in nodule composition and the numbers occupied by each strain might be more even. This would determine whether the mutants initiated formation of nodules as successfully as 3841 but were then displaced by 3841, or whether 3841 has an advantage at the earliest stages of nodule formation. Harvesting the nodules earlier may also result in less instability of the complementing cosmids. This experiment would complement the pre-inoculation experiment suggested above.

The response of nodC of the mutants and 3841 to hesperetin was measured in order to test whether the ability to catabolise myo-inositol is important for signalling between plants and Rhizobium. The responses of the different strains were similar, which suggests that the mutants were not inhibited in their response to plant signals, as did the increased nodule numbers when the mutants were inoculated in excess. However, there are several Nod factors and as it was not practical to study them all, it is possible that other Nod factors are not as effective in the mutants.

It is not known whether the bacteroids were also a mix of two strains, as the composition of undifferentiated bacteria obtained from nodules might not truly reflect the genotype of the bacteroids. A more effective way to study nodule composition would be to use a reporter gene that can be studied in planta, such as green fluorescent protein (GFP). Several types of GFP are available that give different colour responses under UV light. The genes encoding the different types of GFP could be linked to a gene that is constitutively
expressed, either by integration into the chromosome, or on a stable plasmid. The GFP fusions could be introduced into the mutants and 3841, then the nodule contents could be examined under UV light to identify the phenotype of bacteroids and free-living bacteria in the infection threads.

To further investigate the role of myo-inositol in competition for nodulation, mutants in different parts of the degradation pathway need to be isolated and tested. This is crucial in order to determine whether different components of the pathway are required. To determine whether myo-inositol is important for competition in all legume symbioses it would be necessary to see if the effects were in the same in different Rhizobium species and their corresponding host plants. It would also be useful to obtain mutants in pathways that lead to myo-inositol derivatives, such as inositol phosphates or rhizopines to try to determine the specific role of myo-inositol in competition for nodulation.

CHAPTER 5 - REGULATION OF MYO-INOSITOL CATABOLISM .......... 204

5.1 Introduction .......................................................................................... 205

5.2 Results .................................................................................................. 208

5.2.1 Isolation of a myo-inositol Specific Promoter ........................................ 208

5.2.2 Sequence Analysis of the myo-Inositol Inducible Promoter ...................... 209

5.2.3 Induction of the myo-Inositol Inducible Promoter ...................................... 218

5.2.4 Catabolite Repression ............................................................................. 223

5.2.5 Relationship to myo-Inositol Mutants ...................................................... 228

5.2.6 Expression in the Rhizosphere and Nodules ............................................ 230

5.2.7 Stability of pRU601 in the Rhizosphere .................................................. 231
5.2.8 Construction of a Specific *myo*-Inositol Inducible GFP-UV Promoter Probe
........................................................................................................................................ 232

5.2.9 RU360 \( \beta \)-Galactosidase Assay ................................................................. 236

5.2.10 *myo*-Inositol Catabolic Enzyme Assays ....................................................... 238

5.3 Discussion .......................................................................................................................... 241
Chapter 5 - Regulation of myo-inositol catabolism
5.1 Introduction

In order to understand the regulation of the pathway of myo-inositol catabolism, the myo-inositol mutants and 3841 were tested for their ability to utilise the first two enzymes in the proposed myo-inositol catabolic pathway. It had been shown previously that RU360 was not able to induce these enzymes, but that 3841 could (Poole et al., 1994).

To identify additional genes involved in myo-inositol utilisation to those described in Chapter 3, attempts were made to isolate myo-inositol inducible promoters, using Green Fluorescent Protein (GFP) as an indicator. GFP is a small protein that fluoresces when excited by UV or short-wave blue light. It was originally isolated from the bioluminescent jellyfish Aequorea victoria. As no exogenous cofactors are required for fluorescence, the gene for GFP can be placed under the control of a heterologous promoter and expression studied with no apparent interference with cell growth and function (Chalfie et al., 1994). There is limited autofluorescence at the same wavelengths by biological material, enabling study of GFP in the presence of plants or microorganisms.

A R. leguminosarum bv. viciae DNA promoter-probe library consisting of random fragments fused to a promoter-less gfpuv gene in the plasmid pOT1, was previously constructed in this laboratory (Schofield, 1999). Plasmid pOT1 is a 5278bp, broad host range plasmid that
contains a gene encoding gentamycin resistance. It is based on the plasmid pBBR1-MCS5 and contains the promoterless \textit{gfpuv} gene between two unique transcriptional termination sites, to prevent expression of GFP-UV by promoters on the plasmid. The \textit{gfpuv} gene was used because it is 18 times more fluorescent than the wildtype \textit{gfp} gene but retains the same excitation and emission maxima (Crameri \textit{et al}., 1996).

Approximately 15,000 promoter fragments were generated in pOT1, with an average size of 1.56kb. These fragments were estimated to represent approximately 40% of all \textit{Rhizobium} promoters. The fragments were cloned into the \textit{SalI} site of pOT1, which was destroyed in the process. In order to select for promoters that were not induced in standard laboratory conditions, the library was screened on AMA and AMS containing glucose and ammonia as the sole carbon and nitrogen sources respectively. Approximately 71% of the colonies did not express GFP-UV in these conditions and contained a \textit{Rhizobium} chromosomal DNA fragment. These colonies were selected and stocked as a library. This library has previously yielded defined promoter fragments, identified by their ability to induce expression of GFP-UV (Schofield, 1999).

The plasmid pOT1 containing the promoter-less \textit{gfpuv} gene was also used to make defined \textit{myo}-inositol inducible promoter-probe vectors, using the region interrupted in RU360. By fusing the promoter region of
the ioID gene to a promoter-less gfpuv gene it was hoped that induction of myo-inositol catabolic genes in the rhizosphere and nodules could be studied.
5.2 Results

5.2.1 Isolation of a myo-inositol Specific Promoter

Approximately 10,000 colonies of the promoter-probe library in 3841 were plated onto AMA containing 10mM myo-inositol as the sole carbon source. Seventeen colonies expressing GFP-UV were identified on a transilluminator at a wavelength of 420nm. These colonies were streaked onto AMA containing other carbon sources (glucose, pyruvate, mannitol, sorbitol, fructose, glycerol), to determine whether the promoter was specific to myo-inositol. Four colonies were isolated that only expressed GFP-UV in the presence of myo-inositol and therefore contained a myo-inositol specific promoter. Sequencing showed that all four colonies contained the same 2048bp fragment in pOT1. The plasmid containing this fragment was named pRU601 and the promoter contained in the 2048bp fragment was designated iolXp.
5.2.2 Sequence Analysis of the myo-Inositol Inducible Promoter

The \( ioiX_p \) fragment was sequenced in the 5' and 3' direction (Figure 5.1, Figure 5.3) in order to try to identify the promoter responsible for expression of the \( gfpuv \) gene. BlastX analysis of the deduced amino acid sequence revealed a putative ORF of 660bp, which showed homology with the transcriptional regulatory protein of several two-component sensor-regulator systems. The ORF was designated \( orf1 \) (Figure 5.1, Figure 5.3). The closest match was 45% identity (p-value of \( 1 \times 10^{-36} \)) with a putative transcriptional regulator from \( Bordetella pertussis \). There was also 39% identity (p-value of \( 3 \times 10^{-31} \)) with the FeuP protein of \( R. leguminosarum \) bv. \( vicieae \). FeuP is involved in the regulation of iron uptake (Yeoman \textit{et al.}, 1997). The \( orf1 \) gene encodes a predicted protein of 220 amino acids, with a molecular weight of 23988 and a P.I of 5.4.

Regulatory proteins such as FeuP are known to feature an acid pocket (Parkinson and Kofoid, 1992). The residues in FeuP that correspond to glutamic acid (amino acid 7), aspartic acid (amino acids 8, 9, 51) are thought to form part of the acid pocket into which the side chain of lysine (amino acid 101) protrudes (Yeoman \textit{et al.}, 1997). These residues are conserved among similar regulatory proteins (Parkinson and Kofoid, 1992), including in the putative Orf1 protein (Figure 5.3). The \( orf1 \) gene is transcribed divergently to the promoter-less \( gfpuv \)
gene. Therefore, it is not possible that the promoter for orf1 is responsible for expression of the \textit{gfpuv} gene.

The remainder of the \textit{iolX}_p fragment did not have homology with any known sequences in the Genbank or EMBL databases. Therefore, testcode analysis was performed. This indicated that apart from a small region of approximately 100bp, the sequence is likely to be coding (Figure 5.2). Three further putative ORFs were subsequently identified.

An ORF of 636bp was designated \textit{orf2} and is transcribed in the same direction as the promoter-less \textit{gfpuv} gene (Figure 5.1, Figure 5.3). The \textit{orf2} gene encodes a predicted protein of 212 amino acids, with a molecular weight of 22986.2 and a P.I of 5.6. An ORF of 729bp was designated \textit{orf3} and overlaps \textit{orf2} in a different frame, but is also transcribed in the same direction as the promoter-less \textit{gfpuv} gene (Figure 5.1, Figure 5.3). The \textit{orf3} gene encodes a predicted protein of 243 amino acids, with a molecular weight of 25965.2 and a P.I of 10.3. The fourth putative ORF is 88bp and was designated \textit{orf4}. It is at the end of the 2048bp insert and is not complete. It is immediately upstream of the promoter-less \textit{gfpuv} gene and is transcribed in the same direction (Figure 5.1, Figure 5.3).
Figure 5.1  pRU601 Insert

pRU601 7326 bp

orf4
Pac I (4012)
Pst I (3999)
rbs-gfp

orf3
Pac I (4012)
Pst I (4737)

orf2
Pst I (4737)

orf1
Not I (320)
Pst I (3240)

Eco R I (3061)
Eco R I (3240)
Pac I (6090)
Sst I (3260)

termination (omega)
Pac I (6090)

termination (pharmacia)
Pst I (3999)

GFP-UV

mobilisation

replication

gentamycin resistance
Testcode analysis of *iolXp*. Sequence that is in the top portion of the graph is likely to be coding. Sequence in the middle portion is ambiguous and sequence in the bottom portion is unlikely to be coding.

*T* = Testcode.
Figure 5.3 DNA Sequence of pRU601 Insert

1  TCCTCATGGC CGTCCCAGAC GATAGCCGAG CCCGCGCTCG GTCTCGATCA
   AGG
   AGT
   ACCG GCAGGGTCTG CTATCGGCTC GGGCGCGAGC CAGAGCTAGT
-2  P R G L R Y G L R E T E I

51  CATGAGTACC GAGCTTCTTG CGCAACCGGC TCACATGGAC CTCGATGGCG
   GTACTCATGG CTCGAAGAAC GCGTTGGCCG AGTGTACCTG GAGCTACCGC
-2  H T G L K K R S V H V E I A

101  TTGCTGTGCA TTTCCTATTC GAAGGAATAG AGCCGCTCTT CGAGCTGCGC
   AACGACAGCT AAAGGGATAG CTTCCTTATC TCGGCGAGAA GCTCGACGCC
-2  N S D I E R D F S Y L R E E L Q A

151  CTTGGAAGGA GACGTTGCAG GAAGGCCTCA AACAGGCCAG
   GAAGCTTTCTCT TCAGGCGGCC CTGCACAGCT CTACCGAGCT TTGCTCTGGG
-2  K S L L P R Q L F A E F L V

201  ATTCCGGCCGC GTTACAGACCA ATGGCCCTTC CATTGAGTCTG
   GAACTCTGCT TACCGGAGAC GTAAGTCTCA CTACGAAGAC
-2  T R A L T V A K G N L I S R

251  GGCGCGAGAT CGATGCGGAG CGGGCCAGGC GTGATGCTTA AACAGGATCG
   CGGGCGCTTCA GGTACAGCAG CTCGCGACAG TCCTAGTACT GGCCGCAGTG
-2  A A L D I A L P G L T I N P N P N

301  GCCGCGCTAA CGGCGCGCA ACGACGCAGT CGGCGCCGAA GCTCCGGCAA
   CCGCGGTATT TCCGCCTGAT TTGGCGGTTC CACTACAACC CTAACCTTA
-2  A R A V S G I R A S L E A

351  GATCGGAAAGG CTTCAACGTA TAATCGTCAG CGCGCGCGCTCGA CAATCGGCGA
   ATTACAGCTC GGCGCCTGTA GGTAGCGCTG
-2  D F P K V M Y D D A G A N L G A

401  ATACGGTCA GAGCTTCCCTGTTG TAGATCGATG GCCGCGTCAC
   TATGCGGACTG CTAGGCGGAC AGGGATGCTA GCCGCGCGTT TCGAGGTTT
-2  G A Y R A V S G I R A S L E A

451  GTCGGCGGCT CGACGCGCGAC CCGGCGCAGT CGGCGCGGAA GCTCCGGCAA
   CCGCGGTATT TCCGCCTGAT TTGGCGGTTC CACTACAACC CTAACCTTA
-2  D F P K V M Y D D A G A N L G A

501  GCAGCAGTCAA ATCGACGTCA ACCAGCAGCC CGAGATTGGCA
   CAGCGGCGCGG TGGGCTGCGG CTGCTGGTGG TTGCTCGCGG TTGCTACAGT
-2  L M L D L L V L D Y A A A V I A

551  TCGCCCGCTT GGTCCCGCAG GCTTCCGCTAA TCGAGGGAAT GCCGCGTCG
   AGCGCGGAGA CGACGCTGCG CATGCTGCGG GTAATTGGTTC ATGCGCTACGCC
-2  D G A Q D L R N V W D V S H G D A
orf1

$$\text{TCTCCTTTGT}$$

orf2

$$\text{MK} \at \text{TG} \at \text{TL} \at \text{PI} \at \text{L} \at \text{AL} \at \text{AV} \at \text{MA}$$

orf3

$$\text{PA} \at \text{PE} \at \text{SA} \at \text{AA} \at \text{AG} \at \text{AV} \at \text{PA} \at \text{AG} \at \text{YQ} \at \text{E}$$
Letters in blue or red indicate deduced amino acids. For meaning of abbreviations, see Appendix 3.1. Start and stop codons for each ORF are bold and underlined. Putative SD sequences are bold. The conserved residues of the deduced amino acid sequence of orf1 that correspond to glutamic acid (amino acid 7), aspartic acid (amino acids 8, 9, 51) and lysine (amino acid 101) are indicated by asterisks.
5.2.3 Induction of the myo-Inositol Inducible Promoter

The ioI/Xp fragment was isolated because it was responsible for myo-inositol dependent induction of expression of GFP-UV. In order to determine whether this induction is linked to myo-inositol catabolism, expression was measured after the plasmid was conjugated into the mutants RU360/pRU601, RU361/pRU601 and RU307/pRU601. These strains, along with 3841/pRU601, were tested for GFP-UV expression after overnight growth in the presence of different carbon sources (Figure 5.4, Figure 5.5).

GFP-UV expression was calculated in relation to the growth of the cultures, giving specific fluorescence (V). V was calculated using the equation A-X/B-Y where A is the fluorescence of the sample, X is fluorescence of the blank well, B is the OD_{630nm} of the sample and Y is the OD_{630nm} of the blank well. For all strains tested, GFP-UV expression was barely detectable when grown on pyruvate or pyruvate plus myo-inositol, glucose, mannitol, sorbitol and fructose (Figure 5.4, 5.5, 5.6). Induction of expression of GFP-UV only occurred in 10mM myo-inositol for 3841/pRU601 (Figure 5.4) and RU307/pRU601 (Figure 5.6). The expression by RU307/pRU601 was only a quarter of the expression of 3841/pRU601 (specific fluorescence of 11314 and 447735 respectively) (Figure 5.6).
These results indicate that there is a clear need for a functional myo-inositol catabolic pathway in order to induce expression of GFP-UV and that the promoter is specifically induced by myo-inositol.
Figure 5.4 Expression of GFP-UV by 3841/pRU601

P = 20mM pyruvate, PI = 20mM pyruvate + 10mM myo-inositol, I = 10mM myo-inositol, G = 10mM glucose, M = 10mM mannitol, S = 10mM sorbitol, F = 10mM fructose.

Each value represents the mean of three cultures, with standard deviation.
Figure 5.5  Expression of GFP-UV by RU360/pRU601 and RU361/pRU601

P = 20mM pyruvate, PI = 20mM pyruvate + 10mM myo-inositol, I = 10mM myo-inositol, G = 10mM glucose, M = 10mM mannitol, S = 10mM sorbitol, F = 10mM fructose. Each value represents the mean of three cultures, with standard deviation. RU360 = dark blue, RU361 = light blue.
Figure 5.6 Expression of GFP-UV by RU307/pRU601

P = 20mM pyruvate, PI = 20mM pyruvate + 10mM myo-inositol, I = 10mM myo-inositol, G = 10mM glucose, M = 10mM mannitol, S = 10mM sorbitol, F = 10mM fructose.

Each value represents the mean of three cultures, with standard deviation.
5.2.4 Catabolite Repression

Induction of GFP-UV expression by 10mM myo-inositol was repressed in 3841/pRU601 by the addition of 20mM pyruvate (Figure 5.4). Therefore, the effect of combinations of myo-inositol and other carbon sources (pyruvate, malate, succinate, glucose) on induction of GFP-UV expression was tested. (Figure 5.7). The results indicate that catabolite repression occurs. The extent of repression was dependent on the concentration of the additional substrate. Addition of each of the four carbon sources repressed induction of GFP-UV expression at a concentration of 1mM, but the inhibition was considerably less than at 10mM.
Figure 5.7  Catabolite Repression of myo-Inositol Inducible GFP-UV Expression in 3841/pRU601

I = 10mM myo-inositol, P = pyruvate, S = succinate, M = malate, G = glucose.

Each value represents the mean of three cultures, with standard deviation.
The kinetics of induction of GFP-UV expression after exposure to myo-inositol were measured in 3841/pRU601. Cells grown overnight in glucose were resuspended in different carbon sources and fluorescence was measured every 10 minutes for 830 minutes (Figure 5.8). Induction of expression of GFP-UV was considered to occur when the specific fluorescence was in excess of 1500V, as this is higher than expression in the absence of myo-inositol (Figure 5.4). Induction of expression of GFP-UV began at 150 minutes for pRU601 in 3841 in the presence of 10mM myo-inositol and 1mM myo-inositol.

When there was a combination of 10mM myo-inositol with 1mM of a different carbon source, induction of expression of GFP-UV was severely delayed. Expression of GFP-UV began after 200 minutes in 10mM myo-inositol plus 1mM pyruvate, 330 minutes in 10mM myo-inositol plus 1mM malate, 360 minutes 10mM myo-inositol plus 1mM succinate and 380 minutes in 10mM myo-inositol plus 1mM glucose. With all carbon sources, the level of expression increased over the time period measured.

There was no further increase in expression in 1mM myo-inositol after 830 minutes, due to the culture no longer growing, presumably because of substrate exhaustion. There was much higher overall expression in 10mM myo-inositol alone than when in combination with other carbon sources, although at the end of the time period, the amount of expression in the presence of 1mM pyruvate was almost as
much as in 1mM myo-inositol alone. This was probably due to exhaustion of pyruvate in the growth medium, which is a three-carbon compound.

There was no appreciable induction of expression of GFP-UV over the time measured when 3841/pRU601 was grown in 10mM pyruvate, glucose, succinate and malate as the sole carbon source. There was also no appreciable expression when there was a combination of 10mM amounts of the above carbon sources and 10mM myo-inositol. For clarity on the graph, only the data for 10mM pyruvate is shown (Figure 5.8).
Figure 5.8  GFP-UV Expression Over Time by 3841/pRU601

I = myo-inositol, P = pyruvate, G = glucose, S = succinate, M = malate.

Each set of data represents the mean of three cultures of 3841/pRU601, with standard deviation.
5.2.5 Relationship to myo-Inositol Mutants

The iolX_p fragment was extracted from the pOT1 vector as a 2078bp PacI fragment with flanking pOT1 DNA. This fragment was then used as a probe for Southern hybridisation studies with the complementing cosmids pRU3078, pRU3079 and pRU3111 digested with three different restriction enzymes, EcoRI, PstI and SalI. Plasmid pRU601 was digested with PacI, giving a 2078bp fragment and a 5246bp fragment, containing the pOT1 vector (Figure 5.9).

The probe bound to both pRU601 fragments that had been PacI digested. The binding of the probe to the larger fragment may have been due to partial digestion by PacI of pRU601. Incomplete digestion with PacI has been observed previously in the laboratory (P.S. Poole, pers. comm.).

There was no hybridisation of the fragment to the cosmids. These data indicate that iolX_p is a distinct locus to the myo-inositol utilisation loci identified in this project. This was not a surprise, as the genes mutated in RU360 and RU361 are distinct loci and there are clearly other genes involved in myo-inositol utilisation in Rhizobium that have yet to be identified.
Figure 5.10 Southern Hybridisation of pRU601

0.8% agarose gel showing DNA bands stained with ethidium bromide visualised on a UV transilluminator.

Southern blot with of above agarose gel.

1 = 1kb ladder, 2 = pRU3078 EcoRI, 3 = pRU3079 EcoRI, 4 = pRU3111 EcoRI, 5 = pRU3078 PstI, 6 = pRU3079 PstI, 7 = pRU3111 PstI, 8 = pRU3078 SalI, 9 = pRU3079 SalI, 10 = pRU3111 SalI, 11 = pRU601 Paci.
5.2.6 Expression in the Rhizosphere and Nodules

There is a clear pattern of induction of the myo-inositol inducible promoter \textit{in vitro}. To investigate the importance of myo-inositol in the rhizosphere, expression of GFP-UV by the \textit{inoXp} fragment was studied in the rhizosphere and nodules. \textit{P. sativum} and \textit{V. sativa} plants were inoculated with pRU601/3841 at a density of $10^6$ cfu per plant. Three individual plants were harvested 1, 2, 3 and 4 days post-inoculation. The rhizosphere bacteria were isolated and examined under a microscope for expression of GFP-UV. Less than 1% of bacteria observed in the rhizosphere expressed GFP-UV, indicating that the promoter was not induced by conditions in the rhizosphere. The bacteria that were thought to fluoresce might have been contaminating plant particles that autofluoresced and were mistaken for GFP-UV-expressing rhizobia.

Four weeks post-inoculation, five or six nodules were harvested from three \textit{P. sativum} plants, surface sterilised and crushed under a microscope to examine expression of GFP-UV in bacteroids. In contrast to the results gained from free-living bacteria in the rhizosphere, virtually all bacteroids observed expressed GFP-UV, indicating widespread induction of the promoter in nodules. However, it was not possible to quantify the amount of GFP-UV expression, so it is difficult to conclude that massive induction of expression occurred.
5.2.7 Stability of pRU601 in the Rhizosphere

It was not practical to apply gentamycin to the rhizosphere to guarantee the retention of pOT1 plasmid by 3841. It was therefore necessary to assess the stability of the plasmid in the absence of selective pressure. The 16 nodules harvested from three *P. sativum* plants in Section 5.2.6 were streaked onto TY agar, with and without gentamycin. Bacteria from 13 of the 16 nodules were gentamycin resistant (81%). Six individual colonies were streaked from each of the TY agar plates without gentamycin, onto AMA containing 10mM glucose or 10mM myo-inositol, plus gentamycin to select for pOT1. Of the resulting 78 colonies, 37 (47.5%) retained gentamycin resistance and expressed GFP-UV when grown on myo-inositol but not when grown on glucose (Table 5.1).

<table>
<thead>
<tr>
<th>plant</th>
<th>nodule</th>
<th>number of gentamycin resistant colonies (of six)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>
5.2.8 Construction of a Specific myo-Inositol Inducible GFP-UV Promoter Probe

The gene interrupted in RU360 was selected as a candidate for a defined promoter probe, in order to study the induction of myo-inositol catabolic genes in the rhizosphere and nodules. The promoter probe was made by PCR amplification of the complementing cosmid pRU3078 (Figure 5.10) and cloned into pOT1 containing the promoter-less gfpuv gene. This vector was chosen because GFP-UV expression had been successfully studied previously in vitro and in vivo.

Primers p213 and p214 were designed to amplify a 743bp fragment containing the beginning of iolD along with 350bp upstream of the putative start codon of iolD. It was hoped that this would contain the promoter region for the gene. The restriction enzyme sites PstI and SpeI were included in the primers p213 and p214 respectively, to enable the fragment to be cloned into pOT1 immediately upstream of and in the same direction as the gfpuv gene. The resulting plasmid was named pRU703. Plasmids pRU706 and pRU713 were also cloned in the same direction as the gfpuv gene, so that they too could be used for promoter probe studies (c.f. Section 3.2.5.1.1). Sequencing confirmed the correct orientation of the PCR fragments in pRU703, pRU706 and pRU713.
Plasmids pRU703, pRU706 and pRU713 were conjugated into RU360, RU361, RU307 and 3841 to examine induction of the *gfpuv* gene. None of the promoter probe plasmids expressed GFP-UV in the presence of *myo*-inositol or pyruvate (Figure 5.11). There is clearly a functional promoter present in the probes, as pRU713 complemented RU360, restoring the ability to grow on *myo*-inositol. Therefore, there may have been a problem with the cloning strategy. Co-workers in the laboratory have also found that the *gfpuv* gene is not expressed in plasmids constructed by cloning into the SpeI site of the pOT1 plasmid (P.S. Poole, pers. comm.). This site is immediately adjacent to the ribosomal binding site of the *gfpuv* gene so perhaps ribosome binding is disrupted when DNA is cloned into this site in the plasmid.

**Figure 5.10 Promoter-Probe Vectors of *iolD***

![Diagram of plasmids and vectors](image_url)
The 5.606kb region sequenced from pRU3078 with arrows detailing the sequence encompassed by the subclones.
Figure 5.11 Expression of GFP-UV in Promoter Probe Vectors

Each series represents the mean of three replicates. Plasmids in 3841 and RU307 were grown on 10mM myo-inositol, plasmids in RU360 and RU361 on 10mM myo-inositol plus 20mM pyruvate.
5.2.9 RU360 β-Galactosidase Assay

The lack of induction of GFP-UV in the promoter-probes pRU703, pRU706 and pRU713 suggests either that the genes interrupted in RU360 are not inducible by myo-inositol or that there was a problem with the cloning strategy. It was possible to test whether the genes are myo-inositol inducible because Tn5-lacZ in RU360 contains the promoter-less reporter gene lacZ, which codes for β-galactosidase and it is transcribed in the same direction as the interrupted genes.

It was postulated that the promoter for iolD would induce expression of the lacZ gene in the presence of myo-inositol. However, when β-galactosidase activity was measured in RU360, the fusion was not active (Poole et al., 1994). It was decided to repeat this assay and also to test whether the fusion was active in RU360 when complemented by pRU3078 and pRU3079 (Table 5.2). Strain RU361 was not tested, as Tn5-lacZ is transcribed in the opposite direction to the interrupted gene.

The data show that β-galactosidase activity was very low in RU360 grown on 20mM pyruvate. The presence of 10mM myo-inositol in the growth medium did not cause induction of expression of lacZ. When complemented with the cosmids pRU3078 and pRU3079, there were 2-fold and 5-fold increases respectively in the activity of lacZ when myo-inositol was present in addition to pyruvate. When RU360/pRU3078 and RU360/pRU3079 were grown on myo-
inositol as the sole carbon source, there was an increase in expression of *lacZ* of 12-fold and 20-fold respectively. These data indicate that the genes interrupted by Tn5-*lacZ* in RU360 are *myo*-inositol inducible, but that a functional catabolic pathway is required for induction and pyruvate represses induction.

Table 5.2  β-Galactosidase Activity of RU360

<table>
<thead>
<tr>
<th>strain</th>
<th>activity (nmol/mg/min)</th>
<th>pyruvate</th>
<th>pyruvate + <em>myo</em>-inositol</th>
<th><em>myo</em>-inositol</th>
</tr>
</thead>
<tbody>
<tr>
<td>RU360</td>
<td>160.3 ± 15.5</td>
<td>132.2 ± 4.1</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>RU360/pRU3078</td>
<td>125.74 ± 5466</td>
<td>264.80 ± 42466</td>
<td>1453.59 ± 3777</td>
<td></td>
</tr>
<tr>
<td>RU360/pRU3079</td>
<td>90.0 ± 4.3</td>
<td>459.4 ± 35.3</td>
<td>1871.41 ± 315.90 ± 69</td>
<td></td>
</tr>
</tbody>
</table>

Each value represents the mean of three cultures with standard deviation.
5.2.10 *myo*-Inositol Catabolic Enzyme Assays

It has been shown previously that the *myo*-inositol catabolic enzymes, *myo*-inositol dehydrogenase and 2-keto-*myo*-inositol dehydratase are not induced in free-living RU360 or bacteroids of 3841 (Poole *et al*., 1994). To determine whether the enzymes are inducible in free-living RU361 and RU307, this experiment was carried out with these two mutants grown overnight in 20mM pyruvate or 20mM pyruvate plus 10mM *myo*-inositol. RU360 and 3841 were tested as controls (Table 5.3).

Bacteroids of 3841 were also tested. Poole *et al.* (1994) isolated bacteroids under air by differential centrifugation. However, the nitrogenase enzyme is inactivated in the presence of oxygen and it is possible that the bacteroids were damaged by the isolation procedure. Therefore, bacteroids were used that had been isolated anaerobically by a co-worker in the laboratory.

The data indicate that the first two enzymes in the *myo*-inositol breakdown pathway were not induced in 3841 or the mutants when grown on pyruvate as the sole carbon source. There was an 18-fold increase in *myo*-inositol dehydrogenase activity and a 50-fold increase in 2-keto-*myo*-inositol dehydratase activity by 3841 when grown on *myo*-inositol. There was some catabolite repression of the dehydratase, as when grown on pyruvate plus *myo*-inositol, activity was only 45% of the activity when grown on *myo*-inositol alone (Table 5.3).
There was a 7-fold increase in activity of both enzymes in RU307 when grown on myo-inositol as the sole carbon source. There was a reduction when RU307 was grown on pyruvate plus myo-inositol as the activities were only 51% and 52% of the activities when grown on myo-inositol alone. Overall, the activities in RU307 were 43% and 14% of the activities in 3841.

There was no induction of myo-inositol dehydrogenase by RU360 and RU361 of the dehydrogenase when grown on pyruvate plus myo-inositol. There was some induction of 2-keto-myoinositol dehydratase by RU360 and RU361 when grown on pyruvate plus myo-inositol, but the activities were only 16% and 15% respectively of the amount of activity of 3841 grown on pyruvate plus myo-inositol. There was no induction in the bacteroids. The bacteroids were positive for alanine dehydrogenase activity, which indicated that the bacteroids were not damaged by the isolation process.
### Table 5.3 Activity of myo-Inositol Catabolic Enzymes

<table>
<thead>
<tr>
<th>Strain</th>
<th>carbon source</th>
<th>myo-inositol dehydrogenase</th>
<th>2-keto-myoinositol dehydratase</th>
</tr>
</thead>
<tbody>
<tr>
<td>RU360</td>
<td>pyruvate</td>
<td>0.04 ± 0.001</td>
<td>0.014 ± 0</td>
</tr>
<tr>
<td></td>
<td>pyruvate + myo-inositol</td>
<td>0.01 ± 0.002</td>
<td>0.094 ± 0.005</td>
</tr>
<tr>
<td>RU361</td>
<td>pyruvate</td>
<td>0.04 ± 0.006</td>
<td>0.038 ± 0.011</td>
</tr>
<tr>
<td></td>
<td>pyruvate + myo-inositol</td>
<td>0.05 ± 0.017</td>
<td>0.085 ± 0.003</td>
</tr>
<tr>
<td>RU307</td>
<td>pyruvate</td>
<td>0.025 ± 0.001</td>
<td>0.025 ± 0.01</td>
</tr>
<tr>
<td></td>
<td>pyruvate + myo-inositol</td>
<td>0.09 ± 0.018</td>
<td>0.096 ± 0.004</td>
</tr>
<tr>
<td></td>
<td>myo-inositol</td>
<td>0.175 ± 0.01</td>
<td>0.184 ± 0.01</td>
</tr>
<tr>
<td>3841</td>
<td>pyruvate</td>
<td>0.023 ± 0.005</td>
<td>0.026 ± 0.02</td>
</tr>
<tr>
<td></td>
<td>pyruvate + myo-inositol</td>
<td>0.404 ± 0.101</td>
<td>0.579 ± 0.104</td>
</tr>
<tr>
<td></td>
<td>myo-inositol</td>
<td>0.409 ± 0.11</td>
<td>1.295 ± 0.446</td>
</tr>
<tr>
<td>3841</td>
<td>N/A</td>
<td>0.003 ± 0.001</td>
<td>0.013 ± 0.005</td>
</tr>
<tr>
<td>bacteroids</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Mean of at least two cultures assayed in triplicate, with standard deviation.
5.3 Discussion

The *iolXp* fragment contains a promoter that is specifically induced by *myo*-inositol. It is not possible that the promoter for the regulatory gene *orf1* induces expression of the GFP-UV gene, as it is transcribed divergently. Therefore, a promoter from one of the other three ORFs must be responsible for the *myo*-inositol dependent induction of GFP-UV. In order to test this hypothesis promoter probes could be made. Nested deletion could be used to identify the smallest region necessary for induction of GFP-UV expression.

The ORFs could be tested to determine whether they encode genes involved in *myo*-inositol utilisation by transposon mutagenesis of each individual ORF. It is not expected that the ORFs encode genes involved in *myo*-inositol utilisation, because there was no homology with known *myo*-inositol utilisation genes, although *orf4* is too small a region to be certain, but they could encode regulators, the structure of which are not known. If *orf4* were shown to be responsible for the *myo*-inositol dependent expression of GFP-UV, then *iolXp* could be used as a probe to obtain a larger clone. Inverse PCR could be used to identify the entire gene.

Based on sequence homology of the deduced protein, *orf1* is thought to encode the response regulator component of a two-component regulatory system. Such systems allow bacteria to make adaptive
responses to changes in their environment. The most homologous protein was FeuP of *R. leguminosarum* bv. *viciae*, which is involved in the regulation of iron uptake (Yeoman *et al.*, 1997). The deduced product of *orf1* contained conserved residues that are characteristic of similar regulatory proteins, although there was only 39% identity to FeuP. Therefore, *orf1* is not likely to encode a protein that is functionally equivalent to FeuP. To deduce the function of *orf1*, a transposon mutant could be made so that the phenotype could be determined. It is postulated that downstream of *orf1* will be a gene encoding the sensor regulator component. This could be ascertained by constructing a primer to read further downstream in the chromosomal DNA. To test whether this regulatory system affects *myo*-inositol metabolism, a mutant would be required.

When conjugated into the *myo*-inositol mutants, the promoter present in *iolXp* was not induced in any of the strains grown on pyruvate or pyruvate plus *myo*-inositol. There was induction by RU307/pRU601 grown on *myo*-inositol although expression was only a quarter that of 3841/pRU601. The promoter is probably not induced by *myo*-inositol alone, as some expression would also be expected in all the strains grown on pyruvate plus *myo*-inositol, including RU360/pRU601 and RU361/pRU601 because they are able to transport some *myo*-inositol into the cell.
It can be hypothesised that the promoter responds to a breakdown product of myo-inositol. The growth rate of RU307 on myo-inositol is approximately a quarter the rate of 3841, due to reduced uptake of myo-inositol by RU307 into the cell. Strain 3841 has an inducible system for myo-inositol uptake (c.f. Section 3.2.5). If induced by a breakdown product of myo-inositol, this would explain why the induction of expression of GFP-UV was lower in RU307 than 3841. To ascertain the compound that induces iolXp, the induction of expression of GFP-UV by 3841/pRU601 could be measured following addition of intermediates in the myo-inositol breakdown pathway.

The myo-inositol dependent induction of expression of GFP-UV was repressed by other carbon compounds. This suggests that other carbon sources are utilised preferentially to myo-inositol and that catabolite repression occurs. When measured over time, the level of GFP-UV expression increased, which may reflect the catabolite repressors being removed from the growth media and utilisation of myo-inositol beginning.

The iolXp promoter could be a powerful tool in the identification of genes involved in myo-inositol utilisation. If iolXp were conjugated into a 3841 transposon library and the resulting conjugants spread onto AMA containing 10mM myo-inositol as the sole carbon source, it would be expected that most colonies would glow. Any that did not would be likely to be mutated in genes involved in induction of the promoter.
This simple screening method could also be used to identify mutants that have escaped catabolic control. If the colonies were plated onto AMA containing myo-inositol plus other carbon sources such as glucose and pyruvate, then the colonies would not be expected to glow, due to catabolite repression. Any colonies that glow would have escaped this and therefore, may be mutated in genes encoding global regulators of myo-inositol catabolism.

The presence of myo-inositol in root exudates and soil has been reported (Lynch et al., 1958, McKercher and Anderson, 1968, Sulochana, 1962, Yoshida, 1940, Wood and Stanway, 2000). However, there was no induction of GFP-UV expression in the rhizosphere of P. sativum or V. sativa plants, despite evidence that pRU601 was retained by 3841 in 81% of nodules and so was probably stable in the rhizosphere. There are obviously other carbon sources released into the rhizosphere by P. sativum and V. sativa, as the myo-inositol mutants RU360, RU361 and RU307 were able to grow as well as 3841 in the rhizosphere of V. sativa plants (c.f. Section 4). Therefore, there may be catabolite repression of the myo-inositol inducible promoter occurring in the rhizosphere. It is also possible that myo-inositol was not released into the rhizosphere in significant amounts during this experiment. However, the rhizosphere was not analysed for carbon compounds released.
In contrast to the lack of expression in the rhizosphere, there was widespread expression of GFP-UV by bacteroids in pea nodules. The presence of myo-inositol, derivatives and other cyclitols in nodules on several legumes, including *P. sativum* has been widely documented (Davis and Nordin 1983, Kouchi and Yoneyama, 1986, Kouchi and Yoneyama 1984, Skøt and Egsgaard, 1984, Streeter, 1987; Streeter and Saminen, 1986). Carbon dioxide labelling studies, oxygen consumption and enzyme activity assays have shown that myo-inositol and other cyclitols are probably not metabolised in nodules (Davis and Nordin 1983, Kouchi and Yoneyama, 1986, Streeter and Salminen, 1986, Poole *et al.*, 1994). However, the first two enzymes in the myo-inositol catabolic pathway were not expressed in pea nodules containing 3841 (c.f. Section 5.2.10, Poole *et al.*, 1994). Therefore, myo-inositol catabolism appears to be switched off in the nodule. This might be due to repression of catabolic enzymes by other carbon sources, particularly dicarboxylates, which are the preferred carbon source by bacteroids in nodules.

The lack of induction of the first two enzymes in the myo-inositol catabolic pathway was confirmed for free-living RU360 and RU361. Sequencing of the mutated region in the mutants did not reveal any putative genes with homology to those encoding the first two enzymes in the myo-inositol catabolic pathway. Therefore, there could not have been a polar effect of the transposon insertion on these genes. This indicates that there is a global system of regulation of myo-inositol catabolic genes, where the final
breakdown products of the pathway induce expression of the first enzymes. Strain RU307 did not induce the first two enzymes in the pathway when grown on pyruvate and myo-inositol, but when grown on myo-inositol alone, the enzymes were induced at similar rates to that of 3841, indicating that there was no deficiency in the myo-inositol breakdown pathway.

There are clearly two different systems of regulation that occur in free-living rhizobia and in bacteroids. There may be a difference in the regulation of iolX_p in bacteroids as the iolXp promoter is expressed in nodules, even though the catabolic enzymes are not. The product that induces expression of iolX_p might be produced via an alternative pathway that does not utilise the first two enzymes in the catabolic pathway. Or, there might be another compound present in the nodule that has not been tested that can also cause expression of the promoter, such as a derivative of myo-inositol. However, an alternative explanation relates to the sensitivity of the microscope used to study GFP-UV expression. It was noted that although RU360/pRU601 does not appear to express GFP-UV in solid or laboratory media, when cells that had been left to grow for a week were examined under a microscope, some GFP-UV expression was observed, although nowhere near as much as 3841/pRU601 (P.S. Poole, pers.comm.). The level of GFP-UV expression observed by bacteroids in nodules was not quantified, therefore, it is not possible to conclude that widespread expression was occurring. Therefore, it may have just been background expression and nothing specific after all.
One situation where there is a difference in gene regulation between free-living rhizobia and bacteroids concerns rhizopines. The expression of rhizopine synthesis (mos) genes in *S. meliloti* is controlled by the symbiotic nitrogen-fixation regulatory system NifA/NtrA (Murphy *et al.*, 1988). In contrast, free-living rhizobia express rhizopine catabolism (moc) genes, which are not expressed by bacteroids. Rhizopines can be utilised by some rhizobia as the sole carbon and nitrogen source, which may give them a competitive advantage over non-utilising strains. Alternatively, they might be involved in signalling between bacteroids and free-living rhizobia. Strain 3841 is not a rhizopine producing strain, so this process could not be responsible for induction of *iolX*<sub>p</sub>, but there might be other *myo*-inositol derivatives that have not yet been identified that may cause expression of GFP-UV in nodules.

A future experiment could examine whether there is induction of expression of GFP-UV in bacteroids of the *myo*-inositol catabolic mutants. If RU360/pRU601, RU361/pRU601 or RU307/pRU601 express GFP-UV in bacteroids, this would indicate that the regulation of *iolX*<sub>p</sub> is different in nodules to free-living rhizobia in the rhizosphere and that either *myo*-inositol is being utilised by a different pathway, or other compounds induce the promoter. However, as less *myo*-inositol is transported into the mutants, it might be difficult to view GFP-UV expression *in vivo*. 
The complementation of the RU360 mutation by pRU713 (c.f. Section 3.2.5.1.1) indicates that this fragment contains a functional promoter. The failure of the promoter probe vectors to induce expression of GFP-UV was unexpected. Sequencing of the plasmids indicated that there was no apparent deletion of part of the promoter-less GFP-UV gene although the insertion site of the PCR fragments in the SpeI site of the pOT1 plasmid was very close to the start of the gfpuv gene. The presence of a terminator at the end of pRU713 might explain the lack of GFP-UV expression for this fragment, but not why there was no expression in pRU703 and pRU706. The β-galactosidase assay, measuring activity of the lacZ gene of Tn5-lacZ in RU360 indicated that the interrupted gene was induced by myo-inositol, but only when the catabolic pathway was complete. Therefore, induction of expression of 3841/pRU703 and 3841/pRU706 would have been expected.

The most likely explanation is that there was a problem with the cloning strategy, as co-workers in the laboratory have also found that the gfpuv gene is not expressed in plasmids constructed by cloning into the SpeI site of the pOT1 plasmid (P.S.Poole, pers. comm.). This site is immediately adjacent to the ribosomal binding site of the gfpuv gene so perhaps ribosome binding is disrupted when DNA is cloned into this site in the plasmid. To overcome this, the fragments would need to be cloned into a different restriction enzyme site further upstream of the ribosomal binding site, such as the SalI site. This is known to produce active fusions as this was the site used for construction of the Rhizobium library that was used in this project.
The lacZ activity of RU360 when complemented by the cosmids pRU3078 and pRU3079 demonstrates that myo-inositol, or one of its breakdown products, induces the gene interrupted by Tn5-lacZ. Uptake assays demonstrated that myo-inositol is taken into the cells by RU360 but at a low rate. There was no induction in RU360 grown on pyruvate plus myo-inositol, yet there was a small amount of induction when complemented by the cosmids pRU3078 and pRU3079. The lower activity by RU360/pRU3078 and RU360/pRU3079 when grown in the presence of pyruvate was probably due to catabolite repression. Therefore, once again, a breakdown product of myo-inositol is thought to induce the promoter.
Chapter 6 - Final Discussion
6.1 Conclusions

myo-Inositol is ubiquitous throughout nature and is an essential component of many compounds. Several environmental micro-organisms have been identified that can utilise myo-inositol as the sole carbon source, including *B. subtilis*, *K. aerogenes*, *Pseudomonas* species, *S. meliloti* and *R. leguminosarum*. It is likely that there is a specific myo-inositol catabolic pathway that is conserved between these species.

During this project, three *R. leguminosarum* bv. *viciae* myo-inositol mutants were characterised. Strains RU360 and RU361 were mutated in putative genes encoding enzymes in the catabolic pathway. Strain RU360 was mutated in acetolactate synthase (*iolD*) and RU361 in malonic semialdehyde oxidative decarboxylase (*iolA*). Strain RU307 was mutated in a transport system for myo-inositol (*intA*). It is clear that the regulation of myo-inositol catabolism is under control of a global regulatory system, as the catabolic genes are not arranged in one single operon, yet mutation of the final genes in the pathway prevented induction of the first enzymes and the transport system (Int). Based on the probable function of *iolD*, additional steps were added to the proposed catabolic pathway. The work in this thesis shows that myo-inositol catabolism is not complete until acetolactate or a further product has been formed.

It is not possible to assign roles for *iolE* or *iolB*, the two putative genes downstream of *iolD*. Complementation of RU360 required *iolDEB*, showing
that one or both of the *iolEB* genes are required for catabolism. *IolE* and *IolB* of *B. subtilis* are both required for *myo*-inositol catabolism, but their functions are unknown.

The high homology of the genes identified in this project with *B. subtilis* *myo*-inositol catabolic genes indicates that the genes originated from the same common ancestor. However, the genes are not arranged in similar patterns, as the *B. subtilis* genes are arranged in one operon, whereas the *R. leguminosarum* bv. *viciae* genes are located at more than three loci. The three loci identified are not likely to be widely separated, since cosmids containing approximately 30kb of DNA were not able to cross-complement. It is not currently known whether all the genes of *R. leguminosarum* bv. *viciae* are chromosomal as not all of the genes have been identified. There has only been limited research into *myo*-inositol utilisation in rhizobia, but the *myo*-inositol dehydrogenase gene (*idh*) in *S. meliloti* is known to be isolated on the chromosome. In contrast, in *R. trifolii*, some, if not all, of the genes are plasmid-encoded, as are the rhizopine catabolic genes, which are located on the Sym plasmid in *R. leguminosarum* bv. *viciae* and *S. meliloti*.

The reasons for the separation of the genes are not known. Perhaps the genes were originally arranged in an operon in a common ancestor, but were in a promiscuous region, which led to their separation. Alternatively, they may have all originally been on a plasmid and gradually became integrated into the chromosome. It is also possible that the genes were separate and came together in *B. subtilis*. The genes are likely to be subject to different types and
levels of regulation and may even function in more than pathway, which would also explain their separation. It is not possible to speculate on the reasons for the spatial separation of the genes, based on the current evidence, but if the genes are identified in more organisms it might become possible to determine their evolutionary path.

There are at least two transport systems for myo-inositol in *R. leguminosarum* bv. *viciae*, one of which is inducible and highly specific for myo-inositol uptake. The presence of a specific pathway suggests that myo-inositol is an important compound. This system might also be used to transport rhizopines, because of the high homology of *intB* with *mocB* and the fact that only a binding protein, but not a complete transport system has been identified amongst the known *moc* genes. However, uptake was repressed by growth in the presence of pyruvate, indicating that myo-inositol is not preferentially utilised as a carbon source by *Rhizobium*. myo-Inositol has many functions in cells and so it would be expected that a high steady-state intracellular concentration would be required to induce the catabolic pathway. Otherwise, myo-inositol might be used as a carbon source when it is actually required for biosynthesis of other compounds.

Based on the growth rates *in vitro* and *in vivo*, the transposon insertions in RU360, RU361 and RU307 did not impair growth relative to the wild type. All three mutants were also able to nodulate plants and fix nitrogen at rates similar to that of the wild type and the dry weights of the plants were all similar. However, when co-inoculated with the wild type, RU360 and RU361
were at a massive competitive disadvantage for nodulation. However, RU307 was very competitive with the wild type. The disadvantage was thought to be unlikely to be due to the ability to utilise myo-inositol as a carbon source, because RU307 grows much more slowly than the wild type in vitro. However, in the infection threads, it is not known what other factors might be limiting that enable RU307 to compete successfully with the wild type. In the rhizosphere, the ability to utilise myo-inositol offered no advantage to 3841 when co-inoculated with the mutants and other carbon compounds may repress the myo-inositol catabolic enzymes, as occurred in the in vitro studies. It is possible that in situ, when there are many different types of organisms competing for resources, the ability to utilise myo-inositol or derivatives such as rhizopines will be important, as was suggested by the work of Heinrich et al. (1999). However, the strain used in these experiments is not a rhizopine utiliser.

Clearly, the ability to respond to myo-inositol is crucial to nodule competitiveness. It is postulated that the ability to respond to the presence of myo-inositol might be part of a signalling mechanism between the host plant and rhizobia as there are likely to be large quantities of myo-inositol and derivatives in areas of developing nodules. myo-Inositol and derivatives have been consistently reported in nodules. Their role is thus far unknown, but the evidence reported in this work and elsewhere indicates that are probably not utilised as a carbon source. One possibility is that these compounds are involved in regulating osmotic potential, which is crucial for bacteroids in nodules.
myo-Inositol or derivatives induce the first two enzymes in the catabolic pathway, as well as *iolD* and *iolXp*. The first two enzymes in the catabolic pathway are not induced in the *myo*-inositol catabolic mutants RU360 and RU361, indicating that one or more of the final products induce the first enzymes, in a feedback mechanism. The expression of Tn5-*lacZ* also required a complete pathway and was repressed by pyruvate, indicating that *iolD* might be regulated differently to the first two enzymes in the pathway. Expression of *iolXp* was also repressed in the catabolic mutants and by the presence of different carbon compounds in 3841, again indicating that it is a breakdown product rather than pure *myo*-inositol that induces these promoters.

In nodules, bacteroids of 3841/pRU601 were seen to be expressing GFP-UV, indicating induction of *iolXp*. The first two enzymes in the catabolic pathway were not induced in bacteroids. This suggests that there a change in regulation of *myo*-inositol genes in bacteroids compared with in free-living bacteria. However, these results must be treated with caution, as the microscope is extremely sensitive and so there may not have been true induction. There are likely to be many other carbon compounds present in the nodule, including C₄-dicarboxylic acids, which are the major energy source of bacteroids. In free-living bacteria, malate and succinate inhibited *myo*-inositol dependent induction of *iolXp*, but if GFP-UV expression is being induced in the nodule, the lack of repression in the nodule could be due to a high flux and low concentration of C₄-dicarboxylic acids, which are used as an energy
source. There might also be a change in regulation of the genes during symbiosis that results in myo-inositol being channelled into different biosynthetic pathways. This could be tested by mutation of key nitrogen regulatory genes such as nifA and then observing whether that has an effect on expression of GFP-UV.

To conclude, during this project, it was shown that the ability to utilise myo-inositol is crucial for nodulation competition in _R. leguminosarum_ bv. _viciae_. The specific role of myo-inositol remains unknown. Given the plethora of roles that myo-inositol plays in cell function in all organisms, it would seem unlikely that myo-inositol is only important as a carbon source in _Rhizobium_. Possible explanations are that it is utilised as a carbon source in infection threads, that it is involved in signalling between the host plant and bacterium and that derivatives are involved in regulating osmotic potential in the nodule.
6.2 Future Work

There are several experiments that have been identified in each chapter that are important to further ascertain the role of myo-inositol in competition for nodulation. However, due to time constraints within this project, the experiments were not carried out. The experiments are summarised below.

It is crucial to identify all the genes involved in myo-inositol utilisation in order to elucidate the full catabolic pathway. Although the inducible expression of the first two enzymes in the pathway is suggestive of rhizobia sharing the same pathway as *B. subtilis*, it is not unequivocal proof that the entire pathway is the same. The mechanism of regulation also needs to be discovered. It is not currently known whether the genes studied in this project are induced by myo-inositol or by derivatives. The uptake system was inhibited by other sugars, as was the *iolXp* promoter and Tn5-*lacZ* in RU360. To determine whether all the genes are chromosomal, 3841 could be cured of plasmids and then tested to see if it retains the ability to utilise myo-inositol. Another way to identify genes would be to make PCR probes from the *B. subtilis iol* genes and from the *S. meliloti idh* gene.

It is also important to determine the exact regions required to restore the ability to utilise myo-inositol to the mutants and to eliminate the possibility of polar effects on downstream genes caused by transposon insertion. This is particularly crucial for RU307, as the area around the insertion has not been fully sequenced, due to the lack of a complementing cosmid.
It would also be useful to confirm the specificity of the myo-inositol uptake system, using myo-inositol derivatives, such as rhizopines, cyclitols and plant myo-inositol compounds. The whole region around the uptake system also needs to be elucidated, by complementation of RU307, or by constructing primers to extend the sequence further using chromosomal DNA. This should also help to determine the nature of the glutamic acid gamma hydrazide phenotype. By identifying the cause of the toxic escape phenotype, it should be possible to eliminate it as the cause of the competitiveness of RU307.

It is also crucial to discover the function of the gene(s) that the iolXp promoter regulates or is a part of. If the myo-inositol dependent expression of GFP-UV is regulated differently in free-living rhizobia and bacteroids, this may indicate a novel role for the gene(s) regulated by this promoter. Determining which genes are regulated could lead to further understanding of the role of myo-inositol in nodules. This could be achieved by making mutants and screening for loss of the ability to induce expression of iolXp when grown on myo-inositol, or by screening for switching on of iolXp when grown on other carbon sources.

To ascertain whether 3841 has an advantage in nodulation because it induces nodulation more quickly, analysis could be carried out in planta, using staining techniques to study nodule initiation. Timed inoculation experiments could also be carried out. The nodule
occupancy studies should also be repeated, using larger sample sizes, to see if RU307 was at some disadvantage. By making complementing subclones of the cosmids that are stable in the rhizosphere, it should be possible to test more accurately whether restoration of the ability to utilise myo-inositol restored competitive ability to the mutants. This would also eliminate the possibility that independent regions of the cosmids compensated for the lack of the ability to utilise myo-inositol.

Strain RU307 also needs to be tested to see if it is able to utilise rhizopines as the sole carbon source when the moc genes are introduced into this strain. It is postulated that growth will be very slow, if at all, because in RU307, the second, low affinity system may not necessarily transport rhizopines as well as myo-inositol.

By studying the earlier processes of infection of plants by rhizobia, it should be possible to track the progress of the mutants and 3841, possibly by using GFP-UV. The aim would be to identify the exact stage at which 3841 gains a competitive advantage over the myo-inositol mutants. It is also necessary to further check the competitiveness of RU307, to determine whether it is equally competitive as 3841, as that would be an indicator that it is not the ability to utilise myo-inositol as a carbon source.
Ultimately, the work carried out on *R. leguminosarum* bv. *viciae* could be repeated in other species of Rhizobium, in order to determine whether myo-inositol is important in all rhizobia.
Bibliography


Rhizobium meliloti with an extra copy of dctABD and/or modified nifA expression.

*Applied and Environmental Microbiology*. 60: 3815-3852.


iron-molybdenum cofactor of nitrogenase from *Klebsiella pneumoniae*.  


Li, P-L., Farrand, S.K. (2000). The replicator of the nopaline-type Ti plasmid pTiC58 is a member of the *repABC* family and is influenced by the TraR-dependent quorum-sensing regulatory system. *Journal of Bacteriology*. 182: 179-188.


Rossbach, S., Rasul, G., Schneider, M., Eardly, B., deBruijn, F.J. (1995). Structural and functional conservation of the rhizopine catabolism (moc) locus is limited to selected Rhizobium meliloti strains and unrelated to their geographical origin. Molecular Plant-Microbe Interactions. 4: 549-559.


a proline dehydrogenase mutant (*putA*) from *Bradyrhizobium japonicum*.  


APPENDIX
## Appendix 3.1 The Genetic Code

<table>
<thead>
<tr>
<th>codon</th>
<th>amino acid</th>
<th>letter</th>
</tr>
</thead>
<tbody>
<tr>
<td>ttt, ttc</td>
<td>phenylalanine</td>
<td>F</td>
</tr>
<tr>
<td>tta, ttg, ctt, ctc, cta, ctg</td>
<td>leucine</td>
<td>L</td>
</tr>
<tr>
<td>tct, tcc, tca, tcg, agt, cgc</td>
<td>serine</td>
<td>S</td>
</tr>
<tr>
<td>tat, tac</td>
<td>tyrosine</td>
<td>Y</td>
</tr>
<tr>
<td>taa, tag, tga</td>
<td>TERMINATION</td>
<td>N/A</td>
</tr>
<tr>
<td>tgt, tgc</td>
<td>cysteine</td>
<td>C</td>
</tr>
<tr>
<td>cct, ccc, cca, ccg</td>
<td>proline</td>
<td>P</td>
</tr>
<tr>
<td>tgg</td>
<td>tryptophan</td>
<td>W</td>
</tr>
<tr>
<td>cat, cac</td>
<td>histidine</td>
<td>H</td>
</tr>
<tr>
<td>caa, cag</td>
<td>glutamine</td>
<td>Q</td>
</tr>
<tr>
<td>cgt, cgc, cga, cgg, aga, agg</td>
<td>arginine</td>
<td>R</td>
</tr>
<tr>
<td>att, atc, ata</td>
<td>isoleucine</td>
<td>I</td>
</tr>
<tr>
<td><strong>atg</strong></td>
<td>methionine</td>
<td>M</td>
</tr>
<tr>
<td>act, acc, aca, acg</td>
<td>threonine</td>
<td>T</td>
</tr>
<tr>
<td>aat, aac</td>
<td>asparagine</td>
<td>N</td>
</tr>
<tr>
<td>aaa, aag</td>
<td>lysine</td>
<td>K</td>
</tr>
<tr>
<td>gtt, gtc, gta, <strong>gtg</strong></td>
<td>valine</td>
<td>V</td>
</tr>
<tr>
<td>gct, gcc, gca, gcg</td>
<td>alanine</td>
<td>A</td>
</tr>
<tr>
<td>gat, gac</td>
<td>aspartic acid</td>
<td>D</td>
</tr>
<tr>
<td>gaa, gag</td>
<td>glutamic acid</td>
<td>E</td>
</tr>
<tr>
<td>ggt, ggc, gga, ggg</td>
<td>glycine</td>
<td>G</td>
</tr>
</tbody>
</table>

Start codons are highlighted in bold.
Appendix 3.2 Analysis of variance table for myo-inositol uptake by 3841, RU360, RU361 and RU307

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>VR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>4687.295</td>
<td>9</td>
<td>520.811</td>
<td>212.41</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Within Groups</td>
<td>46.586</td>
<td>19</td>
<td>2.452</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>4733.881</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

L.S.D. = 2.676

Appendix 3.3 Analysis of variance table for glucose uptake by 3841, RU360, RU361 and RU307

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>VR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>517.84</td>
<td>9</td>
<td>57.54</td>
<td>1.78</td>
<td>0.161</td>
</tr>
<tr>
<td>Within Groups</td>
<td>453.06</td>
<td>14</td>
<td>32.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>970.90</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

L.S.D. = 9.96
Appendix 3.4  Analysis of variance table for myo-inositol uptake by 3841 in the presence of excess sugar

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>VR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>3969.23</td>
<td>7</td>
<td>567.03</td>
<td>13.32</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Within Groups</td>
<td>680.87</td>
<td>16</td>
<td>42.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>4650.1</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

L.S.D. = 11.29

Appendix 3.5  Analysis of variance table for glucose uptake by 3841 in the presence of excess sugar

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>VR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>879.57</td>
<td>7</td>
<td>125.65</td>
<td>5.56</td>
<td>0.002</td>
</tr>
<tr>
<td>Within Groups</td>
<td>361.31</td>
<td>16</td>
<td>22.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1240.88</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

L.S.D. = 8.225
Appendix 4.1  Analysis of variance table for nodule number on *V. sativa* plants inoculated with 3841, RU360, RU361 and RU307.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>VR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>2.94</td>
<td>3</td>
<td>0.98</td>
<td>0.08</td>
<td>0.972</td>
</tr>
<tr>
<td>Within Groups</td>
<td>341</td>
<td>27</td>
<td>12.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>343.94</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

L.S.D. = 3.646

Appendix 4.2  Analysis of variance table for dry weight of *V. sativa* plants inoculated with 3841, RU360, RU361 and RU307

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>VR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>0.0083</td>
<td>4</td>
<td>0.0021</td>
<td>7.39</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Within Groups</td>
<td>0.0089</td>
<td>32</td>
<td>0.0003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.0172</td>
<td>36</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

L.S.D. = 0.0184
### Appendix 4.3 Analysis of variance table for Acetylene Reduction Assay of *P. sativum* plants inoculated with 3841, RU361 and RU307

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>VR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>0.191E-05</td>
<td>2</td>
<td>0.956E-06</td>
<td>2.63</td>
<td>0.088</td>
</tr>
<tr>
<td>Within Groups</td>
<td>0.116E-04</td>
<td>32</td>
<td>0.363E-06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.135E-04</td>
<td>34</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

L.S.D. = 0.00055

### Appendix 4.4 Analysis of variance table for dry weight of *P. sativum* plants inoculated with 3841, RU361 and RU307

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>VR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>0.06109</td>
<td>2</td>
<td>0.03054</td>
<td>1.27</td>
<td>0.316</td>
</tr>
<tr>
<td>Within Groups</td>
<td>0.28841</td>
<td>12</td>
<td>0.02403</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.34950</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

L.S.D. = 0.2136
Appendix 4.5  Analysis of variance table for nodule number on *P. sativum* plants inoculated with 3841, RU361 and RU307

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>VR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>2784.6</td>
<td>2</td>
<td>1392.3</td>
<td>12.48</td>
<td>0.001</td>
</tr>
<tr>
<td>Within Groups</td>
<td>1339.2</td>
<td>12</td>
<td>111.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>4123.8</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

L.S.D. = 14.56

Appendix 4.6  Analysis of variance table for nodule mass of *P. sativum* plants inoculated with 3841, RU361 and RU307

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>VR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>0.000085</td>
<td>2</td>
<td>0.000042</td>
<td>0.07</td>
<td>0.933</td>
</tr>
<tr>
<td>Within Groups</td>
<td>0.007246</td>
<td>12</td>
<td>0.000604</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.007333</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

L.S.D. = 0.03387
Appendix 4.7 Analysis of variance table for nodule numbers of *V. sativa* when inoculated with 3841, RU360, RU360/pRU3078 and RU360/pRU3079

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>VR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within Groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L.S.D. =</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Appendix 4.8 Analysis of variance table for nodule numbers of *V. sativa* when inoculated with 3841, RU361, RU361/pRU3111

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>VR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>588.75</td>
<td>6</td>
<td>98.13</td>
<td>4.04</td>
<td>0.002</td>
</tr>
<tr>
<td>Within Groups</td>
<td>1215.56</td>
<td>50</td>
<td>24.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1808.32</td>
<td>56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L.S.D. =</td>
<td>5.114</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Appendix 4.9  Analysis of variance table for nodule numbers of *V. sativa* when inoculated with 3841 and RU307

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>VR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>1373.09</td>
<td>4</td>
<td>343.27</td>
<td>11.69</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Within Groups</td>
<td>1585.48</td>
<td>54</td>
<td>29.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2958.58</td>
<td>58</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

L.S.D. = 3.435

### Appendix 4.10  Analysis of variance table for dry weight of *V. sativa* plants when inoculated with 3841, RU360, RU360/pRU3078 and RU360/pRU3079

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>VR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>0.033277</td>
<td>9</td>
<td>0.003698</td>
<td>11.13</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Within Groups</td>
<td>0.016272</td>
<td>49</td>
<td>0.000332</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.049551</td>
<td>58</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

L.S.D. = 0.02218
Appendix 4.11  Analysis of variance table for dry weight of *V. sativa* plants when inoculated with 3841, RU361, RU361/pRU3111

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>VR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>0.031418</td>
<td>7</td>
<td>0.004488</td>
<td>19.36</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Within Groups</td>
<td>0.009039</td>
<td>39</td>
<td>0.000231</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.040457</td>
<td>46</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

L.S.D. = 0.01865

Appendix 4.12  Analysis of variance table for dry weight of *V. sativa* plants when inoculated with 3841 and RU307

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>VR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>0.050328</td>
<td>5</td>
<td>0.010066</td>
<td>17.96</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Within Groups</td>
<td>0.021863</td>
<td>39</td>
<td>0.000561</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.072192</td>
<td>44</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

L.S.D. = 0.01321
Appendix 4.13  Analysis of variance table for colony numbers of 3841, RU360, RU361 and RU307 recovered from the rhizosphere

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>VR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>0.8493</td>
<td>3</td>
<td>0.2831</td>
<td>2.81</td>
<td>0.075</td>
</tr>
<tr>
<td>Within Groups</td>
<td>1.5105</td>
<td>15</td>
<td>0.1007</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2.2943</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

L.S.D. = 0.4278

Appendix 4.14  Analysis of variance table for colony numbers of RU360 and 3841 recovered from the rhizosphere when co-inoculated.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>VR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>2.4151</td>
<td>3</td>
<td>.805</td>
<td>3.59</td>
<td>.041</td>
</tr>
<tr>
<td>Within Groups</td>
<td>3.1418</td>
<td>14</td>
<td>.2244</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>5.3866</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

L.S.D. = 0.682
**Appendix 4.15 Analysis of variance table for colony numbers of RU361 and 3841 recovered from the rhizosphere when co-inoculated.**

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>VR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>4.95223</td>
<td>3</td>
<td>1.65074</td>
<td>21.42</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Within Groups</td>
<td>1.23285</td>
<td>16</td>
<td>.07705</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>6.18508</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

L.S.D. = 0.3722

**Appendix 4.16 Analysis of variance table for colony numbers of RU307 and 3841 recovered from the rhizosphere when co-inoculated.**

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>VR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>2.04601</td>
<td>3</td>
<td>0.682</td>
<td>9.99</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Within Groups</td>
<td>1.09195</td>
<td>16</td>
<td>0.06825</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3.13796</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

L.S.D. 0.3503
Appendix 4.17 Analysis of variance table for alkaline phosphatase assay of RU360, RU361, RU307 and 3841 containing a *nodC-phoA* fusion.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>VR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>1111087</td>
<td>8</td>
<td>1388859</td>
<td>4.05</td>
<td>0.018</td>
</tr>
<tr>
<td>Within Groups</td>
<td>3767624</td>
<td>11</td>
<td>342511</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1487849</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

L.S.D. = 1175.9
Chapter 3 graphs added by PSP as an appendix
Strain uptake of glucose nmol/mg/min
strain uptake of myo-inositol nmol/mg/min

- 3841 P
- 3841 PI
- 3841 I
- 360 P
- 360 PI
- 361 P
- 361 PI
- 307 P
- 307 PI
- 307 I
myo-inositol  mannitol  sorbitol  glucose  fructose  pyruvate  glutamic acid gamma-hydrazide  galactose

uptake of myo-inositol nmol/mg/min

carbon compound
uninoculated 3841 RU360 RU361 RU307
strain

average weight per plant (g)
micro moles ethylene h\(^{-1}\) per plant

<table>
<thead>
<tr>
<th>Strain</th>
<th>3841</th>
<th>RU361</th>
<th>RU307</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.5</td>
<td>2.5</td>
<td>3.0</td>
</tr>
</tbody>
</table>
The graph shows the average plant weight (g) for different strains:

- **3841**
- **RU361**
- **RU307**

RU307 has the highest average plant weight, followed by RU361, and then 3841.
average total nodule number per plant

3841  
RU361  
RU307

strain
average total nodule mass per plant

3841    RU361    RU307
strain

0.01 0.02 0.03 0.04 0.05 0.06 0.07 0.08 0.09 0.1
average no. nodules per plant

inoculum

3841
RU360
RU360/pRU3078
RU360/pRU3079
3841 + RU360
3841 + RU360/pRU3078
3841 + RU360/pRU3079
3841 + RU360 10x
3841 + RU360 100x
average number of nodules per plant

inoculum

3841
RU307
3841 + RU307
3841 + RU307 10x
3841 + RU307 100x
average plant dry weight (g)

inoculum

uninoculated  3841  RU360  RU360/pRU3078  RU360/pRU3079  3841 + RU360  3841 + RU360/pRU3078  3841 + RU360/pRU3079  3841 + RU360 10x  3841 + RU360 100x
average plant dry weight (g)

inoculum

uninoculated, 3841, RU361, RU361/pRU3111, 3841 + RU361, 3841 + RU361/pRU3111, 3841 + RU361 10x, 3841 + RU361 100x
The graph shows the average plant dry weight (g) under different inoculum conditions. The conditions include uninoculated, 3841, RU307, 3841 + RU307, 3841 + RU307/10x, and 3841 + RU307/100x. The average plant dry weight ranges from 0.02 to 0.16 g.
Graphs chapter 4 added by PSP as an appendix
Day of harvest

Log10 CFU per rhizosphere

RU307 - 3841
RU307 of RU307-3841
3841 of RU307-3841
Graphs Chapter 5 added by PSP as an appendix
specific fluorescence (V) vs. carbon source
specific fluorescence (V)