Characterisation of the *fixABCX* operon in symbiotic nitrogen fixation

Isabel Ursula Colette Webb

This thesis is submitted in fulfilment of the requirements of the degree of Doctor of Philosophy at the University of East Anglia

Department of Molecular Microbiology

John Innes Centre

Norwich, UK

September 2016

DECLARATION

I declare that the work contained in this thesis is an account of my own research and has not been submitted for a degree at any other university. Reference to material from other sources has been fully acknowledged where appropriate.

Signed

Isabel Webb

ABSTRACT

The *fixABCX* genes are essential for nitrogen fixation in *R. leguminosarum* bv. *viciae* 3841. Comparison of FixABCX to homologous proteins across the Kingdoms of Life suggests a role in electron transport to nitrogenase, which requires eight electrons per molecule of dinitrogen fixed. Mutation of this operon leads to bacteroids unable to fix nitrogen in symbiosis with *P. sativum* (pea). Electron microscopy revealed a drastically altered bacteroid morphology in *fixAB* mutants, revealing insights into the developmental response of both plant and bacteria to a lack of nitrogen fixation. Observations from electron microscopy were coupled to data obtained using single-cell Raman microscopy in order to understand metabolite production in nitrogen fixing and non-fixing bacteroids.

The promoter controlling *fixA* has been characterised to a minimal region consisting of binding sites for NifA, the general transcriptional activator of nitrogen fixation, and RpoN (σ^{54}), its cognate sigma factor. Mutation analysis reveals that *fixABCX* is part of a larger operon including the *nifA* gene. Promoter analysis of the downstream genes has identified a set of basal promoters found within the *fixCX* region, which control expression of the *nifA* gene. Control of nitrogen fixation occurs at the post-transcriptional level, whereby NifA is able to activate nitrogen fixation genes, including the *fixABCXnifA* operon, autoregulating its own expression under nitrogen-fixing conditions.

Pull-down assays have revealed protein-protein interactions between FixAB and nitrogenase, as well as an interaction with both pyruvate dehydrogenase and 2-oxoglutarate dehydrogenase. FixAB may interact with these dehydrogenases and via electron bifurcation couple the exergonic reduction of the quinone pool to the endergonic reduction of ferredoxin and subsequently nitrogenase. Furthermore, FixAB, nitrogenase and pyruvate dehydrogenase and 2-oxoglutarate dehydrogenase may form a supra-molecular complex within nitrogen fixing bacteroids.

ACKNOWLEDGEMENTS

First and foremost I would like to thank Professor Philip Poole for all the support and instruction he has provided me over the past four years. This project would not have been possible without his extensive knowledge of rhizobia, particularly their biochemistry and genetics. I would also like to thank my co-Supervisor Dr Nicholas Watmough for his invaluable guidance surrounding electron transfer biochemistry. Thanks also to Professor Mark Banfield, for his advice on protein expression, as well as his patience in submitting reports and forms from afar.

It has been a pleasure to work with every member of the Poole lab family. In particular I owe a huge debt to Karunakaran Ramakrishnan (KK), for all his teaching and patience as I found my way around the lab. I would also like to thank Rob Green, for friendship and advice throughout my time in Norwich (and since). Optimisation of protein protocols would not have been the same without the moral support and solidarity provided by Carmen Sanchez-Cañizares. Finally I would like to thank Barney Geddes for discussions about cloning strategies and Alison East, for her support and advice. I am also grateful to the support staff at the John Innes Centre and Oxford Plant Sciences, in particular Helen Prescott.

Outside of the Poole lab I would like to thank Elaine Barclay (JIC) and Euan James (JHI) for their skills in electron microscopy. I would also like to thank the Oxford Advanced Proteomics Facility for training and advice. Thanks go to Jiabao Xu and Wei Huang for collaboration using Raman microscopy.

I would like to thank the BBSRC for funding this project, and the John Innes Centre and related institutions for use of their facilities. I would like to extend my gratitude to the University of Oxford, including the Department of Plant Sciences and Somerville College, for treating me as one of their own for two years. I would also like to thank my internship host, the Centre for Science and Policy, Cambridge, particularly Rob Doubleday, Jackie Ouchikh and my fellow interns.

Finally I would like to thank my family and friends who have supported me for the past four years and been so understanding when experiments took priority.

ABBREVIATIONS

ABC	ATP-binding cassette
ACN	Acetonitrile
ADP	Adenosine diphosphate
AIM	Auto induction media
AMP	Adenosine monophosphate
Amp	Ampicillin
Amp^r	Ampicillin resistant
ATP	Adenosine triphosphate
BNF	Biological nitrogen fixation
BLAST	Basic local alignment search tool
BSA	Bovine serum albumin
BTH	Bacterial two-hybrid assay
cAMP	Cyclic adenosine monophosphate
CAP	Catabolite activator protein
CCaMK	Calcium calmodulin-dependent kinase
CDS	Coding sequence
Cfu	Colony-forming units
Chl	Chloramphenicol
Chl^r	Chloramphenicol resistant
C-IAA	ChloroIodoacetamide
Co-IP	Co-immunoprecipitation
Cps	Counts per second
CV	Column volume
cv	Cultivar
dpi	Days post inoculation
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTT	1,4-Dithiothreitol
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EM	Electron micrograph
EPS	Exopolysaccharides
ETF	Electron transfer flavoprotein
ETF-QO	Electron transfer flavoprotein-quinone oxidase

FAD	Flavin adenine dinucleotide
Fd	Ferredoxin
Fix^+	Nitrogen fixing phenotype
Fix^-	Non-nitrogen fixing phenotype
Fix^{red}	Reduced nitrogen fixation phenotype
GC	Gas chromatography
Gm	Gentamycin
Gm^r	Gentamycin resistant
GOGAT	Glutamate oxoglutarate amidotransferase
GS	Glutamine synthetase
HABA	4'-hydroxyazobenzene-2-carboxylic acid
Hr	Hour
HRP	Horseradish peroxidase
IGR	Intergenic region
IRLC	Inverted repeat lacking clade
IT	Infection thread
LB	Luria Bertani broth
K_m	Michaelis constant
Km	Kanamycin
Km^r	Kanamycin resistant
MAMP	Microbe-associated molecular pattern
MCS	Multiple cloning site
ME	Malic enzyme
MOPS	3-(N-morpholino)propanesulfonic acid
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
m/z	Mass-to-charge ratio
NAD^+	Nicotinamide adenine dinucleotide (oxidised form)
NADH	Nicotinamide adenine dinucleotide (reduced form)
$NADP^+$	Nicotinamide adenine dinucleotide phosphate (oxidised form)
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced form)
NCR	Nodule-specific cysteine rich peptides
Nm	Neomycin
Nm^r	Neomycin resistant
OD	Optical density
ODH	2-oxyglutarate dehydrogenase
ONPG	Ortho-Nitrophenyl-β-galactoside

ORF	Open reading frame
PAMP	Pathogen-associated molecular pattern
PBS	Phosphate buffered saline
PBS-T	Phosphate buffered saline plus Tween-20
PCA	Principal component analysis
PCR	Polymerase chain reaction
PDH	Pyruvate dehydrogenase
PEG	Poly(ethylene glycol)
PEP	Phosphoenolpyruvate
PEPCK	Phosphoenolpyruvate carboxykinase
PHA	Polyhydroxyalkanoate
PHB	Polyhydroxybutyrate
P_i	Inorganic phosphate
PolyP	Polyphosphate
r	Resistant
Rlu	Relative light units
Rlv3841	*Rhizobium leguminosarum* bv. *viciae* 3841
RNA	Ribonucleic acid
ROS	Reactive oxygen species
SCRM	Single-cell Raman microscopy
sfGFP	Superfolder green fluorescent protein
Spc	Spectinomycin
Spc^r	Spectinomycin resistant
Str	Streptomycin
Str^r	Streptomycin resistant
SUMO	Small ubiquitin-like modifier
TBS	Tris-buffered saline
TBS-T	Tris-buffered saline plus Tween-20
TCA	Tricarboxylic acid
Tc	Tetracycline
Tc^r	Tetracycline resistant
TCEP	Tris(2-carboxyethyl)phosphine
TEAB	Tetraethylammonium bromide
TEM	Transmission electron microscopy
TEMED	Tetramethylethylenediamine
TFA	Trifluoroacetic acid
T_m	Melting temperature

TY	Tryptone yeast
UAS	Upstream activator sequence
UMS	Universal Minimal Salts
UQ	Ubiquinone
UV	Ultraviolet
v/v	Volume of solute/volume of solution
WT	Wild-type
w/v	Mass of solute/volume of solution
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

CONTENTS

CHAPTER 1

Introduction

1.1 The legume-rhizobia symbiosis

Several nutrients are essential for life, including sulphur, phosphorus and nitrogen. All organisms require nitrogen to make key organic molecules including nucleic acids (~5% of all plant nitrogen), amino acids (~85% of all plant nitrogen) (Maathuis 2009), coenzymes, and many secondary products, with nitrogen making up 1.5% of plant dry mass (Taiz and Zeiger 2010). Although dinitrogen makes up around 80% of the atmosphere, it is in an inert form that is not biologically accessible due to the high activation energy of the triple bond between the two nitrogen atoms. Plants are able to take up nitrogen from the soil typically as either nitrate (NO_3^-) or ammonia (NH_4^+) depending on the soil conditions, carrying nitrogen into the base of the food chain. In agricultural systems nitrogen tends to be applied as part of chemical fertilisers. Nitrate production for chemical fertiliser occurs via the Haber-Bosch process, which is estimated to use 1-2% of the world's energy supply due to its high temperature and pressure demand (Smil 2011, Scheibel and Schneider 2012). This process is the highest input cost in crop production (Xu, Fan et al. 2012). Global demand for chemical fertilisers is continually growing with the increasing world population. Application of chemical fertilisers is accompanied by an additional environmental cost due to leaching of applied nitrate into water systems, affecting river and marine ecosystems (Schlesinger 2009), as well as release of nitrous oxide (N_2O), a greenhouse gas, into the atmosphere (Shcherbak, Millar et al. 2014).

Much of the nitrogen entering the food chain is converted from dinitrogen into a biologically accessible form by nitrogen-fixing organisms known as diazotrophs. Some diazotrophs are able to carry out biological nitrogen fixation (BNF) as free-living organisms, whilst others are only able to fix nitrogen in symbiosis with plants. The most common nitrogen fixing symbiosis is between plants of the Fabaceae family (legumes) and a group of diazotrophic bacteria called rhizobia. Within this system the legume develops root nodules, an organ which is able to host the invading bacteria and provide an optimal environment for BNF. Once inside the plant the bacteria will differentiate, switching off metabolic processes involved in reproduction and motility and turning on processes involved in BNF. Within nodules the bacteria provides fixed nitrogen in return for plant organic acids, a source of energy (White, Prell et al. 2007, Terpolilli, Hood et al. 2012, Udvardi and Poole 2013). Evolution of nitrogen fixing symbioses is thought to have occurred ~8 times, explaining the diversity in infection, nodule form and host-specificity (Werner, Cornwell et al. 2014). It is estimated that BNF from legumes input 21 Tg $year^{-1}$ fixed nitrogen to the nitrogen cycle (Herridge, Peoples et al. 2008). This relationship makes legumes such as bean, pea and soybean important crops globally. The *Rhizobium*-legume symbiosis is an extremely

important process within the global nitrogen cycle and understanding the processes involved will help to reduce our dependence on chemical nitrogen fertilisers.

1.2 Initialising the symbiosis

1.2.1 Plant host recognition: host-symbiont signalling

The interaction between rhizobia and their legume host is tightly controlled, requiring signal generation and perception by both partners across all stages of symbiosis, from initial recognition to nodule senescence (Lang and Long 2015). The nature of the signalling molecules contributes to the tight host-specificity that can occur in the *Rhizobium*-legume system. The first step in this process is the release of flavonoids by the plant roots into the surrounding soil. Flavonoids are cyclic plant secondary metabolites, with different chemical modifications that lead to a spectrum of compounds contributing to host-specificity. Flavonoids act both as chemoattractants and as inducers for production of rhizobial signalling molecules, Nod factors. The recognition of flavonoids occurs via NodD, a LysR transcriptional regulator that works in a complex with flavonoids to act as a general transcriptional activator for other *nod* genes by binding a *nod*-box upstream of Nod factor synthesis genes (Spaink 2000, Novak, Lisa et al. 2004).

Nod factors are released by rhizobia and recognised by the plant to initiate rhizobial infection processes. Nod factors are specific to host legumes, working in a 'lock-and-key' mechanism to ensure tight control over host-specificity. *nodABC* are core *nod* genes, encoding the Nod factor lipo-chito-oligosaccharide backbone (Roche, Maillet et al. 1996). Different rhizobial species have variation in the *nodABC* genes which can allow for structural alterations, such as different *nodC* genes which lead to different length backbones (Kamst, Pilling et al. 1997). Other *nod* genes encode enzymes that can modify the core Nod factor to create the diversity needed for the lock-and-key recognition, as well as encoding transporters to allow Nod factor efflux (Cardenas, Dominguez et al. 1996, Fernandez-Lopez, D'Haeze et al. 1996).

Nod factors are recognised by plant receptor-like kinases with extracellular LysM domains that can bind chito-oligosaccharides (Limpens, Franken et al. 2003, Madsen, Madsen et al. 2003). The recognition of Nod factors initiates calcium oscillations from within the nucleus which acts as a second-messenger to carry the recognition signal and begin the nodulation process (Charpentier and Oldroyd 2013). The resultant calcium spikes are perceived by a calcium-calmodulin-dependent protein kinase (CCaMK) that phosphorylates proteins at the start of a nodulation signalling pathway (Murray 2011). This pathway leads to expression of

several transcription factors which promote nodule organogenesis (Oldroyd and Downie 2004, Murray 2011). Many components of this pathway are part of a common symbiosis signalling pathway shared with the arbuscular mycorrhizal symbiosis (Bonfante and Genre 2010). Calcium-mediated signalling is not unique to legume plants, and calcium spikes can be seen in nodulating non-legumes such as in actinorrhizal interactions with *Frankia spp*. (Granqvist, Sun et al. 2015). Untangling this common pathway to understand how plants differentiate between formation of nodules or establishment of a mycorrhizal symbiosis is an active research area.

1.2.2 Plant host recognition: plant immune responses

In addition to Nod factor responses, rhizobia also elicit some plant immunity mechanisms. This includes transient induction of plant-defence related genes, as seen in soybean after inoculation with *Bradyrhizobium japonicum* (Libault, Farmer et al. 2010), *Medicago truncatula* after inoculation with *Sinorhizobium meliloti* (Lohar, Sharopova et al. 2006) and *Lotus japonicum* after infection with *Mesorhizobium loti* (Kouchi, Shimomura et al. 2004). Production of reactive oxygen species (ROS) has also been proposed to have a role in rhizobial infection, and H_2O_2 has been suggested to play a role in nodule function in both *Medicago sativa* and *M. truncatula* (Jamet, Mandon et al. 2007, Marino, Andrio et al. 2011, Andrio, Marino et al. 2013).

Plant immunity responses are elicited by microbe-associated molecular patterns (MAMPs) or by microbial-secreted proteins known as effectors (Zipfel 2014); no rhizobial MAMPs have been detected with effects on legumes. MAMP-triggered immunity has been shown to affect nodulation, both by application of plant-defence phytohormones and application of the Flg22 MAMP (Stacey, McAlvin et al. 2006, Lopez-Gomez, Sandal et al. 2012). It is thought that succinoglycan, an expopolysaccharide, may suppress MAMP-triggered immunity during early stages of infection by *S. meliloti* (Jones and Walker 2008) via calcium chelation (Aslam, Newman et al. 2008). In addition to MAMPs, there is now a growing body of data suggesting a role for effector-triggered immunity in the symbiosis, with impact on rhizobial host range (Gourion, Berrabah et al. 2015). The tight interplay between defence responses and symbiosis has been further demonstrated in recent studies showing that presence of plant pathogens is inhibitory to nodule development, as reviewed by de Souza et al. (2016).

1.2.3. Infection thread formation

Rhizobia must travel through the root epidermis and cortex to reach the developing nodule. Infection most commonly occurs via bacterial attachment to root hairs (Lhuissier, De Ruijter et al. 2001). Rhizobial surface attachment to legume root hairs is enhanced by both plant lectins and bacterial rhicadhesin (Smit, Swart et al. 1992). Root hair response to attached rhizobia requires Nod factor perception, but does not involve the downstream calcium-signalling pathway (Miwa, Sun et al. 2006), demonstrating a dual response to Nod factors. On recognition of Nod factors, root hair growth is interrupted and apical growth arrests (Lhuissier, De Ruijter et al. 2001). This is followed by cell wall deformation and a change in the direction of root tip growth (Vanbatenburg, Jonker et al. 1986) causing the root hair to curl back on itself, trapping the bacteria in an 'infection pocket' (Geurts, Fedorova et al. 2005). The direction of root hair curling is dependent on the location of Nod factor detection, and spot-inoculation of Nod factor onto root hairs causes the hair to curl where inoculum is applied (Esseling, Lhuissier et al. 2003).

Once inside the curled root hair trapped bacteria continue to divide, forming infection foci. From these foci an infection thread (IT) forms, an intercellular invagination of the root hair cell down which the bacteria can travel. The IT is derived from plant cell wall components, requiring cell wall degradation, as well as provision of membrane vesicles to the growing IT tip (Gage 2004, Monahan-Giovanelli, Pinedo et al. 2006). The cytoskeleton plays a key role in developing the IT and directing its inward growth, which is predicted by dense cytoplasmic regions known as 'pre-infection threads' (Brewin 2004). The bacteria divide at the growing IT tip (Fournier, Timmers et al. 2008) providing a forward force for growth. The bacterial carbon storage compound poly-β-hydroxybutyrate (PHB) accumulates in bacterial cells in the IT, suggesting that ample carbon is available for growth in ITs (Lodwig, Hosie et al. 2003, Lodwig, Leonard et al. 2005). IT formation tends to select out one bacterium, resulting in a clonal infection (Gage 2002). Mixed infections can occur, although the frequency of such events in nature is unknown (Denison 2000), and frequencies calculated under laboratory conditions vary wildly (Johnston and Beringer 1975, Gage 2002). The infection thread development continues below the root epidermis through the cortical cells and into the developing nodule (Figure 1.1).

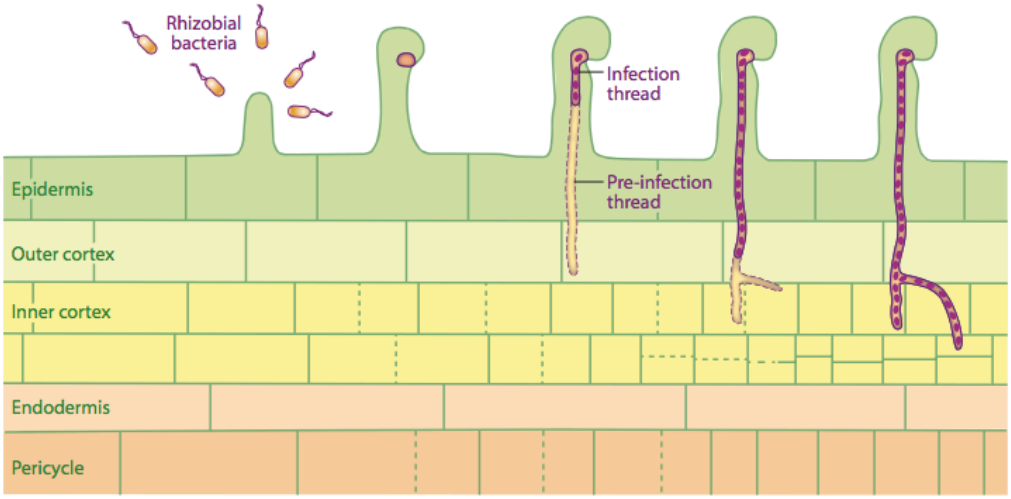


Figure 1.1 Nodulation and infection thread formation through the plant epidermis and cortex. The route of the infection thread is predicted by dense cytoplasmic subdomains with aligned cytoskeleton, known as pre-infection threads. Reproduced from Oldroyd, Murray et al. (2011).

1.2.4 Nodule organogenesis

Development of nodules originates in root cortical cells as a result of Nod factor perception (Brewin 2004). Plant hormones cytokinin and auxin (involved in lateral root development) are both thought to play a role in this signalling, initiated as a result of Nod factor recognition (Frugier, Kosuta et al. 2008, Oldroyd, Murray et al. 2011, Breakspear, Liu et al. 2014). Activation of organogenesis in cortical cells results from a signal at the initial contact site at the root surface. Regulators of the cell cycle play an important role in establishment of a nodule primordium as the cell cycle is reactivated (Cebolla, Vinardell et al. 1999). The cells first undergo anticlinal divisions, and then divide periclinally as the primordium becomes established (Scheres, Vanengelen et al. 1990). The position of these cell divisions and so the direction of the primordium is determined by a gradient of the plant hormone ethylene (Heidstra, Yang et al. 1997). Plant cells undergo several rounds of endoreduplication, resulting in large cells hosting thousands of nitrogen-fixing bacteroids (Cebolla, Vinardell et al. 1999, Kondorosi, Roudier et al. 2000, Suzaki, Ito et al. 2014).

Legume nodules may take one of two forms, determinate or indeterminate, determined by the host plant rather than the nodulating rhizobial strain. Indeterminate nodules are elongated with a persistent meristem, and are found on temperate legumes such as *Pisum sativum* (pea) and *Trifolium spp.* (clover). Plants producing indeterminate nodules are found within the

invert repeat-lacking clade (IRLC). This meristematic form leads to indeterminate nodules having several zones, ranging from the young meristem to actively nitrogen-fixing tissue to senescent tissue near the base of the nodule. Determinate nodules are spherical, lacking this meristem, and found on tropical legumes such as *Glycine max* (soybean) and *L. japonicus*. Indeterminate nodules contain terminally differentiated bacteroids that have lost their ability to reproduce, whilst bacteroids of determinate nodules are able to revert to their free-living state once released from nodules.

In determinate nodules the nodule meristem ceases functioning once the nodule primordium is formed. In indeterminate nodules the meristem persists, and the nodule has several developmental zones. The active meristem forms zone I, followed by the invasion zone (zone II), an interzone II-III, then nitrogen-fixing zone (zone III) and the senescence zone (zone IV). Zones I-II and the interzone maintain a constant size, whilst zones III and IV grow as the nodule ages (Vasse, de Billy et al. 1990, Kereszt, Mergaert et al. 2011). An additional zone (zone V), proximal to zone IV, has been described in alfalfa, which contains rhizobia living off decaying nodule tissue (Timmers, Soupene et al. 2000). The differing morphologies of determinate and indeterminate nodules can be seen in Figure 1.2.

Bacteria from both determinate and indeterminate nodules are able to recolonise the rhizosphere upon nodule senescence. Bacteroids from determinate nodules do this by de-differentiating and regaining their free-living state (Muller, Wiemken et al. 2001). Bacteroids in indeterminate nodules are terminally differentiated, but some undifferentiated rhizobia are present within the nodules that are able to colonise the rhizosphere once released (Paau, Bloch et al. 1980).

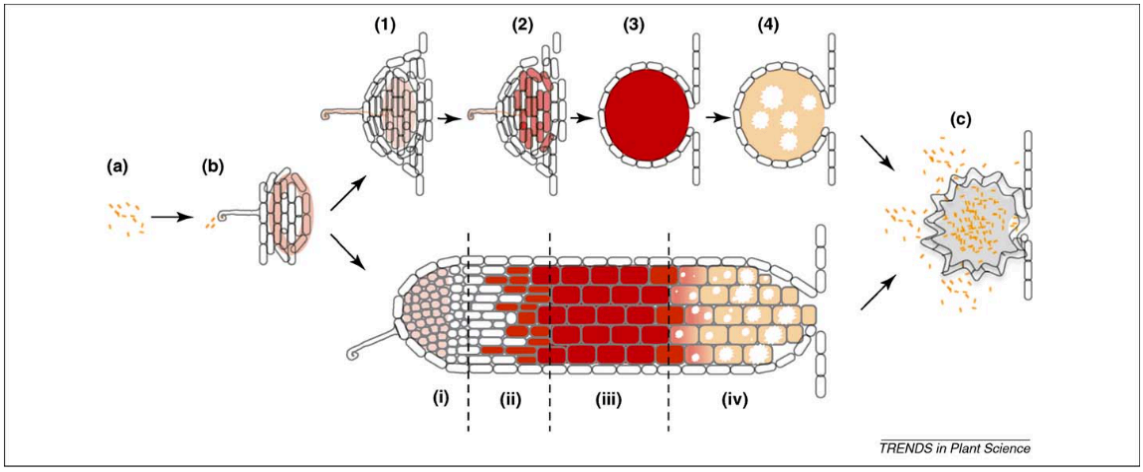


Figure 1.2 Development of determinate and indeterminate nodules. (a) Rhizobia exist as saprophytes in the rhizosphere. (b) Rhizobia respond to flavonoids, releasing Nod factors and moving towards the plant root hair. Steps (1)-(4) represent a determinate nodule, such as those formed on *Glycine max* (soybean). Steps (i)-(iv) represent an indeterminate nodule such as those formed on *Pisum sativum* (pea). Nod factors induce root hair curling, allowing rhizobia to travel down an infection thread into the developing nodule primordial (1 and i). Rhizobia are then released into the plant cell cytosol (2 and ii). Rhizobia multiply, and the plant cell undergoes several rounds of endoreduplication to result into a symbiosome containing nitrogen-fixing bacteroids (3 and iii). Mature nodules eventually senesce (4 and iv). Decay of determinate nodules releases bacteroids which de-differentiate. Decay of indeterminate nodules releases undifferentiated rhizobia which can then recolonise the rhizosphere. Reproduced from Schumpp and Deakin (2010).

Nodule organogenesis is accompanied by huge changes in plant gene expression involving hundreds of genes. Plant genes that are highly upregulated or exclusively expressed in nodule development are termed 'nodulins' (Benedito, Torres-Jerez et al. 2008). Early nodulins are induced during nodule development, and late nodulins are induced as nitrogen fixation begins. Recent advances in expression profiling tools have allowed large scale identification of nodulins (Kuster, Vieweg et al. 2007). Over 50% of nodule-specific genes identified in *M. truncatula* nodules encode secretory proteins, supporting the theory that nodule-specific proteins are targeted to the symbiosome (Maunoury, Redondo-Nieto et al. 2010).

Once the nodule primordia has been established, the infection thread grows towards the dividing plant cells. Bacteria are released from the tip of the infection thread into an unwalled infection droplet in the host cell cytoplasm (Brewin 2004). The trigger for this

release of bacteria is unknown. The bacteria become endocytosed by a plant membrane, called the peribacteroid membrane (or symbiotic membrane), which forms a structure known as a 'symbiosome' (Roth, Jeon et al. 1988, Parniske 2000). Within the symbiosome the bacteria continue to divide before differentiating into bacteroids. In the case of indeterminate nodules, a single bacterium enters a symbiosome. In determinate nodules 8-12 bacteria may share a symbiosome (Brewin 2004).

1.2.5 NCR peptides

Bacteroids in determinate nodules resemble free-living bacteria in size, reproductive ability and genomic content. Conversely, bacteroids in indeterminate nodules undergo irreversible morphological changes. These changes include cell elongation, genome amplification and the loss of reproductive ability (Kondorosi, Roudier et al. 2000). Since this terminal differentiation only occurs within legumes that can produce indeterminate nodules, it was concluded that this fate is controlled by plant factors (Mergaert, Uchiumi et al. 2006). A candidate for host factors produced exclusively by the IRLC clade was nodule-specific cysteine-rich (NCR) peptides. IRLC plants produce large numbers of these peptides; a recent estimate suggests that *M. truncatula* encodes over 700 different NCR peptides (Mergaert, Nikovics et al. 2003, Graham, Silverstein et al. 2004, Alunni, Kevei et al. 2007, Silverstein, Moskal et al. 2007). Microarray analysis of *M. truncatula* shows that expression of the large range of NCRs is tightly controlled temporally and spatially in successive waves across the zones of nodules (Guefrachi, Nagymihaly et al. 2014). Van de Velde et al. (2010) demonstrated targeting of these plant peptides to the bacteroid membrane and cytosol. Obstruction of NCR transport or expression correlates with absence of terminal bacteroid differentiation (Maunoury, Redondo-Nieto et al. 2010, Van de Velde, Zehirov et al. 2010). A gain-of-function effect can also be seen by application of NCR peptides to *in vitro* cultures, or transforming non-IRLC clade legume *L. japonicus* with NCR genes. These gain-of-function experiments resulted in bacteria with features of terminal differentiation (Van de Velde, Zehirov et al. 2010). NCR peptides resemble antimicrobial peptides, and so show an adoption of immune system effectors to control endosymbiotic bacteria. Recently NCR peptides have been identified in plants not belonging to the inverted repeat-lacking clade; *Aeschynomene spp* from the Dalbergioid lineage produce elongated nodules containing terminally differentiated bacteria. These *Aeschynomene* produce a suite of NCR peptides, providing evidence for convergent evolution of bacteroid differentiation (Czernic, Gully et al. 2015).

The key processes coordinated by NCR peptides include endoreduplication and changes in cell size and shape. Change of morphology from rod-shaped to Y-shaped was first identified in 1888 by Beijerinck, with the Y-shaped cells being four to seven times larger than the free-living cells in *S. meliloti* (Oke and Long 1999, Mergaert, Uchiumi et al. 2006). Cells uncouple DNA replication and cell cycle in order to form large, polyploid cells with 8-20 copies of the genome (Mergaert, Uchiumi et al. 2006, Nakabachi, Koshikawa et al. 2010, Maroti and Kondorosi 2014). This may occur by downregulation of a master regulator of cell cycle processes, such as CtrA in *S. meliloti*, depletion of which leads to bacteroid-like changes (Pini, De Nisco et al. 2015).

1.2.6 Bacterial response to NCR peptides: BacA

In addition to the plant-derived NCR peptides, a key microbial protein has been identified that is involved in bacteroid development. The protein BacA is part of the ATP-binding cassette (ABC) superfamily, with homology to very long chain fatty acid (VLCFA) exporters. BacA plays a role in peptide import and addition of VLFCAs to lipopolysaccharide (Ardissone, Kobayashi et al. 2011). BacA, a cytoplasmic membrane protein, is also required for bacteroid differentiation in IRLC-clade legumes. Mutants of either *S. meliloti* or *Rhizobium leguminosarum* bv. *viciae* lacking BacA are unable to differentiate into bacteroids on alfalfa and pea respectively, instead bacteria lyse and die following invasion (Glazebrook, Ichige et al. 1993, Karunakaran, Haag et al. 2010). In addition to the altered bacteroid phenotype, *bacA* mutants also have altered responses to antimicrobial peptides, including increased resistance to bleomycin and Bac7 (Ichige and Walker 1997, Karunakaran, Haag et al. 2010). Haag et al. (2011) demonstrated a critical role for BacA in protecting *S. meliloti* from the antimicrobial action of NCR peptides by reducing membrane permeabilisation and bacterial death.

1.3 Bacteroid development

1.3.1 Transcriptional and metabolic changes in bacteroids

When bacteria differentiate into bacteroids they undergo many metabolic changes, focusing their metabolism towards the purpose of nitrogen fixation and behaving more like plant organelles than individual organisms. Recent advances in genetic techniques have allowed transcriptomic profiling to understand exactly what changes occur during bacteroid development. These changes include cessation of processes involved in growth, chemotaxis,

motility, ribosomal protein synthesis and nucleic acid synthesis, whilst expressing the genes required for a functional symbiosis (Ampe, Kiss et al. 2003, Becker, Berges et al. 2004, Colebatch, Desbrosses et al. 2004, Uchiumi, Ohwada et al. 2004, Chang, Franck et al. 2007, Karunakaran, Ramachandran et al. 2009, Tsukada, Aono et al. 2009). Genes dramatically upregulated during symbiosis include the *nif* and *fix* genes, such as *nifHDK,* encoding the nitrogenase subunits and *fixABCX,* encoding electron transport flavoproteins thought to transfer electrons to nitrogenase. In *R. leguminosarum* bv. *viciae* 3841, a large number of bacteroid-specific genes are carried on the pRL10 plasmid, also known as the symbiotic plasmid. Of those genes over two-fold upregulated in mature bacteroids, approximately 20% are involved in central bacteroid metabolism, including nitrogen fixation, whilst 20% encode proteins involved in transport. Genes downregulated at least two-fold include ribosomal proteins and those encoding proteins involved in small molecule metabolism such as amino acid synthesis (Karunakaran, Ramachandran et al. 2009).

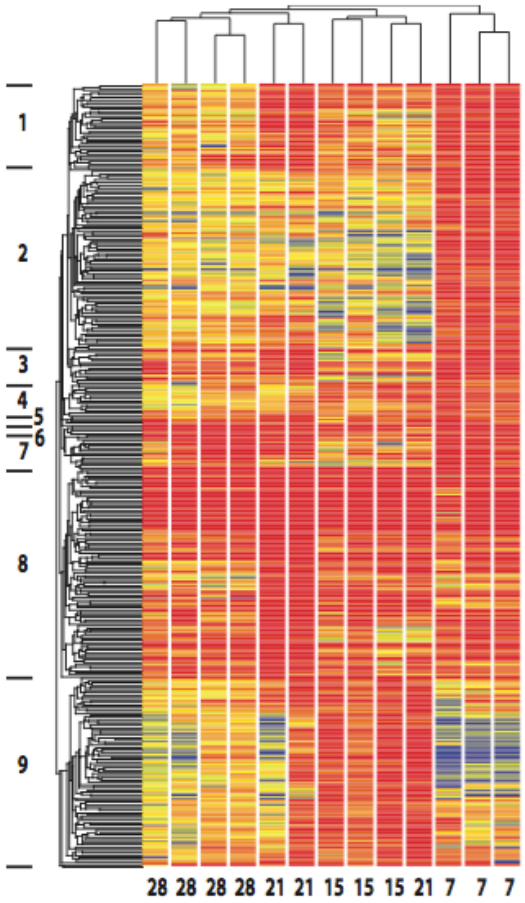


Figure 1.3 Hierarchical gene tree showing changes in gene expression over the course of bacteroid development in peas from 7 to 28 days post inoculation with *R. leguminosarum* bv *viciae* 3841. Red indicates highly expressed genes, yellow intermediate and blue low. Nine gene clusters were identified with similar regulatory patterns, indicated by the numbers to the left of the clustering. Reproduced from Karunakaran, Ramachandran et al. (2009).

1.3.2 Nodule oxygen environment

A key requirement for bacteroid function is a controlled oxygen environment due to oxygen-inactivation of nitrogenase (Wong and Burris 1972, Hartmann and Burris 1987). Oxygen tension above 57 nM will inhibit nitrogenase activity in soybean nodules (Kuzma, Hunt et al. 1993). Balance of free oxygen levels is essential in order to maintain nitrogenase activity whilst meeting the oxygen requirement for respiration and ATP synthesis; the nitrogen fixation reaction requires sixteen ATP molecules per molecule of dinitrogen. This balance is established using a plant protein (Santana, Pihakaski-Maunsback et al. 1998), leghaemaglobin, to buffer oxygen concentrations and maintain an oxygen diffusion barrier (Appleby 1984). Leghemoglobin is the protein that gives nodules their characteristic pink

colour, in a similar manner to animal myoglobins. Leghemoglobins have a 20-fold higher affinity for oxygen than animal myoglobins, allowing them to maintain free oxygen concentrations around 7-11 nM (Downie 2005). The Km of pea leghaemaglobin is 65 nM. The association rate constant (k_{on}) of pea leghaemaglobin is 250 μM s^{-1}, and the dissociation rate constant (k_{off}) is $16s^{-1}$ (Becana and Klucas 1992). This is much higher affinity than myoglobins, as can be seen in comparison to the constants for human myoglobin; k_{on} = 19 μM s^{-1} and k_{off} = 22 s^{-1} (Migita, Matera et al. 1998). Whilst leghemoglobins are essential for nitrogen fixation in nodules, they do not play any other essential roles in plant development (Ott, van Dongen et al. 2005).

An additional mechanism that allows the bacteroid to deal with respiration in a low oxygen environment is expression of a bacteroid-specific heme-copper terminal cbb_3-oxidase, FixNOQP, which has a low operational K_m for oxygen (Mandon, Kaminski et al. 1994, Preisig, Zufferey et al. 1996, Delgado, Bedmar et al. 1998). Though initially identified in *B. japonicum* these high-affinity oxidases have since been identified in other bacteria such as *P. stutzeri*, designated as CcoNOQP (Pitcher and Watmough 2004). Whilst rhizobia encode other oxidases, only the cbb_3-type oxidase has a low enough K_m to operate in microaerobic bacteroid conditions. The cbb_3-type oxidases have been shown to translocate protons (Raitio and Wikstrom 1994); this maximises protonmotive force and ATP production, important in the energy-demanding nitrogen-fixation process. Mutants in *fixNOQP* display a Fix^- phenotype in *B. japonicum* (Kahn, David et al. 1989, Preisig, Zufferey et al. 1996). Some rhizobia, such as *R. leguminosarum* bv. *viciae* VF39 and *R. leguminosarum* bv. *viciae* 3841 encode more than one copy of *fixNOQP*. In VF39 mutation of a single copy of *fixNOQP* reduces nitrogen fixation ability, but a full Fix^- phenotype is only seen when both copies of the operon are knocked out (Schluter, Patschkowski et al. 1997).

The expression of many key bacteroid genes is inhibited under high oxygen. The mechanism of regulation at this genetic level varies depending on the rhizobial species, including whether they act as a free-living nitrogen fixer, or can only fix nitrogen in a symbiotic state. Control of key genes for symbiotic nitrogen fixation, including the nitrogenase subunits *nifHDK* and electron transfer proteins *fixABCX* comes from a general transcriptional regulator, NifA, and its cognate sigma factor RpoN. NifA expression is controlled in response to the oxygen levels in the cell. In symbiotic species, this often involves an oxygen-reponsive regulator cascade consisting of FixLJ, as seen in *S. meliloti* and *B. japonicum* (Dixon and Kahn 2004). The FixLJ cascade regulates a range of processes, including respiration, stress responses and exopolysaccharide production. FixL has a sensory PAS domain, which is found in many sensor proteins including those responsive to oxygen, light

and redox potential (Taylor and Zhulin 1999). This PAS domain undergoes a conformational, inhibitory change when bound reversibly to oxygen, and only in the absence of oxygen is autophosphorylation of the transmitter domain possible. This transmitter domain is able to carry out a phosphotransfer from ATP to response regulator FixJ in an oxygen-independent manner (Tuckerman, Gonzalez et al. 2002).

The downstream responses to FixLJ are transduced through a variety of secondary regulatory systems (Figure 1.4), such as the FixK protein, a member of the Crp-Fnr superfamily of bacterial transcriptional regulators (Korner, Sofia et al. 2003). FixK acts as a positive regulator for NifA, the general transcriptional activator of nitrogen fixation as demonstrated in *S. meliloti* and *A. caulinodans* (Batut, Daveranmingot et al. 1989, Ratet, Pawlowski et al. 1989) as well as controlling other genes such as *fixNOQP*. In *B. japonicum* a second level of regulation comes from the RegR-RegS redox-responsive system (Bauer, Kaspar et al. 1998, Sciotti, Chanfon et al. 2003).

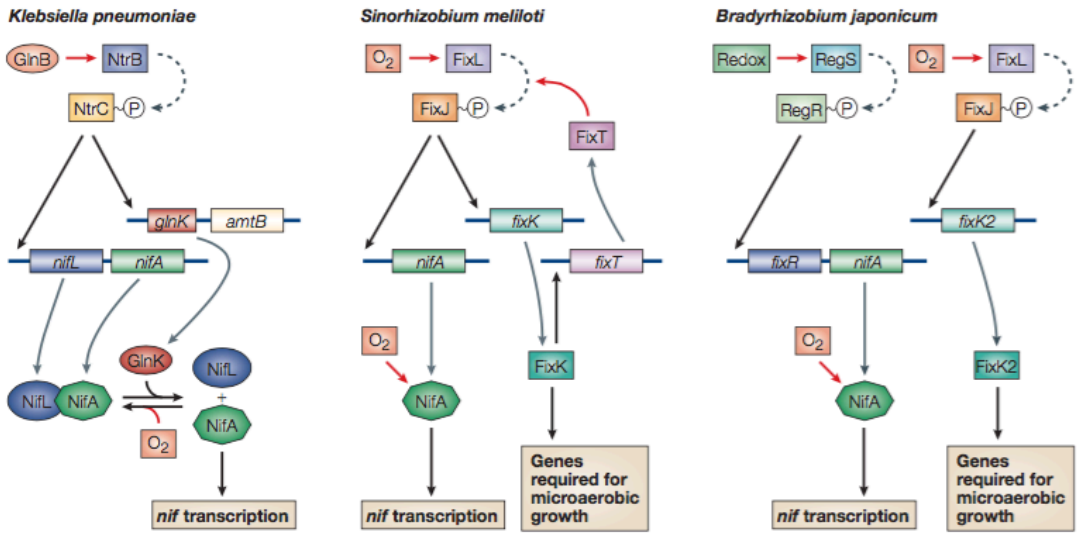


Figure 1.4 Regulatory cascades involved in oxygen response. Reproduced from Dixon and Kahn, 2004

In *R. leguminosarum* oxygen sensing and regulation differs even at the strain level. A soluble FixL can be found in *R. leguminosarum* bv. *viciae* VF39 (Patschkowski, Schluter et al. 1996) comprising a sensor domain covalently linked to a receiver domain. No separate *fixJ* gene can be identified in *R. leguminosarum* bv. *viciae.* Boesten and Priefer (2004) have proposed a model in which the FixL histidine kinase domain becomes transiently autophosphorylated before transferring the phosphoryl groups to the C-terminal receiver domain, activating its regulatory function. This phosphorylated FixL is then able to induce the promoter of a

second protein from the Crp-Fnr family, FnrN. FnrN shows structural and functional homology to *S. meliloti* FixK (Colonna-Romano, Arnold et al. 1990). In strain VF39 a mutation in either *fixL* or *fixK* generated a Fix^+ phenotype, though fixation levels were at ~60% of wild-type rates. An *fnrN* deletion showed further reduction in fixation rates (~35%). The *fixNOQP* operon is under the control of *fnrN,* and mutants in either *fixL* or *fnrN* showed reduced expression of these genes (Schluter, Patschkowski et al. 1997). A *fixK/fnrN* double mutant resulted in a Fix^- phenotype, which could be explained by two hypotheses (Terpolilli, Hood et al. 2012). The first is that FixK acts as a positive regulator of FnrN expression, but is not essential for *fnrN* expression. The second is that FixK and FnrN work in parallel, coordinating expression of additional nitrogen fixation genes.

In *R. leguminosarum* bv. *viciae* UPM791 a different suite of genes can be found. Whilst UPM791 has no FixK homolog, it encodes two copies of FnrN, one found on the chromosome (*fnrN1*) and the other on the symbiotic plasmid (*fnrN2*) (Gutierrez, Hernando et al. 1997). Mutations in either result in Fix^+ phenotypes, whilst a double *fnrN1*/*fnrN2* mutant is Fix^-. FnrN is only functional when its oxygen-labile Fe-S cluster is in the reduced state, and so is only active under microaerobic conditions (Colombo, Gutierrez et al. 2000, Jervis and Green 2007). FnrN has been suggested to autoregulate its own activity, binding to its distal anaerobox under microaerobic conditions. Once FnrN protein accumulates it will bind its proximal anaerobox, negatively regulating its own expression (Colombo, Gutierrez et al. 2000). This autoregulation model explains the lack of FixLJ homologs in UPM791.

An additional protein, FxkR, has been identified in *R. etli* CFN42 which transduces the signal from FixL to FixK in the absence of a FixJ homolog. (Zamorano-Sanchez, Reyes-Gonzalez et al. 2012). FxkR is a regulator belonging to the OmpR/PhoB family, and is thought to be post-transcriptionally regulated by FixL in order to direct *fixK* expression. This protein has a homolog in *R. leguminosarum* bv. *viciae* 3841 which may play a similar role.

In 3841 it has been reported that an *fnrN* mutant fixes nitrogen at ~10% of wild-type levels, confirming its likely role as a main regulator of oxygen response. Rlv3841 has two copies of *fixL*, on the chromosome (*fixLc*) and on the symbiotic plasmid pRL10 (*fixLp*). A triple mutant in *fnrN, fixLc* and *fixLp* is unable to fix nitrogen. There are *fixNOQP* operons found on both pRL9 and pRL10 in *R. leguminosarum* and a Fix^- phenotype is only seen if both operons are disrupted(Schluter, Patschkowski et al. 1997). Both are thought to play a role in symbiotic nitrogen fixation, and both are thought to be under the control of two pathways, one dependent on FixK and the other dependent on FnrN (Hood 2013). More research is needed to form a full model for oxygen regulation in strain 3841, though an incomplete

model has been proposed by Hood (2013), as shown in Figure 1.5. Chapter 6 of this thesis provides an updated version of this model.

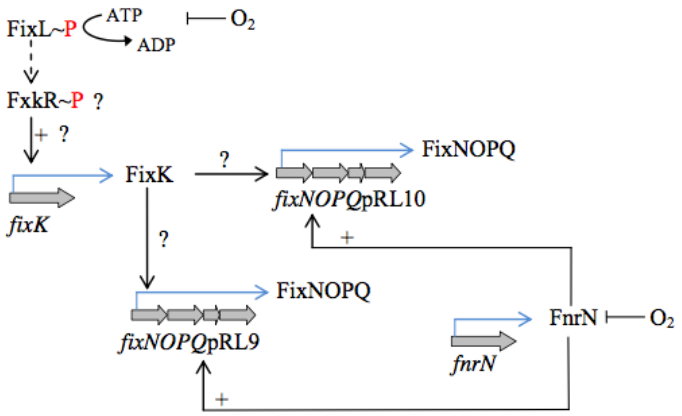


Figure 1.5 Incomplete model showing regulation of *fixNOQP* (pRL9) and *fixNOQP* (pRL10) in R. leguminosarum. Grey arrows represent genes, black arrows denote DNA binding, blue arrows indicate transcription, + specifies positive regulation and dotted arrows show phosphorylation.

Control of cbb_3-type oxidases by Fnr is not unique to rhizobia. Induction of these oxidases by the Crp-Fnr system is seen as a response to an aerobic/anaerobic switch in several species, such as *Rhodobacter capsulatus* (Swem and Bauer 2002), *Rhodobacter sphaeroides* (Mouncey and Kaplan 1998). In *E. coli* Fnr acts as a sensor-regulator, whilst in the rhizobial system there are additional levels of regulatory control over *fixNOPQ* expression.

1.3.3 Symbiotic auxotrophy

R. leguminosarum encodes two broad specificity amino acid transporters, AapJQMP and BraDEFGC, both of which are integral membrane ATP-binding proteins belonging to the ABC family (Walshaw and Poole 1996, Hosie, Allaway et al. 2002). AapJ and BraC are the solute-binding proteins of these complexes, binding amino acids in bacterial periplasm and delivering them to the membrane complex to be translocated into the cell. The *aapQM* and *braDE* genes encode integral membrane proteins, and *aapP*, *braF* and *braG* encode ATP-binding cassettes. Inoculation of peas with single mutantions in *aap* or *bra* result in healthy

plants indistinguishable from those inoculated with wild-type rhizobia. A double *aap bra* mutant showed yellowing of the plant, and nitrogen fixation rates 32-55% of wild-type, though *aap bra* mutants demonstrate prototrophy under free-living conditions (Lodwig, Hosie et al. 2003). This suggests a requirement for amino acid transport for effective nitrogen fixation in symbiosis with peas. An initial explanation for this was an amino acid cycle in which plants provided amino acids such as glutamate in return for aspartate or alanine from the fixing nodules (Prell and Poole 2006). Labelling studies have previously shown export of alanine and aspartate from bacteroids of pea and soybean during nitrogen fixation (Rosendahl, Dilworth et al. 1992, Waters, Hughes et al. 1998, Allaway, Lodwig et al. 2000). Subsequent mutation of a second solute binding protein, BraC3, which restricts transport to branched amino acids led to plants unable to fix nitrogen, so provision of only branched amino acids is essential for nitrogen fixation (Prell, White et al. 2009). This was named 'symbiotic auxotrophy' since the need for provision of branched amino acids is only essential during symbiosis with plants. The reason behind evolution of symbiotic auxotrophy is not yet known, but three reasons have been suggested. The first is that auxotrophy is induced by the plants; the second that it may be energetically favourable for the bacteroid; and finally that supply of branched amino acids by the bacteroid may be due to the nature of bacteroid metabolism limiting key intermediates (Terpolilli, Hood et al. 2012).

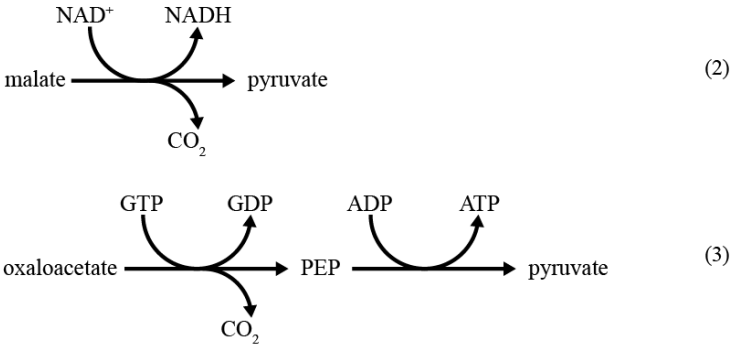
Ammonia produced by nitrogenase is exported to plants, where it is assimilated into amino acids. This begs the question whether rhizobia also assimilate ammonia, or whether provision of amino acids from plants is sufficient. Rhizobia possess the glutamate oxoglutarate amidotransferase (GOGAT) enzyme, mutation of which inhibits amino acid transport via Aap and Bra. The GS (glutamine synthase)/GOGAT pathway and alanine dehydrogenase pathway are principal routes for ammonia assimilation in *Rhizobium* (Allaway, Lodwig et al. 2000). A single mutant in *aldA* (alanine dehydrogenase) is still able to assimilate ammonia, and is not affected in its nitrogen fixation ability, whilst GOGAT activity is essential for bacteroid development, due to an inability to synthesise amino acids during bacteroid differentiation (Mulley, White et al. 2011).

1.3.4 Carbon metabolism during symbiosis

Whilst some *Bradyrhizobium* are photosynthetic (Fleischman and Kramer 1998), the majority of *Rhizobium* are heterotropic, able to grow on a wide range of carbon sources in the free-living state. These carbon sources are metabolised through three routes: the Entner-Doudoroff pathway; the pentose-phosphate pathway; and the tricarboxylic acid (TCA) cycle. Once inside nodules, bacteroids obtain carbon from the plant symbiont. Plants provide

dicarboxylic acids, usually malate, to the bacteroid in order to provide energy for nitrogen fixation. Dicarboxylic acids including malate and succinate are imported by the high-affinity dicarboxylic acid transport system, encoded by *dct* genes (Glenn, Poole et al. 1980, Ronson, Littleton et al. 1981, Finan, Hirsch et al. 1985). In *R. leguminosarum* the Dct system comprises of three proteins. DctA works as a H^+ symporter (Bhandari and Nicholas 1985) to move dicarboxylic acids across the membrane. DctBD act as a two-component sensor to activate *dctA* expression in the presence of dicarboxylates (Ronson, Astwood et al. 1987, Jording, Sharma et al. 1993). Mutants in any of the *dct* genes are unable to transport C_4-dicarboxylates in free-living culture, and form ineffective nodules on plants.

The C_4 dicarboxylates are metabolised via the TCA cycle, producing carbon dioxide. Mutational studies of TCA cycle enzymes show that the entire TCA cycle is generally essential for nitrogen fixation. The cycle may operate below its full potential due to the microaerobic environment in the nodule (Mcdermott, Griffith et al. 1989, Day and Copeland 1991, Streeter 1991). Acetyl-CoA is the input for the TCA cycle, produced from pyruvate via the pyruvate dehydrogenase complex (PDH). Pyruvate may be produced by one of two pathways. The first pathway utilises malic enzyme (ME) to produce pyruvate from malate (Reaction 1.1). The second pathway converts oxaloacetate to phosphoenolpyruvate (PEP) using PEP carboxykinase (PEPCK). PEP is then converted to pyruvate by pyruvate kinase (PYK) (Reaction 1.2).



(2)

(3)

Reaction 1.1 Production of pyruvate via malic enzyme.

Reaction 1.2 Production of pyruvate via PEP.

PEPCK is essential for free-living rhizobial growth on succinate, but not on glucose since PEPCK plays a role in gluconeogenesis. It has been shown in *S. meliloti* that the NAD^+-dependent form of ME, encoded by *dme*, is essential for symbiotic nitrogen fixation, whilst

the $NADP^+$-dependent form, encoded by *tme*, is not (Driscoll and Finan 1997). PEPCK activity and protein has not been detected in *S. meliloti* bacteroids, reinforcing dependence on ME (Finan, Mcwhinnie et al. 1991, Djordjevic 2004). In constrast, PEPCK activity has been detected in *R. leguminosarum* bv. *viciae* bacteroids during symbiosis both in biochemical assays (Mckay, Glenn et al. 1985) and in transcriptomic analyses where its expression has been seen to be increased in nodules (Karunakaran, Ramachandran et al. 2009). Mulley, Lopez-Gomez et al. (2010) demonstrated a role for both pathways in bacteroids of *R. leguminosarum* bv. *viciae* 3841, and only a *dme pckA* double mutant showed a Fix^- phenotype. Single mutants both were able to fix nitrogen, though the Dme pathway supported a higher rate of N_2 than the PEPCK pathway, suggesting this is the main pathway for pyruvate production. Figure 1.6 outlines the TCA cycle and interacting pathways in *Rhizobium* bacteroids.

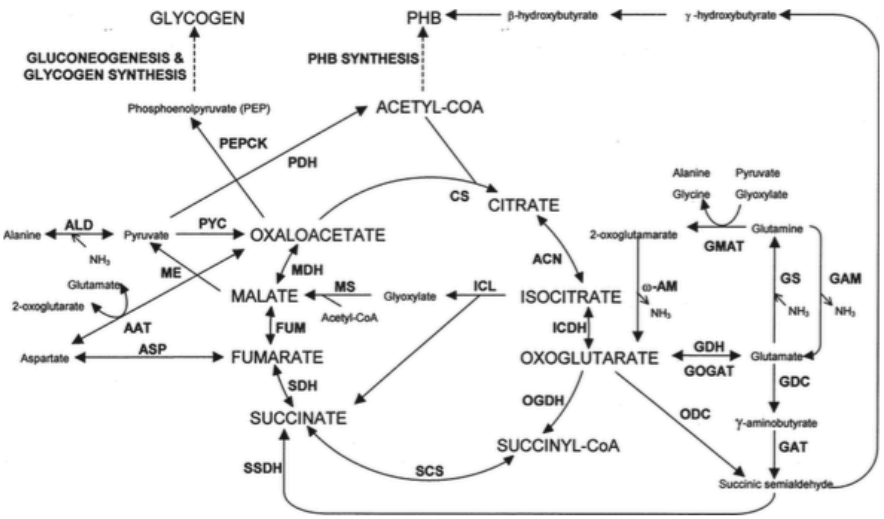


Figure 1.6 The TCA cycle and possible pathways of regulation in a *Rhizobium* bacteroid. AAT: aspartate aminotransferase; ACN: aconitase; ALD: alanine dehydrogenase; AM: amidase; ASP: aspartase; CS: citrate synthase; FUM: fumarate; GAM: glutaminase; GMAT: glutamine aminotransferase; GAT: γ-aminobutyrate aminotransferase; GDC: glutamate decarboxylase; GDH: glutamate dehydrogenase; GOGAT: glutamate synthase; GS: glutamine synthetase; ICDH: isocitrate dehydrogenase; ICL: isocitrate lyase; ODC: 2-oxoglutarate decarboxylase; OGDH: 2-oxoglutarate dehydrogenase; MDH: malate dehydrogenase; ME: malic enzyme; MS: malate synthase; PDH: pyruvate dehydrogenase; PEPCK: phosphoenolpyruvate carboxykinase; PYC: pyruvate carboxylase; SCS: succinyl-CoA synthetase; SDH: succinate dehydrogenase; SSDH: succinic semialdehyde dehydrogenase (Lodwig and Poole 2003).

1.3.5 Changes in carbon allocation and storage

Excess carbon in bacteroids is converted into polymers for use as redox stores. This includes poly-β-hydroxybutryate (PHB) (Lodwig, Leonard et al. 2005), a polyhydroxyalkanoate that can be used to generate acetyl-CoA under carbon starvation (Anderson and Dawes 1990). PHB is produced by many bacteria in nutrient-limited but carbon-rich environments (Madison and Huisman 1999). PHB has been identified as a storage compound in several rhizobial strains, including *R. leguminosarum* bv. *viciae* (Lodwig, Leonard et al. 2005, Wang, Saldanha et al. 2007). PHB is produced by a biosynthetic pathway, including a PHB synthase encoded by gene *phaC*, which can be found conserved across many PHB-producing bacteria (Rehm and Steinbuchel 1999). A role has been suggested for PHB in both the free-living and symbiotic state. In *R. leguminosarum* bv. *viciae* 3841 there are two copies of PHB synthase gene *phaC*. The *phaC1* gene is found on the chromosome, encoding a type I PHB synthase. A type III PHB synthase encoded by *phaC2* is found on pRL10, under the control of NifA, the general transcriptional regulator of nitrogen fixation. (Terpolilli, Masakapallic et al. 2016) As well as its role as a carbon store protecting from starvation (Ratcliff, Kadam et al. 2008), the presence of PHB has been shown to provide the cell protection from other stresses including heat and osmotic shock (Kadouri, Jurkevitch et al. 2003). PHB metabolism is thought to be linked to redox state; it has been proposed that in *Azotobacter* PHB synthesis acts as an alternative electron acceptor under conditions of oxygen limitation to relieve NAD(P)H inhibition of the TCA cycle (Senior, Beech et al. 1972, Page and Knosp 1989). In *A. caulinodans* PHB has been proposed to be required to maintain the redox state (Mandon, Michel-Reydellet et al. 1998). This redox role is important to consider due to the microaerobic conditions seen inside legume nodules during symbiosis.

Not all rhizobia produce PHB in bacteroids. This PHB production appears to be dependent on nodule morphology. Bacteroids in determinate nodules, such as *R. etli* in symbiosis with bean, accumulate large amounts of PHB during nitrogen-fixation (Wong and Evans 1971, Bergersen and Turner 1990); PHB can comprise up to 50% of *B. japonicum* bacteroid dry weight (Wong and Evans 1971, Karr, Waters et al. 1983). These PHB stores may be utilised by the plant during carbon-limitation (Gerson, Patel et al. 1978, Bergersen and Turner 1990). In contrast PHB is not accumulated in bacteroids of plants forming indeterminate nodules, though it is seen in early stages of infection. PHB can be clearly seen in electron micrographs as large electron-transparent granules (Craig and Williamson 1972, Hirsch, Bang et al. 1983) and is seen in high quantities in developing infection threads in *R. leguminosarum* bv. *viciae* before being broken down during bacteroid development; this suggests a role for accumulated PHB in fueling bacteroid differentiation (Lodwig, Leonard et al. 2005). A hypothesis for the difference in PHB accumulation between determinate and

indeterminate nodules is for provision of a competitive advantage upon release to the rhizosphere during nodule senescence. From indeterminate nodules these are undifferentiated cells in the infection thread, whilst in determinate nodules these are the nitrogen-fixing bacteria (Denison 2000). Accumulation of PHB in nodule sections seen by electron microscopy is shown in Figure 1.7, including strains mutated in *phaC1* and *phaC2* of *R. leguminosarum* bv. *viciae*. These mutations show production of infection thread PHB from PhaC1, whilst bacteroid PHB is produced by NifA-controlled PhaC2.

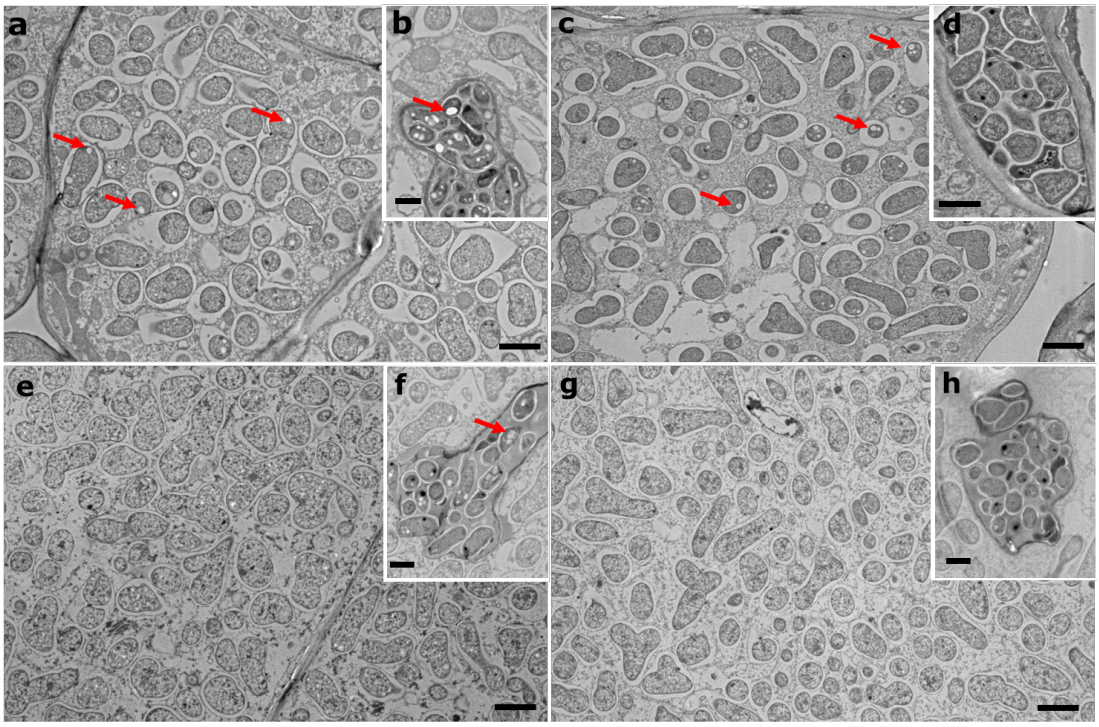


Figure 1.7 PHB can be seen in electron micrographs of *P. sativum* nodules inoculated with strains of *R. leguminosarum* bv. *viciae* at 28 days post-inoculation. Inset images show infection threads. a,b: Rlv3841; c,d: RU137 (*phaC1*); e,f: LMB814 (*phaC2*); g,h: LMB816 (*phaC1C2*). Scale bars a,c,e,g: 2 µm; b,d,f,h: 1 µm. Reproduced from Terpolilli, Masakapallic et al. (2016).

Glycogen has also been shown to play a role in carbon storage in rhizobia, though it has been less studied than PHB. Glycogen is produced simultaneously with PHB under limiting conditions such as nitrogen-limitation, suggesting it plays a similar role to PHB (Zevenhuizen 1981). Glycogen can also be detected via transmission electron microscopy, and was first identified in *R. leguminosarum* bv. *trifolii* (Dixon 1967). Mutation of the

glycogen synthase *glgA* in *R. tropici* led to increased nodulation and symbiotic efficiency (Marroqui, Zorreguieta et al. 2001), though *glgA* mutation in *R. leguminosarum* showed no effect on nitrogen-fixation ability (Lodwig, Leonard et al. 2005). Mutation of *glgA* in *S. meliloti* in symbiosis with *M. truncatula* showed lower than wild-type levels of nitrogen fixation (Wang, Saldanha et al. 2007). Cross-talk between PHB and glycogen has been suggested in free-living conditions by Cevallos, Encarnacion et al. (1996) as successive subculture of *R. etli phaC* mutants showed increased glycogen accumulation. Increases in glycogen content in PHB mutants was also seen in *R. tropici* in the free-living state (Povolo and Casella 2002). Though PHB and glycogen are thought to play similar roles in redox protection, double mutants of *phaC glgA* in *R. leguminosarum* also show wild-type nitrogen fixation levels (Lodwig, Leonard et al. 2005).

1.4 Nitrogen fixation

1.4.1 The nitrogenase enzyme

Nitrogen fixation is catalysed by the nitrogenase enzyme encoded by *nifHDK*. The reduction of nitrogen is an energy demanding process, requiring sixteen ATP and eight electrons per molecule of dinitrogen fixed. This reaction occurs at ambient temperature and pressure in rhizobia, in stark contrast to the high temperatures and pressures required for the chemical Haber-Bosch Process.

$$N_2 + 16H_2O + 16ATP + 10H^+ + 8e^- \rightarrow 2NH_4{}^+ + H_2 + 16ADP + 16P_i$$

Reaction 1.3 Reaction catalysed by nitrogenase.

There are three families of metalloenzyme nitrogenase proteins, distinguished from each other mainly by the presence of Mo-Fe, V-Fe, or Fe only in the active site (Igarashi and Seefeldt 2003, Barney, Lee et al. 2006). The Mo-Fe nitrogenase is by far the best characterised, though recent studies have begun elucidating the mechanism of the V-Fe and Fe-Fe nitrogenases (Hinnemann and Norskov 2004, Hu, Lee et al. 2012). All three nitrogenase families include a smaller Fe-containing component protein, containing a [4Fe-4S] cluster which requires MgATP to deliver electrons to the larger protein complex (Burgess and Lowe 1996). The MoFe subunit contains two redox active clusters, the [8Fe-7S] 'P-cluster' and a [7Fe-9S-Mo-C-homocitrate] 'FeMo cofactor'. The FeMo cofactor is the

site of nitrogen binding and reduction (Shah, Stacey et al. 1983). Electrons are delivered one at a time (Hageman and Burris 1978) from the Fe-protein to the P-cluster (Chan, Christiansen et al. 1999) in a process coupled to hydrolysis of two MgATP molecules (Howard and Rees 1994, Seefeldt and Dean 1997). Eight association/dissociation events between the Fe and Mo-Fe proteins are required per N_2 fixed (Barney, Lee et al. 2006). Whilst a lot is known about the nitrogenase mechanism, there is still much to discover despite 40 years worth of research (Hoffman, Dean et al. 2009). This includes full understanding of the inter- and intra-molecular electron transfer events occurring during nitrogen-fixation, substrate binding and activation, and understanding how nucleotide binding and hydrolysis are linked to electron transfer within the nitrogen-fixation reaction. Advances in chemical and spectroscopic techniques have allowed better understanding of processes involved, and opened doors to understanding more about these processes.

1.4.2 Homocitrate

The nitrogenase enzyme complex includes a Fe-Mo cofactor, with an essential requirement for homocitrate (Hoover, Imperial et al. 1989). In free-living diazotrophs such as *K. pneumoniae* and *A. vinelandii* homocitrate is generated from acetyl-CoA and 2-oxoglutarate via homocitrate synthase, encoded by the *nifV* gene (Zheng, White et al. 1997). Whilst *K. pneumoniae nifV* mutants are able to fix nitrogen, they do so at lower rates than the wild-type, though rates can be returned to their wild-type level if homocitrate is provided (Hoover, Imperial et al. 1988). Most symbiotic diazotrophs lack a copy of *nifV*. The absence of homocitrate synthase in these bacteroids can be compensated by provision of homocitrate synthase by the host plant. This was shown in the *M. loti-L. japonicus* symbiosis, in which mutants in the plant *FEN1* gene were unable to fix nitrogen, though they formed nodules (Hakoyama, Niimi et al. 2009). Expression of *L. japonicus FEN1* or *A. vinelandii nifV* genes in *M. loti* restored a nitrogen fixation phenotype on *FEN1* mutant plants (Hakoyama, Niimi et al. 2009). Rhizobial species such as *A. caulinodans* which fix nitrogen in both free-living culture and in symbiosis carry *nifV* genes (Lee, De Backer et al. 2008). The reason behind plant provision of homocitrate is unknown, though it may be an example of plant control over nitrogen fixation.

1.4.3 Other genes essential for nitrogen fixation

In addition to *nifHDK* there are a host of other genes encoding key proteins involved in nitrogen fixation. The *nif* genes were originially identified in *K. pneumoniae* as having a role in nitrogen fixation. The *K. pneumoniae* major *nif* cluster was first characterised by Dixon, Eady et al. (1980) and is shown in Figure 1.8

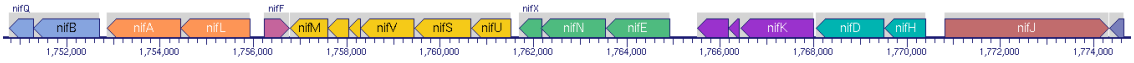


Figure 1.8 The *K. pneumoniae* major *nif* gene cluster (biocyc.org).

Additional nitrogen fixation genes identified in rhizobia were given the *fix* designation. The *fix* genes include the terminal oxidase *fixNOQP,* and oxygen-regulatory genes *fixLJK* (Fischer 1994). Function has since been assigned to other *nif* genes. *nifFJ* encode electron transferring proteins, involved in electron transfer to the *K. pneumoniae* nitrogenase (Hill and Kavanagh 1980). These genes are not conserved across all nitrogen-fixing species; Most symbiotic diazotrophs have alternative electron transfer proteins e.g. FixABCX (Edgren and Nordlund 2004). In addition to these, a ferredoxin, FdxN is also required in some species for functional nitrogen fixation (Klipp, Reilander et al. 1989, Riedel, Jouanneau et al. 1995, Edgren and Nordlund 2005). NifA is a positive regulator of nitrogen fixation (Beynon, Williams et al. 1988, Salazar, Diaz-Mejia et al. 2010), and NifL is a negative regulator under oxygen control (Dixon 1998). *nifB, nifNE, nifQ, nifV, nifZ, nifU, nifX* and *nifS* are required for nitrogenase cluster biosynthesis, and *nifM* and *nifY* are required for nitrogenase processing (Beynon, Cannon et al. 1988, Ludden 1993, Fu, Jack et al. 1994, Hu, Fay et al. 2007). Some *nif* proteins, such as *nifW* and *nifT* are yet to be functionally characterised.

NifA is known as the general transcriptional regulator of nitrogen fixation. NifA is a member of the enhancer-binding protein (EBP) family, which works in concert with its cognate sigma factor (σ^{54}) RpoN to drive expression of key nitrogen fixation genes including *nifHDK* and *fixABCX* (Salazar, Diaz-Mejia et al. 2010, Sullivan, Brown et al. 2013) as well as expression of other systems involved in legume-rhizobia symbioses such as uptake hydrogenases (Brito, Martinez et al. 1997) and rhizopine synthesis (Heinrich, Ryder et al. 2001). The NifA-RpoN regulon acts on -24/-12 type promoters, recognising an upstream activator sequence (UAS) in the promoter region of nitrogen fixation genes. Studies in several organisms have shown that the *nifA* gene is essential for symbiotic nitrogen fixation across rhizobial species (Fischer, Alvarezmorales et al. 1986, Hirsch and Smith 1987, Labes, Rastogi et al. 1993).

1.4.4. The *fixABCX* genes

One of the most upregulated operons in bacteroids is the *fixABCX* gene cluster. In *R. leguminosarum* bv. *viciae fixA* is the most upregulated gene in mature bacteroids, with transriptomic studies showing *fixA* to be upregulated 120-fold 28 days post-inoculation, double that of the nitrogenase subunit *nifH* (60-fold). The other proteins in the cluster, *fixBCX* are all upregulated ~20-fold in mature bacteroids, and *nifHDK* are upregulated 60-100-fold. (Karunakaran, Ramachandran et al. 2009). The top 100 upregulated genes from this study are listed in Table 1.1. Studies in other rhizobia have shown similar upregulation; the *fixA* gene as listed in the top 25 upregulated genes in *M. haukii* 7653R bacteroids (Peng, Hao et al. 2014); *fixA* is upregulated in *Bradyrhizobium* sp. ORS278 bacteroids (Delmotte, Mondy et al. 2014) and *fixABCX* are also highly upregulated in *S. meliloti* bacteroids in symbiosis with *M. truncatula* (Becker, Berges et al. 2004).

Table 1.1 The 100 most highly upregulated genes in 28-day *R. leguminosarum* bv *viciae* 3841 pea bacteroids compared with succinate-grown free-living cells. Gene annotation from Poole lab, September 2016. Adapted from Karunakaran, Ramachandran et al. (2009).

Gene	Fold induction	Gene name	Gene annotation
pRL100067	18		Putative racemase/decarboxylase
pRL100092	5		Putative transcriptional regulatory protein, pseudogene
pRL100097	12		Conserved hypothetical protein
pRL100098	14		Conserved hypothetical protein
pRL100099	19		Conserved hypothetical protein
pRL100100	65		Putative nitrogenase iron protein, pseudogene (has start of *nifH*)
pRL100101	105		Putative acetolactate synthase, pseudogene
pRL100103	13		Putative hydroxyvalerate/ alcohol dehydrogenase
pRL100104	39	*phaE*	Putative subunit of polyhydroxyalkanoate (PHA) synthase
pRL100105	10	*phaC2*	Polyhydroxyalkanoate (PHA) synthase subunit C
pRL100106	21		Hypothetical protein
pRL100117	7		Putative iron-containing alcohol dehydrogenase, pseudogene
pRL100119	34		Putative propionate CoA transferase
pRL100120	9		Phasin-like protein, often associated with PHB or other polyhydroxyalkanoate (PHA)
pRL100121	6	*acsA3*	Putative acetyl-coA synthetase
pRL100137	25	*metx*	Putative homoserine O-acetyltransferase
pRL100148	19	*thiC*	Putative thiamine biosynthesis protein, pseudogene
pRL100154	5		Hypothetical protein
pRL100156	9	*fdxB*	Putative ferredoxin
pRL100157	12		NifX protein, pseudogene
pRL100158	23	*nifN*	Putative nitrogenase iron-molybdenum cofactor biosynthesis protein
pRL100159	46	*nifE*	Putative nitrogenase iron-molybdenum cofactor biosynthesis protein
pRL100160	94	*nifK*	Nitrogenase molybdenum-iron protein beta chain
pRL100161	71	*nifD*	Nitrogenase molybdenum-iron protein alpha chain
pRL100162	60	*nifH*	Nitrogenase iron protein
pRL100169	6	*rhiA*	Rhizosphere-induced protein
pRL100192	9		Hypothetical protein
pRL100193	9		Conserved hypothetical protein
pRL100194	9		Hypothetical protein
pRL100195	10	*nifB*	FeMo cofactor biosynthesis protein
pRL100197	17	*fixX*	Ferredoxin-like protein
pRL100198	24	*fixC*	Nitrogen fixation protein
pRL100199	28	*fixB*	Electron transfer protein
pRL100200	120	*fixA*	Electron transfer protein
pRL100201	26		Conserved hypothetical protein
pRL100205	27	*fixN1*	Putative transmembrane cytochrome oxidase subunit
pRL100206	45	*fixO1*	Putative cytochrome oxidase subunit
pRL100206	24	*fixQ1*	Putative component of cytochrome oxidase
pRL100207	38	*fixP1*	Putative cytochrome oxidase subunit

pRL100210A	7	*fixS1*	Putative component of cation pump
pRL100444	103		Putative oxidoreductase
pRL110199	62		Conserved hypothetical protein
pRL120027	15		Putative aldolase
pRL120227	10		Putative NAD-dependent epimerase/dehydratase
pRL120254	5		Conserved hypothetical protein
pRL120632	31		Putative dehydrogenase
pRL120644	7		Conserved hypothetical protein
pRL90016	32	*fixP2*	Putative cytochrome oxidase component cytochrome c protein
pRL90016A	27	*fixQ2*	Putative component of cytochrome oxidase
pRL90017	21	*fixO2*	Putative transmembrane cytochrome c oxidase protein
pRL90018	11	*fixN2*	Putative cytochrome oxidase transmembrane component
pRL90021	14	*azuP*	Putative pseudoazurin protein
pRL90031	10		Conserved hypothetical protein
pRL90039	6		Conserved hypothetical protein
pRL90046	5		Hypothetical protein
pRL90047	14		Putative universal stress protein
pRL90194	5		Putative UPF0261 domain protein
pRL90322	18		Hypothetical protein, pseudogene
RL0018	11		Putative transmembrane protein
RL0054	9	*glcB*	Putative malate synthase
RL0102	7	*gabT*	Putative 4-aminobutyrate aminotransferase
RL0540	7		Putative two-component sensor/regulator histidine kinase
RL0680	35	*secDF2*	Putative transmembrane export SecD/F family protein
RL0702	7	*fliM*	Putative flagellum motor switch protein
RL0913	11		Putative PRC family protein
RL1145	12		Putative conjugated bile salt hydrolase
RL1148	33		Hypothetical protein
RL1149	6		Putative exported arylsulfatase
RL1150	7		Hypothetical protein
RL1151	7		Conserved hypothetical protein
RL1173	10		Putative transmembrane protein
RL1249	5		Hypothetical protein
RL1259	6		Conserved hypothetical protein
RL1423	6		Putative transmembrane protein
RL1606	5		Putative esterase (beta-lactamase family)
RL1694	22		Putative aldo-keto reductase
RL1860	6	*phhA*	Putative phenylalanine-4-hydroxylase
RL1868	9		Putative universal stress-related protein
RL1870	7	*dksA*	Putative DNAK suppressor protein
RL1871	5		Putative transmembrane cation ATPase transporter
RL1873	5		Putative pyridoxine oxidase protein
RL1876	5	*adh*	Putative alcohol dehydrogenase
RL1877	5		Putative protease
RL1878	13		Putative peptidoglycan binding protein
RL1883	56	*hspF*	Putative small heat shock protein
RL2271	30		Conserved hypothetical protein
RL2272	19		Conserved hypothetical protein
RL2323	6		Putative GFO/IDH/MocA dehydrogenase
RL2561	7		Putative transmembrane efflux/MDR protein

RL3016	137	*pcaH2*	Protocatechuate 3,4-dioxygenase beta chain (3,4-pcd)
RL3130	14		Putative phosphatase protein
RL3274	5	*prkA*	Putative PrkA protein kinase
RL3366	14		Putative flavoprotein
RL3495	5		Conserved hypothetical protein
RL3533	5	*qatX1*	Putative solute-binding component of ABC transporter, QAT family
RL3682	9		Putative transmembrane component of ABC exporter
RL4089	21	*ibpA*	Putative heat shock protein A
RL4244	7		Putative transmembrane component of ABC transporter, CUT1 family
RL4279	5	*clpB*	Putative chaperone ClpB (heat shock protein)
RL4530	6		Putative orphan ATP-binding component of ABC transporter, family unclassified

It has been known for several decades that the *fix* operon is essential for symbiotic nitrogen fixation (Gubler and Hennecke 1986, Kaminski, Norel et al. 1988). The *fix* operon has conserved regulation across the *Rhizobiacae*, with *fixABCX* being under the control of NifA, the general transcriptional regulator of nitrogen fixation (Beynon, Williams et al. 1988, Salazar, Diaz-Mejia et al. 2010, Sullivan, Brown et al. 2013). There is some evidence that the *nifA* gene is autoregulated from an upstream region, potentially within or upstream of the *fixABCX* region (Kim, Helinski et al. 1986, Martinez, Palacios et al. 2004).

Figure 1.9 Genetic context of the *fixABCX* operon in *R. leguminosarum* bv. *viciae* 3841.

The *fixABCX* operon encodes a set of four proteins thought to play a role in electron transfer (Arigoni, Kaminski et al. 1991). FixA and FixB show homology to the beta and alpha subunits respectively of mammalian electron transfer flavoproteins (ETFs) (Arigoni, Kaminski et al. 1991, Tsai and Saier 1995). FixCX show homology to the ETF's cognate electron acceptor electron transfer flavoprotein-quinone oxidase (ETF-QO) (Edgren and Nordlund 2004, Watmough and Frerman 2010), with FixX being a ferredoxin-like protein (Earl, Ronson et al. 1987). Roles have been shown for the *fixABCX* genes in carrying electrons to nitrogenase in free-living diazotrophs such as *Rhodospirillum rubrum* (Edgren and Nordlund 2004), and a similar role has been proposed in symbiosis (Gubler and Hennecke 1986, Earl, Ronson et al. 1987).

1.4.5 Electron-transfer flavoproteins

The ETF protein family is found across the Kingdoms of Life with roles in energy metabolism. The first ETF was identified in mammals in 1956 (Crane and Beinert 1956), involved in mitochondrial β-oxidation of long-chain fatty acids. ETFs are heterodimeric FAD-containing proteins mediating electron transport (Buckel and Thauer 2013). The mammalian ETF is found in mitochondria, where it carries one electron from nine primary flavoprotein dehydrogenases, including four acyl-CoA dehydrogenases in the mitochondrial matrix to the inner membrane where it passes electrons to an electron-transfer flavoprotein-quinone oxidase (ETF-QO) (Crane and Beinert 1956, Roberts, Frerman et al. 1996). The human ETF is important for meeting the energy requirements of heart and skeletal muscle

(Bartlett and Eaton 2004) and has been well-studied due to the consequences of inherited mutations in metabolic diseases (Bartlett and Pourfarzam 1998, Olsen, Andresen et al. 2003). The structure of human ETF has been resolved to 2.1Å, showing three distinct domains. Two of these domains, facilitating FAD binding and flavoprotein-flavoprotein interactions, are found in the alpha-subunit (homologous to FixB) with the other domain, binding 5'-AMP found in the beta-subunit (homologous to FixA). The role of 5'-AMP in ETF function has not yet been characterised; structural analysis of the human ETF does not indicated a clear role for 5'AMP binding(Sato, Nishina et al. 1993, Roberts, Frerman et al. 1996). Alignments with other known ETF-alpha proteins including a rhizobial FixB have shown conservation in the FAD-binding domain (Roberts, Frerman et al. 1996). The structure of a mammalian ETF-QO has been resolved to 2.1Å, also consisting of three functional domains binding FAD, iron-sulfur cluster and ubiquinone (Zhang, Frerman et al. 2006). These three domains appear to be conserved across Kingdoms of Life, including prokaryotes (*R. sphaeroides)* (Watmough and Frerman 2010).

1.4.6 Flavin-based electron bifurcation

In the last decade a number of studies have demonstrated an electron bifurcating mechanism carried out by ETF complexes. Electron bifurcation couples an exergonic electron transfer reaction to an endergonic one, so conserving energy. The concept of electron bifurcation is not a recent one; it occurs in the bc_1 complex during respiration (Mitchell 1975, Brandt 1996). However, the idea of flavin-based bifurcation was only recently suggested (Herrmann, Jayamani et al. 2008). In these flavin-based bifurcation reactions the endergonic reduction of ferredoxin (Fd) (E°' = -500mV) is coupled to an exergonic redox reaction in order to overcome the energetic requirement for Fd reduction (Buckel and Thauer 2013). Flavin-based bifurcation was initially proposed for a step of butyric acid synthesis in the anaerobe *Clostridium kluyveri*, in a reaction catalysed by the butyryl-CoA dehydrogenase-EtfCB complex (Figure 1.10). In *C. kluyveri* each Fd has two Fe-S centres, and so these ferredoxins are able to carry two electrons each. The ETF in this case contains two molecules of FAD (only one is found in mammalian and rhizobial ETFs). In this reaction Fd reduction with NADH is coupled to crotonyl-CoA reduction (Reaction 1.4) (Li, Hinderberger et al. 2008). Bifurcation has been demonstrated using the same proteins from both *Clostridium tetanomorphum* and *Acidaminococcus fermentans* (Li, Hinderberger et al. 2008, Chowdhury, Mowafy et al. 2014) .

Fd_{ox} +2NADH + crotonyl-CoA ⟶ $Fe^{2-}{}_{red}$ + $2NAD^+$ + butyryl-CoA

$\Delta G° = -44 kJ/mol.$

Reaction 1.4 Reaction catalysed by butyryl-CoA dehydrogenase.

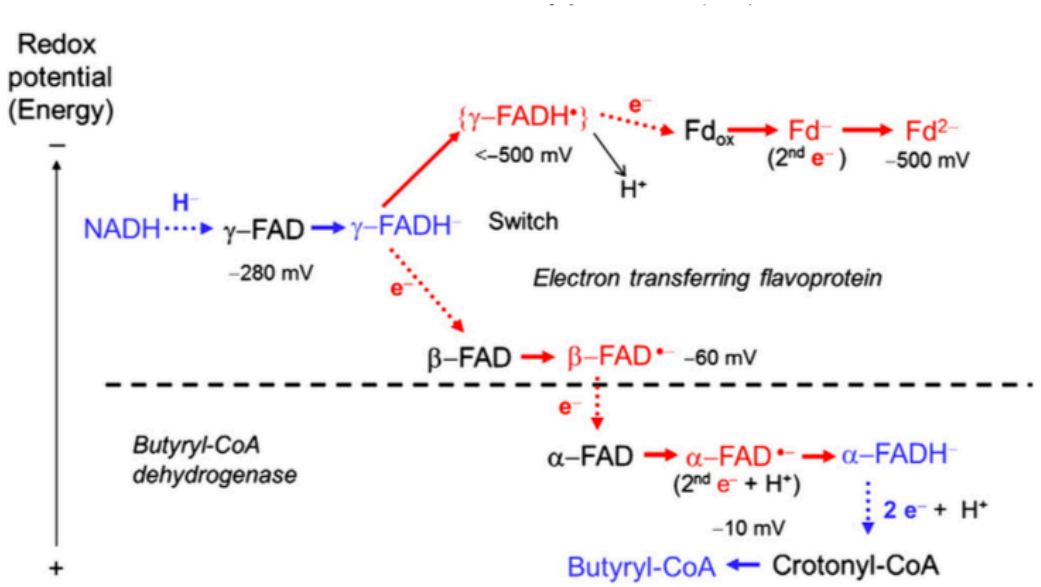


Figure 1.10 Proposed mechanism of electron bifurcation by the butyryl-CoA-EtfBC complex. Black species are oxidised, blue species are reduced by $2e^-$, red species carry unpaired e^-. Dotted arrows indicate electron transfers. Reproduced from Buckel and Thauer (2013).

Flavin-based electron bifurcation has been demonstrated in multiple other systems, such as the reduction of caffeyl-CoA in the acetogenic bacterium *Acetobacterium woodii*, where caffeyl-CoA reduction with NADH is coupled to Fd reduction by the caffeyl-CoA reductase-Etf complex. In this study the in vitro reduction of caffeyl-CoA increased 10-fold upon addition of ferredoxin (Bertsch, Parthasarathy et al. 2013). Another example is the bifurcation by the hydrogenase HydABC from *Thermotoga maritima*, which catalyses evolution of hydrogen with reduced ferredoxin as an electron donor when NADH is present (Verhagen, O'Rourke et al. 1999).

Based on experimental data, electron bifurcation can be predicted for several further ETF-based reactions, as reviewed by Buckel and Thauer (2013) and (Herrmann, Jayamani et al. 2008). Herrmann, Jayamani et al. (2008) hypothesises that the nitrogen fixation FixAB

proteins function in this manner, where two electrons from NADH are bifurcated, with one being passed to nitrogen fixation via ferredoxin, whilst the other is passed to FixCX and onwards to ubiquinone in the respiratory chain. Since FixAB has only one flavin, there may be a role for ubiquinone as acting as a flavin in this system. There is currently a lack of experimental evidence for FixAB functioning in this bifurcating role, and no biochemical evidence for interaction between FixAB and a dehydrogenase.

1.5 Biotechnology and future directions for nitrogen fixation research

As global populations grow so will demand for food and the inputs required for food production. Chemical inputs of nitrogen already put a strain on the environment due to pollution and leaching into water systems. Additionally the energy requirements for production of chemical fertilisers are huge, and lead to high costs that are impractical for poorer regions of the world. Given current trends in population, agriculture and energy use, humans are likely to be responsible for doubling turnover rates of the Earth's terrestrial nitrogen cycle (Gruber and Galloway 2008).

Advances in biology and biotechnology provide potential for manipulation of the rhizobia-legume symbioses to lower pressures on the nitrogen cycle by reducing chemical nitrogen inputs. This could be by improving our current symbioses or by engineering a nitrogen symbiosis into new systems and new organisms (Mus, Crook et al. 2016). Improving our understanding of key processes and enzymes in nitrogen fixation will allow better use of synthetic biology to work towards these aims. A role for synthetic biology in symbiosis was proposed several decades ago, and it is only now that technology is enabling this to become a reality.

One such aim is the development of nitrogen-fixation ability by new organisms such as cereal crops (Oldroyd and Dixon 2014). This requires expression of a functional nitrogenase enzyme in eukaryotes. Direct *nif* gene transfer has proven difficult due to the complex assembly of the enzyme, as well as the inhibitory action of oxygen (Curatti and Rubio 2014). Recent work has demonstrated expression of a functional nitrogenase Fe protein in yeast by use of the mitochondrial matrix to provide the required low-oxygen environment (Lopez-Torrejon, Jimenez-Vicente et al. 2016). This was achieved by provision of not only the *nifH* gene, but also the other *nif* genes required for correct assembly. In this experiment *nifS* and *nifU* were not required for assembly of a functional NifH. This proof-of-concept in yeast

mitochondria will lead to the engineering of a fully-functional nitrogenase in plant mitochondria.

An alternative to nitrogen-fixing cereal crops is to engineer the existing symbiosis into one that allows partnership between nitrogen-fixing bacteria and a cereal, rather than a legume. This requires dissection of several pathways, including signalling, nodule formation, bacterial infection and development of the required environment for nitrogen-fixing genes to function (Oldroyd and Dixon 2014). Just as the legume-rhizobia symbiosis operates a tight host-symbiont signalling system, a synthetic symbiosis would require tight signalling to allow colonisation of the root and establishment of a successful relationship (Mus, Crook et al. 2016). Development of the signalling pathway may be aided by the presence of the mycorrhizal symbiosis in several crops including rice. Mycorrhizal and rhizobial symbioses share a common signalling pathway (Bonfante and Genre 2010), so only the rhizobia-specific components must be engineered. Full characterisation of Nod factor recognition and downstream effects will be important for developing this pathway (Oldroyd, Murray et al. 2011).

As synthetic biology improves, so does the suite of tools and parts available for engineering (Purnick and Weiss 2009, Brophy and Voigt 2014). Many such tools are available in bacteria, which are far easier to manipulate than plants due to their fast generation times and smaller genomes. Development of new symbiotic relationships requires better understanding of plant and eukaryote genetic systems and development of novel tools to allow controlled gene expression in cereals (Mus, Crook et al. 2016). Technologies such as refactoring (rewriting the coding sequence to maintain amino acid sequence whilst altering DNA enough to not retain any native regulatory elements) will aid the transfer of bacterial operons and allow new regulatory systems to be put in place. The *K. oxytoca nif* cluster has already been refactored, demonstrating the potential for this as a synthetic tool (Temme, Zhao et al. 2012).

Reaching a full understanding of the processes involved in the legume-rhizobia symbioses will require cross-disciplinary approaches, and the emergence of new technologies will provide new information about the symbiotic relationship that will help reduce dependence on chemical nitrogen fertilisers.

1.6 Research aims

To date the FixABCX proteins of symbiotic rhizobia have not been biochemically characterised. In this study we set out to investigate these proteins, and the genes encoding them using the model rhizobium *R. leguminosarum* bv. *viciae* 3841. This strain was chosen due to the volume of transcriptomic and physiological knowledge available in the Poole lab. Additionally, the symbiont of 3841, *P. sativum* (pea), yields large nodules which provides a substantial amount of working material.

Initial bioinformatics (Chapter 3) was used to better understand *fixABCX* and its gene products in *R. leguminosarum*. The project then began with characterisation of polar and in-frame mutants of genes encoding the *fix* operon, as well as other putative genes encoding electron transport proteins (Chapter 4). In Chapter 5 transmission electron microscopy of FixABCX mutants revealed evidence of strategies for redox protection in the absence of nitrogen fixation. A novel microscopy technique was used to investigate the biochemical composition of these and other mutants at the single-cell level. Chapter 6 considers the regulation of the *fixABCX* operon as well as regulation of the region immediately downstream which encodes the general transcriptional regulator of nitrogen fixation, *nifA*. Finally Chapter 7 uses a proteomics approach to investigate interactions between FixABCX and other proteins in *R. leguminosarum*. This data has provided better characterisation of the *fix* operon, as well as more general insights into the functioning of the nitrogen fixation symbiosis between legumes and rhizobia.

CHAPTER 2

Materials and Methods

2.1. Media, antibiotics and other chemicals

2.1.1 Media

E. coli strains were grown at 37°C in Luria Bertani broth (LB) (10g L^{-1} tryptone, 5 g L^{-1} yeast extract, 5 g L^{-1} NaCl). Cultures were shaken at 200 rpm unless otherwise stated. For growth on solid media, agar was added 1.4 w/v prior to autoclaving.

R. leguminosarum and *A. caulinodans* were grown on tryptone yeast (TY) media (5 g L^{-1} tryptone, 3 g L^{-1} yeast extract, 6 mM $CaCl_2$) (Beringer 1974). For growth on solid media, agar was added 1.75 w/v before autoclaving.

When a minimal growth media was required, rhizobia were grown in Universal Minimal Salts (UMS). UMS contains 0.5 mM K_2HPO_4, 0.5 g L^{-1} $MgSO_4.7H_2O$, 0.2 g L^{-1} NaCl, 4.19 g L^{-1} MOPS, adjusted to pH 7.0 and autoclaved. Following autoclaving media was supplemented with 1 mL of iron stock solution (12 g L^{-1} FeSO4.7H2O), 1 mL of calcium stock solution (75 g L^{-1} CaCl2.2H2O) and 1 mL of a stock solution containing 1 g L^{-1} thiamine hydrochloride, 2 g L^{-1} D-Pantothenic acid calcium salt and 100 mg L^{-1} biotin. For *R. leguminosarum* growth UMS was supplemented with filter-sterilised 10 mM glucose, 10 mM NH_4Cl and 1 µL of 1000x vitamin solution. One litre of vitamin stock contains 0.375 g L^{-1} EDTA-Na_2, 0.16 g L^{-1} ZnSO4.7H2O, 0.2 g L^{-1}, $NaMoO_4$, 0.25 g L^{-1} H_3BO_3, 0.2 g L^{-1} $MnSO_4.4H_2O$, 0.02 g L^{-1} $CuSO_4.5H_2O$, 1 g L^{-1} $CoCl_2.6H_2O$. *R. leguminosarum* was grown at 28°C, with 200 rpm shaking for liquid media. *A. caulinodans* ORS571 growth media was supplemented with 20mM succinate, 10 mM NH_4Cl plus vitamin solution. *A. caulinodans* is auxotrophic for nicotinate, so 30 µM nicotinic acid was added to the growth medium.

All bacterial strains were stocked in 15% glycerol and snap-frozen in liquid nitrogen for storage at -80°C.

2.1.2 Antibiotics and other chemicals

Appropriate antibiotics were added to media in concentrations as shown in Table 2.1. *R. leguminosarum* bv. *viciae* 3841 (Rlv3841) has chromosomal resistance to streptomycin, and this was used as a selection antibiotic. Ampicillin or nitrofuratoin was used for selection of *A. caulinodans* ORS571. In cases of *E. coli* transformation where blue-white screening could be used for screening colonies, X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) was added to LB agar at 40 µg mL^{-1}.

Table 2.1 Working concentrations of antibiotics used in this work in μg mL^{-1}. a: higher concentration used for selection of interposon mutants.

	E. coli	***R. leguminosarum* bv. *viciae* 3841**	***A. caulinodans* ORS571**
Ampicillin	100		100
Chloramphenicol	10		
Gentamycin	10	20-50[a]	25
Kanamycin	20	40-160	
Neomycin	20	80/250[a]	80
Nitrofurantoin		5	10
Spectinomycin	50	100	
Streptomycin	25	250	
Tetracycline	10	5	

2.2 Bacterial strains, plasmids and primers

2.2.1 Strains

All bacterial strains used for this thesis are listed in Table 2.2 (*R. leguminosarum*), Table 2.3 (*A. caulinodans*) and Table 2.4 (*E. coli)*.

Table 2.2 *R. leguminosarum* strains used in this work.

Strain	Description	Reference
Rlv3841	*Rhizobium leguminosarum* bv. *viciae*; Str^r derivative of strain 300	Johnston and Beringer (1975)
D5250	Rlv3841 [pIJ11282]; Tc^r	Frederix, Edwards et al. (2014)
LMB151	Rlv3841 [pJP2]; Tc^r	Prell, Boesten et al. (2002)
LMB767	Rlv3841 [pLMB811]; Tc^r	This work
LMB768	Rlv3841 [pLMB810]; Str^r Tc^r	This work
LMB769	Rlv3841 [pLMB802]; Str^r Tc^r	This work
LMB770	Rlv3841 [pLMB809]; Str^r Tc^r	This work
LMB771	Rlv3841 *fixAB*::Ω; Str^r Spc^r	This work
LMB777	Rlv3841 Δ*fixAB*; Str^r	This work
LMB779	Rlv3841 [pLMB779]; Str^r Gm^r	This work
LMB780	Rlv3841 [pLMB780]; Str^r Gm^r	This work
LMB781	Rlv3841 [pLMB781]; Str^r Gm^r	This work
LMB782	Rlv3841 [pLMB782]; Str^r Gm^r	This work
LMB783	Rlv3841 [pLMB783]; Str^r Gm^r	This work
LMB784	LMB777 [pLMB811]; Str^r Tc^r	This work
LMB785	LMB777 [pLMB810]; Str^r Tc^r	This work
LMB786	LMB777 [pLMB802]; Str^r Tc^r	This work
LMB787	LMB777 [pLMB809]; Str^r Tc^r	This work

LMB788	LMB771 [pLMB811]; Str^r Spc^r Tc^r	This work
LMB789	LMB771 [pLMB810]; Str^r Spc^r Tc^r	This work
LMB790	LMB771 [pLMB802]; Str^r Spc^r Tc^r	This work
LMB791	LMB771 [pLMB809]; Str^r Spc^r Tc^r	This work
LMB792	Rlv3841 [pLMB826]; Str^r Tc^r	This work
LMB793	LMB771 [pLMB826]; Str^r Spc^r Tc^r	This work
LMB794	LMB777 [pLMB826]; Str^r Tc^r	This work
LMB816	Rlv3841 *phaC1*::Tn5 *phaC2*::Ω; St^r Nm^r Spc^r	Terpolilli, Masakapallic et al. (2016)
LMB819	Rlv3841 *fixB*::Ω; Str^r Spc^r	This work
LMB820	Rlv3841 [pLMB833]; Str^r Tc^r	This work
LMB823	Rlv3841 *fixC*::pK19mob; Str^r Nm^r	This work
LMB824	Rlv3841 *fixX*::pK19mob; Str^r Nm^r	This work
LMB825	Rlv3841 Δ*fixB*; Str^r	This work
LMB827	Rlv3841 *fixC*::Ω; Str^r Sp^r	This work
LMB829	Rlv3841 *fixX*::Ω; Str^r Spc^r	This work
OPS0280	Rlv3841 Δ*fixC*; Str^r	This work
OPS0282	LMB819 [pLMB826]; Str^r Spc^r Tc^r	This work
OPS0283	LMB825 [pLMB826]; Str^r Tc^r	This work
OPS0284	LMB827 [pLMB826]; Str^r Spc^r Tc^r	This work
OPS0285	OPS0280 [pLMB826]; Str^r Tc^r	This work
OPS0287	Rlv3841 [pOPS0129]; Str^r Gm^r	This work
OPS0288	Rlv3841 [pOPS0130]; Str^r Gm^r	This work
OPS0289	Rlv3841 [pOPS0131]; Str^r Gm^r	This work
OPS0290	Rlv3841 [pOPS0132]; Str^r Gm^r	This work
OPS0328	Rlv3841 [pOPS0124]; Str^r Tc^r	This work
OPS0354	Rlv3841 [pOPS0126]; Str^r Tc^r	This work
OPS0355	LMB777 [pOPS0124]; Str^r Tc^r	This work

OPS0357	OPS0280 [pOPS0126]; Str^r Tc^r	This work
OPS0372	LMB771 [pOPS0126]; Str^r Spc^r Tc^r	This work
OPS0373	LMB777 [pOPS0126]; Str^r Tc^r	This work
OPS0469	Rlv3841 Δ*fixX*; Str^r	This work
OPS0487	Rlv3841 [pOPS0151]; Str^r Tc^r	This work
OPS0488	Rlv3841 [pOPS0152]; Str^r Tc^r	This work
OPS0489	Rlv3841 [pOPS0153]; Str^r Tc^r	This work
OPS0490	Rlv3841 [pOPS0154]; Str^r Tc^r	This work
OPS0491	Rlv3841 [pOPS0155]; Str^r Tc^r	This work
OPS0492	Rlv3841 [pOPS0156]; Str^r Tc^r	This work
OPS0493	Rlv3841 [pOPS0157]; Str^r Tc^r	This work
OPS0494	Rlv3841 [pOPS0158]; Str^r Tc^r	This work
OPS0495	Rlv3841 [pOPS0159]; Str^r Tc^r	This work
OPS0500	Rlv3841 *fdxB*::Ω; Str^r Spc^r	This work
OPS0536	OPS0469 *fdxB*::Ω; Str^r Spc^r	This work
OPS0651	LMB829 [pLMB826]; Str^r Spc^r Tc^r	This work
OPS0652	OPS0469 [pLMB826]; Str^r Tc^r	This work
OPS0706	Rlv3841 *ppk*::Ω; Str^r Tc^r	This work
OPS0707	LMB777 *ppk*::Ω; Str^r Tc^r	This work
OPS0709	Rlv3841 [pOPS0271]; Str^r Tc^r	This work
OPS0710	Rlv3841 [pOPS0272]; Str^r Tc^r	This work
OPS0711	OPS0280 [pOPS0271]; Str^r Tc^r	This work
OPS0712	OPS0280 [pOPS0272]; Str^r Tc^r	This work
OPS0713	Rlv3841 [pOPS0273]; Str^r Tc^r	This work
OPS0714	Rlv3841 [pOPS0274]; Str^r Tc^r	This work
OPS0715	OPS0469 [pOPS0273]; Str^r Tc^r	This work
OPS0716	OPS0469 [pOPS0274]; Str^r Tc^r	This work
OPS0739	LMB777 *phaC2*::Ω; Str^r Spc^r	This work

OPS0740	Rlv3841 [pOPS0300]; Str^r Tc^r	This work
OPS0741	Rlv3841 [pOPS0279]; Str^r Tc^r	This work
OPS0742	Rlv3841 [pOPS0301]; Str^r Tc^r	This work
OPS0744	LMB137 *ppk*::Ω; Str^r Tc^r	This work
OPS0746	Rlv3841 [pOPS0803]; Str^r Nm^r	This work
OPS0747	LMB829 [pOPS0803]; Str^r Spc^r Nm^r	This work
OPS0748	OPS0469 [pOPS0803]; Str^r Nm^r	This work
OPS0749	Rlv3841 [pOPS0325]; Str^r Nm^r	This work
OPS0770	LMB816 glgA::Tn*B60*; Str^r Nm^r	This work
OPS0771	OPS0744 glgA::Tn*B60*; Str^r Nm^r Tc^r	This work
OPS0796	Rlv3841 *dmtH*::pK19mob; Str^r Nm^r	This workRlv3841
RU1448	RlvA34 glgA::Tn*B60*; Nm^r	Lodwig, Leonard et al. (2005)
RU1478	RlvA34 glgA::Tn*B60 phaC*::Ω; Nm^r	Lodwig (2001)
RU1893	RU1978 [pRU1113]; Str^r Tc^r	Poole lab, unpublished
RU1978	Rlv3841 Δ*aapJ*; Str^r	Poole lab, unpublished
RU3940	Rlv3841 *nifH*::Ω; Str^r Spc^r	Karunakaran, Ramachandran et al. (2009)

Table 2.3 *Azorhizobium caulinondans* strains used in this work.

Strain	Description	Reference
ORS571	*Azorhizobium caulinodans* wild-type strain	Dreyfus and Dommergues (1981)
OPS0501	ORS571 *fixAB*::Ω; Str^r Spc^r	This work
OPS0614	ORS571 Δ*fixAB*; Str^r	This work
OPS0623	ORS571 [pOPS0124]; Tc^r	This work
OPS0625	OPS0614 [pOPS0124]; Tc^r	This work
OPS0626	ORS571 [pOPS0129]; Gm^r	This work
OPS0628	OPS0614 [pOPS0129]; Gm^r	This work
OPS0629	ORS571 [pOPS0131]; Gm_r	This work
OPS0631	OPS0614 [pOPS0131]; Gm^r	This work
OPS0846	ORS571 [pLMB779]; Gm_r	This work
OPS0847	OPS0614 [pLMB779]; Gm_r	This work

Table 2.4 *E. coli* strains used in this work.

Strain	Description	Reference
DH5alpha	F- *deoR endA1 recA1 relA1 gyrA96 hsdR17*(rk- mk+) *supE44 thi*-1 - *phoA* Δ(*lacZYA-argF*)U169 Φ80*lacZ*ΔM15 λ	Bioline
S17-1	Conjugative donor strain; *recA pro hsdR* RP4-2-Tc::Mu-Km::Tn7 integrated into the chromosome	ATCC47055
BTH101	Bacterial two-hybrid expression strain; *cya* mutant that requires plasmid-derived adenylate cyclase activity for the activation of the lac operon; F^-, *cya-99, araD139, galE15, galK16, rpsL1, hsdR2, mcrA1, mcrB1*; Str^r	Karimova, Pidoux et al. (1998)

2.2.2 Plasmids

All plasmids used in this thesis are listed in table 2.5.

Table 2.5 Plasmids used in this work.

Plasmid	Description	Reference
pIJ11268	Promoterless reporter vector containing luciferase genes; Tc^r	Frederix, Edwards et al. (2014)
pIJ11282	pIJ11268 with luciferase genes under control of the *nptII* promoter; Tc^r	Frederix, Edwards et al. (2014)
pJET1.2/blunt	Positive selection cloning vector derived from pUC19; Amp^r	Fermentas
pJP2	Stable broad-host-range cloning vector, pTR102 GUS with artificial MCS; Amp^r Tc^r	Prell, Boesten et al. (2002)
pHP45Ω-Km	pRB322 derivative carrying Ω interposon Km resistance cassette; Amp^r Spc^r	Fellay, Frey et al. (1987)
pHP45Ω-Spc	pRB322 derivative carrying Ω interposon Spc resistance cassette; Amp^r Spc^r	Fellay, Frey et al. (1987)
pHP45Ω-Tc	pRB322 derivative carrying Ω interposon Tc resistance cassette; Amp^r Spc^r	Fellay, Frey et al. (1987)
pJQ200SK	Suicide vector, pACYC derivative, P15A origin of replication; *lacZ sacB traJ* Gm^r	Quandt and Hynes (1993)
pK19mob	Mobilisable vector used for integration mutagenesis; pMB1 replicon, RP4 mob, l*acZ*α. Km^r Nm^r	Schafer, Tauch et al. (1994)
pL0V-C	Level 0 destination vector for Golden Gate cloning C module; Spc^r	Weber, Engler et al. (2011)
pL0V-PU	Level 0 destination vector for Golden Gate cloning PU module; Spc^r	Weber, Engler et al. (2011)
pL0V-SC	Level 0 destination vector for Golden Gate cloning SC module; Spc^r	Weber, Engler et al. (2011)
pL0V-SC1	Level 0 destination vector for Golden Gate cloning SC1 module; Spc^r	Weber, Engler et al. (2011)
pLMB509	Expression plasmid with a taurine-inducible promoter; Gm^r	Tett, Rudder et al. (2012)

pUT18C	C-terminal fusion of the target protein to the T18 fragment of adenylate cyclase; Amp^r	Karimova, Pidoux et al. (1998)
pUT18	N-terminal fusion of the target protein to the T18 fragment of adenylate cyclase; Amp^r	Karimova, Pidoux et al. (1998)
pKNT25	N-terminal fusion of the target protein to the T25 fragment of adenylate cyclase; Km^r	Karimova, Pidoux et al. (1998)
pKT25	C-terminal fusion of the target protein to the T25 fragment of adenylate cyclase; Chl^r	Karimova, Pidoux et al. (1998)
pRK2013	Helper plasmid; triparental conjugation; Km^r	Ditta, Stanfield et al. (1980)
pOPINF	High-throughput cloning vector for use in recombinant protein expression with an N-terminal His_6-tag; Amp^r	Berrow, Alderton et al. (2007)
pOPINS3C	High-throughput cloning vector for use in recombinant protein expression with an N-terminal His_6-*SUMO* tag; Amp^r	Bird (2011)
pRU1113	*aapJ*-FLAG cloned into pJP2 *XhoI/HindIII* site; Tc^r	Prell, Mulley et al. (2012)
pLMB721	pr1440/pr1441 PCR product of *fixAB* from Rlv3841 cloned into pJET1.2/blunt; Amp^r	This work
pLMB779	pr1507/pr1508 PCR product of Rlv3841 *fixAB* cloned into pLMB509; Gm^r	This work
pLMB780	pr1507/pr1509 PCR product of Rlv3841 *fixAB* cloned into pLMB509; Gm^r	This work
pLMB781	pr1507/pr1510 PCR product of Rlv3841 *fixAB* cloned into pLMB509; Gm^r	This work
pLMB782	pr1507/pr1511 PCR product of Rlv3841 *fixAB* cloned into pLMB509; Gm^r	This work
pLMB783	pr1507/pr1512 PCR product of Rlv3841 *fixAB* cloned into pLMB509; Gm^r	This work
pLMB788	pr1485/pr1486 PCR product of Rlv3841 *fixAB* cloned into pJET1.2/blunt; Amp^r	This work
pLMB789	pr1485/pr1487 PCR product of Rlv3841 *fixAB* cloned into pJET1.2/blunt; Amp^r	This work
pLMB790	pr1485/pr1488 PCR product of Rlv3841 *fixAB* cloned into pJET1.2/blunt; Amp^r	This work

pLMB791	pr1485/pr1489 PCR product of Rlv3841 *fixAB* cloned into pJET1.2/blunt; Amp^r	This work
pLMB800	pr1512/pr1513 PCR product of Rlv3841 *fixAB* cloned into pOPINF; Amp^r	This work
pLMB802	XbaI/BamHI fragment of pLMB790 ligated into pJP2; Tc^r	This work
pLMB809	XbaI/BamHI fragment of pLMB791 ligated into pJP2; Tc^r	This work
pLMB810	XbaI/BamHI fragment of pLMB789 ligated into pJP2; Tc^r	This work
pLMB811	XbaI/BamHI fragment of pLMB788 ligated into pJP2; Tc^r	This work
pLMB814	pr1424/pr1425 inverse PCR product of pLMB721 with *EcoRI* deleted *fixAB*; Amp^r	This work
pLMB815	pLMB815 containing the Ω interposon in *EcoRI* site; Amp^r, Spc^r	This work
pLMB817	*XbaI/BamHI* fragment of pLMB814 ligated into pJQ200SK; Gm^r	This work
pLMB818	*XbaI/BamHI* fragment of pLMB815 ligated into pJQ200SK; Gm^r Spc^r	This work
pLMB820	pr1620/pr1621 PCR product of Rlv3841 *fixABCX* cloned into pJET1.2/blunt; Amp^r	This work
pLMB822	pr1512/pr1513 PCR product of Rlv3841 *fixAB* cloned into pOPINS3C *KpnI/HindIII*; Amp^r	This work
pLMB826	*XbaI/BamHI* fragment of pLMB820 ligated into pJP2; Tc^r	This work
pLMB833	pr1642/pr1643 PCR product of Rlv3841 p*fixA* ligated into pIJ11268 *KpnI/BamHI*; Tc^r	This work
pLMB845	pr1626/pr1627 PCR product of pLMB721 with *EcoRI* deleted *fixB*, Amp^r	This work
pLMB846	pLMB845 containing the Ω interposon in *EcoRI* site; Amp^r Spc^r	This work
pLMB850	*XbaI/BamHI* fragment of pLMB846 ligated into pJQ200SK; Gm^r Spc^r	This work
pLMB856	pr1661/pr1662 PCR product of Rlv3841 *fixABCX*	This work

	cloned into pJET1.2/blunt; Amp^r	
pLMB857	pr1663/pr1664 PCR product of Rlv3841 *fixABCX* cloned into pJET1.2/blunt; Amp^r	This work
pLMB862	*XbaI/HindIII* fragment of pLMB856 ligated into pK19mob; Km^r	This work
pLMB863	*XbaI/HindIII* fragment of pLMB857 ligated into pK19mob; Km^r	This work
pLMB867	*XbaI/BamHI* fragment of pLMB845 ligated into pJQ200SK; Gm^r	This work
pLMB869	pr1701/pr1702 PCR product of Rlv3841 *fixC* cloned into pJET1.2/blunt; Amp^r	This work
pLMB870	pr1707/pr1708 PCR product of Rlv3841 *fixX* cloned into pJET1.2/blunt; Amp^r	This work
pLMB873	pr1709/pr1710 PCR product of pLMB870 with *EcoRI* deleted *fixX*; Amp^r	This work
pLMB875	pr1703/pr1704 PCR product of pLMB870 with *EcoRI* deleted *fixC*; Amp^r	This work
pLMB876	pLMB875 containing the Ω interposon in *EcoRI* site; Amp^r Spc^r	This work
pLMB877	*XbaI/BamHI* fragment of pLMB873 ligated into pJQ200SK; Gm^r	This work
pLMB878	*XbaI/BamHI* fragment of pLMB881 ligated into pJQ200SK; Gm^r Spc^r	This work
pLMB881	pLMB873 containing the Ω interposon in *EcoRI* site; Amp^r Spc^r	This work
pLMB882	*XbaI/BamHI* fragment of pLMB881 ligated into pJQ200SK; Gm^r Spc^r	This work
pOPS0116	*XbaI/BamHI* fragment of pLMB873 ligated into pJQ200SK; Gm^r	This work
pOPS0121	oxp0365/oxp0366 PCR product of Rlv3841 *fdxB* cloned into pJET1.2/blunt; Amp^r	This work
pOPS0122	oxp0367/oxp0368 PCR product of pLMB870 with *EcoRI* deleted *fdxB*; Amp^r	This work
pOPS0123	pOPS0122 containing the Ω interposon in *EcoRI* site; Amp^r Spc^r	This work

pOPS0124	oxp0355/oxp0356 PCR product of *fixAB* from Rlv3841 cloned into pJP2 *XbaI/BamHI* site; Tc^r	This work
pOPS0126	oxp0355/oxp0357 PCR product of *fixABC* from Rlv3841 cloned into pJP2 *XbaI/BamHI* site; Tc^r	This work
pOPS0129	oxp0259/oxp0199 PCR product of *fixAB* from Rlv3841 cloned by HD into pLMB509; Gm^r	This work
pOPS0130	oxp0259/oxp0198 PCR product of *fixAB* from Rlv3841 cloned by HD into pLMB509; Gm^r	This work
pOPS0131	oxp0196/oxp0258 PCR product of *fixAB* from Rlv3841 cloned by HD into pLMB509; Gm^r	This work
pOPS0132	oxp0196/oxp0243 PCR product of *fixAB* from Rlv3841 cloned by HD into pLMB509; Gm^r	This work
pOPS0134	*XbaI/BamHI* fragment of pOPS0123 ligated into pJQ200SK; Gm^r Spc^r	This work
pOPS0151	oxp0371/pr1643 PCR product of Rlv3841 p*fixA* ligated into pIJ11268 *KpnI/BamHI*; Tc^r	This work
pOPS0152	oxp0374/pr1643 PCR product of Rlv3841 p*fixA* ligated into pIJ11268 *KpnI/BamHI*; Tc^r	This work
pOPS0153	oxp0375/pr1643 PCR product of Rlv3841 p*fixA* ligated into pIJ11268 *KpnI/BamHI*; Tc^r	This work
pOPS0154	pr1642/pr1643 PCR product of Rlv3841 p*fixA* with distal NifA site deleted by overlap PCR using oxp0494/oxp0495 ligated into pIJ11268 *KpnI/BamHI*; Tc^r	This work
pOPS0155	pr1642/pr1643 PCR product of Rlv3841 p*fixA* with proximal NifA site deleted by overlap PCR using oxp0496/oxp0497 ligated into pIJ11268 *KpnI/BamHI*; Tc^r	This work
pOPS0156	pr1642/pr1643 PCR product of Rlv3841 p*fixA* with RpoN site deleted by overlap PCR using oxp0498/oxp0499 ligated into pIJ11268 *KpnI/BamHI*; Tc^r	This work
pOPS0157	oxp0382/oxp0385 PCR product of Rlv3841 p*nifA* ligated into pIJ11268 *KpnI/BamHI*; Tc^r	This work
pOPS0158	oxp0383/oxp0385 PCR product of Rlv3841 p*nifA* ligated into pIJ11268 *KpnI/BamHI*; Tc^r	This work

pOPS0159	oxp0384/oxp0385 PCR product of Rlv3841 p*nifA* ligated into pIJ11268 *KpnI/BamHI*; Tc^r	This work
pOPS0160	pr1642/oxp0385 PCR product of Rlv3841 p*nifA* ligated into pIJ11268 *KpnI/BamHI*; Tc^r	This work
pOPS0161	oxp0428/oxp0429 PCR product of ORS571*fixAB* cloned into pJET1.2/blunt; Amp^r	This work
pOPS0162	oxp0430/oxp0431 PCR product of pOPS0161 with *EcoRI* deleted *fixAB*; Amp^r	This work
pOPS0163	pOPS0162 containing the Ω interposon; Amp^r Spc^r	This work
pOPS0168	*XbaI/BamHI* fragment of pOPS0163 ligated into pJQ200SK; Gm^r Spc^r	This work
pOPS0169	*XbaI/BamHI* fragment of pOPS0162 ligated into pJQ200SK; Gm^r	This work
pOPS0204	oxp0500/oxp0501 PCR product of Rlv3841 *fixA* ligated into pKNT25 *KpnI/BamHI*; Km^r	This work
pOPS0205	oxp0502/oxp0503 PCR product of Rlv3841 *fixB* ligated into pKNT25 *KpnI/BamHI*; Km^r	This work
pOPS0206	oxp0504/oxp0505 PCR product of Rlv3841 *pdhA* ligated into pKNT25 *KpnI/BamHI*; Km^r	This work
pOPS0207	oxp0506/oxp0507 PCR product of Rlv3841 *pdhB* ligated into pKNT25 *KpnI/BamHI*; Km^r	This work
pOPS0208	oxp0508/oxp0509 PCR product of Rlv3841 *pdhC* ligated into pKNT25 *KpnI/BamHI*; Km^r	This work
pOPS0209	oxp0510/oxp0511 PCR product of Rlv3841 *lpdH* ligated into pKNT25 *KpnI/BamHI*; Km^r	This work
pOPS0210	oxp0500/oxp0501 PCR product of Rlv3841 *fixA* ligated into pUT18C *KpnI/BamHI*; Amp^r	This work
pOPS0211	oxp0502/oxp0503 PCR product of Rlv3841 *fixB* ligated into pUT18C *KpnI/BamHI*; Amp^r	This work
pOPS0212	oxp0504/oxp0505 PCR product of Rlv3841 *pdhA* ligated into pUT18C *KpnI/BamHI*; Amp^r	This work
pOPS0213	oxp0506/oxp0507 PCR product of Rlv3841 *pdhB* ligated into pUT18C *KpnI/BamHI*; Amp^r	This work
pOPS0214	oxp0508/oxp0509 PCR product of Rlv3841 *pdhC*	This work

	ligated into pUT18C *KpnI/BamHI*; Amp^r	
pOPS0215	oxp0510/oxp0511 PCR product of Rlv3841 *lpdH* ligated into pUT18C *KpnI/BamHI*; Amp^r	This work
pOPS0216	oxp0510/oxp0511 PCR product of Rlv3841 *fixB* ligated into pUT18 *KpnI/BamHI*; Amp^r	This work
pOPS0217	oxp0510/oxp0511 PCR product of Rlv3841 *lpdH* ligated into pUT18 *KpnI/BamHI*; Amp^r	This work
pOPS0218	oxp0502/oxp0503 PCR product of Rlv3841 *fixB* ligated into pKT25 *KpnI/BamHI*; Chl^r	This work
pOPS0219	oxp0502/oxp0503 PCR product of Rlv3841 *lpdH* ligated into pKT25 *KpnI/BamHI*; Chl^r	This work
pOPS0232	oxp0600/oxp0601 PCR product of Rlv3841 *dmtH* ligated into pJET1.2/blunt; Amp^r	This work
pOPS0233	oxp0600/oxp0602 PCR product of Rlv3841 *dmtH* ligated into pJET1.2/blunt; Amp^r	This work
pOPS0234	oxp0603/oxp0617 PCR product of Rlv3841 *ppk* cloned into pJET1.2/blunt; Amp^r	This work
pOPS0248	oxp0605/oxp0606 PCR product of pOPS0234 with *SmaI* deleted *ppk*; Amp^r	This work
pOPS0265	pOPS0248 containing the Ω interposon in *SmaI* site; Amp^r Tc^r	This work
pOPS0266	*XbaI/BamHI* fragment of pOPS0265 ligated into pJQ200SK; Gm^r Tc^r	This work
pOPS0249	Commercially synthesised region 1kb downstream of Rlv3841 *dmtH* in pUC57 (synthesis by GeneWhizz); Amp^r	This work
pOPS0250	Golden gate assembly of Rlv3841 *dmtH::ΩKm* mutagenesis plasmid in pOGG028 backbone; Gm^r Km^r	This work
pOPS0251	Golden gate assembly of Δ*dmtH* mutagenesis plasmid in pOGG028 backbone; Gm^r	This work
pOPS0271	Golden gate assembly expressing Rlv3841 *fixC* under an *nptII* promoter with pharmacia terminator in pOGG041 backbone; Tc^r	This work
pOPS0272	Golden gate assembly expressing Rlv3841 *fixC* with C-terminal 3xFLAG *tag* under an *nptII*	This work

	promoter with pharmacia terminator in pOGG041 backbone; Tc^r	
pOPS0273	Golden gate assembly expressing Rlv3841 *fixX* under an *nptII* promoter with pharmacia terminator in pOGG041 backbone; Tc^r	This work
pOPS0274	Golden gate assembly expressing Rlv3841 *fixX* with C-terminal 3xFLAG tag under an *nptII* promoter with pharmacia terminator in pOGG041 backbone; Tc^r	This work
pOPS0275	pOPS0250 with ΩKm cassette removed and ΩSpc cassette ligated using *NotI*; Spc^r	This work
pOPS0276	pOPS0250 with ΩKm cassette removed and ΩTc cassette ligated using *NotI*; Tc^r	This work
pOPS0277	Golden gate assembly expressing Rlv3841 *fixC* under an *nptII* promoter with pharmacia terminator in pOGG026 backbone; Nm^r	This work
pOPS0278	Golden gate assembly expressing Rlv3841 *fixCX* under an *nptII* promoter with pharmacia terminator in pOGG026 backbone; Nm^r	This work
pOPS0279	oxp0792/oxp0385 PCR product of Rlv3841 p*nifA* ligated into pIJ11268 *KpnI/BamHI*; Tc^r	This work
pOPS0300	oxp0791/oxp0385 PCR product of Rlv3841 p*nifA* ligated into pIJ11268 *KpnI/BamHI*; Tc^r	This work
pOPS0301	oxp0382/oxp0385 PCR product of OPS0469 p*nifA* ligated into pIJ11268 *KpnI/BamHI*; Tc^r	This work
pOPS0302	oxp0801/oxp0808 PCR product of Rlv3841 *fixXnifAB* with *BsaI* sites removed by Gibson assembly with oxp0802-807 ligated into pJET1.2/blunt; Amp^r	This work
pOPS0303	Golden Gate assembly of oxp0793/oxp0794 PCR product of pOPS0302 under p*fixA* promoter with pharmacia terminator in pOGG026 backbone; Nm^r	This work
pOPS0325	Golden Gate assmbly of GFP under p*fixA* promoter with pharmacia terminator in pOGG026 backbone; Nm^r	This work
pOPS0347	oxp1204/oxp1205 PCR product of Rlv3841 *dmtH* ligated into pK19mob; Km^r	This work

pOGG001	*NptII* promoter cloned into pL0V-PU by Golden Gate cloning; Spc^r	ENSA
pOGG003	*Pharmacia* terminator cloned into pL0V-T by Golden Gate cloning; Spc^r	ENSA
pOGG026	Minimal version of pJP2 constructed as a level 1 Golden Gate cloning vector; Nm^r	Geddes et al., in production
pOGG028	pJQ200SK constructed as a level 1 Golden Gate cloning vector domesticated of all Type IIS restriction sites; Gm^r	Geddes et al., in production
pOGG037	ΩKm cassette designed for use in a Cre/Lox system in a Golden Gate level 0 cloning vector; Km^r	Poole lab, unpublished
pOGG041	pJP2 constructed as a level 1 Golden Gate cloning vector domesticated of all Type IIS restriction sites; Tc^r	Geddes et al., in production
pOGG035	3xFLAG tag cloned into pL0V-C2 by Golden Gate cloning; Spc^r	ENSA
pOGG095	*pfixA* cloned into pL0V-PU by Golden Gate cloning; Spc^r	This work

2.2.3. Primers

Primers used in this thesis are described in Table 2.6.

Table 2.6 Primers used in this work.

Primer	Sequence	Description
M13 uni (-21)	TGTAAAACGACGGCCAGT	Sequencing/mapping primer for *lacZ*-containing vectors (pK19mob, pJQ200SK, pLMB509)
M13 rev (-29)	CAGGAAACAGCTATGACC	Sequencing/mapping primer; *lacZ*-containing vectors (pK19mob, pJQ200SK, pLMB509)
pJET1.2 for	GACTCACTATAGGGAGAGCGGC	Sequencing/mapping primer; pJET1.2
pJET1.2 rev	AAGAACATCGATTTTCCATGGCAG	Sequencing/mapping primer; pJET1.2
pOT forward far	GACCTTTTGAATGACCTTTA	Mapping primer; interposon mutagenesis
pLMB509 for	CGCCCAACTGGACTCATCTAACTTC	Sequencing/mapping primer; pLMB509
pLMB509 rev	CGCAGTCGGCCTATTGGTTAAAA	Sequencing/mapping primer; pLMB509
pIJ11268 for	CCATCTTTGCCCTACCGTAT	Sequencing/mapping primer; pIJ11268
pIJ11268 rev	AAACCGACGCCATCACCCAG	Sequencing/mapping primer; pIJ11268
pOPIN for	CCAGCCACCACCTTCTGATA	Sequencing/mapping primer; pOPIN
pOPIN rev	CCAGCCACCACCTTCTGATA	Sequencing/mapping primer; pOPIN
pr0096	TCGTAAATGCTGGACCCGATGG	Sequencing/mapping primer; pJP2
p611	GCGATCCAGACTGAATGCCC	Sequencing/mapping primer; pJP2
oxp0715	CAAACCGCCTCTCCCCG	Sequencing/mapping primer; pL0V
oxp0716	AAAGTGCCACCTGACGTCT	Sequencing/mapping primer; pL0V

pr1440	TTGTAGACATGTCGGCAACC	Sense primer for amplification of Rlv3841 *fixAB* region plus *XbaI* site
pr1441	AGAGAGATTGCGCTTCGTCA	Antisense primer for amplification of Rlv3841 *fixAB* region plus *BamHI* site
pr1485	TCTAGATTGTCGGCAACCCTTC TGCCTG	Sense primer for amplification of Rlv3841 *fixAB* including promoter region plus *XbaI* site
pr1486	GGATCCTTCTTACCCTTGGTCA TCTAGCCTTC	Antisense primer for amplification of Rlv3841 *fixAB* including promoter region plus *BamHI* site
pr1487	GGATCCTGGAGCCATCCGCAGT TTGAAAAATGATTCCCTTGGTC ATCTAGCCTTC	Antisense primer for amplification of Rlv3841 *fixAB* plus single Strep tag plus *BamHI* site
pr1488	GGATCCCTTGTCGTCGTCATCC TTGTAGTCTGATTCCCTTGGTC ATCTAGCCTTC	Antisense primer for amplification of Rlv3841 *fixAB* plus single FLAG tag plus *BamHI* site
pr1489	GGATCCCATCATCATCATCATC ATTGATTCCCTTGGTCATCTAG CCTTC	Antisense primer for amplification of Rlv3841 *fixAB* plus His_6 tag plus *BamHI* site
pr1506	AGGAGGAAGAACATATGATGC ACATCGTGGTCTGTAT	Sense primer for amplification of Rlv3841 *fixAB* for HD cloning into pLMB509
pr1507	TGGTGATGATGCATATGCCCTT GGTCATCTAGCCTTC	Antisense primer for amplification of Rlv3841 *fixAB* for HD cloning into pLMB509
pr1508	TGGTGATGATGCATATGTGGAG CCATCCGCAGTTTGAAAAACCC TTGGTCATCTAGCCTTC	Antisense primer for amplification of Rlv3841 *fixAB* minus stop codon plus single Strep tag for HD cloning into pLMB509
pr1509	TGGTGATGATGCATATGTGGAG CCATCCGCAGTTTGAAAAATGA CCCTTGGTCATCTAGCCTTCC	Antisense primer for amplification of Rlv3841 *fixAB* plus single Strep tag for HD cloning into pLMB509 plus pLMB509
pr1510	TGGTGATGATGCATATGCTTGT CGTCGTCATCCTTGTAGTCCCC TTGGTCATCTAGCCTTC	Antisense primer for amplification of Rlv3841 *fixAB* minus stop plus single FLAG tag for HD cloning into pLMB509
pr1511	TGGTGATGATGCATATGCTTGT	Antisense primer for amplification

	CGTCGTCATCCTTGTAGTCTGACCCTTGGTCATCTAGCCTTC	of Rlv3841 *fixAB* plus single FLAG tag for HD cloning into pLMB509
pr1512	AAGTTCTGTTTCAGGGCCCGATGCACATCGTGGTCTGTATCAAACAAG	Antisense primer for amplification of Rlv3841 *fixAB* for HD cloning into the POPIN vectors
pr1523	ATGGTCTAGAAAGCTTTATTACCCTTGGTCATCTAGCCTTCCTCG	Sense primer for amplification of Rlv3841 *fixAB* for HD cloning into the POPIN vectors
pr1607	TTTGACCGTTCTAACGATGGGA	Sense primer for amplification of internal region of Rlv3841 *fixA*
pr1608	TTAGGCCGCATTTGGTCAACTC	Antisense primer for amplification of internal region of Rlv3841 *fixA*
pr1609	TTTGGTAGAGTCGCCCCTCCTC	Sense primer for amplification of internal region of Rlv3841 *fixB*
pr1610	TTCGTCGTTGCCAGGTTTCTCA	Antisense primer for amplification of internal region of Rlv3841 *fixB*
pr1620	TTTCTAGACATGTCGGCAACCCTTCTGC	Sense primer for amplification of Rlv3841 *fixABCX* including upstream IGR plus *XbaI* site
pr1621	TTGGATCCCTCATCCGAATTTGAAGAGG	Antisense primer for amplification of Rlv3841 *fixABCX* plus *BamHI* site
pr1622	TTTTTGAATTCATGGACCGAGGAAGGCTAGA	Sense inverse PCR primer for deletion of Rlv3841 *fixAB* plus *EcoRI* site
pr1623	TTTTTGAATTCCACTTGTTTGATACAGACCA	Antisense inverse PCR primer for deletion of Rlv3841 *fixAB* plus *EcoRI* site
pr1626	TTTTTGAATTCTTGGACCGAGGAAGGCTAGATG	Sense inverse PCR primer for deletion of Rlv3841 *fixB* plus *EcoRI* site
pr1627	TTTTTGAATTCTTAGTTCCATGAAGACCCAGAC	Antisense inverse PCR primer for deletion of Rlv3841 *fixB* plus *EcoRI* site
pr1637	GGATTATTCGGCGTTGAAA	Sense primer for mapping PCR of Rlv3841 *fixAB*
pr1638	CAGATGCCGTTCAAGAGGA	Antisense primer for mapping PCR

		of Rlv3841 *fixAB*
pr1642	TTGGTACCCGAGAAAAGGCTATTCGAAT	Sense primer for amplification of p*fixA* plus *KpnI* site
pr1643	TTGGATCCTGGGTTAATGATGGTCGGTA	Antisense primer for amplification of p*fixA* plus *BamHI* site
pr1659	ACAATTGACGGCGACACGGC	Sense primer for mapping PCR of Rlv3841 *fixB*
pr1660	CGGCTCGAGTTCCAAGCAGT	Antisense primer for mapping PCR of Rlv3841 *fixB*
pr1661	TTTCTAGAGAACTGAAGCCAAACCGGTA	Sense primer for amplification of internal region of Rlv3841 *fixC* plus *XbaI* site
pr1662	TTAAGCTTATCGAGGGATGTCGCTTGAA	Antisense primer for amplification of internal region of Rlv3841 *fixC* plus *HindIII* site
pr1663	TTTCTAGAGATGAAGGCGACCATCATTG	Sense primer for amplification of internal region of Rlv3841 *fixX* plus *XbaI* site
pr1664	TTAAGCTTCTCACCGAATTTGAAGAGG	Antisense primer for amplification of internal region of Rlv3841 *fixX* plus *HindIII* site
pr1667	TTAATGCGGCCTACGAGGATCG	Sense primer for mapping PCR of Rlv3841 *fixB*
pr1668	TTGATGCCGTTCAAGAGGAGCGTC	Antisense primer for mapping PCR of Rlv3841 *fixB*
pr1701	AATGGATCCTTAAGTGTTTCGAAGATCAC	Sense primer for amplification of Rlv3841 *fixC* region plus *BamHI* site
pr1702	AAAGGATCCTAGCTAATACCCTCTTGAGC	Antisense primer for amplification of Rlv3841 *fixC* region plus *BamHI* site
pr1703	TTTGAATTCGTTGCCGGACATACCGGCGC	Sense inverse PCR primer for deletion of Rlv3841 *fixC* plus *EcoRI* site
pr1704	TTTGAATTCGTCCGCTCCGCCGTATCTTG	Antisense inverse PCR primer for deletion of Rlv3841 *fixC* plus *EcoRI* site

pr1707	TTTGGATCCGCGCCCAGTTTGACAAGTGG	Sense primer for amplification of Rlv3841 *fixX* region plus *XbaI* site
pr1708	TTTGGATCCTGGCCAGTTCGAAACGTCCA	Antisense primer for amplification of Rlv3841 *fixX* region plus *BamHI* site
pr1709	TTTGAATTCCATCGAATTTTCCTTTAACG	Sense inverse PCR primer for deletion of Rlv3841 *fixX* plus *EcoRI* site
pr1710	TTTGAATTCGTCCTCTTCAAATTCGGATG	Antisense inverse PCR primer for deletion of Rlv3841 *fixX* plus *EcoRI* site
pr1711	AGGCTCTAAAAATGTTCAAGGG	Sense primer for mapping PCR of Rlv3841 *fixX*
pr1712	ATAGAGGTCGGCTCTGAACT	Antisense primer for mapping PCR of Rlv3841 *fixX*
oxp0196	AGGAGGAAGAACATATGATGGACTACAAGGACCACGACGGTGACTACAAGGACCACGACATCGACTACAAGGACGACGACGACAAGCACATCGTGGTCTGTAT	Sense primer for amplification of Rlv3841 *fixAB* plus 3xFLAG tag for HD cloning into pLMB509
oxp0198	TGGTGATGATGCATATGCTTGTCGTCGTCGTCCTTGTAGTCGATGTCGTGGTCCTTGTAGTCACCGTCGTGGTCCTTGTAGTCCCCTTGGTCATCTAGCCTTC	Antisense primer for HD cloning of Rlv3841 *fixAB* into pLMB509 plus 3xFLAG tag
oxp0199	TGGTGATGATGCATATGTCACTTGTCGTCGTCGTCCTTGTAGTCGATGTCGTGGTCCTTGTAGTCACCGTCGTGGTCCTTGTAGTCCCCTTGGTCATCTAGCCTTC	Antisense primer for amplification of Rlv3841 *fixAB* minus stop codon plus 3xFLAG tag for HD cloning into pLMB509
oxp0243	TCATTTTTCGAACTGCGGGTGGCTCCAAGCGCTACCTCCCGATCCACCTCCGGAACCTCCACCTTTTTCGAACTGCGGGTGGCTCCAAGCGCTCCCTTGGTCATCTAGCCTTC	Antisense primer for amplification of Rlv3841 *fixAB* plus twin-Strep tag for HD cloning into pLMB509
oxp0258	AGGAGGAAGAACATATGATGAGCGCTTGGAGCCACCCGCAGTTCGAAAAAGGTGGAGGTTCCGGAGGTGGATCGGGAGGTAGCGCTTGGAGCCACCCGCAGTTCGAA	Antisense primer for amplification of Rlv3841 *fixAB* minus stop codon plus twin-Strep tag for HD cloning into pLMB509

	AAACACATCGTGGTCTGTAT	
oxp0259	TGGTGATGATGCATATGTTTTT CGAACTGCGGGTGGCTCCAAG CGCTACCTCCCGATCCACCTCC GGAACCTCCACCTTTTTCGAAC TGCGGGTGGCTCCAAGCGCTCC CTTGGTCATCTAGCCTTC	Sense primer for amplification of Rlv3841 *fixAB* plus twin-Strep tag for HD cloning into pLMB509
oxp0336	TTATCTGGCGCGGTCCAACATC	Sense primer for mapping PCR of Rlv3841 *fixC*
oxp0337	TTGGGAACAATCGCCACCTGAC	Antisense primer for mapping PCR of Rlv3841 *fixC*
oxp0338	TTCACACTGACAGTCACAACTT	Sense primer for mapping PCR of Rlv3841 *fixX*
oxp0339	TTAATCATGCCTACAAACAGGG	Antisense primer for mapping PCR of Rlv3841 *fixX*
oxp0355	TTTCTAGAGTCGGCAACCCTTC TGCC	Sense primer for amplification of Rlv3841 *fixAB* including promoter region plus *XbaI* site
oxp0356	TTGGATCCCTTGTCGTCGTCGT CCTTGTAGTCGATGTCGTGGTC CTTGTAGTCACCGTCGTGGTCC TTGTAGTCCCCTTGGTCATCTA GCCTTCC	Antisense primer for amplification of Rlv3841 *fixAB* plus 3xFLAG tag plus *BamHI* site
oxp0357	TTGGATCCCTTGTCGTCGTCGT CCTTGTAGTCGATGTCGTGGTC CTTGTAGTCACCGTCGTGGTCC TTGTAGTCACGCCAAGATACGG CGGAGC	Antisense primer for amplification of Rlv3841 *fixABC* plus 3xFLAG tag plus *BamHI* site
oxp0358	TTGGATCCTTTTTCGAACTGCG GGTGGCTCCAAGCGCTACCTCC CGATCCACCTCCGGAACCTCCA CCTTTTTCGAACTGCGGGTGGC TCCAAGCGCTCCCTTGGTCATC TAGCCTTCC	Antisense primer for amplification of Rlv3841 *fixAB* plus twin-Strep tag plus *BamHI* site
oxp0359	TTGGATCCTTTTTCGAACTGCG GGTGGCTCCAAGCGCTACCTCC CGATCCACCTCCGGAACCTCCA CCTTTTTCGAACTGCGGGTGGC TCCAAGCGCTACGCCAAGATAC GGCGGAGC	Antisense primer for amplification of Rlv3841 *fixABC* plus twin-Strep tag plus *BamHI* site

oxp0365	TTTCTAGAAAGAGGTGGGCGG AAAATGC	Sense primer for amplification of Rlv3841 *fdxB* region plus *XbaI* site
oxp0366	TTGGATCCTCTCGGTGCGACGA TGCA	Antisense primer for amplification of Rlv3841 *fdxB* region plus *BamHI* site
oxp0367	TTTTTGAATTCCCCGTCGCGAG TCACGGATG	Sense inverse PCR primer for deletion of Rlv3841 *fdxB* plus *EcoRI* site
oxp0368	TTTTTGAATTCGTCCATCATCGT TCATTGCACG	Antisense inverse PCR primer for deletion of Rlv3841 *fdxB* plus *EcoRI* site
oxp0374	TTGGTACCCCCGGGGGATTATT CGGC	Sense primer for amplification of Rlv3841 p*fixA* plus *KpnI* site
oxp0375	TTGGTACCGTGATCGGGGCGGA TTGC	Sense primer for amplification of Rlv3841 p*fixA* plus *KpnI* site
oxp0382	TTGGTACCCTAGATGACCAAGG GTAAGTTCGA	Sense primer for amplification of p*nifA* plus *KpnI* site
oxp0383	TTGGTACCTCGATGAAGGCGAC CATCATT	Sense primer for amplification of p*nifA* plus *KpnI* site
oxp0384	TTGGTACCGGAGTCCCTACCTC CGGC	Sense primer for amplification of p*nifA* plus *KpnI* site
oxp0385	TTGGATCCCGTTCATCGCAGCC TTCAGG	Antisense primer for amplification of p*nifA* plus *BamHI* site
oxp0428	TTTCTAGATCTATGACATCGAC GCCAGC	Sense primer for amplification of ORS571 *fixAB* region plus *XbaI* site
oxp0429	TTGGATCCGATGGACGGCATTG TTGAGC	Antisense primer for amplification of ORS571 *fixAB* region plus *BamHI* site
oxp0430	TTTTTGAATTCTGCAGACGACG ATGTGCATT	Sense inverse PCR primer for deletion of ORS571 *fixAB* plus *EcoRI* site
oxp0431	TTTTTGAATTCGCGACCGGATC GCCAGCTGA	Antisense inverse PCR primer for deletion of ORS571 *fixAB* plus *EcoRI* site
oxp0432	CCCTTCGCGTTTCAGGGC	Sense primer for mapping PCR of ORS571 *fixAB*

oxp0433	CATCCAGAAGCGCTGTTCCA	Antisense primer for mapping PCR of ORS571 *fixAB*
oxp0494	CATTGTCATCACCTTAGCGTTC CTGTTCTCAATCG	Sense primer for overlap PCR deletion of Rlv3841 p*fixA* distal NifA binding site
oxp0495	GAGAACAGGAACGCTAAGGTG ATGACAATGTCGCA	Antisense primer for overlap PCR deletion of Rlv3841 p*fixA* distal NifA binding site
oxp0496	TATGCGAAGGGGTTTTGTCGTG TCGCAATCCGCCC	Sense primer for overlap PCR deletion of Rlv3841 p*fixA* proximal NifA binding site
oxp0497	GATTGCGACACGACAAAACCC CTTCGCATAGCGAC	Antisense primer for overlap PCR deletion of Rlv3841 p*fixA* proximal NifA binding site
oxp0498	ACCCGGCTAACTCTGATCAGGT TGCGCCGGTAAGC	Sense primer for overlap PCR deletion of Rlv3841 p*fixA* RpoN binding site
oxp0499	CCGGCGCAACCTGATCAGAGTT AGCCGGGTCCCTA	Antisense primer for overlap PCR deletion of Rlv3841 p*fixA* RpoN binding site
oxp0500	TTTGGATCCCATGCACATCGTG GTCTGTATC	Sense primer for amplification of Rlv3841 *fixA* plus *BamHI* site for BTH
oxp0501	TTTGGTACCCGCCGATTGCCAT GGGACGTGA	Antisense primer for amplification of Rlv3841 *fixA* plus *KpnI* site for BTH
oxp0502	TTGGATCCCTGGTCGCTAGAAA AAATGAA	Sense primer for amplification of Rlv3841 *fixB* plus *BamHI* site for BTH
oxp0503	TTTGGTACCCGCCCTTGGTCAT CTAGCCTTC	Antisense primer for amplification of Rlv3841 *fixB* plus *KpnI* site for BTH
oxp0504	TTTTCTAGACATGGCGCCGCGA AAGACCGC	Sense primer for Rlv3841 *pdhA* plus *BamHI* site for BTH
oxp0505	TTTGGTACCCGGAGCAGAATGT CGGTATAGA	Antisense primer for amplification of Rlv3841 *pdhA* plus *KpnI* site for BTH
oxp0506	TTGGATCCCATGCCTATCGATA	Sense primer for amplification of Rlv3841 *pdhB* plus *BamHI* site for

	TCCTCAT	BTH
oxp0507	TTTGGTACCCGTTTGTAGCAAA CAGCCTTCA	Antisense primer for amplification of Rlv3841 *pdhB* plus *KpnI* site for BTH
oxp0508	TTGGATCCCATGCCGATCAATA TCACGATGCC	Sense primer for amplification of Rlv3841 *pdhC* plus *BamHI* site for BTH
oxp0509	TTTGGTACCCGGACAAGCATGC CCATCGGGT	Antisense primer for amplification of Rlv3841 *pdhC* plus *KpnI* site for BTH
oxp0510	TTGGATCCCATGGCTGAATCCT ACGACGT	Sense primer for amplification of Rlv3841 *lpdH* plus *BamHI* site for BTH
oxp0511	TTTGGTACCCGAGCGTTCAGCA CGCGGCCGT	Antisense primer for amplification of Rlv3841 *lpdH* plus *KpnI* site for BTH
oxp0570	ATCCAGAGGACAAGCTTGGC	Sense primer for mapping PCR of Rlv3841 *fdxB*
oxp0571	GCCATGCTTGACGGACATTC	Antisense primer for mapping PCR of Rlv3841 *fdxB*
oxp0600	TTTTGGTCTCAGGAGCGATGAA CCTCACGA	Sense primer for Golden Gate cloning of Rlv3841 *dmtH* left flanking region
oxp0601	TTTTGGTCTCTTTATCATCAAAT CACCTTCTG	Antisense primer for amplification of Rlv3841 *dmtH* left flanking region for Golden Gate cloning
oxp0603	TTTCTAGATGTTTGCGCTCTTTG CCTTC	Sense primer for amplification of Rlv3841 *ppk* region plus *BamHI* site
oxp0617	TTTCTAGATCGCACTCAACACC TCCTTC	Antisense primer for Rlv3841 *ppk* region plus *BamHI* site
oxp0618	AATCACCCGGTCTCTTGCTG	Sense primer for mapping PCR of Rlv3841 *ppk*
oxp0619	GAACGGGACATGCCTTCGTA	Antisense primer for mapping PCR of Rlv3841 *ppk*
oxp0620	TTCCCGGGTCTGCGACTGCGCT ATCCAT	Sense inverse PCR primer for deletion of Rlv3841 *ppk* plus *SmaI*

		site
oxp0621	TTCCCGGGAAGGCCGCAACAA CAAGTAA	Antisense inverse PCR primer for deletion of Rlv3841 *ppk* plus *SmaI* site
oxp0670	TTTTGAAGACAAAGGTATGACC AAGGGTAAGTTCGACG	Sense primer for amplification of Rlv3841 *fixC* for Golden Gate cloning in pL0V-C
oxp0671	TTTTGAAGACAAAATGATGACC AAGGGTAAGTTCGACG	Sense primer for amplification of Rlv3841 *fixC* for Golden Gate cloning in pL0V-SC
oxp0672	TTTTGAAGACAAAGGTATGAA GGCGACCATCATTGAG	Sense primer for amplification of Rlv3841 *fixX* for Golden Gate cloning in pL0V-C
oxp0673	TTTTGAAGACAAAATGATGAA GGCGACCATCATTGAG	Sense primer for amplification of Rlv3841 *fixX*for Golden Gate cloning in pL0V-SC
oxp0674	TTTTGAAGACAACACCACGCCA AGATACGGC	Antisense primer for amplification of Rlv3841 *fixC* for Golden Gate cloning in pL0V-SC1
oxp0675	TTTTGAAGACAAAAGCACGCC AAGATACGGC	Antisense primer for amplification of Rlv3841 *fixC* for Golden Gate cloning in pL0V-SC
oxp0676	TTTTGAAGACAACACCTCCGAA TTTGAAGAGGACCCC	Antisense primer for amplification of Rlv3841 *fixX* for Golden Gate cloning in pL0V-SC1
oxp0677	TTTTGAAGACAAAAGCTCCGAA TTTGAAGAGGACCCC	Antisense primer for amplification of Rlv3841 *fixX*for Golden Gate cloning in pL0V-SC
oxp0756	TCAAAAATATGGTATTGATAAT CCTGA	Sense primer for mapping PCR of Rlv3841 *dmtH*
oxp0757	TTTAATCGCGGCCTCGAGC	Antisense primer for mapping PCR of Rlv3841 *dmtH*
oxp0791	TTTGGTACCGGTCGCCAAGCCC GAAC	Sense primer for amplification of Rlv3841 p*nifA* plus *KpnI* site
oxp0792	TTTGGTACCCTTGGAATGCGGC ACATGC	Sense primer for p*nifA* plus *KpnI* site
oxp0793	TTTTCTAGAATGAAGGCGACCA	Sense primer for amplification of Rlv3841 Rlv3841 *fixX* plus *XbaI*

	TCATTGAG	site
oxp0794	TTTGGATCCTCACTCCTTCTTCACATCGA	Antisense primer for amplification of Rlv3841 *nifA* plus *BamHI* site
oxp0795	TTTTGGATCCTCAATTAGAGGGGCCTAAAGC	Antisense primer for Rlv3841 *nifB* plus *BamHI* site
oxp0801	ATGAAGGCGACCATCATTGAG	Sense primer for Gibson Assembly of Rlv3841 *fixXnifAB* domesticated of *BsaI* sites - fragment 1
oxp0802	CGGATTCCAGAACGGTTTCCGACAGCGC	Antisense primer for Gibson Assembly of Rlv3841 *fixXnifAB* domesticated of *BsaI* sites - fragment 1
oxp0803	GGAAACCGTTCTGGAATCCGAATTGTTTGGC	Sense primer for Gibson Assembly of Rlv3841 *fixXnifAB* domesticated of *BsaI* sites - fragment 2
oxp0804	TCACGCTCAGTCAGACCCGGTGCCTCCA	Antisense primer for Gibson Assembly of Rlv3841 *fixXnifAB* domesticated of *BsaI* sites - fragment 2
oxp0805	CCGGGTCTGACTGAGCGTGATCGACTGATC	Sense primer for Gibson Assembly of Rlv3841 *fixXnifAB* domesticated of *BsaI* sites - fragment 3
oxp0806	TCAATTAGAGGGGCCTAAAG	Antisense primer for Gibson Assembly of Rlv3841 *fixXnifAB* domesticated of *BsaI* sites - fragment 3
oxp0807	TGAATGCGCCGCGATCTCTGATCCATCG	Sense primer for Gibson Assembly of Rlv3841 f*ixXnifAB* domesticated of *BsaI* sites - fragment 4
oxp0808	CAGAGATCGCGGCGCATTCATGCACAATG	Antisense primer for Gibson Assembly of Rlv3841 *fixXnifAB* domesticated of *BsaI* sites - fragment 4
oxp0819	CACTCTGTGGTCTCAAATGATGAAGGCGACCATCATTGAG	Sense primer for amplification of Rlv3841 *fixXnifAB* plus *BsaI* site for Golden Gate cloning
oxp0820	CACTTCGTGGTCTCAAAGCGCGGTCACTCCTTCTTCACA	Antisense primer for amplification of Rlv3841 *fixXnifA* plus *BsaI* site for Golden Gate cloning

oxp0821	CACTTCGTGGTCTCAAAGCTCA ATTAGAGGGGCCTAAAGC	Antisense primer for amplification of Rlv3841 *fixXnifAB* plus *BsaI* site for Golden Gate cloning
oxp0822	TTTTGAAGACAAGGAGGTGATC GGGGCGGATTGC	Sense primer for amplification of Rlv3841 p*fixA* in pL0V-PU for Golden Gate cloning
oxp0823	TTTTGAAGACAACATTCGTCGT TTAAGCTCCCAATCG	Antisense primer for amplification of Rlv3841 p*fixA* in pL0V-PU for Golden Gate cloning
oxp1204	TGATTACGCCAAGCTCGATCTC AAGGTGCAGGC	Sense primer for amplification of Rlv3841 *dmtH* for HD cloning into pK19mob
oxp1205	GCAGGCATGCAAGCTTCGGCGT AGGAATCGGA	Antisense primer for amplification of Rlv3841 *dmtH* for HD cloning into pK19mob

2.2.4 Bacteriophage

Bacteriophage RL38 (Buchanan-Wollaston 1979) was used to transduce mutations between *R. leguminosarum* bv. *viciae* strains.

2.3 Molecular techniques

2.3.1 DNA isolation

Genomic DNA was isolated from *R. leguminosarum* and *A. caulinodans* cultures using DNeasy Blood and Tissue Kit (Qiagen), following the manufacturers protocol. This involved pre-treatment for Gram-negative bacteria, using $1x10^9$ cells per isolation.

Plasmid DNA was extracted from *E. coli* using either the GeneJET Plasmid Minirep Kit (Thermo Scientific) or QIAprep Spin Miniprep Kit (Qiagen). Both kits were used following the manufacturer protocols.

When DNA purification was required following restriction digest or PCR either GeneJET PCR Purification kit (Thermo Scientific) or QIAquick PCR Purification kit (Qiagen) was used, according to manufacturers protocols. DNA isolation from electrophoresis gels was carried out using GENEJet Gel Extraction Kit (Thermo Scientific) or QIAquick Gel Extraction Kit (Qiagen) according to manufacturers protocols. Isopropanol was used to improve efficiency of purification of small products.

2.3.2 DNA amplification by PCR

PCR primers were designed using Vector NTI 11.0 or Geneious R8. Primers for Gibson assembly were designed using NEBuilder (NEB). Primers for HD (formerly BD) cloning were designed using the ClonTech InFusion™ online tool. Primers were synthesised by Eurofins MWG Operon.

Mapping/screening PCR reactions were carried out in 10 µL reactions using GoTaq® Green Master Mix (Promega). Amplification PCR reactions were carried out using in 25 µL or 50 µL reactions using Phusion® High-fidelity PCR Master Mix (Thermo Scientific). When difficult-to-amplify or high-GC regions of DNA were to be amplified a Phusion® reaction was set up using GC buffer with DMSO included in the reaction. Thermocycler conditions were set according to the manufacturers instructions for polymerases, with T_m determined by Vector NTI 11.0 or Geneious R8. PCR amplification for cloning or sequencing was carried

out using genomic DNA at 5 pmol. Screening of *E. coli* transformants was carried out using colony PCR, where a single colony was transferred into the reaction using a sterilie pin or pipette tip. Where large-scale screening of rhizobial colonies was required colonies were transferred into 10 µL sterile H_2O using a sterile pipette tip, and this was heated to 60°C for five minutes. The resulting suspension (1 µL) was used in the PCR reaction.

When custom DNA sequences were required synthesis was carried out by GeneWiz (Takeley, UK).

2.3.3 DNA Gel Electrophoresis

GeneRuler 1 kb or 100 bp (Thermo Scientific) was used as a marker ladder for DNA electrophoresis. Electrophoresis was used to separate PCR products and restriction digests. Electrophoresis was carried out in 0.9% agarose (Sigma Aldrich) in TAE buffer (400 mM Tris acetate, 1 mM EDTA). Gels were run at 120 mV for 20-30 minutes. Agarose percentage was raised for better resolution when isolating small DNA fragments, and running time was altered accordingly. DNA was stained using either ethidium bromide or Sybr® Safe (Invitrogen). Gels were incubated in ethidium bromide (0.5 µg mL^{-1}) after electrophoresis for fifteen minutes and visualized using a UV transilluminator. Sybr® Safe was added to molten agar (1:10000) before gels were cast. Stained agarose gels were visualized using the GelDoc EZ System (BioRad).

2.3.4 Restriction digest

Restriction enzyme digests were carried out on purified DNA using restriction endonucleases and respective buffers (Roche, Thermo Scientific or NEB) following manufacturers instructions. Fragmented DNA was separated by gel electrophoresis, and when required DNA was purified from the restriction reaction or agarose gels.

2.3.5 DNA ligation

DNA ligation was carried out using T4 DNA Ligase and 10X T4 DNA Ligase Buffer (Thermo Scientific). Reactions were carried out according the manufacturers protocol. Ligations into the pJET1.2/blunt vector were carried out at room temperature for 30 minutes. All other ligations were carried out at room temperature overnight.

2.3.6 HD cloning

HD cloning (previously known as BD cloning) was carried out using InFusion® HD Cloning Kit (ClonTech). Primers were designed with a 15bp overhang corresponding to the ends of the destination vector, designed using the online Primer Design Tool provided by ClonTech and tested *in silico* using Geneious. HD cloning was carried out in a thermocycler according to the manufacturers protocols.

2.3.7 Golden Gate cloning

Golden Gate cloning uses Type IIS restriction enzymes *BsaI* and *BpiI* to allow direct cloning of several fragments into a defined order. Type IIS restriction sites cleave at a defined distance from their recognition site, leaving a specific overhang. These overhangs can be designed as part of primer and vector design and so fragments can be efficiently cloned in a chosen order. This also allows all fragments to be assembled in a single reaction. A modular cloning (MoClo) system has subsequently been developed with a standardized set of overhangs for parts used in the Golden Gate system. This relies on a level-based system. Level 0 carries the individual parts required for an assembly, such as promoters or coding sequences. Level 0 modules can then be assembled into a transcriptional unit, known as a Level 1 module. These Level 1 modules can be assembled into a vector and used for expression. Alternatively, multiple level 1 modules can be combined together in a bector to form a Level 2 multigene construct (Weber, Engler et al. 2011). The MoClo system is illustrated in Figure 2.1.

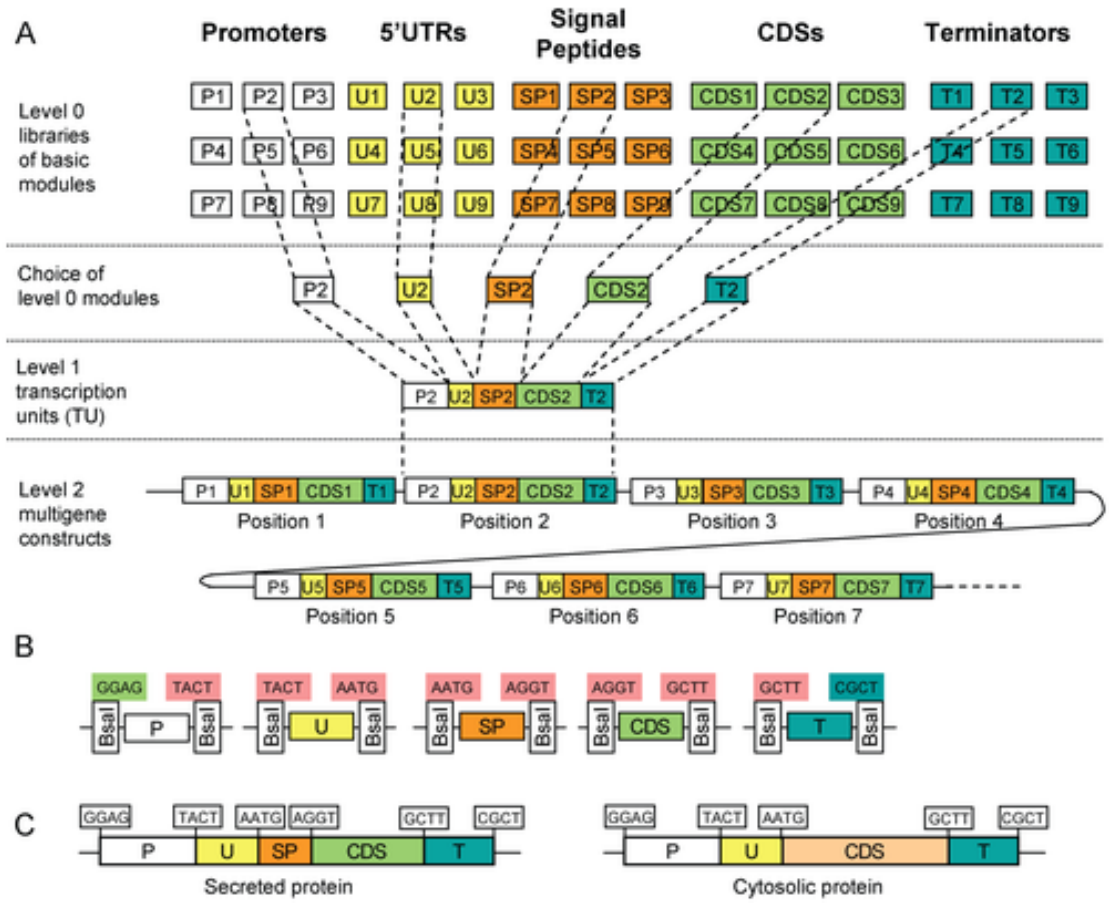


Figure 2.1 Overview of the modular cloning system used in Golden Gate assembly. A: Level 0 modules are assembled to make a transcriptional unit (level 1). These level 1 units can then be assembled into multigene units (level 2); B: Each module has a unique overhang; C: These different modules can then be used to design different transcriptional unit e.g. with and without secretion tags. Reproduced from Weber, Engler et al. (2011).

A set of standardised Level 0 vectors are available for cloning parts of interest. Once a Level 0 module has been constructed it can be used in a single reaction to assemble a transcriptional unit. Parts are assembled into Level 0 vectors using Type IIS enzyme *BpiI* (Thermo Scientific). Each Level 0 vector contains recognition sites for Type IIS enzyme *BsaI*, along with a standardised overhang. A Level 1 assembly reaction can be carried out with *BsaI* (Thermo Scientific) in order to assemble the DNA parts in the desired order.

Golden Gate reactions were carried out using 1 µL of the desired Type IIS enzyme, 1 µL T4 DNA ligase, 1.5 µL DNA ligase buffer (Thermo Scientific), 1.5 µL BSA (2 mg mL^{-1}) and 40 fmol of the required parts and vector backbone, made up to 15 µL with H_2O. A series of vector backbones based on conventional cloning vectors have now been developed in the Poole lab (Geddes et al., in production). These vectors are still being optimised and tested.

If DNA fragments of interest contain recognition sites for Type IIS enzymes, these sites must be 'domesticated'. Domestication involves changing nucleotides whilst conserving codons. This may be carried out by overlap PCR or Gibson Assembly. In cases where domestication is difficult, DNA fragments may be synthesised commercially.

2.3.8 Overlap PCR

Overlap Extension PCR can be used to direct mutagenesis within DNA regions. This may be used for domestication, or for targeted alterations or deletions within DNA fragments used in conventional cloning. Overlap PCR primers are designed using complementary 15 bp primer overhangs. These overhangs can be altered to allow deletion or mutation. All overlap PCR primers were designed using Geneious R8.

2.3.9 Gibson Assembly

Gibson Assembly was used to domesticate Type IIS restriction sites from regions of DNA used in Golden Gate cloning. This cloning strategy assembles multiple regions of DNA in a single Gibson reaction, according to regions of homology designed as primer overlaps. Domesticated DNA regions were designed *in silico* and primers for Gibson Assembly were subsequently designed against these new regions using NEBuilder (NEB). Gibson cloning was carried out using Gibson Assembly® Cloning Kit (NEB). Fragments assembled by Gibson cloning were subsequently cloned into pJET1.2/blunt.

2.3.10 Transformation

Chemically competent *E. coli* DH5α cells were used for transformations. These were either obtained commercially or generated in the laboratory. DH5α was inoculated in 500 mL LB for 3 hours. Once the culture reached OD_{600} 0.3-0.4 cells were incubated on ice for 10 minutes before being pelleted at 4000 rpm at 4°C for 10 minutes. The cell pellet was resuspended in 10 mL ice cold 0.1 M $CaCl_2$ 15% glycerol (v/v) and incubated on ice for 15-30 minutes. Suspensions were pelleted at 4°C 4000 rpm for 10 minutes and $CaCl_2$ solution decanted. Pellets were resuspended in 20 mL ice cold 0.1 M $CaCl_2$ 15% glycerol and incubated at 4°C for 30 minutes. Aliquots (50-200 µL) were dispensed into chilled Eppendorf tubes and snap frozen in liquid nitrogen. Competent cells were stored at -80°C. S17.1 and BTH101 competent cells were generated using the same treatment. For Golden

Gate cloning and Gibson Assembly commercial Gold standard competent cells (Bioline) were used for high efficiency.

For transformation into *E. coli* 50 µL of cells were thawed on ice and 5 µL of plasmid or ligation reaction added. In the case of Golden Gate reactions 1 µL of reaction was added to 10 µL of competent cells. Cells were incubated on ice for 30 minutes before being heat-shocked at 42°C for 1 minute. Cells were incubated on ice for a further two minutes. SOC medium (400 µL) (Melford) was added, and cells shaken at 37°C for 1 hour at 200 rpm. Cells were spread-plated on LB agar containing appropriate antibiotics and plates incubated overnight at 37°C.

2.3.11 Conjugation from *E. coli* to *R. leguminosarum*

Triparental mating was used to conjugate chromosomal integration plasmids from *E. coli* into *R. leguminosarum* bv. *viciae* 3841. Helper plasmid pRK2013 (Ditta, Stanfield et al. 1980) was used in conjugations. For low efficiency conjugations biparental mating was used. Plasmids were transformed into competent S17.1 *E. coli* cells, and pRK2013 was not used. Three days prior to conjugation the recipient *Rhizobium* strain was grown on a slope containing relevant antibiotics. One day prior to conjugation, the donor *E. coli* strain containing the plasmid and *E. coli* strain containing pRK2013 were inoculated in 10 mL LB containing appropriate antibiotics. Overnight cultures were subcultured 500 µL into 5 mL fresh LB containing appropriate antibiotics, and grown for 3-4 hours at 37°C shaking at 100 rpm until OD_{600} 0.4-0.6. *E. coli* strains were pelleted at 4000 rpm for 10 minutes and gently washed with TY medium three times in order to remove traces of antibiotics. The rhizobial slope was washed with 3-5 mL of TY. Recipient rhizobia (400 µL) was mixed with 400 µL donor *E. coli* and 200 µL pRK2013. This suspension was centrifuged at 6000 rpm for 5 minutes. The resultant pellet was resuspended in 30 µL SOC medium (Melford) and put on a sterile nitrocellulose filter on a TY plate, incubated at 28°C overnight. A sterile loop was then used to streak bacteria from the filter onto a TY plate containing appropriate antibiotics, and incubated for 3-4 days at 28°C. Streptomycin at 500 µg mL^{-1} was used to select against *E. coli*.

Conjugation of constructs for expression and complementation was carried out using a patch mating protocol. A sterile loop was used to mix single colonies from the recipient rhizobial strain, donor *E. coli* strain and pRK2013 on a TY plate. The plate was incubated at 28°C overnight. A sterile loop was used to streak from the plate onto TY contaning with appropriate selection antibiotics.

2.3.12 Conjugation from *E. coli* to *A. caulinodans*

Triparental mating using pRK2013 was used for conjugation into *A. caulinodans*. All conjugations were carried out using patch mating due to a failure of *A. caulinodans* to grow normally on nitrocellulose filters. Nitrofurantoin (5 μg mL^{-1}) was used to select against *E. coli.*

2.4 Mutagenesis techniques

2.4.1 Mutagenesis by pK19mob-integration (single cross-over)

Generation of integration mutants using pK19mob (Schafer, Tauch et al. 1994) allows a high-throughput production of mutants. However, there is potential for instability using the pK19mob system, and so this was used for a faster, initial check of mutant phenotypes.

The pK19mob vector was digested using *HindIII* and purified. A 300 bp internal fragment of the gene of interest was amplified using PCR. PCR primers were designed to add a 15bp overhang with homology to the digested pK19mob vector. InFusion® HD cloning (2.3.6) was used to clone the internal gene fragment into the digested pK19mob vector. The product of the InFusion® reaction was transformed into chemically competent DH5α, and plated on LB containing kanamycin and X-gal to allow blue-white selection. Correct colonies were confirmed using restriction digest and sequencing (Eurofins MWG Operon). M13 uni (-21) and M13 rev (-29) primers were used for the sequencing reaction.

Once correct colonies had been obtained, the correct pK19mob plasmid was conjugated into *R. leguminosarum* bv. *viciae* 3841 (2.3.11). Conjugants were selected on TY containing 500 μg mL^{-1} streptomycin and 250 μg mL^{-1} neomycin. Genomic DNA was made from conjugants, and the insertion confirmed by sequencing (MWG Operon) using M13 *uni* (-21) and M13 *rev* (-29).

2.4.2 Mutagenesis by omega interposon insertion (double cross-over)

When a stable mutant was required, a strategy was used to insert an Ω cassette into the gene of interest. Depending on the background strain, Ω interposon plasmids were designed containing either ΩSpc, ΩTc or ΩKm. A suite of Ω cassettes allow for appropriate selection during construction of double and triple mutants. Once an Ω interposon mutant has been

generated, the same strategy can be used to create an in-frame deletion mutant by conjugation into the omega background strain.

Targeted mutagenesis was carried out using suicide vector pJQ200SK (Gm^r) (Quandt and Hynes 1993). The sequence of the gene of interest plus 1kb flanking regions was amplified using primers designed with restriction sites to allow cloning into the pJQ200SK MCS (typically *XbaI/BamHI)*. The resulting PCR product was cloned into pJET1.2/blunt (Amp^r) (CloneJET^{TM}). Inverse PCR of the resultant plasmid was carried out using primers facing outwards from the edge of the desired gene. The inverse PCR primers were designed with *EcoRI* restriction site overhangs, and designed to maintain codons, allowing an in-frame deletion of the gene of interest. The resultant inverse PCR product was digested using *EcoRI.* The omega cassette was excised from vector pHP45Ω (Fellay, Frey et al. 1987) using *EcoRI.* The *EcoRI* sticky ends were used to clone the omega cassette into the inverse PCR product. Additionally, the *EcoRI* sticky ends were ligated together with no cassette to create a second plasmid containing a clean (Δ) deletion. The region of interest containing the omega cassette was cloned into pJQ200SK using *BamHI/XbaI*. The same reaction was used to clone the in-frame version into pJQ200SK.

The cassette-containing plasmids was conjugated into *R. leguminosarum* bv. *viciae* 3841 or relevant background strain using triparental filter-mating (2.3.11). A single recombination event was selected for using gentamycin. These colonies were then grown on TY slopes, and then a serial dilution plated on UMS plates containing appropriate antibiotics, 10 mM NH_4Cl and 15% sucrose to select for double recombinants. The *sacB* gene in pJQ200SK means that the suicide vector will be lost when grown on sucrose, and so only chromosomal integrations of the Ω cassette will be maintained on selective media. Successful colonies were patched on gentamycin and a replicate plate containing the relevant Ω antibiotic in order to confirm loss of pJQ200SK. Genomic DNA was extracted from putative mutants (2.3.1) and mutations were confirmed by PCR using primers designed to the region, and the pOT forward_far primers (bind within the interposon cassette), and correct PCR products sent for sequencing. The pJQ200SK clone containing the in-frame deletion was subsequently conjugated into the Ω mutant by the same conjugation procedure to result in in-frame deletion. Primers, plasmids and resultant mutant strains are outlined in section 2.2.

2.4.3 Phage transduction between *R. leguminosarum* strains

Bacteriophage RL38 (Buchanan-Wollaston 1979) was used to transduce mutations containing antibiotic resistance markers between *R. leguminosarum* strains. Phage

transduction was used to create double and triple mutants once single mutants had been obtained.

The donor strain (containing the mutation to be transferred) was grown on a TY slope containing appropriate antibiotics at 28°C for 3 days. The slope was diluted in 3-5 mL of sterile H_2O. Aliquots of 100 µL were added to a serial dilution of phage RL38 stock (10^0-10^{-6}), along with appropriate no-phage and no-bacteria controls. These dilutions were added to melted TY agar in a 1:1 mix with TY broth, and this poured onto a TY agar plate. Plates were incubated at 28°C until a phage lawn formed. Once a lawn has formed, plates were incubated in 10 mL sterile H_2O at room temperature with gentle rocking. After two hours, the eluted phage solution was filtered through a 0.22 µm filter (Millipore) and 4 µL chloroform added to remove any bacteria that was able to pass through the filter. The phage solution was then stored at 4°C.

The recipient strain for the transduction was then grown on a TY slope containing appropriate antibiotics for three days at 28°C. The slope was diluted in 3-5 mL TY broth. A serial dilution of bacterial suspension (200 µL) with phage (0.1-100 µL) was incubated at 28°C for one hour before plating on TY plates containing appropriate antibiotics. Single colonies were obtained after 3-5 days, and checked using colony PCR.

2.5 Plant Experiments

2.5.1 Growth of *P. sativum*

P. sativum cv. Aveola seeds were sterilised in 95% alcohol for 30-60 seconds and immediately washed in sterile H_2O. Seeds were then immersed in 2% sodium hypochlorite for 5 minutes. Seeds were then washed 5 times in sterile H_2O. Seeds were transferred to a sterile flask and washed 5 more times in sterile H_2O to remove any traces of sodium hypochlorite.

Pots (1 L) were filled with medium vermiculite, and 400 mL of a nitrogen-free rooting solution was added to each pot. Nitrogen-free rooting solution comprised 4 mM Na_2HPO_4, 3.7 mM K_2PO_4, 1 mM $CaCl_2$, 800 µM $MgSO_4$, 100 µM KCl, 35 µM H_3BO_3, 10 µM Fe EDTA, 9 µM $MnCl_2$, 0.8 µM $ZnCl_2$, 0.5 µM Na_2MoO_4, 0.3 µM $CuSO_4$. Pots were autoclaved prior to planting.

Strains for inoculations were grown on TY slopes containing appropriate antibiotics. Slopes were diluted with 3-5 mL of sterile water, and diluted 1 mL in 100 mL of sterile H_2O. Two

seeds were planted per pot, and each seed inoculated with 1 mL of the diluted bacterial suspension. Pots were covered with cling-film to reduce bacterial contamination. After 3-6 days a sterile scalpel was used to cut holes in the cling-film for emerging shoots, and sterile forceps used to remove one of the two emerging seedlings to result in one pea per pot. All plants were harvested at 21-28 days post-inoculation. Whole-plant images were taken by John Baker (Department of Plant Sciences, Oxford) or Andrew Davis (JIC). Nodules were photographed using a Leica M165 FC stereo microscope with a Leica DFC310 FX digital camera.

2.5.2 Acetylene reduction assay

An acetylene reduction assay was used to determine the activity of nitrogenase. This assay takes advantage of the promiscuity of the nitrogenase enzyme, which is able to break triple bonds in several chemical species. Nitrogenase breaks the triple carbon-carbon bond in acetylene to result in ethylene. Reaction 2.1 demonstrates the reaction carried out by nitrogenase on acetylene.

$$HC{\equiv}CH + 2H^{+} + 2e^{-} \rightarrow CH_2{=}CH_2$$

Reaction 2.1 Conversion of acetylene to ethylene by nitrogenase.

Plants were placed in 250 mL Duran bottles (total volume 320 mL) with a small piece of wet tissue to prevent drying out during the assay. Airtight neoprene lids were used on the Duran bottles to allow for injection whilst not breaking the gas seal. Air (8 mL) was removed from each bottle using a syringe and needle, and replaced by 6.4 mL acetylene to result in 2% acetylene per bottle. Bottles were incubated at room temperature for 1 hour. After an hour 1 mL samples of the gas inside the bottle were taken. Needles were left in rubber bungs to prevent air escaping from the syringes during sampling. Each 1 mL sample was run through a Clarus® 480 (Perkin-Elmer) in order to determine acetylene and ethylene levels in the bottles, and this was converted to a rate of nitrogenase activity $plant^{-1}$ $hour^{-1}$.

2.5.3 Nodule collection and re-isolation of nodule bacteria

After plants were harvested, roots were gently washed to remove vermiculite. All nodules were then picked by hand and collected in Falcon tubes for use in future assays if required. Nodules were stored at -80°C

Undifferentiated bacteria can be isolated from nodules in order to check the identity of the strain and confirm plasmid retention. Nodules (5-10 per plant) were collected and sterilised in 2% sodium hypochlorite for 2 minutes, and subsequently washed 5-10 times with sterile H_2O. Each nodule was individually crushed using a small spatula sterilised with 70% EtOH. The resulting brei was streaked onto TY plates using a sterile loop. Nodule streaks were incubated at 28°C for 3-4 days until single colonies grew. These single colonies were patched onto relevant antibiotics to determine if the correct markers were retained. In the case of pK19mob mutants or unstable plasmids, ratios of resistance to sensitivity can be used to determine if the phenotype was due to plasmid or mutation loss.

2.5.4 Isolation of bacteroids

In order to carry out biochemical investigation on bacteroids, they must be isolated from plant nodules. Harvested nodules were thawed (if required) and then crushed using a pestle and mortar in 2 mL isolation buffer per 1 g nodules (8 mM K_2HPO_4, 2 μM KH_2PO_4, 30mM sucrose, 2 mM $MgCl_2$ made up in sterile H_2O). The resulting brei was filtered through two layers of muslin. Brei was pelleted at 1000 rpm for 5 minutes to pellet heavier plant materials. The supernatant from this spin was then pelleted at 6000 rpm for 5 minutes to pellet the bacteroids. Isolated bacteroids were stored at -80°C.

Where nodule extract was required, the crushed nodule brei was filtered through two layers of muslin and then boiled at 100°C for 10 minutes to kill any bacteroids and break open cells. The sample was then filtered through a 22 μm filter to remove cell debris for use in assays.

2.6 Assays

2.6.1 Growth of *A. caulinodans* in modified UMS media

Growth of *A. caulinodans* was determined under both nitrogen-fixing and non-nitrogen fixing conditions. Under non-fixing conditions, UMS media was supplemented with 20 mM

succinate, 10 mM NH_4Cl, 30 μM nicotinic acid and vitamin solution (2.1.1). *A. caulinodans* strains were grown on TY ampicillin slopes for two days at 37°C. Slopes were resuspended in 3-5 mL UMS. The bacterial suspension was pelleted at 4000 rpm and washed 3-5 times in UMS in order to remove traces of rich media. UMS (50 mL) was inoculated with a suspension of the washed bacteria to an OD_{600} 0.01. Cultures were grown at 37°C at 200 rpm. Optical density was recorded over a 24-48 hour period.

Growth under nitrogen-fixing conditions was carried out in a controlled environment cabinet (Belle). The cabinet was flushed with nitrogen until oxygen was measured as 3%. A heat bar was used to set the cabinet to 37°C. Cultures were grown in baffled conical flasks, and a fish pump was used to bubble gas through the cultures. Cultures were shaken at 200 rpm. Optical density was measured over a 24-48 hour period.

Samples were taken at several optical densities to determine protein concentration and dry weight per OD_{600}. Protein concentration was determined by bicinchoninic acid assay and used to standardise subsequent calculations.

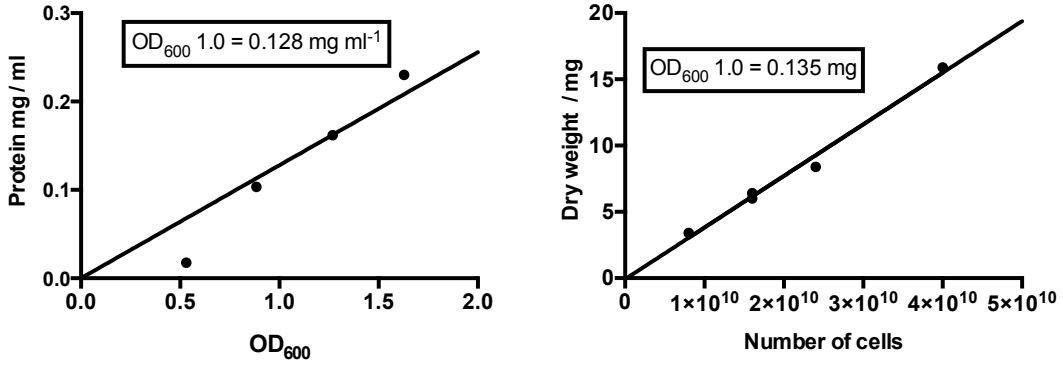


Figure 2.2 Standard curves of protein and dry weight for *A. caulinodans* ORS571.

2.6.2 Growth of *R. leguminosarum* in rich media in a plate reader

Growth curves of *R. leguminosarum* were obtained using a FLUOStar Omega Microplate Reader (BMG). *R. leguminosarum* was grown on TY slopes plus appropriate antibiotics for 3 days. Slopes were washed with 3-5 mL TY and diluted to OD_{600} 0.01 (where OD_{600} 1.0 = $1x10^9$ cfu). Cultures (400 μL per well) were grown in black 24-well plates (Vision Plate™ 24, 4titude) with transparent, gas permeable seals (Gas Permeable Moisture Barrier Seal, 4titude). Plates were shaken at 400 rpm double orbital shaking. OD_{595} was measured every 30 minutes for 24-48 hours. Data was processed using MARS (BMG) data analysis software.

Where required, an Atmospheric Control Unit (BMG) was used to control gas levels. Carbon dioxide levels were maintained at 0.1%. Oxygen levels were brought down to 1% for microaerobic growth of *R. leguminosarum* by purging with nitrogen.

Growth of *A. caulinodans* was attempted in the plate reader. However, cultures flocculated and grew poorly under all conditions, including growth at 3-10% O_2. *A. caulinodans* grew better under orbital shaking than double orbital, but still had some flocculation.

2.6.3 Measuring luciferase activity in *R. leguminosarum*

A luciferase-based reporter system was described by Frederix, Edwards et al. (2014). This system allows *in vivo* measurement promoter activity by detection of luciferase expression. A promoter or reporter region of interest is cloned upstream of the *luxCDABE* operon. The pJP2 vector is used as backbone for this system, allowing stability *in planta*. Promoter regions of interest were amplified by PCR from genomic DNA. *KpnI/BamHI* sites were added to the ends of the PCR product, and the product was cloned into the MCS ahead of the *lux* operon. Primers pIJ11268 fw and pIJ11268 rv were used to confirm constructs by PCR mapping and subsequently by sequencing. Correctly constructed *lux* plasmids were conjugated into *R. leguminosarum* bv. *viciae* 3841 by patch conjugation (2.3.11). Plasmid pIJ11268 in the WT background (LMB542) was used as an empty vector control. Plasmid pIJ11282 comprising the *lux* cassette under a constitutive *nptII* promoter in the wild-type Rlv3841 background (D5250) was used as a positive control.

Growth on agar plates was measured in a NightOwl II camera (Berthold Technologies). Plates were imaged with a 20 second exposure. To determine promoter activity *in planta*, strains were inoculated on plants. Plants were harvested at 21 days and washed to remove all vermiculite. Plants were imaged with a 120 second exposure. Expression in nodules was quantified by manual selection of 20 nodules per plant, and expression in luminescence counts per second (cps) per nodule. Alternative quantification methods were measurement of the entire image or manual selection of the entire plant root system, and measurement of cps mm^{-2}. Absolute values varied between quantification method, however the pattern of expression between strains was maintained independent of method used.

Quantification of luminescence in free-living cultures was determined using a FLUOStar Omega Plate Reader. Cultures were grown as described in section 2.6.2. A script was used to measure both OD_{600} and luminescence in relative light units (RLU). Maximum rlu was determined for each strain and expressed as RLU OD_{595}^{-1}.

2.6.4 Acetylene reduction assay of *A. caulinodans* under nitrogen fixing conditions

A. caulinodans was grown under nitrogen-fixing conditions, as described in 2.6.1. After 24 hours, when cultures reached OD_{600} ~0.3-0.5, 10 mL samples of culture were aliquoted (in triplicate) into universal tubes within the oxygen cabinet at 3% O_2. Gas impermeable silicon lids were used to allow injection into the universal tubes whilst maintaining gas composition. Sealed tubes were removed from the low-oxygen cabinet. Air (0.67 mL) was removed from each universal tube and 0.56 mL acetylene was added to create a 2% acetylene concentration. Tubes were incubated at 37°C at 200 rpm shaking for two hours. Gas samples (1 mL) were removed and run through a Clarus® 480 (Perkin-Elmer) to determine conversion of acetylene to ethylene. Nitrogenase activity was expressed as μmol ethylene produced per mg protein per hour.

2.6.5 Determination of protein concentration

Two protocols were used for determination of protein concentration. Qubit quantification was used to determine protein concentration for Western blot analysis. The Qubit™ Protein Quantification Kit (Thermo Scientfic) was used according to manufacturers protocol.

For quantification to express biochemical values per mg protein, a bicinchoninic acid assay was used. Sample (10-25 μL) was made up to 25 μL in sterile water. When required, samples were further diluted to obtain a concentration 0.2-10 mg mL^{-1} to fall within the linear range of the assay. NaOH (25 μL 2N) was added to each tube, and samples incubated at 90-100°C for 10 minutes. Cu_2SO_4 (20 μL of 4% stock solution) and 1 mL bicinchoninic acid was added to each sample. Samples were vortexed thoroughly and incubated at 60 °C for 30 minutes. Samples were then cooled, and absorbance measured at 562 nm. BSA 0-1 mg mL^{-1} was used to construct a standard curve.

2.6.6 Polyphosphate quantification

DAPI fluorescence was used to quantify polyphosphate in bacteroids (Kulakova, Hobbs et al. 2011). Bacteroids harvested from 0.5 g nodules was pelleted at 6000 rpm for 10 minutes. The pellet was snap-frozen in liquid nitrogen and defrosted at room temperature. The pellet was then resuspended in 500 μL DAPI assay buffer (150 mM KCl, 20 mM HEPES-KOH pH 7, DAPI solution 10 μM) and incubated for 10 minutes at room temperature. Fluorescence was measured in black 96-well plates (Greiner bio-one) at an excitation wavelength of 390

nm and emission wavelength of 535 nm. Due to issues acquiring a polyphosphate standard (sodium phosphate glass, Sigma Aldrich) values were expressed as relative fluorescence and not absolute values of polyphosphate.

2.6.7 PHB quantification

Polyhydroxybutyrate (PHB) levels in *R. leguminosarum* were quantified as described by Lodwig, Leonard et al. (2005). Bacteroid pellets from 1 g of nodules were resuspended in 1 mL technical grade Na hypochlorite and incubated for 37°C for 1 hour. The sample was then spun at 13000 rpm for 30 minutes. The pellet was sequentially washed with 1 mL H_2O, 1mL acetone and 1 mL ethanol. Between washes the sample was centrifuged for 5 minutes at 13000 rpm. The polymer was then extracted from the pellet in 100 µL boiling chloroform at 61°C. The chloroform was evaporated and the extraction step repeated twice. H_2SO_4 (5 mL) was added and heated in a water bath for 20 minutes at 100°C. Once the reaction cooled absorbance was measured against a sulphuric acid blank at 235 nm.

2.6.8 Glycogen quantification

Glycogen was extracted from bacteroids and quantified using Anthrone reagent (Chun and Yin 1998). The protocol was optimised for *R. leguminosarum* by Lodwig, Leonard et al. (2005). Cells were resuspended to OD_{600} 0.1 in H_2O. KOH (200 µL 30% w/v) was added to 50 µL of cells and incubated at 100°C for 20 minutes to lyse cells. Once cooled, 600 µL 100% ethanol was added to precipitate glycogen. Samples were pelleted at 4000 g for 15 minutes and the pellet resuspended in 500 µL H_2O. Sulphuric acid (1 mL 75%) was added to digest glycogen. Anthrone reagent (2 mL) (Sigma Aldrich) was added and incubated at 100°C for 10 minutes before cooling on ice. Once cooled absorbance was measured at 620 nm. Glycogen was quantified against a standard curve created using glycogen 0-200 µg.

2.6.9 Differential centrifugation to obtain membrane fractions

Where membrane fractions were required for assays a differential ultracentrifugation was used. Bacteroid pellets were defrosted on ice and resuspended in 1-2 mL 40 mM HEPES pH 7.2 plus 2 mM DTT and 20% glycerol. Cells were then lysed in a ribolyser at speed 6.0 for 30 seconds and then incubated on ice to cool. Samples were centrifuged at 1000 rpm to pellet the ribolyser beads and the supernatant pipetted into an Eppendorf tube. Cell extracts were

spun at 15000 rpm for 30 minutes to pellet the cytoplasmic fraction. The remaining supernatant was then spun in an ultracentrifuge at 50000 rpm for 2 hours at 5°C to pellet the membrane fraction. The supernatant was discarded and the remaining pellet resuspended in 100-200 µL PBS plus protease inhibitor (Roche, one tablet per 10 mL PBS) to give a membrane fraction.

2.7 Protein biochemistry

2.7.1 Bacterial two-hybrid assay

A bacterial two-hybrid (BTH) assay as described by Karimova, Pidoux et al. (1998) was used to test interaction between *R. leguminosarum* proteins. All strains and plasmids were obtained from Euromedex, and protocols were based on the manufacturers protocol (Bacterial Adenylate Cyclase-based Two Hybrid System).

The bacterial two-hybrid system is an *in vivo* system for determining protein-protein interaction, based on reconstitution of the adenylate cyclase system in *E. coli*. Vectors are used containing T18 and T25 fragments of adenylate cyclase from *Bordetella pertussis*, corresponding to amino acids 1-224 and 225-399 respectively. Genes of interest are expressed in these vectors, creating a recombinant protein displaying the T18 or T25 peptide on either the N- or C-terminus. These fragments are not functional when physically separated, but once interaction is established are able to form a functional catalytic unit of adenylate cyclase. Functional adenylate cyclase produces cyclic-AMP (cAMP), which interacts with catabolite activator protein (CAP). This cAMP/CAP complex forms a transcriptional activator able to activate several host genes, including the *lac* and *mal* operons (Figure 2.3). This means that lactose and maltose can then be used as selective media to indicate the level of protein interaction. Activation of the *lac* operon allows for β-galactosidase activity as a determinant of interaction in blue-white selection on X-gal-containing media. Additionally, a β-galactosidase assay allows quantitative determination of interactions.

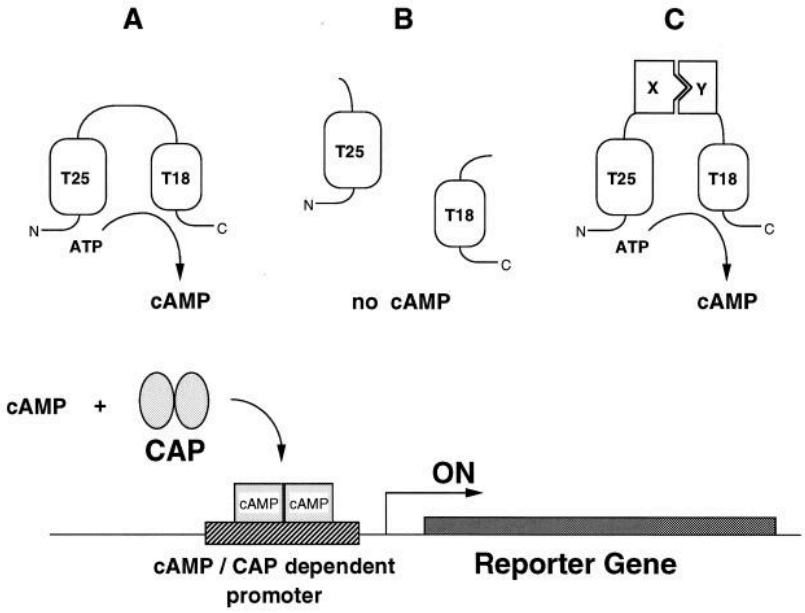


Figure 2.3 Diagram of the BACTH system (Karimova, Pidoux et al. 1998). A: The T18 and T25 subunits interact to form a functional adenylate cyclase, able to produce cAMP; B: When T25 and T18 are distant within the cell, no cAMP is produced; C: When two recombinant proteins containing T25 and T18 interact, the subunits can interact to produce cAMP; cAMP works with CAP to activate a cAMP/CAP-dependent promoter, expressing reporter genes which can be used to determine protein-protein interactions.

Genes of interest are cloned into vectors pUT18, pUT18C, pKT25 or pKNT25 (described in section 2.2). pUT18 is derived from high-copy number plasmid pUC19, with an MCS at the N-terminal end of T18. pUT18C differs from pUT18 that the MCS is found at the C-terminal end of T18. pKT25 is a low-copy number plasmid derived from pSU40, with the MCS at the C-terminal end of T25. pKNT25 has the MCS at the N-terminal end of T25. The gene of interest was amplified by PCR from genomic DNA and cloned into the MCS using *KpnI/BamHI*.

E. coli strain BTH101 was used as the host for testing protein-protein interactions. BTH101 lacks a functional *cya* gene which encodes adenylate cyclase. Competent cells were generated according to the protocol outlined in 2.3.10. The high copy number T18 clone (1 µL) was transformed into 100 µl competent BTH101 cells, and selected for on LB plus ampicillin. Selected colonies were inoculated in 5ml LB plus ampicillin overnight at 37°C at 200 rpm. Strains were then subinoculated in LB ampicillin, and grown at 37°C at 200 rpm until OD_{600} 0.3-0.5. Cells were then pelleted at 4000 rpm. Pellets were resuspended in 0.5 mL TSB buffer (100 mL LB, 5 mL 1M $MgSO_4$, 5 mL DMSO, 10 g PEG 6000) and

incubated on ice for 10 minutes. The low-copy number T25 clone (1 µL) was then transformed into 100 µL of these TSB-treated cells. The resulting transformation was plated on LB plates containing relevant antibiotics and X-gal. Plates were incubated at 28°C for 3-5 days until colour began forming. Plates were then scored for colour as an initial determination of interaction, before being quantified using a β-galactosidase assay.

For quantitation of interaction, single colonies from these plates were inoculated in 5 mL LB plus relevant antibiotics overnight at 28°C with 200 rpm shaking. Three colonies from each strain were used to inoculate separate cultures. Cultures were then subinoculated 500 µL into 5 mL LB 1% glucose plus antibiotics. At OD_{600} 0.4 cultures were induced with 0.5 mM IPTG for six hours up to overnight. These cultures were then used for a quantitative β-galactosidase assay. Cells (100 µL) were aliquoted into a 2 mL Eppendorf tube. The remainder of the sample was kept on ice to measure OD_{600}. Z-buffer (850 µl) (0.06 M Na_2HPO_4, 0.04 M $NaH_2PO_4.2H_2O$, 0.01 M KCl, 0.001 M $MgSO_4.7H_2O$ pH7.0) plus 50 µL 0.1% SDS per sample was added. Cells were permeabilised by addition of 50 µL chloroform to each tube followed by vortexing and subsequent incubation for 5 minutes at 28°C. For the β-galactosidase assay 200 µL of ONPG (4 mg mL^{-1} in Z-buffer) was added to each sample, and starting time was measured. When sufficient yellow colour was formed 500 µL of stop solution (1 M Na_2CO_3) was added and stop time was noted. Yellow colour was measured by absorbance at OD_{420}. Enzyme activity was calculated in Miller units (Miller units = 1000 x $[(OD_{420})/(T \times V \times OD_{600})]$).

Positive controls of Rlv3841 PtsN1 and Rlv3841 KdpD as described by Prell, Mulley et al. (2012) were used to optimise the BACTH system.

2.7.2 SDS-Page electrophoresis

SDS-PAGE gels were used to separate proteins by molecular weight for visualisation using staining or Western blotting. Gels were either 1.5 mm (staining) or 0.75 mm (Western blot) thick. Separating gels were made up with 12% acrylamide (per 10 mL gel: 2.5 mL 1.5 M Tris HCL pH 8.8, 0.1 mL 10% SDS, 4.3 mL H_2O, 3 mL 40% acrylamide, 0.1 mL 10% APS, 5 µL TEMED). Once separating gel had solidified, a stacking gel (per 2.5 mL gel: 0.315 mL 1.5 M Tris HCL pH 6.8, 25 µL 10% SDS, 1.816 mL H_2O, 0.3125 mL 40% acrylamide, 25 µL 10% APS, 10 µL 0.2% bromophenol blue, 1.25 µL TEMED) was poured with 10 sample wells. Gels were stored at 4°C.

Protein samples were prepared in 20 µL volumes using a 2x loading buffer (62.5 mM Tris-HCl pH 6.8, 2% SDS, 10% glycerol, 0.024% bromophenol blue, 0.1 M DTT, 4 mL H_2O). Where appropriate, samples were diluted with H_2O before adding loading buffer. Either 20 µL (0.75 mm gel) or 40 µL (1.5 mm gel) of sample was loaded. BLUE Wide Range Prestained Protein Ladder (2.5 µL) (GeneFlow) was used for molecular weight estimation. Protein gels were run in a running buffer, prepared at 10x concentration (30.285 g Tris base, 144 g glycine, 10 g SDS in 1 L H_2O). Gels were run at 180 mV for 45-50 minutes.

InstantBlue (Expedeon) was used to stain proteins in SDS-PAGE gels. Gels were incubated in 15 mL InstantBlue for at least one hour, up to overnight. Gels were imaged using a GelDoc (BioRad) on the Coomassie Blue protocol. When higher resolution of proteins was required, SYPRO Ruby (Thermo Scientific) staining was used. Gels were incubated for at least one hour in fixing solution (50 % methanol, 7% acetic acid). Samples were then incubated in the dark in SYPRO Ruby solution overnight. Following overnight incubation samples were incubated in wash solution (10% methanol, 7% acetic acid) for 30-60 minutes before imaging on the GelDoc system using the SYPRO Ruby protocol.

2.7.3 Western blotting

In order to test protein samples with antibodies, proteins must first be transferred from SDS-PAGE gels to a nitrocellulose membrane (Amersham™ Protran™ 0.2 µm Nitrocellulose Blotting Membrane, GE Healthcare). The SDS-PAGE gel and a 6x9 cm nitrocellulose membrane were sandwiched between two layers of dampened thick filter paper (when no thick filter paper was available two layers of thinner paper were used. These were sandwiched between two sponges. Protein was run from negative to positive at 100 mV for two hours (or at 30 mV overnight).

Following transfer membranes were incubated in 20 mL blocking buffer (1 g milk powder in 20 mL TBS-T) at room temperature for 1 hour. Membranes were subsequently washed three times for 5 minutes in TBS-T (8 g NaCl, 20 mL 0.5 M Tris pH 7.5, 200 µL Tween-20 in 1 L H_2O). Membranes were then incubated in primary antibody overnight at 4°C. Primary antibodies against protein tags were diluted 1:10000 in blocking buffer whilst monoclonal antibodies against specific proteins were diluted 1:3000. After incubation with primary antibody, membranes were washed 3x in TBS-T for 15 minutes. Membranes were then incubated with secondary antibody (1:20000 in TBS-T) for an hour before 1x wash with TBS-T for 5 minutes followed by 2x washes with TBS (8 g NaCl, 20 mL 0.5 M Tris pH 7.5 in 1 L H_2O). All incubation steps were carried out with shaking at 100 rpm. Secondary

antibodies were conjugated to horseradish peroxidase (HRP), allowing detection of antibodies by chemiluminescence. Chemiluminescence was detected using Clarity™ Western ECL Substrate (BioRad), used according to manufacturers instructions. Membranes were incubated in ECL substrate for one minute before imaging in a GelDoc (BioRad) using the High Sensitivity Chemiluminescence protocol. Antibodies used in this work are shown in Table 2.7.

Western blotting against Strep-tag® proteins follows an alternative protocol. Membranes were blocked in BSA (0.6 g BSA in 20 mL PBS-T). Membranes were washed three times for 5 minutes in PBS-T (400 µL Tween-20 in 1 L PBS). Anti-Strep® antibody is directly conjugated to HRP, and so no secondary antibody step is required. Anti-Strep® antibody was added to the membrane 1:20000 in PBS-T, and membranes were incubated for 1 hour with shaking at room temperature. The membranes were then washed 1x in PBS-T and 2x in PBS before detection using the ECL substrate.

Table 2.7 Table of antibodies used in this work. a: the anti-Strep antibody is conjugated directly to HRP; b. Antibodies from GenScript were raised for the purposes of this project.

Epitope	**Secondary conjugant**	**Source**
FLAG-tag	Mouse	Sigma Aldrich
Strep-tag®	n/a[a]	IBA
His-tag	Mouse	Thermo Scientific
FixA	Rabbit	Ruston (2003)
FixB	Rabbit	Ruston (2003)
NifH	Rabbit	GenScript[b]
PdhA	Rabbit	GenScript[b]
LpdH	Rabbit	GenScript[b]

2.7.4 Generating recombinant proteins displaying protein tags

In order to develop a system that allows for *in vitro* investigation of proteins of interest, commercially developed protein tags were fused to proteins of interest. Protein tags were added by adding extensions to cloning primers designed against genes of interest. In the case of His_6-tag, the tag is incorporated in some cloning vectors. In this case, the stop codon was omitted if His_6-tag was desired. Protein tags used in this work are described in Table 2.8. Both the Strep-tag® II and Twin-Strep-tag were used. The Twin-Strep-tag is a tandem version of Strep-tag® II which provides more efficient purification. The same applies for single-FLAG and 3xFLAG tag.

Table 2.8 Protein tags used in this work.

Tag	Amino acid sequence	Source/reference
His_6	HHHHHH	Bornhorst and Falke (2000)
FLAG	DYKDDDDK	Sigma Aldrich
3xFLAG	DYKDHDGDYKDHDIDYKDDDDK	Sigma Aldrich
Strep-tag® II	AWRHPQFGG	IBA
Twin-Strep®	WSHPQFEKGGGSGGGSGGSAWSHPQFEK	IBA

2.7.5 Expression of tagged proteins in *E. coli*

The pOPIN system of high-throughput expression vectors was used (OPPF-UK) for expression in *E. coli*. Cloning into the POPIN system is by InFusion® HD cloning (2.3.6). Primers were designed using the ClonTech tool with complementary overhangs to the pOPIN vectors. The same PCR product can be cloned into all vectors from the pOPIN system. Vectors pOPINF and pOPINS3C were used in this study (obtained from M. Banfield, JIC). pOPINF adds an N-terminal His_6-tag, separated from the protein of interest by a 3C protease site, allowing for cleavage of tags if desired. POPINS3C adds an additional SUMO (small ubiquitin-like modifier) tag between the His_6-tag and the 3C-tag. SUMO tags lead to modifications on the target protein, including increased protein stability.

Expression of pOPIN constructs was carried out using a T7-based expression protocol. *E. coli* strains containing POPIN constructs were grown in LB to OD_{600} 0.5-1.0. Culture (5 mL) was sub-inoculated in 50 mL of Formedium Auto Induction Media (AIM). Cultures were incubated at 28°C 150 rpm for two hours and then at 17°C for two days. After two days cultures were incubated for 5 minutes on ice before 25 mL was pelleted at 4500 rpm for 15 minutes at 4°C. Pellets were then resuspended in 1 ml cold binding buffer (20 mM sodium phosphate, 500 mM NaCl, 20 mM imidazole, pH 7.4) plus 2 mM DTT. Resuspended pellets were homogenised in a ribolyser (FastPrep FP120, Thermo) for 30 seconds at speed 6.5, incubated on ice for 5 minutes and ribolysed for a further 30 seconds. Samples were incubated for 5 minutes on ice and then spun in a microcentrifuge for 5 minutes at 1000 rpm to remove beads. The resulting sample was tested in an anti-His Western Blot to analyse presence of tagged protein.

2.7.6 Expression of tagged proteins in *R. leguminosarum* under native control

The highly-stable vector pJP2 (Prell, Boesten et al. 2002) was used to express *R. leguminosarum* genes under their native promoter. Plasmid pJP2 provides stable expression *in planta*, which is required for genes controlled by bacteroid-specific promoters. Primers were designed to amplify the genes including the upstream region containing the promoter and ribosome-binding site, with a 5' *XbaI* site. The reverse primers bound to the 3'-region of the gene of interest, lacking the stop codon. Extensions were added to the 5'-end of the reverse primer with the reverse complement of the protein tag of interest preceded by a *BamHI* site. PCR products were cloned into pJET1.2/blunt. The desired fragment was then cloned into the MCS of pJP2 using *XbaI/BamHI*.

2.7.7 Expression of tagged proteins in *R. leguminosarum* under inducible control

A taurine-based expression system was used to express bacteroid-specific proteins in the free-living state. This system, developed by Tett, Rudder et al. (2012), utilises the taurine-inducible promoter *tauAp*. Plasmid pLMB509 is a high-copy number plasmid containing the taurine promoter and is amenable to BD cloning. pLMB509 incorporates a C-terminal His_6-tag on cloned products. Cloning into pLMB509 is by InFusion® HD cloning (2.3.6). Primers were designed using the ClonTech online tool with complementary overhangs to the pLMB509 vector. Protein tag sequences were added to the 5'-end of the primer, with the complementary overhang at the 5'-end of this tag sequence. Cloned plasmids were

confirmed by mapping colony PCR using primers pLMB509 fw and pLMB509 rv, and subsequently confirmed by sequencing. Correct plasmids were conjugated into *R. leguminosarum* strains by patch mating (2.3.11).

Taurine-inducible strains were grown in 50 mL UMS with 10 mM glucose, 10 mM NH_4Cl and 10 mM taurine. Cultures were incubated overnight at 28°C at 200 rpm shaking. Cultures were pelleted at 4000 rpm once at OD_{600} 0.4-0.6.

2.7.8 Crude lysis of bacterial cells for protein assays

To prepare cells for protein biochemistry, cell or bacteroid pellets were resuspended in 1 mL HEPES pH 7.0 plus protease inhibitor (Roche). Samples for Strep-tag Western blots were resuspended in sterile PBS (phosphate buffer saline) plus protease inhibitor. The resuspended bacteria were added to ribolyser tubes containing 1 mm silica and 1 mm glass beads. Samples were ribolysed for 30 seconds at a speed of 6.5. Tubes were incubated on ice for 5 minutes before a further 30 seconds of ribolysing. Samples were incubated on ice for 5 minutes before being spun in a microfuge at 1000 rpm for 5 minutes to pellet beads. Samples were carefully pipetted into a sterile microfuge tube ready for use in assays. Samples were stored at -80°C.

2.7.9 Gentle lysis of bacterial cells for protein assays

When required, a gentle cell lysis was used to maintain protein-protein interactions. Cells were grown overnight and 50 mL of culture at OD_{600} ~0.6 was spun at 4000 rpm. Alternatively, bacteroids yielded from 1 g of nodules were spun at 13000 rpm. Pelleted cells were resuspended in a high-salt solution (1.2 M NaCl, 0.8 M sodium citrate) to remove exopolysaccharides (EPS), and incubated at room temperature for 10 minutes before pelleting at 4000 rpm for 20 minutes. Pellets were resuspended in gentle lysis buffer (30 mM Tris-HCL pH 8, 20% sucrose, one protease inhibitor tablet (Roche), 10% glycerol, 0.01% TritonX100 made up in 10 mL H_2O). Lysozyme (50 µL of a 10 mg ml^{-1} stock) was added per 1 mL sample and samples incubated for 15 minutes at room temperature. EDTA (2 µL of a 0.5 M pH 8 stock) was added and samples incubated for 20 minutes at room temperature. Benzonase® (Sigma Aldrich) was added (1 unit per 1 mL) in order to break down nucleic acids. Samples were incubated with benzonase for 30 minutes at 37°C. Tubes were centrifuged at 4°C at 13000 rpm for 30 minutes. Supernatant was collected as lysate to be used in further assays. Samples were used immediately post-lysis and were not stored. The

salt-wash and benzonase steps were added to reduce viscosity and increase flow through the gravity column.

2.7.10 FLAG-tag protein purification

Purification of FLAG-tag proteins was carried out using gravity-flow columns using anti-FLAG M2 beads (Gerace and Moazed 2015). Columns were used to both purify FLAG-tagged proteins from crudely lysed samples, and to carry out pull-down assays from gently lysed samples. BioSpin Chromatography Columns (BioRad) with a volume of 1.2 mL were washed with TBS. A column volume (CV) of 80-100 µL of Anti-FLAG M2 Affinity Gel (Sigma Aldrich) was added to the column, and washed with 3 CV of TBS. This was followed by a wash of 3 CV of glycine-HCL pH 3.5, and a subsequent wash of 5 CV TBS. The sample was then washed over the column. Sample was collected and washed over the column a further two times. The column was then washed with TBS. For protein purification from crudely lysed samples 15 CV washes of TBS were used. For protein pull-downs 40 CV were used. 3xFLAG M2 peptide (Sigma Aldrich) was used to competitively elute any protein bound to the anti-FLAG beads. The elution buffer comprised TBS with 100 µg mL^{-1} FLAG M2 peptide. Elution buffer was passed over the column in five subsequent 1 CV elutions. Eluates 2-5 were pooled and used as the resultant purified sample. The column was washed 3 times with Glycine-HCL and then stored in column storage buffer (TBS 50% glycerol (v/v), 0.02% sodium azide). Columns were stored at 4°C for between uses. Purification was confirmed using a Western blot using an anti-FLAG M2 primary antibody.

2.7.11 Strep-tag protein purification

Strep-tag® proteins were purified via gravity flow in Strep-Tactin® Sepharose columns (IBA). Columns were equilibrated with 1 CV buffer W (100 mM Tris-HCL pH 8.0, 150 mM NaCl, 1 mM EDTA) before the sample was run over the column three times. The column was then washed with 5 CV buffer W. To elute the bound Strep-tagged proteins, six 0.5 CV of buffer E (buffer W plus 15 mM desthiobiotin) was run over the column. Eluates 2-5 were collected as sample. The column was regenerated immediately after use by washing with 15 CV buffer R (buffer W plus 1 mM HABA). Columns were stored at 4°C between uses.

2.7.12 Protein mass spectrometry

Mass spectrometry was carried out in collaboration with the Oxford Advanced Proteomics Facility. Protein samples were prepared by a protocol for Filter-Aided Sample Preparation. Protein samples were transferred onto a Vivacon® 500 10K Da centrifugal concentrator (Viva Products) and centrifuged at 13000 g for 30 minutes. Flow-through was discarded. To remove traces of detergent the column was washed 5x with 200 µL Lysis buffer (8 M Urea, 100 mM TEAB, pH 8), each wash consisting of 20 minutes shaking at room temperature followed by centrifuging at 13000 g for 10 minutes. After washing 200 µL of reduction buffer (8M Urea, 100mM TEAB, 10-20mM TCEP) was added to the filter and samples incubated with shaking at room temperature for 30 minutes. The column was then spun for 30 minutes at 13000 rpm and 200 µL of alkylation buffer (8M Urea, 100mM TEAB, 20-50mM C-IAA) was added to the filter and samples were incubated with shaking in the dark at room temperature. Columns were centrifuged at 13000 rpm for 30 minutes. Samples were washed 2-5x with 200 µL of a Wash buffer (1M Urea, 50mM TEAB pH 8.0) until no detergent could be detected by shaking tubes and looking for foam. Once no detergent could be detected columns were moved to fresh Eppendorf tubes. Trypsin digest was then carried out in two steps. For the first digest, Trypsin was added (5 µg in 200 µL of Wash buffer) and incubated for 3-4 hours at 37°C with gentle shaking. The second step added a further 5 µg trypsin in 100 µL Wash buffer. This was then incubated overnight at 37°C with gentle shaking. The columns were then centrifuged at 13000 g for 30 minutes before washing with 200 µL TFA and centrifuging for a further 30 minutes at 13000 g. The membrane was then washed with 200 µL 50% acetonitrile (ACN) in 0.1 % TFA. The column was then spun at 13000 g for 30 minutes. Tubes were then dried in a vacuum evaporator for 4-5 hours.

Once dry peptides were re-suspended in 5% formic acid and 5% DMSO. Resuspended peptides were separated on an Ultimate 3000 UHPLC system (Thermo Scientific) and electrosprayed directly into a QExactive mass spectrometer (Thermo Scientific) through an EASY-Spray nano-electrospray ion source (Thermo Scientific). The peptides were trapped on a C18 PepMap100 pre-column (300 µm i.d. x 5 mm, 100 Å, Thermo Scientific) using Solvent A (0.1% Formic Acid in water) at a pressure of 500 bar. The peptides were separated on an in-house packed analytical column (75 µm i.d. packed with ReproSil-Pur 120 C18-AQ, 1.9 µm, 120 Å, Dr.Maisch GmbH) using a linear gradient (length: 120 minutes, 7% to 28% solvent B (0.1% formic acid in acetonitrile), flow rate: 200 nL min^{-1}). The raw data was acquired on the mass spectrometer in a data-dependent mode (DDA). Full scan spectra were acquired in the Orbitrap (scan range 350-2000 m/z, resolution 70000, AGC target $3x10^6$, maximum injection time 100 ms). After the scans, the 20 most intense peaks were selected for HCD fragmentation at 30% of normalised collision energy. HCD spectra were also

acquired in the Orbitrap (resolution 17500, AGC target $5x10^4$, maximum injection time 120 ms) with first fixed mass at 180 m/z.

Raw mass spectrometry data was processed by MaxQuant (version 1.5.0.35i) for peak detection and quantification. MS spectra were searched against the *R. leguminosarum* bv. *viciae* 3841 proteome (Uniprot) using the Andromeda search engine with the following search parameters: full tryptic specifity, allowing two missed cleavage sites, fixed modification was set to carbamidomethyl (C) and the variable modification to acetylation (protein N-terminus), oxidation (M). Mass spectra were recalibrated within MaxQuant with precursor error tolerance of 50 ppm and then re-searched with a mass tolerance of 5 ppm. The search results were filtered with a false discovery rate (FDR) of 0.01 for proteins, peptides and peptide spectra matches (PSM).

Data analysis was carried out using Perseus software (version 1.5.2.6). Intensity values were log-transformed. Any protein candidates lacking three true values within a single sample set were removed, as were potential contaminants and any reverse peptides. Missing values were replaced with a basal value of 18 to provide three replicates for a two-tailed t-test to compare samples against a negative control. For analysis, data was filtered to remove non-significant and to remove candidates that were more highly expressed in the negative control than the sample of interest.

2.8 Microscopy

2.8.1 Histochemical staining of nodule bacteria

Nodules were removed from plants at 14 days, 21 days or 28 days post-inoculation. Nodules were sectioned using a microtome and stained with toluidene blue (Elaine Barclay, JIC Bioimaging). Light micrographs were imaged using a Meiji MT4310H light microscope and photographed using a Canon EOS 1100D.

2.8.2 Transmission electron microscopy of sectioned nodules

For analysis of bacteroid morphology, ultrathin sections of nodules were prepared for transmission electron microscopy (TEM). After harvesting, nodules were fixed in a solution of 2.5% (v/v) glutaraldehyde in 0.05 M sodium cacodylate pH7.3 (Gordon, Miller et al. 1963). Samples were washed in 0.05 M sodium cacodylate and post-fixed with 1% (w/v) OsO_4 in 0.05 M sodium cacodylate for 60 minutes at room temperature. Samples were then

washed and dehydrated with ethanol (Beringer, Johnston et al. 1977) before gradual infiltration with LR White resin (London Resin Company, Reading, UK) according to manufacturers protocol. Nodules were then sectioned using a Leica UC6 ultramicrotome (Leica). Ultrathin sections of 90 nm were picked up on 200 mesh hold grids coated in pyroxylin and carbon. Grids were viewed on a FEI Tecnai 20 transmission electron microscope at 2000 kV. Digital images were taken using an AMT XR60B digital camera (Deben). When appropriate the osmium staining step was omitted from the preparation in order to remove lipid staining. In order to determine if lipid was present in bacteroids, nodule sections were pre-treated with hexane before TEM sample preparation.

TEM sample preparation and imaging was carried out by Elaine Barclay (JIC Bioimaging) and Euan James (JHI)

Bacteroid sizes were determined using ImageJ software (NIH). Bacteroid area was measured for 100 bacteroids per strain. Subsequent data analysis was carried out using GraphPad Prism.

2.8.3 Raman microscopy

Raman microscopy was used to investigate the biochemical composition of rhizobia at a single-cell level. Raman microscopy couples a standard confocal microscope to a Raman spectroscope to allow Raman analysis within a laser spot. Raman spectroscopy is commonly used in chemistry and has been adopted for use in biological sciences, such as in biomedical research (Puppels, Demul et al. 1990). Raman spectroscopy relies on scattering of monochromatic light caused by light interacting with molecular vibrations (Nie 2001). The shift of energy may be up or down, and results in a unique fingerprint by which molecules can be identified. Raman microscopy was carried out in collaboration with Wei Huang and Jiabao Xu (University of Oxford).

Bacteroid samples were prepared as discussed in 2.5.4. Free-living samples were grown on TY slopes plus appropriate antibiotics. Slopes were diluted in 3-5 mL TY and 1 mL of suspension inoculated into 50 mL TY, grown overnight at 28 °C at 200 rpm. Cultures were pelleted at 4000 rpm at 4°C. Pellets were stored at 4°C. Samples were washed three times with H_2O to remove traces of media. Cells were diluted to 1000 cells mL^{-1} in order to observe individual cells under a microscope. No fixative was used. Cell suspensions were spread onto a quartz slide and observed under 100x magnification. Individual single cell spectra were acquired using an HR Evolution confocal Raman microscope (Horiba Jobin-

Yvon) equipped with a 532 nm neodymium-yttrium aluminium garnet laser and 600 grooves mm^{-1} diffraction grating. Spectra were acquired in the range of 600 to 1900 cm^{-1}, exposed to a 1 μm laser for 20 seconds.

2.9 Computational Methods

2.9.1 *In silico* cloning

All cloning reactions were tested *in silico*. All work done prior to 2015 was carried out using VectorNTI 11.0. Any work done from January 2015 onwards was carried out using Geneious R8. When appropriate, specialised online tools were used; for Gibson cloning (NEBuilder, NEB) and for HD cloning (ClonTech).

2.9.2 Bioinfomatic analysis

Global nucleotide and protein alignments were carried out using BLASTn and BLASTp (NCBI) respectively. Local alignments were carried out using Geneious R8, using the Geneious alignment algorithm. Sequences were obtained from the NCBI database.

The *R. leguminosarum* bv. *viciae* 3841 genome annotation has been maintained by the Poole lab (rlegDB.org). All other annotations were obtained from NCBI or BioCyc.org. Categorisation of gene/protein function was according to Riley codes assigned in the Poole lab. Riley classifications can be found in Supplementary Table 1. Microarray analysis of gene expression was carried out using existing data from the Poole lab (Karunakaran, Ramachandran et al. 2009).

2.9.3 Statistical analysis and data handling

Data input was into MS Excel. Creation of graphs and appropriate statistical tests were carried out using GraphPad Prism 7.

Protein interaction analysis was carried out using MaxQuant and Perseus software (MPI Biochemistry).

CHAPTER 3

Bioinformatic analysis of the *fixABCX* operon

3.1 Introduction

Electron transfer flavoproteins can be found conserved across the Kingdoms of Life. The *fixAB* genes encode members of the electron transfer flavoprotein family (ETF) (Arigoni, Kaminski et al. 1991, Tsai and Saier 1995), whilst the *fixCX* genes have been proposed to encode an electron transfer flavoprotein-quinone oxidase (ETF-QO). FixX is also a ferredoxin-like protein. These electron transfer roles have led to the hypothesis that the FixABCX proteins participate in electron transfer to nitrogenase. The aim of this study is to investigate the real role of the *R. leguminosarum* FixABCX proteins in symbiotic nitrogen fixation. Improvements in bioinformatics and biotechnology have led to a huge range of data that can be used to understand rhizobial ETF proteins and their role in nitrogen fixation.

3.2 Alignment of ETFs and ETF-QOs with the Fix gene products

3.2.1 ETFs across the Kingdoms of Life

Amino acid sequences of ETF-alpha (FixB) and ETF-beta (FixA) subunits and ETF-QOs (FixCX) were extracted from the NCBI database for a set of model organisms. These were aligned with the *R. leguminosarum fix* gene products, as well as the *fix* gene products from free-living diazotroph *A. caulinodans.* Alignments were made using Geneious bioinformatics software, using the Geneious alignment algorithm of progressive pairwise alignments. The alignment clearly shows conserved residues between both ETF subunits across the Kingdoms of Life (Figures 3.1, 3.2). FixB has an FAD-binding domain (Roberts, Frerman et al. 1996). The Rlv3841 FixB FAD-binding domain is annotated in Figure 3.2 as predicted by the NCBI Conserved Domain Database (Marchler-Bauer, Lu et al. 2011).

The alignment of the Rlv3841 FixCX gene products with ETF-QO sequences confirms the hypothesis that they make up a heterodimer that functions as an ETF-QO (Figure 3.3). The FixCX alignment includes annotations determined by Watmough and Frerman (2010) as determined from the 3D structure of ETF-QO from *R. sphaeroides*. These annotated domains are of the FAD-binding domain, two ubiquinone-binding domains and an iron-sulphur cluster [4Fe-4S]-binding domain. In the case of Rlv3841 this iron-sulphur cluster-binding domain corresponds to the FixX ferredoxin.

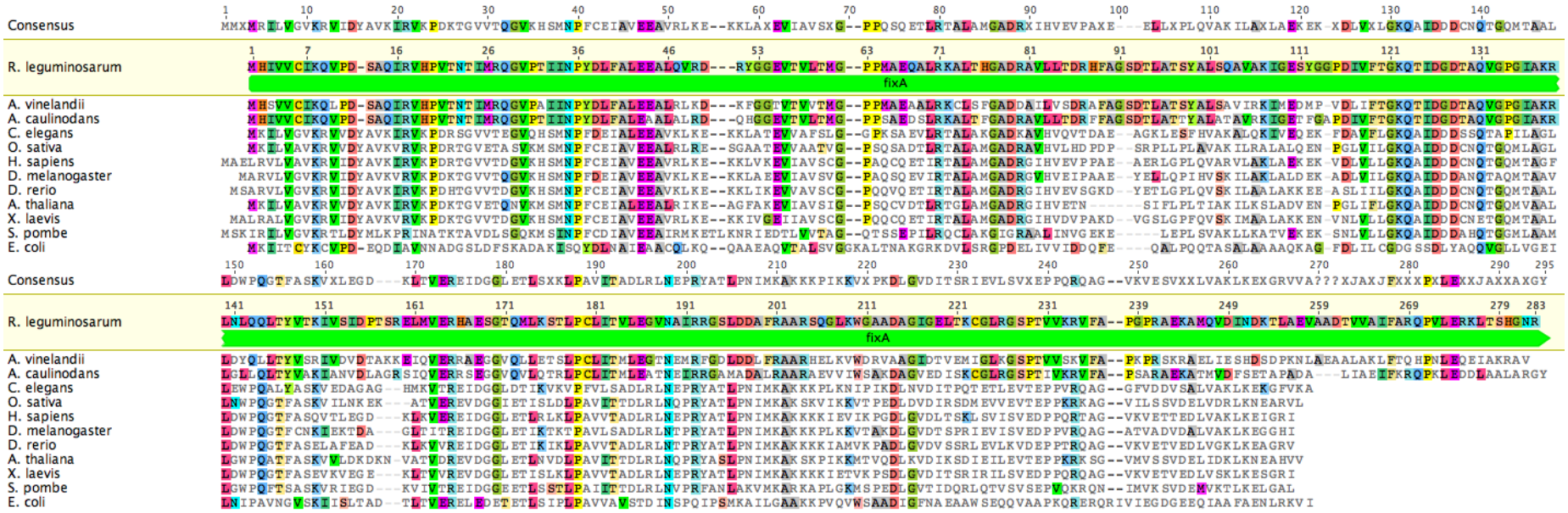


Figure 3.1 Alignment of the *R. leguminosarum* bv. *viciae* 3841 FixA protein with ETF-beta subunits from across the Kingdoms of Life. Alignment created using Geneious. Species listed in order of homology to *R. leguminosarum* bv. *viciae* 3841. *Azotobacter vinelandii* (diazotroph)*, Azorhizobium caulinodans* (diazotroph), *Caenorhabditis elegans* (nematode), *Oryza sativa* (rice), *Homo sapiens* (human), *Drosophila melanogaster* (fruit fly), *Danio rerio* (zebrafish), *Arabidopsis thaliana* (thale cress), *Xenopus laevis* (frog), *S. pombe* (yeast), *Escherischia coli.*

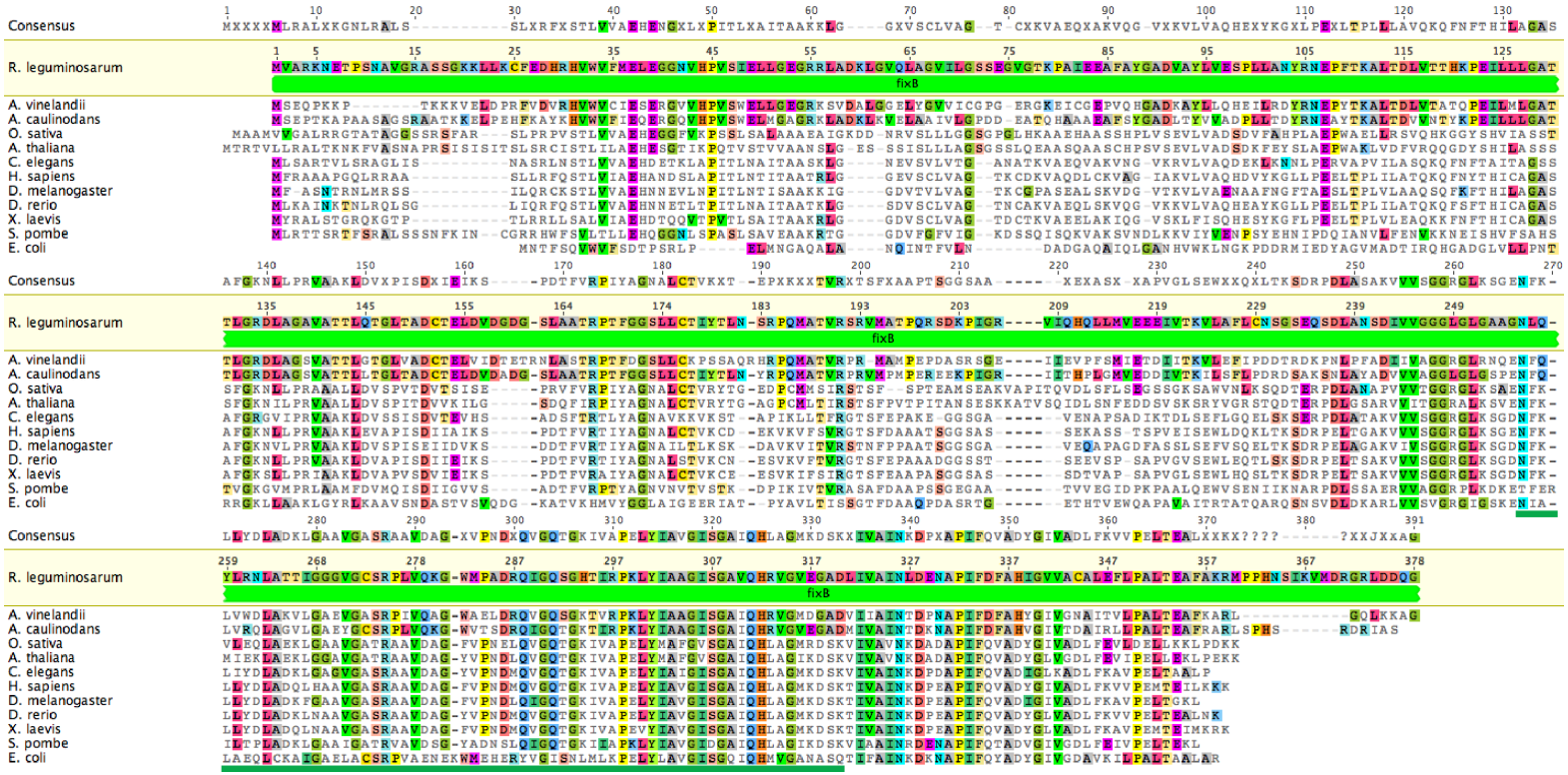


Figure 3.2 Alignment of the *R. leguminosarum* bv. *viciae* 3841 FixB protein with ETF-alpha subunits from across the Kingdoms of Life. Alignment created using Geneious. Annotation determined using NCBI Conserved Domain database; FAD binding (green). Species listed in order of homology to *R. leguminosarum* bv. *viciae* 3841. *Azotobacter vinelandii* (diazotroph), *Azorhizobium caulinodans* (diazotroph), *Caenorhabditis elegans* (nematode), *Oryza sativa* (rice), *Homo sapiens* (human), *Drosophila melanogaster* (fruit fly), *Danio rerio* (zebrafish), *Arabidopsis thaliana* (thale cress), *Xenopus laevis* (frog), *S. pombe* (yeast), *Escherischia coli.*

Figure 3.3 Alignment of the *R. leguminosarum* bv. *viciae* 3841 FixCX protein with ETF-QO proteins from across the Kingdoms of Life. Alignment created using Geneious. Annotations as determined by Watmough and Frerman (2010). FAD-binding (green); ubiquinone-binding (cyan); [4Fe-4S] cluster-binding (red). Species listed in order of homology to *R. leguminosarum* bv. *viciae* 3841. *Azotobacter vinelandii* (diazotroph), *Azorhizobium caulinodans* (diazotroph), *Caenorhabditis elegans* (nematode), *Oryza sativa* (rice), *Homo sapiens* (human), *Drosophila melanogaster* (fruit fly), *Danio rerio* (zebrafish), *Arabidopsis thaliana* (thale cress), *Xenopus laevis* (frog), *S. pombe* (yeast), *Escherischia coli.*

3.2.2 The Fix proteins are conserved across rhizobia

The Fix proteins are relatively well conserved across rhizobia. BLAST analysis was carried out between the *fix* gene products from *R. leguminosarum* bv. *viciae* 3841 and from a set of model rhizobia. The selected strains were *Bradyrhizobium japonicum* USDA 6 and *A. caulinodans* ORS571, which have both free-living and symbiotic N-fixation ability; *S. meliloti* 1021 and *S. fredii* HH103, both *Sinorhizobium*; and two *Rhizobium* strains, IRBG74 and *R. etli* CFN42. Percentage identity is shown in Table 3.1. Genes encoding the FixABCX proteins are found in all of the strains, with the exception of HH103, which has no annotated *fixX* gene. BLAST was used to search for a FixX homolog in *S. fredii* HH103. A region of 79% identity was found, encoded by the intergenic region (IGR) immediately downstream of *fixC*. This region is likely to be a *fixX* gene. All strains showed at least 62% identity between all proteins. This value rises to 72% when only the *Rhizobiacae* are considered. Functional regions such as the FixB FAD-binding domain have much higher homology, agreeing with a conserved function across the rhizobia.

Protein	IRBG74	CFN42	1021	HH103	USDA 6	ORS571
FixA	75.6	78.7	72.8	76.3	65.3	69.2
FixB	73.2	72.7	74.5	73.0	62.4	65.4
FixC	80.2	77.2	80.7	81.1	71.0	71.7
FixX	76.3	75.5	82.7	78.6*	63.3	67.0

Table 3.1 BLAST identity between FixABCX amino acid sequences from Rlv3841 compared with six model rhizobia. *: no annotated FixX is found in HH103, however a homolog can be found encoded by the IGR immediately downstream of *fixC*.

The model bacterium *E. coli* has a set of genes designated *fixABCX*. In *E. coli* these are involved in carnitine metabolism (Eichler, Buchet et al. 1995, Walt and Kahn 2002). The *E. coli* FixABCX proteins are also members of the ETF and ETF-QO family. Alignment of *E. coli* FixA, B, C and X with the Rlv3841 Fix proteins shows 26%, 27%, 34% and 41% respectively. There is no clear conservation between the functional regions such as FAD-binding sites, which suggests that these proteins carry out different functions within the cell. The model diazotroph *Klebsiella oxytoca* has a set of genes designated *fixABCX*; the proteins encoded by these are similar to those seen in *E. coli.*

In *K. oxytoca* a separate set of proteins is involved in electron transfer to nitrogenase, encoded by *nifFJ*. The *nifJ* gene encodes a pyruvate:flavodoxin oxidoreductase (Shah, Stacey et al. 1983), whilst *nifF* encodes a flavodoxin which serves as the electron donor to nitrogenase (Deistung, Cannon et al. 1985). Neither NifF nor NifJ has a homolog in Rlv3841. *K. oxytoca* and *R. leguminosarum* appear to have evolved different strategies for provision of electrons to nitrogenase.

3.2.3 Other ETFs in *R. leguminosarum* bv. *viciae* 3841

The Rlv3841 genome reveals several other proteins from the ETF family. These are described in Table 3.2. BLAST analysis was used to determine identity with their respective Fix protein; the low identity suggest that it is unlikely that these other ETF proteins play the same role as the Fix proteins.

Table 3.2 Other putative electron transfer flavoproteins from *R. leguminosarum* bv. *viciae* 3841. Identity is to respective Fix protein. In the case of ETF-QO, a comparison was made with the *fixCX* gene product.

Locus tag	Gene	Description	Type of ETF	Identity to Fix protein
RL4320	*etfA1*	putative electron transfer flavoprotein alpha-subunit	ETF-α	29.5%
pRL120522	*etfA2*	putative electron transfer flavoprotein alpha-subunit	ETF-α	29.2%
RL4319	*etfB1*	putative electron transfer flavoprotein beta-subunit	ETF-β	29.0%
pRL120521	*etfB2*	putative electron transfer flavoprotein beta-subunit	ETF-β	29.8%
RL1457		putative electron transfer flavoprotein-ubiquinone oxidoreductase	ETF-QO	22.7%
pRL120519		putative flavoprotein oxidoreductase	ETF-QO	4.8%

3.3 Microarray analysis of *fixABCX* expression

Microarray data from experiments carried out by Karunakaran, Ramachandran et al. (2009) can be used to determine expression of the *fix* gene products under different growth conditions, including bacteroids at several time points. All microarray data used is from datasets obtained from Karunakaran, Ramachandran et al. (2009), and desired data was extracted and compared as part of this work. Expression of *nifH* was used as a bacteroid-specific standard. Expression was investigated under a range of growth conditions including secreted plant carbon sources and hormones, as well as in both the rhizosphere and nodules at different ages. Expression changes of the *fix* genes under several conditions are shown in Figure 3.4.

Significant upregulation of gene expression can be seen in bacteroids on both pea and vetch. The most highly upregulated in all bacteroids is *fixA*, which is upregulated even higher than the nitrogenase gene *nifH*. No significant change in gene expression (>2-fold up- or down-regulated) was seen under any other conditions investigated.

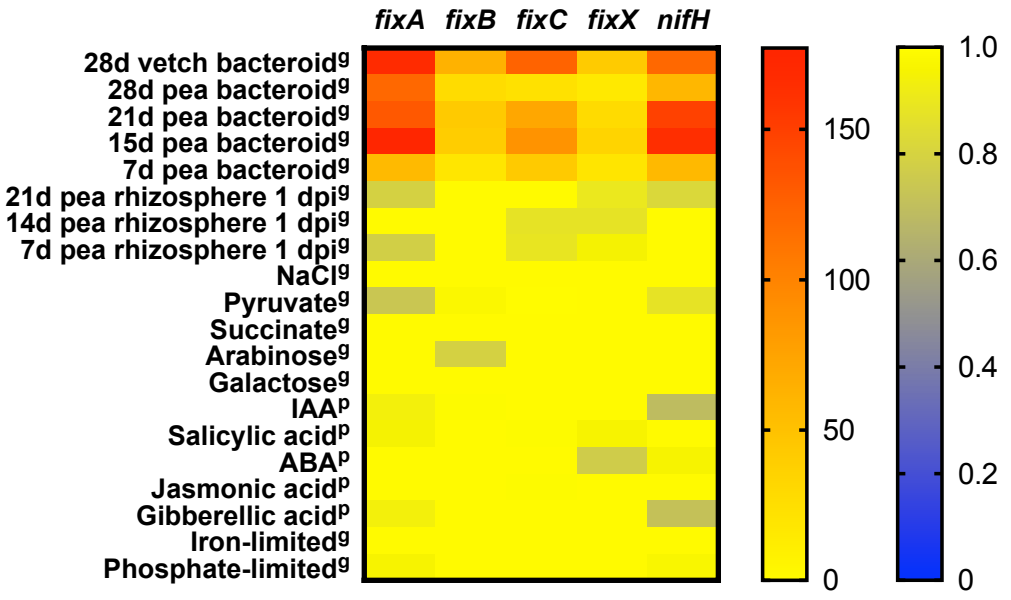


Figure 3.4 Changes in expression of bacteroid genes under different growth conditions. All values are relative to free-living *R. leguminosarum* bv. *viciae* 3841 grown with either glucose (g) or pyruvate (p). dpi: days post inoculation. Red: highly upregulated; yellow: no expression change; blue: downregulated. Microarray data was obtained from Karunakaran, Ramachandran et al. (2009).

3.4 Discussion

Advances in biotechnology and computational power have allowed for a huge amount of data to now be available both locally and globally. This data has allowed an initial investigation into the *fixABCX* genes and their gene products.

The *fixAB* and *fixCX* genes encode electron transfer proteins from the ETF and ETF-QO protein families respectively. These protein families are conserved across the Kingdoms of Life, playing electron transfer roles in many different processes. The FixABCX proteins are found across the rhizobia, suggesting a conserved role. In *R. leguminosarum* bv. *viciae* 3841 the *fixABCX* genes are highly upregulated in mature nodules of both pea and vetch, suggesting a key role in nitrogen fixation processes.

CHAPTER 4

Identification and mutagenesis of key bacteroid electron transfer proteins

4.1 Introduction

The importance of *fixABC* was first demonstrated by Gubler and Hennecke (1986) in *B. japonicum*, where single mutants showed *fixA*, *fixB* and *fixC* to be essential for both symbiotic and free-living microaerobic nitrogen fixation. The same group subsequently identified *fixX* and characterised its requirement for nitrogen fixation (Gubler, Zurcher et al. 1989). The *fixABCX* genes are conserved across rhizobia (Earl, Ronson et al. 1987). This was shown using early hybridisation studies (Earl, Ronson et al. 1987, Gronger, Manian et al. 1987, Iismaa and Watson 1987) and can be subsequently seen in the multitude of sequenced rhizobial genomes. The *fixABCX* genes have now been shown to be essential for nitrogen fixation in several species of diazotrophs, including both in free-living and symbiotic states. In *Rhodospirillum rubrum*, a non-symbiotic diazotroph, a *fixC* mutant strain has reduced but not abolished nitrogen fixation ability (Edgren and Nordlund 2004). Table 4.1 shows a list of phenotypes of published *fix* mutants.

Table 4.1 Nitrogen fixation phenotype of published *fixABCX* mutants in both symbiotic and free-living conditions. N/A: not applicable. Fix^-: no nitrogen fixation. Fix^+: Wild-type nitrogen fixation. Fix^{red}: Significantly reduced nitrogen fixation compared to the wild-type.

Strain	Host plant	Mutation	Symbiotic phenotype	Microaerobic free-living phenotype	Reference
B. japonicum	*G. max*	*fixA*	Fix^-	Fix^-	Gubler and Hennecke (1986)
B. japonicum	*G. max*	*fixB*	Fix^-	Fix^-	Gubler and Hennecke (1986)
B. japonicum	*G. max*	*fixC*	Fix^-	Fix^-	Gubler and Hennecke (1986)
B. japonicum	*G. max*	*fixX*	Fix^-	Fix^-	Gubler, Zurcher et al. (1989)
Bradyrhizobium sp. ORS278	*G. max*	*fixA*	Fix^-	Fix^-	Delmotte, Mondy et al. (2014)
S. meliloti	*M. sativa*	Not specified	Fix^-	N/A	Batut, Terzaghi et al. (1985)
S. meliloti	*M. sativa*	*fixA*	Fix^-	N/A	Hirsch and Smith (1987)
S. meliloti	*M. sativa*	*fixB*	Fix^-	N/A	Hirsch and Smith (1987)
S. meliloti	*M. sativa*	*fixX*	Fix^-	N/A	Dusha, Kovalenko et al. (1987), Hirsch and Smith (1987)
A. caulinodans	*S. rostrata*	*fixA*	Fix^-	Fix^-	Kaminski, Norel et al. (1988)

A. caulinodans	*S. rostrata*	*fixB*	Fix^-	Fix^-	Kaminski, Norel et al. (1988)
A. caulinodans	*S. rostrata*	*fixC*	Fix^-	Fix^-	Kaminski, Norel et al. (1988)
A. caulinodans	*S. rostrata*	*fixX*	Fix^-	Fix^-	Arigoni, Kaminski et al. (1991)
R. rubrum	N/A	*fixC*	N/A	Fix^{red}	Edgren and Nordlund (2004)

To date, no published mutants have confirmed the essential role for *fixABCX* in *Rhizobium* spp. Investigation of the functioning of the *fix* operon in *R. leguminosarum* bv. *viciae* 3841 in symbiosis with its host legume *P. sativum* (pea) will lead to a model for its functioning in *Rhizobium* species. An additional bacteroid-specific ferredoxin was identified in order to elucidate if other ferredoxins were involved in nitrogen fixation. A putative gene was identified with a role in menaquinone biosynthesis in order to determine if menaquinone is involved in bacteroid electron transfer in Rlv3841.

Table 4.2 Electron transfer genes investigated in this chapter.

Accession	Gene	Product
pRL100200	*fixA*	FixA: ETF beta subunit
pRL100199	*fixB*	FixB: ETF alpha subunit
pRL100198	*fixC*	FixC: ETF-QO subunit
pRL100197	*fixX*	FixX: ETF-QO subunit
pRL100156	*fdxB1*	FdxB1: Ferredoxin
RL3587	*dmtH*	DmtH: Putative menaquinone methyltransferase

4.2 Mutation of *fixABCX* in *R. leguminosarum* bv. *viciae*

4.2.1 Construction of *fixABCX* mutants

Omega-interposon mutagenesis was used to create polar and in-frame mutants in *fixAB*, as outlined in section 2.4.2. The same protocol was used to create omega-insertion and in-frame mutants in *fixB*, *fixC* and *fixX.* Mutants in *fixC* and *fixX* were also generated using interposon vector pK19mob. A 300 bp fragment of the gene of interest is typically cloned into pK19mob. Since *fixX* is only 297bp, the entire *fixX* region was cloned into pK19mob. The pK19mob variants were then conjugated into Rlv3841. Mutants constructed in the pK19 system are not stable, and may revert back to wild-type. Primers and plasmids used for construction of all mutants are outlined in Chapter 2. Mutant strains are outlined in Table 4.3.

Strain	Description	Source
LMB771	*fixAB*::Ωspc	This work
LMB777	Δ*fixAB*	This work
LMB819	*fixB*::Ωspc	This work
LMB825	Δ*fixB*	This work
LMB827	*fixC*::Ωspc	This work
OPS0280	Δ*fixC*	This work
LMB829	*fixX*::Ωspc	This work
OPS0469	Δ*fixX*	This work
LMB823	pK19mob*::fixC*	This work
LMB824	pK19mob*::fixX*	This work

Table 4.3 *fix* mutant strains in Rlv3841 background used in this work.

4.2.2 Free-living phenotype of *fixAB* mutants

To determine if deletions in *fixABCX* had an effect on the free-living phenotype of strain 3841, cultures were grown in rich media (TY) for 48 hours in a FLUOstar Omega plate reader (BMG Labtech). No difference was seen in growth in any of the mutants tested, suggesting they play no role in free-living growth in Rlv3841.

Strain	Description	Doubling time /hours
3841	*R. leguminosarum* bv. *viciae* 3841	3.87
LMB777	Rlv3841 Δ*fixAB*	3.59
OPS0280	Rlv3841 Δ*fixC*	3.80
OPS0469	Rlv3841 Δ*fixX*	3.85

Table 4.4 Doubling time for *fixABCX* mutants grown in rich media in a FLUOstar Omega plate reader at 21% O_2, 400 rpm.

4.2.3 Symbiotic phenotype of *fixAB* mutants

In order to investigate the symbiotic phenotype of the *fix* mutants, peas were planted in sterile vermiculite in 1L pots. Each pea was inoculated with $1x10^7$ colony forming units of the chosen strain, or sterile water as a negative control. Plants were grown for 28 days before symbiotic phenotype was recorded.

Peas inoculated with wild-type strain Rlv3841 appear healthy with green leaves. Plants inoculated with either the polar *fixAB*::Ωspc (LMB771) or in-frame Δ*fixAB* (LMB777) mutant had stunted growth and yellowing leaves due to lack of nitrogen. Water-inoculated controls had the same phenotype as the mutant-inoculated plants.

Figure 4.1 Visual phenotype of plants inoculated with sterile water, wild-type Rlv3841, *fixAB* mutants LMB771 (*fixAB*::Ωspc) and LMB777 (Δ*fixAB*). Plants are representative of four replicates. Photographs were taken at 28 days post-inoculation.

Nodules of wild-type plants were elongated, with a characteristic pink colour due to the presence of leghaemaglobin. Nodules of plants inoculated with either LMB771 (*fixAB*::Ωspc) or LMB777 (Δ*fixAB*) were small and lack any pink pigmentation. This is due to the plant terminating nodule development if its symbiont is not providing fixed nitrogen. Strains of all plants were confirmed by crushing the nodules and streaking the resultant brei on selective agar plates. All strains grew on streptomycin, whilst only LMB771 was able to grow on spectinomycin. Both plant and nodule phenotypes resemble the phenotype of NifH (nitrogenase subunit) mutant RU3940 (Karunakaran, Ramachandran et al. 2009).

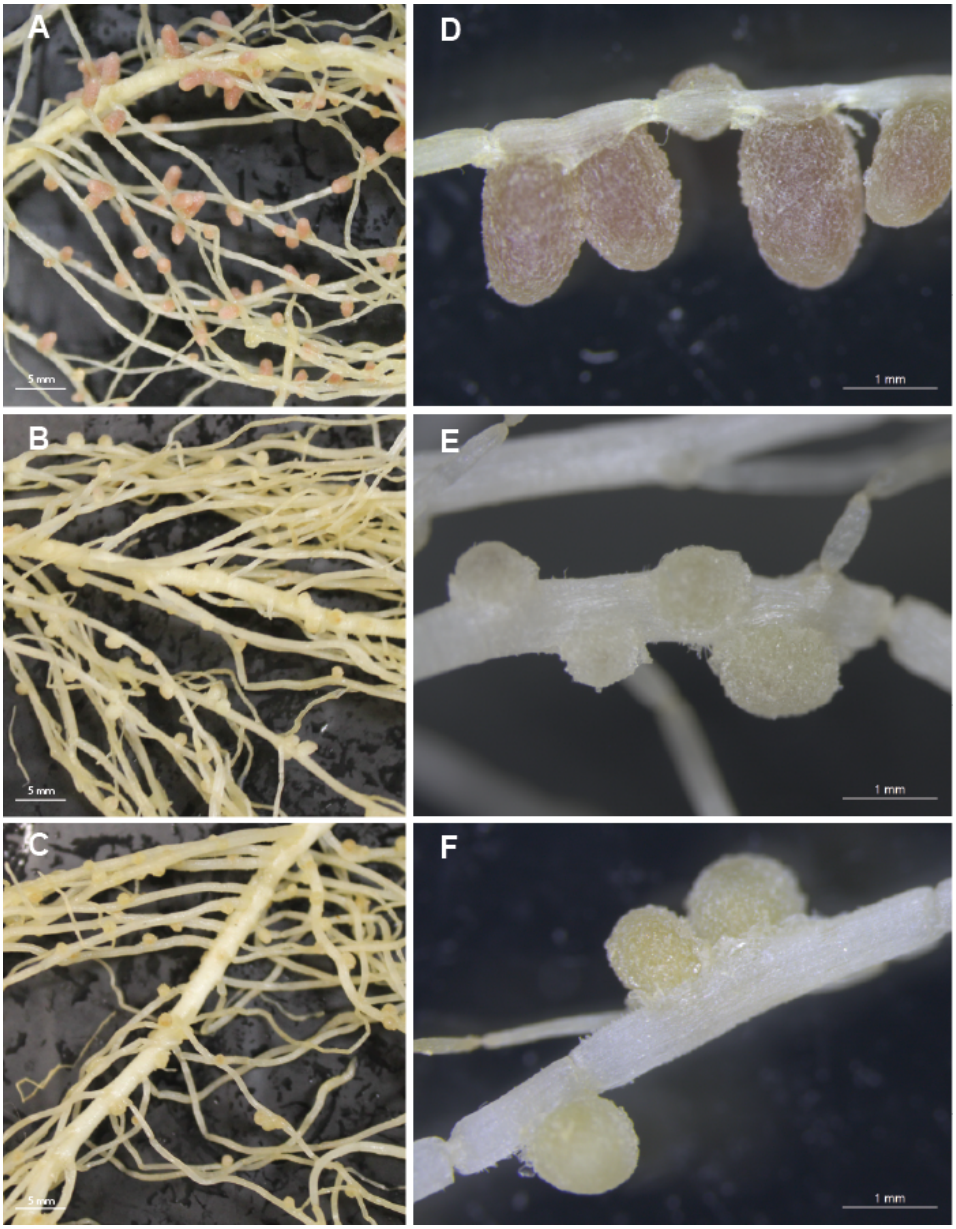


Figure 4.2 Nodule phenotype of plants inoculated with *fixAB* mutants: A,D: Rlv3841; D,E: LMB771 (*fixAB::*Ωspc)**;** C,F: LMB777 (Δ*fixAB)*. Plants are representative of four replicates. Photographs were taken at 28 days post inoculation. Scale bars A,B,C: 5 mm; D,E,F: 1 mm.

Nodule sections were mounted on slides and stained with toluidine blue (high affinity stain that dyes acidic tissue components) in order to visualise the bacterial and plant cells. The nodules from the *fixAB* mutant are notably stunted (0.5-1 mm) compared to the wild-type (2-3 mm), with fewer plant cells accommodating bacteroids.

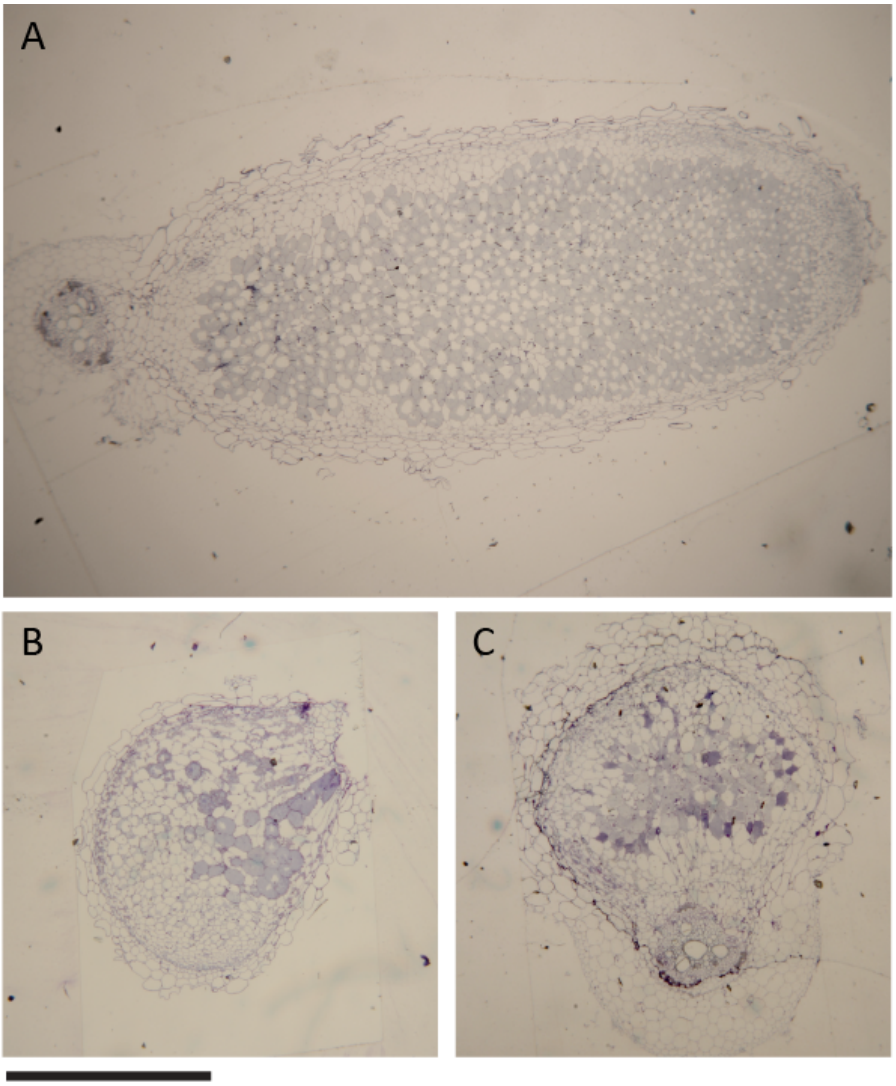


Figure 4.3 Light micrographs of nodule sections stained with toluidene blue. Magnification 40x. A: 3841; B: LMB771 (*fixAB*::Ωspc); C: LMB777 (Δ*fixAB*). Images are representative of five replicates. Scale bar: 1 mm.

Acetylene reduction assays were used to measure nitrogenase activity within nodules. Acetylene reduction assays on plants inoculated with the *fixAB* mutants show no nitrogenase activity. *fixAB* is therefore essential for nitrogen fixation in *R. leguminosarum* bv. *viciae* in symbiosis with peas.

Strain	Description	Symbiotic phenotype	Nodule colour	Acetylene reduction μmol ethylene plant^{-1} hr^{-1}
3841	Wild-type	Nod^{+} Fix^{+}	Pink	3.3 ± 0.1
LMB771	*fixAB*::Ωspc	Nod^{+} Fix^{-}	White	0.0 ± 0.0
LMB777	Δ*fixAB*	Nod^{+} Fix^{-}	White	0.0 ± 0.0
Uninoculated control	n/a	Nod^{-} Fix^{-}	n/a	0.0 ± 0.0

Table 4.5 Symbiotic phenotype of *fixAB* mutant strains on pea (*P. sativum*). Acetylene reduction assay was carried out 28 days post inoculation; ± SEM, n ≥ 6.

4.2.4 Complementation of *fixAB* mutants

The vector pJP2 was used to complement LMB771 (*fixAB*::Ωspc) and LMB777 (Δ*fixAB*). pJP2 is a stable vector developed for expression in rhizobia, including in bacteroids. Two plasmids were created for complementation of these mutants (Figure 4.4). The first comprised *fixAB* under its native promoter (the entire ~1kb intergenic region upstream) (pLMB811). The second construct expressed the entire *fixABCX* region under the same promoter region (pLMB826).

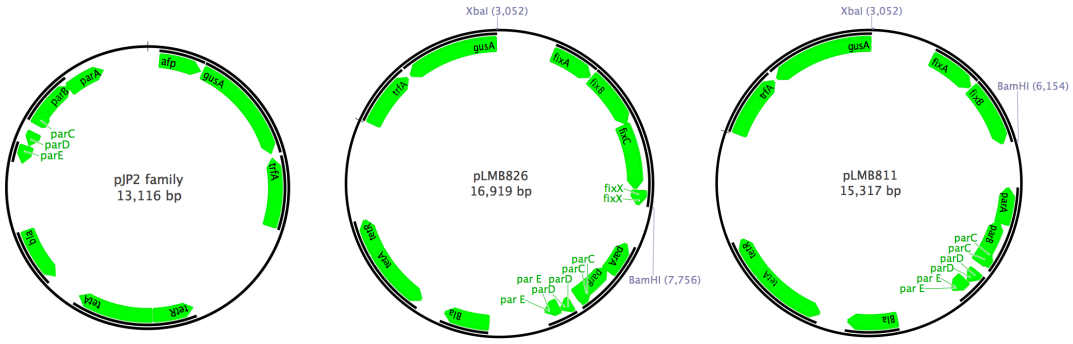


Figure 4.4 Plasmids used for complementation of *fix* mutants.

Background strain	No complement		pJP2:*fixAB*		pJP2:*fixABCX*	
	Symbiotic phenotype	Acetylene reduction	Symbiotic phenotype	Acetylene reduction	Symbiotic phenotype	Acetylene reduction
3841	$Nod^+ Fix^+$	4.1 ± 0.4^a	$Nod^+ Fix^+$	4.3 ± 0.3^a	$Nod^+ Fix^+$	5.2 ± 0.3^b
LMB771 (*fixAB*::Ωspc)	$Nod^+ Fix^-$	0.0 ± 0.0^c	$Nod^+ Fix^-$	0.0 ± 0.0^c	$Nod^+ Fix^+$	3.3 ± 0.3^e
LMB777 (Δ*fixAB*)	$Nod^+ Fix^-$	0.0 ± 0.0^c	$Nod^+ Fix^+$	2.0 ± 0.2^d	$Nod^+ Fix^+$	5.6 ± 0.2^b

Table 4.6 Symbiotic phenotype of complemented *fixAB* mutant strains on pea (*P. sativum*). Acetylene reduction assay was carried out 28 days post inoculation, with values expressed as μmol ethylene $plant^{-1}$ hr^{-1}; ± SEM, n ≥ 3; a, b, c, d, e represent statistically distinct groups ($p \leq 0.05$) determined using one-way ANOVA and Tukey's multiple comparisons test.

The in-frame Δ*fixAB* mutant can be complemented by pLMB811 (pJP2:*fixAB*). This complemented mutant only fixed nitrogen at ~50% of wild-type levels. The polar *fixAB*::Ωspc mutant, however was only complemented by the entire *fix* operon (pLMB826), fixing nitrogen at ~80% wild-type rates. The wild-type and Δ*fixAB* mutant complemented with the entire *fix* operon are able to fix nitrogen at higher rates than the uncomplemented wild-type.

4.2.5 Symbiotic phenotype of *fixBCX* mutants

In addition to mutations in *fixAB*, stable mutants were also generated in *fixB, fixC* and *fixX*. Again, symbiotic phenotypes were observed at 28 days post-inoculation, looking at whole plant and nodule phenotype and using acetylene reduction to assess nitrogenase activity. Plants inoculated with *fixB* (Figure 4.5) and *fixX* (Figure 4.7) mutants displayed the same phenotype as plants inoculated with *fixAB* mutants, with small white nodules and yellowing, stunted plants. The same phenotype was observed for the in-frame Δ*fixC* mutant (Figure 4.6). The polar *fixC*::Ωspc mutant, however, appeared to display an intermediate phenotype, with paler pink, smaller nodules and nitrogen fixation ~70% of wild-type. Since in-frame deletions in *fixC* and in *fixX* were Fix^-, it seems unlikely that a polar *fixC* mutant could fix nitrogen. Mapping primers were used to determine if the mutant had been successfully

generated. A PCR with primers oxp0336 and oxp0337, designed to amplify the entire *fixC* gene resulted in wild-type fragments. A PCR with primers oxp0336 and potfarforward (which binds within the omega cassette) resulted in fragments expected in the omega mutant. This PCR mapping of the *fixC*::Ωspc mutant suggests it carries a duplication, and this duplication may have lead to partially restored phenotype. However, given the complexity of mapping a partial duplication and availability of a defined *fixC* deletion mutant, this was initially not further analysed. The strain has since been sent for whole-genome sequencing in order to elucidate the reason behind this unusual phenotype; data will not be available by the end of this project. Overall, since in-frame deletions of *fixABCX* lead to a Fix$^-$ phenotype, it can be concluded that the entire *fix* operon is essential for symbiotic nitrogen fixation in *R. leguminosarum*.

The *fixC*::pK19mob mutant also showed a Fix$^-$ phenotype with white nodules, as seen in the in frame *fixC* deletion. The *fixX*::pK19mob mutant demonstrated an intermediate phenotype, with smaller pink nodules and slightly yellowed leaves, with ~80% wild-type rates of acetylene reduction. However, due to the small size of *fixX,* the entire gene was used in the pK19mob vector, so it is likely that the interposon mutagenesis was not effective for this gene. Stable in-frame mutants have confirmed that *fixX* is indeed essential.

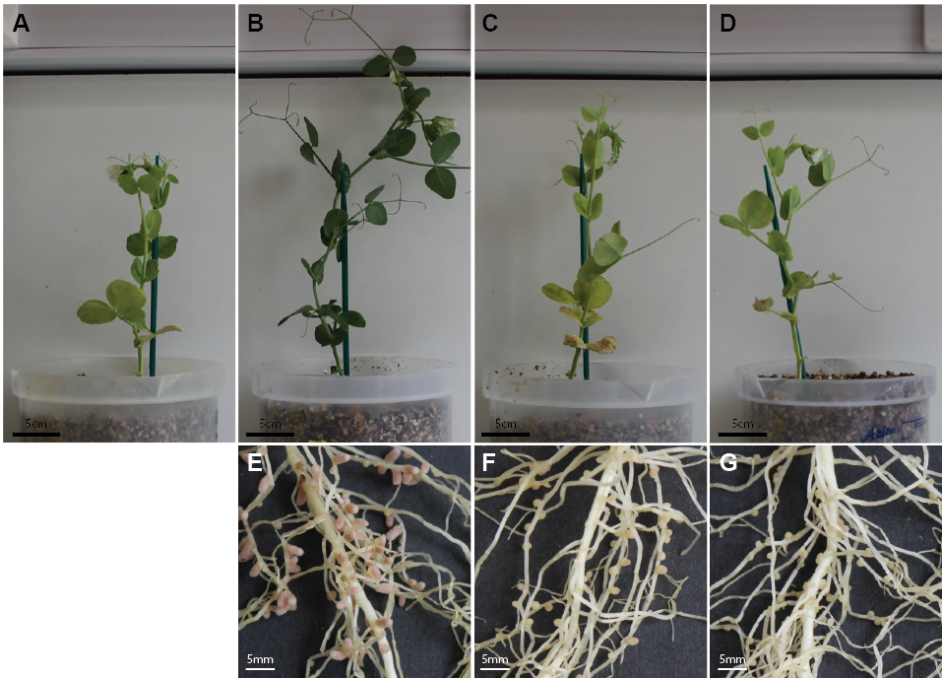


Figure 4.5 Plant phenotype of plants inoculated with *fixB* mutants: A: Uninoculated control; B,E: Rlv3841; C,F: LMB819 (*fixB*::Ωspc); D,G: LMB825 (Δ*fixB*)**.** Plants are representative of four replicates. Photographs taken at 28 days post-inoculation. Scale bars A-D: 5 cm; E-G: 5 mm.

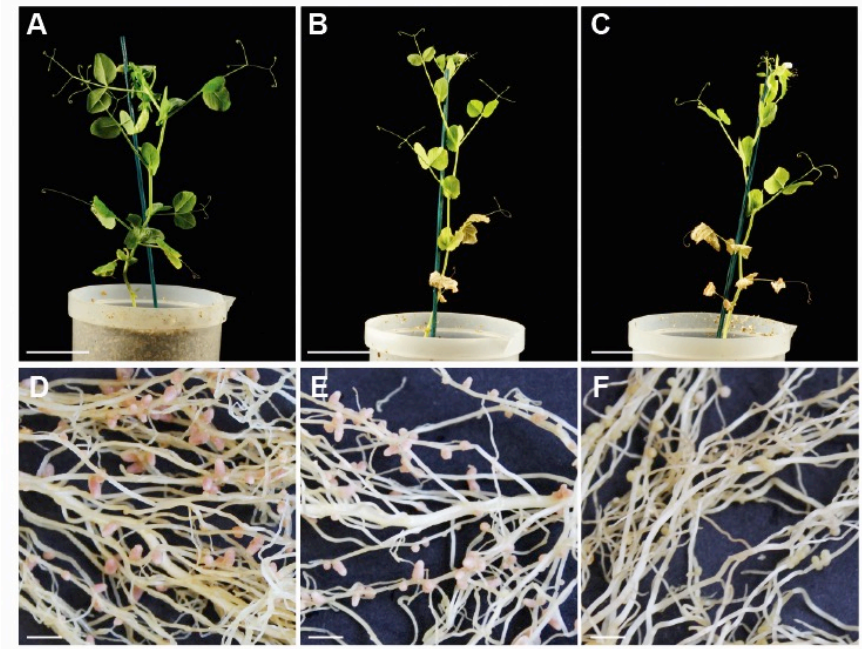


Figure 4.6 Plant phenotype of plants inoculated with *fixC* mutants: A,D: Rlv3841; B,E: LMB827 (fixC::Ωspc); C,F: OPS0280 (ΔfixC). Plants are representative of four replicates. Photographs taken at 28 days post-inoculation. Scale bars A-C: 5 cm; D-F: 5 mm.

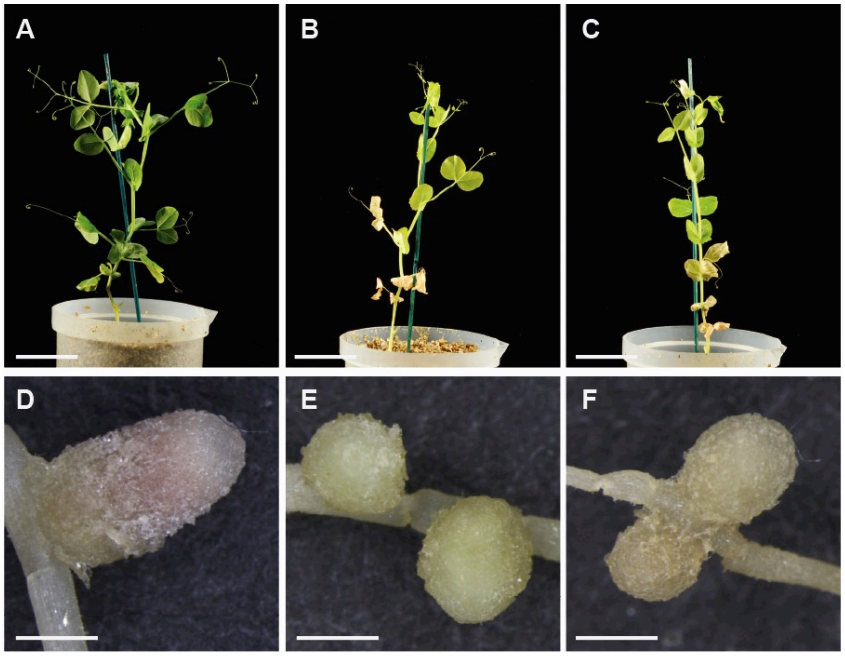


Figure 4.7 Plant phenotype of plants inoculated with *fixX* mutants: A,D: Rlv3841; B,E: LMB829 (fixX::Ωspc); C,F: OPS0469 (ΔfixX). Plants are representative of four replicates. Photographs taken at 28 days post-inoculation. Scale bars A-C: 5 cm; D-F: 1 mm.

Strain	Description	Symbiotic phenotype	Nodule colour	Acetylene reduction µmol ethylene plant^{-1} hr^{-1}
3841	Wild-type	Nod^{+} Fix^{+}	Pink	3.6 ± 0.2^{a}
LMB819	*fixB*::Ωspc	Nod^{+} Fix^{-}	White	0.0 ± 0.0^{b}
LMB825	Δ*fixB*	Nod^{+} Fix^{-}	White	0.0 ± 0.0^{b}
LMB827	*fixC*::Ωspc	Nod^{+} Fix^{red}	Pink	2.5 ± 0.2^{c}
OPS0280	Δ*fixC*	Nod^{+} Fix^{-}	White	0.0 ± 0.0^{b}
LMB829	*fixX*::Ωspc	Nod^{+} Fix^{-}	White	0.0 ± 0.0^{b}
OPS0469	Δ*fixX*	Nod^{+} Fix^{-}	White	0.0 ± 0.0^{b}
LMB823	*fixC::*pK19mob	Nod^{+} Fix^{-}	White	0.0 ± 0.0^{b}
LMB824	*fixX::*pK19mob	Nod^{+} Fix^{red}	Pink	3.0 ± 0.2^{c}
Uninoculated control	n/a	Nod^{-} Fix^{-}	n/a	0.0 ± 0.0^{b}

Table 4.7 Symbiotic phenotype of *fix* mutant strains on pea (*P. sativum*). Acetylene reduction assay was carried out 28 days post-inoculation; ± SEM, n ≥ 4. a, b, c represent statistically distinct groups (p≤0.05) determined using one-way ANOVA and Tukey's multiple comparisons test.

4.2.6 Complementation of *fixBCX* mutants

All Fix mutants were complemented using pLMB826 (the entire *fixABCX* operon under its native promoter).

Complementation of the *fixB* mutants was able to return a Fix^{+} phenotype, though complemented strains showed a reduction in nitrogen fixation ability compared to wild-type (~66%). Since FixA and FixB form a heterodimer, this may be due to the imbalance between FixA and FixB levels in the bacteroid.

Background strain	No complement		pJP2:*fixABCX*	
	Symbiotic phenotype	**Acetylene reduction**	**Symbiotic phenotype**	**Acetylene reduction**
3841	$Nod^+ Fix^+$	3.5 ± 0.2^a	n.d	n.d
LMB819 (*fixB*::Ωspc)	$Nod^+ Fix^-$	0.0 ± 0.0^b	$Nod^+ Fix^+$	2.3 ± 0.2^c
LMB825 (Δ*fixB*)	$Nod^+ Fix^-$	0.0 ± 0.0^b	$Nod^+ Fix^+$	2.3 ± 0.1^c

Table 4.8 Symbiotic phenotype of complemented *fixB* mutant strains on pea (*P. sativum*). Acetylene reduction assay was carried out 28 days post inoculation, with values expressed as μmol ethylene $plant^{-1}$ hr^{-1}; ± SEM, n ≥ 3; a, b, c represent statistically distinct groups ($p \leq 0.05$) determined using one-way ANOVA and Tukey's multiple comparisons test; n.d. not determined.

The *fixC*::Ωspc mutant in this repeated assay fixed nitrogen at a rate lower but not statistically different from the wild-type. Its complemented rate was still lower than wild-type. However, due to the unusual behaviour of the omega mutant, few conclusions can be made from its complementation. Complementation of the in-frame *fixC* mutant resulted in a restored Fix^+ phenotype, though at a reduced rate of ~66% wild-type levels.

Background strain	No complement		pJP2:*fixABCX*	
	Symbiotic phenotype	**Acetylene reduction**	**Symbiotic phenotype**	**Acetylene reduction**
3841	$Nod^+ Fix^+$	3.1 ± 0.1^a	n.d	n.d
LMB827 (*fixC*::Ωspc)	$Nod^+ Fix^+$	2.5 ± 0.1^a	$Nod^+ Fix^+$	1.5 ± 0.1^c
OPS0280 (Δ*fixC*)	$Nod^+ Fix^-$	0.0 ± 0.0^b	$Nod^+ Fix^+$	2.0 ± 0.2^c

Table 4.9 Symbiotic phenotype of complemented *fixC* mutant strains on pea (*P. sativum*). Acetylene reduction assay was carried out 28 days post inoculation, with values expressed as μmol ethylene $plant^{-1}$ hr^{-1}; ± SEM, n ≥ 3; a, b, c represent statistically distinct groups ($p \leq 0.05$) determined using one-way ANOVA and Tukey's multiple comparisons test; n.d. not determined.

In contrast to the *fixB* and *fixC* mutants, a complemented *fixX* strain fixed nitrogen at a higher rate than wild-type, in a similar manner to the complemented Δ*fixAB* strain. This may mean that *fixX* is also limiting for nitrogen fixation in bacteroids. The complemented *fixX*::Ωspc mutant is unable to fix nitrogen. This may be explained by a regulatory effect; in *R. leguminosarum* bv. *viciae* UPM791 (Martinez, Palacios et al. 2004) and *R. etli* CFN42 (Benhassine, Fauvart et al. 2007) *nifA* expression originates from a site within *fixX*. So, a *fixX* omega mutant may be unable to express *nifA*. This does not explain why an in-frame deletion of *fixX* can be complemented. There is evidence for autoregulation of *nifA* from the *fixA* promoter (Martinez, Palacios et al. 2004). This autoregulation may explain why complementation is possible when a clean deletion has been made, however the omega cassette contains terminators, leading to a loss of *nifA* expression which would prevent the complementation in a *fixX*::Ωspc mutant. The regulatory elements found within the *fix* operon are discussed in Chapter 5.

Background strain	**No complement**		**pJP2:*fixABCX***	
	Symbiotic phenotype	**Acetylene reduction**	**Symbiotic phenotype**	**Acetylene reduction**
3841	$Nod^+ Fix^+$	2.5 ± 0.1^a	n.d	n.d
LMB829 (*fixX*::Ωspc)	$Nod^+ Fix^-$	0.0 ± 0.0^b	$Nod^+ Fix^+$	0.0 ± 0.0^b
OPS0469 (Δ*fixX*)	$Nod^+ Fix^-$	0.0 ± 0.0^b	$Nod^+ Fix^+$	3.0 ± 0.2^c

Table 4.10 Symbiotic phenotype of complemented *fixX* mutant strains on pea (*P. sativum*). Acetylene reduction assay was carried out 28 days post inoculation, with values expressed as μmol ethylene $plant^{-1}$ hr^{-1}; ± SEM, n ≥ 4. a, b, c represent statistically distinct groups ($p \leq 0.05$) determined using one-way ANOVA and Tukey's multiple comparisons test.

4.3 Mutation and complementation of *fixAB* in *Azorhizobium caulinodans*

Since the FixAB proteins are conserved across the rhizobia, they may be able to complement across species. *A. caulinodans* is a diazotroph able to fix nitrogen both in free-living conditions and in symbiosis with the tropical legume *Sesbania rostrata*. Previous studies (Kaminski, Norel et al. 1988) have shown that the *fixABCX* genes are essential for nitrogen

fixation in *A. caulinodans* strain ORS571 in both free-living and symbiotic conditions. Since FixABCX are expected to play the same role in ORS571 and Rlv3841 polar and in-frame *fixAB* mutants were constructed in ORS571 (OPS0501 and OPS0614 respectively) (strains and plasmids used are detailed in Chapter 2). BLAST analysis between ORS571 and Rlv3841 shows 70% identity between the FixA proteins and 65% identity between the FixB proteins (Figure 4.8).

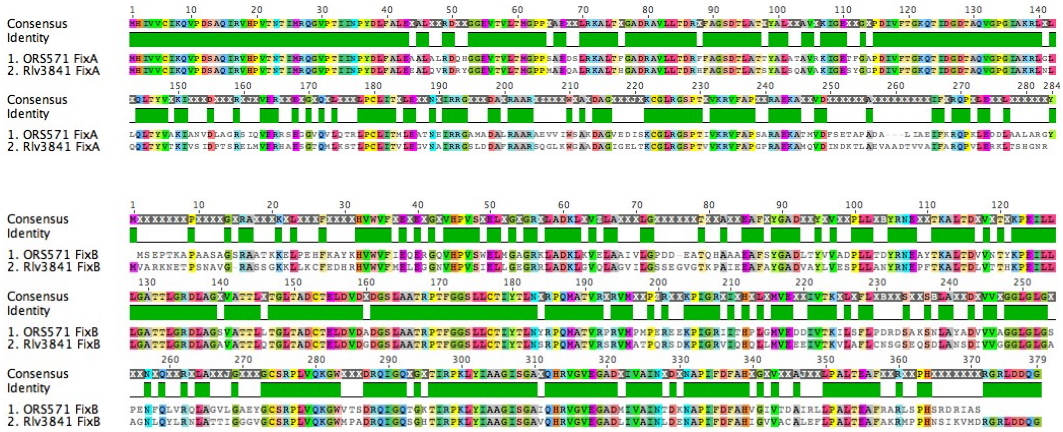


Figure 4.8 Alignment of FixA and FixB from ORS571 and Rlv3841. Alignments made in Geneious.

4.3.1 Growth and nitrogen fixation of *fixAB* mutants under microaerobic and nitrogen-limiting conditions

ORS571 is able to carry out nitrogen fixation under microaerobic conditions. All growth was carried out in an oxygen cabinet set to 3% O_2, with both aeration and shaking in order to maximise gas flow through the growing cultures. Nitrogen fixation assays were carried out after 24 hours, once exponential growth had begun (OD_{600} 0.3-0.5). ORS571 fixed nitrogen under nitrogen-starvation conditions, but not if nitrogen was supplied to the medium in the form of ammonium chloride. The *fixAB* mutants of ORS571 were unable to fix nitrogen under low oxygen, nitrogen-starved conditions. No growth phenotype was seen between the mutants and the wild-type; strains reached similar low OD_{600} ~0.3 after 24 hours, despite no nitrogen fixation in the mutant. The mutant should not be able to grow without nitrogen fixation, suggesting either a yet unknown nitrogen store, or that normal wild-type growth is unable to occur, and these strains are both in lag phase.

Strain	Description	Growth conditions	Free-living phenotype	Acetylene reduction
ORS571	Wild-type	-NH_4Cl	Fix^+	0.70 ± 0.03^a
ORS571	Wild-type	+NH_4Cl	Fix^-	0.00 ± 0.00^b
OPS0501	*fixAB*:Ωspc	-NH_4Cl	Fix^-	0.03 ± 0.00^b
OPS0614	Δ*fixAB*	-NH_4Cl	Fix^-	0.04 ± 0.00^b

Table 4.11 Nitrogen-fixation phenotype of free-living ORS571 mutant strains in UMS 20mM succinate ± 10mM NH_4Cl. Cultures were grown for 24 hours at 3% oxygen with 200rpm shaking. Acetylene reduction assay was carried out for two hours on a 10ml culture, with values expressed as µmol ethylene mg protein^{-1} hr^{-1}; ± SEM, n ≥ 9. a, b represent statistically distinct groups ($p \leq 0.05$) determined using one-way ANOVA and Tukey's multiple comparisons test.

Oxygen-limitation was also attempted in a FLUOstar Omega plate reader (BMG) under several conditions (3, 6 and 10% O_2) (Figure 4.9). Cultures were unable to grow under microaerobic nitrogen-fixing conditions, suggesting that the conditions used in the assay were not optimal for nitrogen fixation. Supplementation with nitrogen (as ammonium chloride) led to characteristic exponential growth.

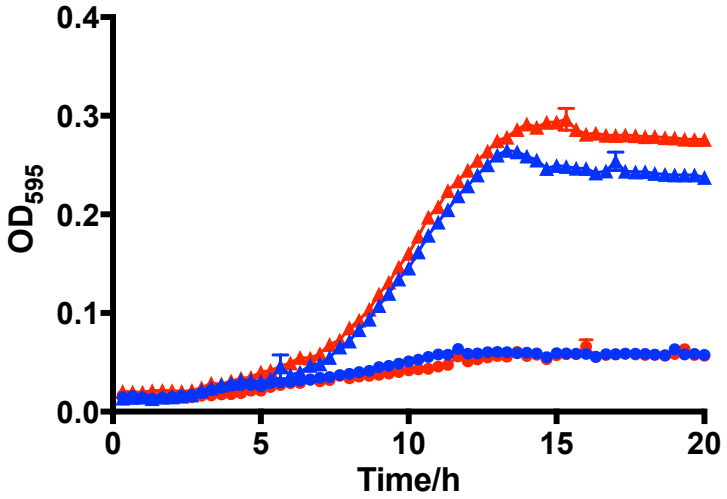


Figure 4.9 Growth of ORS571 (blue) and OPS0614 (red) in a FLUOStar Omega plate reader; growth in minimal media with (triangle) and without (circle) nitrogen (NH_4Cl) at 6% oxygen, with 500rpm orbital shaking.

4.3.2 Cross-species complementation of *fixAB* mutants

Two strategies were used in order to determine if the Rlv3841 proteins could complement *fixAB* mutants in the ORS571 background. The first was plasmids constructed by HD (formerly BD) cloning into pLMB509, which allows gene expression from a taurine-inducible promoter (Tett. A, Rudder et al. 2012). The *fixAB* genes were cloned into pLMB509 using primers pr1506/pr1507 to give pLMB779. A second construct was built using a primers oxp0259/oxp0199, containing *fixAB* with the addition of an N-terminal twin-Strep tag and a C-terminal 3xFLAG-tag (pOPS0129). Table 4.12 outlines the strains used for this complementation. To test for the activity of the taurine promoter in this system, the strains were grown at ambient oxygen at 37°C with and without 10 mM taurine. A Western blot using anti-FLAG M2 antibody showed expression of the FLAG-tagged FixB in this system (Figure 4.10).

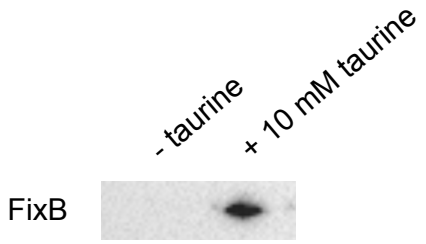


Figure 4.10 Western blot to test taurine induction of FixB in strain OPS0628 (pLMB509-2xStrep-FixA-FixB-3XFLAG in Δ*fixAB* background); anti-Flag M2 antibody used for detection. Strains grown in minimal media supplemented with 20 mM succinate,10 mM NH_4Cl, 30 μM nicotinic acid, 200 rpm.

The second strategy for complementation was *fixAB* expressed from the pJP2 vector under their native promoter, with addition of a C-terminal 3xFLAG-tag on FixB. Use of the Rlv3841 native promoter will determine if ORS571 NifA is able to work on an Rlv3841 NifA-dependent promoter. A Western blot was used to look for the presence of FixAB (Figure 4.11). A native antibody against FixA (which acts against a conserved peptide and so is able to detect both ORS571 and Rlv3841 FixA proteins) shows expression of FixA under nitrogen fixing conditions in the wild-type strain. Strain OPS0631 (pJP2:*fixAB*-3xFLAG in Δ*fixAB* background) did not show expression of any FixA protein. A second Western blot using an anti-FLAG M2 antibody showed no expression of Flag-tagged FixB in OPS0631 under nitrogen fixing conditions.

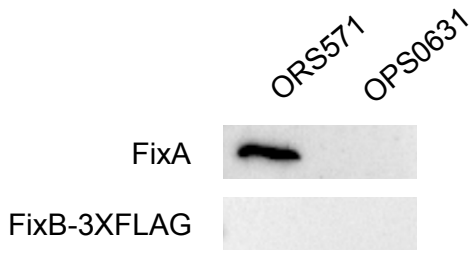


Figure 4.11 Western blot to test native induction of FixAB in strain OPS0631 (pJP2:*fixAB*-3xFLAG in Δ*fixAB* background); strains were grown at 3% O_2 for 24 hours in 50 mL cultures without nitrogen; anti-FixA antibody used for detection of FixA; anti-FLAG M2 antibody used for detection of FLAG-tagged FixB.

All complemented strains were grown under nitrogen fixation conditions to test whether the Rlv3841 proteins could complement an ORS571 mutant.

No expression was seen from the Rlv3841 native promoter. Strains containing plasmid pOPS0124 (pJP2:*fixAB*-3xFLAG) were not affected in nitrogen fixation ability relative to their background strain.

In order to test the effect of taurine under nitrogen fixation conditions, the wild-type ORS571 was grown at 3% O_2 in media supplemented with 10 mM taurine. Nitrogen fixation ability was not affected by the presence of taurine. Strains containing taurine inducible plasmids were then grown under nitrogen fixing conditions in the presence of taurine. These strains were unable to grow, suggesting a toxic effect of Rlv3841 FixAB under nitrogen-fixing conditions. Toxicity was seen in both the wild-type and *fixAB* mutant background. This toxic effect was not seen when these strains were grown under ambient oxygen; as shown above the FixAB proteins can be expressed under the control of taurine at 21% oxygen.

Strain	Back-ground	Plasmid	Description	Free-living phenotype	Acetylene reduction
OPS0846	ORS571	pLMB779	*fixAB* under taurine promoter	No growth	n.d.
OPS0847	OPS0614	pLMB779	*fixAB* under taurine promoter	No growth	n.d.
OPS0626	ORS571	pOPS0129	2xStrep-*fixAB*-3xFLAG under taurine promoter	No growth	n.d.
OPS0628	OPS0614	pOPS0129	2xStrep-*fixAB*-3xFLAG under taurine promoter	No growth	n.d.
OPS0623	ORS571	pOPS0124	*fixAB*-3xFLAG under native promoter	Fix^+	0.96 ± 0.02^a
OPS0631	OPS0614	pOPS0124	*fixAB*-3xFLAG under native promoter	Fix^-	0.03 ± 0.00^b

Table 4.12 Nitrogen-fixation phenotype of free-living ORS571 mutant strains grown in UMS 20mM succinate. Cultures were grown for 24 hours at 3% oxygen with 200rpm shaking. No nitrogen source was added to the media. Acetylene reduction assay was carried out for two hours on a 10ml culture, with values expressed as μmol ethylene mg $protein^{-1}$ hr^{-1}; n.d. not determined; ± SEM, $n \geq 9$; a, b represent statistically distinct groups ($p \leq 0.05$) determined using one-way ANOVA and Tukey's multiple comparisons test.

4.4 Identification of other genes involved in bacteroid electron transport

4.4.1 Identification of a putative bacteroid ferredoxin

Herrmann, Jayamani et al. (2008) hypothesised that FixX acts in a heterodimer with FixC, whilst an additional ferredoxin takes the role of carrying electrons to nitrogenase. There are 13 proteins in Rlv3841 annotated as ferrexoxins, with three (including FixX) encoded on the symbiotic plasmid pRL10.

To investigate whether other ferredoxins are involved in nitrogen fixation, the expression of the 13 putative ferredoxins was investigated. Microarray analysis of Rlv3841 (Karunakaran, Ramachandran et al. 2009) provided values for gene expression in symbiosis with *P. sativum* at several time points post-inoculation (Table 4.13). Analysis of all the putative ferredoxins showed that a ferredoxin found on the symbiotic plasmid pRL10, pRL100156, is highly upregulated in mature bacteroids, with expression nearly 20-fold increased in 21-day bacteroids compared to expression in the free-living state. *fdxB1* (pRL100156) is proceeded by a consensus site for NifA, the general transcriptional regulator of nitrogen fixation, strengthening the hypothesis that it is involved in this symbiosis.

Several other symbiotic diazotrophs have a ferredoxin-like protein encoded by *fdxN* found immediately downstream of *fixABCXnifAB*. In *R. meliloti* and *B. japonicum* FdxN is essential for nitrogen fixation (Klipp, Reilander et al. 1989, Hauser, Pessi et al. 2007). However, a BLAST of the *R. meliloti* FdxN against *R. leguminosarum* bv. *viciae* 3841 found no obvious orthologues.

Locus tag	Gene	Annotation	Fold-induction	
			21d	28d
RL0326		Putative 4Fe-4S ferredoxin protein	0.76	1.07
RL2369		Putative NADPH:ferredoxin reductase	1.57	1.03
RL2652	*fdxB2*	Putative ferredoxin, 2Fe-2S	0.54	0.96
RL4485		Putative ferredoxin containing dehydrogenase	0.78	0.93
fRL4616	*fdxA*	Putative ferredoxin II	1.26	0.82
pRL100075		Putative ferredoxin	2.26	1.11
pRL100156	*fdxB1*	Putative ferredoxin	17.85	9.46
pRL100197	*fixX*	Ferredoxin-like protein FixX	29.44	17.21
pRL110029		Putative ferredoxin	1.35	0.96
pRL110471		Putative flavoprotein/ferredoxin	1.24	1
pRL110621		Putative Rieske-type ferredoxin	1.37	1.42
pRL120179		Putative pyruvate ferredoxin oxidoreductase	0.9	1.2
pRL120537		Putative FAD/NAD/ferredoxin protein	0.89	1.14

Table 4.13 Putative ferredoxins found in Rlv3841. Fold-induction shown for 21 and 28 days post-inoculation of peas relative to free-living Rlv3841. Values >2-fold upregulated shown in green. Microarray data was obtained from Karunakaran, Ramachandran et al. (2009).

The FixX and FdxB1 proteins show little homology (Figure 4.12), so it is unlikely that they play the same role in the bacteroid or that the *fdxB1* gene is the result of a duplication of *fixX*.

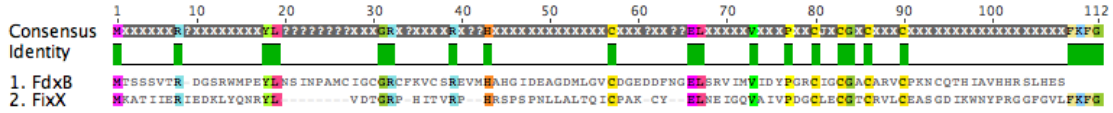


Figure 4.12 Alignment of the FixX and FdxB1 proteins. Alignment constructed using Geneious.

To investigate if FdxB1 plays an essential role in bacteroids, a stable mutant was constructed with an omega-spectinomycin cassette. The *fdxB1::*Ωspc mutant was designated OPS0500. Additionally, a *fixX fdxB1* double mutant (OPS0536) was constructed by conjugating the *fdxB1*::Ωspc pJQ200SK plasmid into OPS0469 (*ΔfixX)*.

A growth assay of OPS0500 on rich medium showed no difference in free-living growth between Rlv3841 and an *fdxB1* mutant; FdxB1 is not essential for free-living growth.

Strain	**Description**	**Doubling time /hours**
3841	*R. leguminosarum* bv. *viciae* 3841	3.87
OPS0500	*fdxB1*::Ωspc	3.63

Table 4.14 Doubling time of *fdxB1* mutant in rich media in an Omega FluoSTAR plate reader; 400 rpm double orbital shaking.

The symbiotic phenotype of OPS0500 (*fdxB1*::Ωspc) resembled that of wild-type, with healthy green plants and pink nodules. An acetylene reduction assay showed that an *fdxB1* mutant had wild-type nitrogenase activity. FdxB1 is therefore not essential for nitrogen fixation in symbiosis with pea. An *fixX fdxB1* double mutant was unable to fix nitrogen in symbiosis and the plant had stunted growth with white nodules.

Strain	Description	Symbiotic phenotype	Nodule colour	Acetylene reduction μmol ethylene $plant^{-1}$ hr^{-1}
3841	Wild-type	Nod^+ Fix^+	Pink	3.5 ± 0.3
OPS0500	*fdxB1*::Ωspc	Nod^+ Fix^+	Pink	3.6 ± 0.1
OPS0536	Δ*fixX fdxB1*::Ωspc	Nod^+ Fix^-	White	0.0 ± 0.0
Uninoculated control	n/a	Nod^- Fix^-	n/a	0.0 ± 0.0

Table 4.15 Symbiotic phenotype of *fdxB1* mutant strains on pea (*P. sativum*). Acetylene reduction assay was carried out 28 days post inoculation; ± SEM, n ≥ 3.

4.4.2 Identification of a menaquinone biosynthesis gene

Xie, Cheng et al. (2011) identified a bacteroid-specific electron carrier menaquinone, required for effective symbiosis of *Mesorhizobium huakuii* 7653R with *Astragalus sinicus*. The gene *dmtH* was identified by screening of a transposon-insertion mutant library., and encodes a dimethylmenaquinone methyltransferase catalyzing the final step in the menaquinone biosynthesis pathway. FixCX is hypothesised to pass electrons to the ubiquinone pool in bacteroids, but Xie, Cheng et al. (2011) suggest that menaquinone may play the role of ubiquinone in *M. huakuii* bacteroids. To identify if this is the case in *R. leguminosarum*, putative genes involved in menaquinone biosynthesis were identified. A BLAST against the 7653R *dmtH* gene product identified one protein with 89% identity. Additionally, three putative methyltransferases were identified from the Rlv3841 gene annotation. Expression changes in bacteroids were investigated in microarrays. RL3587 showed 2/3-fold and 3/4-fold expression in 21- and 28-day bacteroids respectively compared to the free-living state. The other three genes were upregulated in bacteroids, though all less than two-fold.

Locus tag	Gene	Annotation	Fold induction		% homology between gene product and *M. haukuii* DmtH
			21d	28d	
RL0371	*ubiE*	Ubiquinone/menaquinone biosynthesis C-methyltransferase	1.75	1.74	9.7%
RL1838	*ubiE*	Putative ubiquinone/menaquinone biosynthesis methyltransferase	1.69	1.36	16.3%
RL2712		Putative ubiquinone/menaquinone biosynthesis methyltransferase	1.06	1.24	11.4%
RL3587		Putative polysaccharide deacetylase family protein	0.66	0.75	89.2%

Table 4.16 Putative menaquinone methyltransferases in 3841. Fold-induction shown for 21- and 28-days post inoculation of peas relative to free-living Rlv3841 grown on glucose. Microarray data was obtained from Karunakaran, Ramachandran et al. (2009).

Alignments were made of these putative methyltransferases along with bacterial homologues identified by Xie, Cheng et al. (2011) (Figure 4.13). A tree of these alignments was generated to determine which of the Rlv3841 homologues was most likely to act as a menaquinone methytransferase (Figure 4.14). Only the RL3587 gene product showed clear homology to the *M. huakii* DmtH. Additionally, only RL3587 clusters on the same branch as *dmtH*, suggesting that it is the homolog found in Rlv3841.

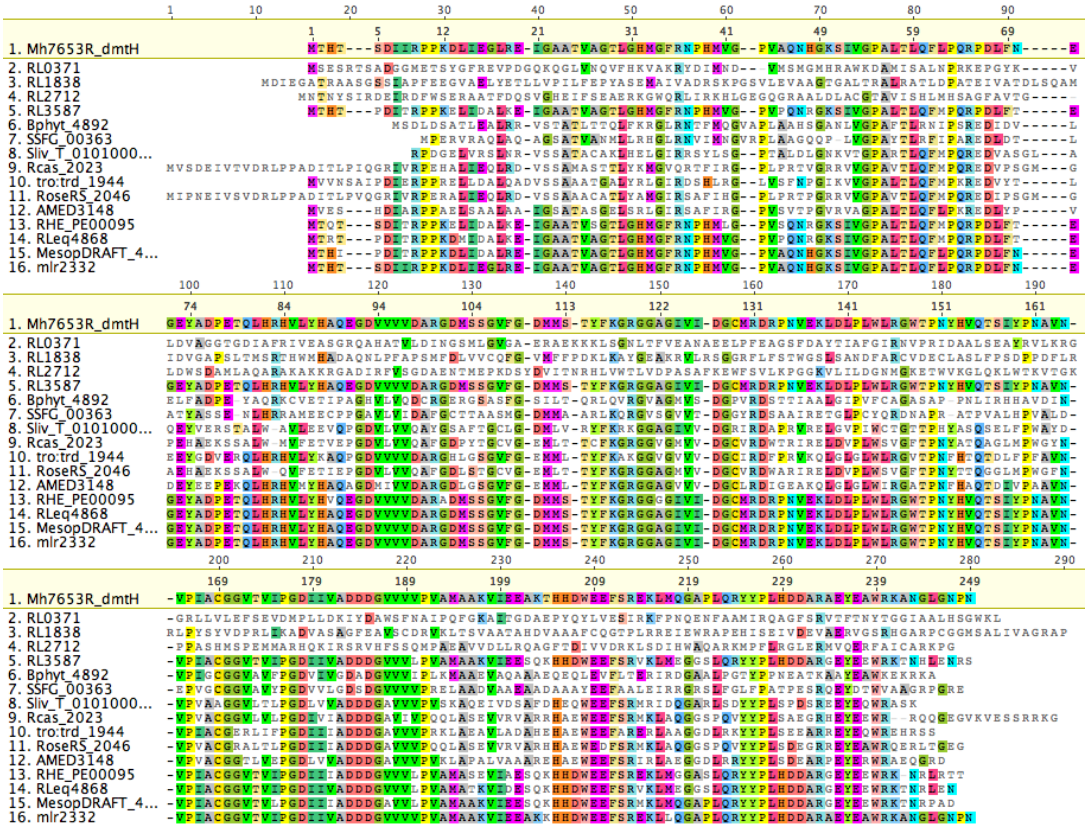


Figure 4.13 Alignment of putative menaquinone methyltransferases from Rlv3841 with *M. huakuii* DmtH and bacterial homologs identified by Xie, Cheng et al. (2011). Alignment generated using Geneious.

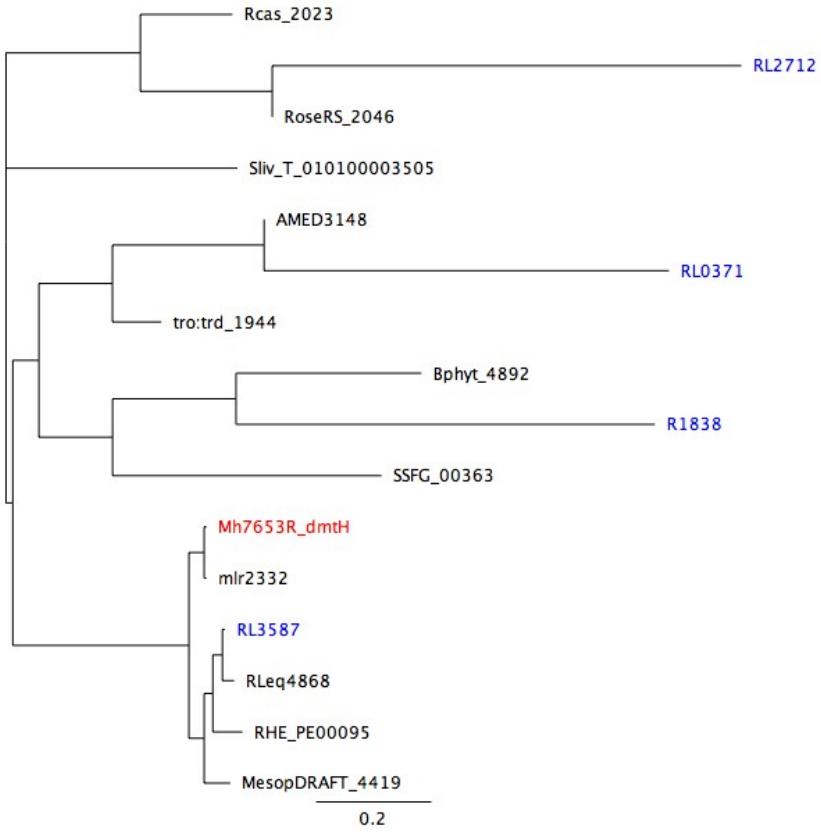


Figure 4.14 Tree of putative menaquinone methyltransferases from Rlv3841 (blue) and *M. huakuii* DmtH (red) with bacterial homologs identified by Xie, Cheng et al. (2011). Tree generated by Geneious.

A Golden Gate based strategy was chosen to construct a *dmtH* mutant. Primers oxp0600 and oxp0601 were designed to amplify the left hand flanking region of RL3587. The right hand flanking region contained a single *BsaI* site, and so a version of the flanking region was constructed which removed this *BsaI* site. This involved introducing a single nucleotide change that conserves the amino acid sequence whilst removing the enzyme recognition site. This right hand fragment was synthesized by GeneWiz® for use in the Golden Gate reaction. A Golden Gate assembly was carried out using the left and right hand flanking regions and a kan/lox cassette module (pOGG029). These were assembled into level 1 Golden Gate vector pOGG028, a domesticated version of the pJQ200SK vector. Whilst the vector was able to assemble correctly, the subsequent conjugation steps resulted in colonies that retained gentamycin resistance after sucrose selection. Restriction digest was used to remove the kan/lox cassette and replace it with a conventional omega cassette, but the plasmid was still retained after sucrose selection. The same phenomenon was seen in other attempts to generate mutants with pOGG028 and the kan/lox cassette. A stable RL3587 mutant was

therefore never constructed. This work was carried with kanamycin at 80 μg ml^{-1}. Subsequent work using this cassette in the Poole lab has suggested that lower kanamycin concentrations may lead to successful mutants.

A pK19 insertion mutant was constructed. Primers oxp1204 and oxp1205 were used to amplify a 300 bp region within the *dmtH* gene. This fragment was cloned into pK19mob *HindIII* using HD InFusion cloning (ClonTech) resulting in plasmid pOPS0347. This plasmid was conjugated into 3841 resulting in strain OPS0796. Strain OPS0796 was inoculated on peas and nitrogenase activity measured by acetylene reduction. Plants were Fix^+ with nodules resembling that of the wild-type.

Strain	**Description**	**Symbiotic phenotype**	**Nodule colour**	**Acetylene reduction** μmol ethylene $plant^{-1}$ hr^{-1}
3841	Wild-type	Nod^+ Fix^+	Pink	2.5 ± 0.1^a
OPS0796	*dmtH::*pK19mob	Nod^+ Fix^+	Pink	2.7 ± 0.2^a
Uninoculated control	n/a	Nod^- Fix^-	n/a	0.0 ± 0.0^b

Table 4.17 Symbiotic phenotype of *dmtH* mutant strain on pea (*P. sativum*). Acetylene reduction assay was carried out 28 days post inoculation; ± SEM, $n \geq 3$; a, b represent statistically distinct groups ($p \leq 0.05$) determined using one-way ANOVA and Tukey's multiple comparisons test.

4.5 Discussion

The *fixABCX* genes are conserved across the rhizobia. Mutants have been constructed and characterised in *R. leguminosarum* bv. *viciae* across the *fix* operon to confirm the essential role for all Fix proteins in the *Rhizobium*-legume symbiosis. In-frame deletion mutants constructed in *fixAB*, *fixB*, *fixC*, and *fixX* showed a clear symbiotic phenotype lacking detectable nitrogenase activity. Plants inoculated with non-fixing strains displayed a stunted, yellowing phenotype characteristic of nitrogen deficiency. Nodules of plants inoculated with *fix* mutants were white and distinctly smaller than in the wild-type, indicating that the plant has stopped providing leghaemaglobin to maintain a microaerobic environment. These nodules lacking leghaemaglobin suggest a feedback mechanism, where the plant detects a lack of fixed nitrogen and responds by reducing its own input into the partnership. Nodule sections viewed under light microscopy also indicate a lack of bacteroid occupancy within the nodule, suggesting additional mechanisms blocking development of symbiosomes. Bacteroid morphology in these mutants is further discussed in Chapter 5. This lack of nitrogen fixation ability supports the hypothesis that the FixABCX proteins carry electrons to the nitrogenase enzyme; the nitrogen fixation reaction requires eight electrons per molecule of dinitrogen fixed. As expected, there was no growth phenotype seen in these mutants in the free-living state; the *fixABCX* operon is not required for growth in free-living culture.

In-frame *fix* deletion mutants could be complemented by *fixABCX* under its native promoter to return nitrogen fixation ability within the symbiosis. The return of nitrogenase activity is accompanied by restoration of plant phenotype to green leaves and pink nodules. Complementation of *fixAB* returned mutants to varying levels of nitrogen fixation. An in-frame Δ*fixAB* mutant could be complemented by *fixAB* under the control of a native promoter, but fixation only returned to ~50% of wild-type fixation rates. This may be explained by the differing expression of *fixAB* vs *fixCX* due to expression from the native plasmid compared to expression from a multi-copy plasmid. When the same strain or the wild-type was complemented with the whole *fixABCX* operon on a multi-copy plasmid, nitrogen fixation rates were higher than wild-type; this indicates that *fixABCX* expression may be a limiting factor in nitrogen fixation. However, when the polar *fixAB*::Ωspc mutant was complemented by *fixABCX* it fixed at lower levels than wild-type. In *R. leguminosarum* bv. *viciae* it has been shown that *nifA* (found immediately downstream of *fixABCX*) is under the control of a promoter found upstream of *fixA*, with only a basal promoter found immediately upstream of the *nifA* gene. This has been shown in multiple strains, including 3841 (Chapter 6, this work) and UPM791 (Martinez, Palacios et al. 2004). If this were so in

R. leguminosarum bv. *viciae* 3841, the spectinomycin cassette would block the transcription of *nifA* from the upstream promoter. Since NifA is the general transcriptional regulator of nitrogen fixation, a reduction in its expression would lead to reduced fixation rates. Complementation of *fixB* led to a similar reduced fixation rate, as did complementation of the in-frame *fixC* mutant.

Whilst the in-frame *fixX* could be complemented with *fixABCX*, the polar mutant of *fixX*::Ωspc could not be complemented. This may be due to the autoregulation of NifA mentioned above. The effect of mutation on *nifA* gene expression is discussed in Chapter 6.

An in-frame deletion of *fixC* and a pK19mob interposon in *fixC* both showed that FixC is essential for nitrogen fixation. However, a polar *fixC*::Ωspc mutant did not display the expected phenotype, as it was able to fix nitrogen at lower levels than wild-type (70%-80%). Additionally, it had a phenotype mid-way between the wild-type and a Fix^- mutant. Attempts at mapping this mutant suggested the presence of a partial duplication leading to a restored phenotype. Whilst this mutant was assumed incorrect, it has become useful as a stable mutant with lowered nitrogen fixation without effect on other processes within the bacteroid. This reduced fixation mutant has potential for better understanding of competition and nutrient allocation within microbial ecology. This strain has now been sent for whole-genome sequencing, and the results may provide new insights into its Fix^{red} phenotype.

In addition to *R. leguminosarum*, mutants in *fixAB* were also constructed in *A. caulinodans*, a diazotroph able to fix nitrogen both in symbiosis and in microaerobic free-living growth (Dreyfus, Elmerich et al. 1983). Growth conditions were optimised for nitrogen fixation in strain ORS571. Optimal fixation was only obtained using both sparging and shaking in baffled flasks. This suggests a high demand for dissolved oxygen during nitrogen fixation, despite the requirement for microaerobic growth.

No nitrogen fixation was seen in *fixAB* mutants in the ORS571 background. This confirms work by Kaminski, Norel et al. (1988) characterising the nitrogen-fixation requirement for the *fix* operon in this strain. When the wild-type ORS571 and these mutants were grown under nitrogen fixing conditions, no growth phenotype was seen; similar optical densities were reached after 24 hours. It was thought that the nicotinic acid may be providing enough nitrogen for growth during nitrogen fixation. Nicotinic acid must be supplemented to the media because *Azorhizobium spp.* are auxotrophic for NAD^+ biosynthesis (Elmerich, Dreyfus et al. 1983, Kitts, Lapointe et al. 1992). However nicotinic acid was only added to media at a concentration of 30 μM, so this is unlikely to be the reason for growth in Fix^- mutants. It may be that *A. caulinodans* stores nitrogen in some form, which it is able to use

as a nitrogen source to supplement nitrogen fixation under nitrogen limiting conditions. These mutants have not yet been tested in symbiosis with legume partner *S. rostrata*.

A. caulinodans mutants were constructed to investigate cross-species complementation using Rlv3841 FixAB. BLAST analysis shows 65-70% identity between the ORS571 and Rlv3841 FixAB proteins. Two strategies were used for expression of the Rlv3841 *fixAB* genes; control under the Rlv3841 native promoter and inducible control under a taurine promoter.

The native promoter complement was unable to express Rlv3841 *fixAB* under either aerobic or microaerobic, nitrogen-limited conditions. Both strains have *fixAB* under the control of NifA, there may not be enough structural similarity between Rlv3841 NifA and ORS571 NifA to express from the Rlv3841 promoter. BLAST analysis shows only 45% identity between these two proteins. Identity between these proteins is particularly low in the region of the GAF domain, which is involved in regulation of the AAA+ domain. The AAA+ domain, which is better conserved between these two strains, is an ATPase, which catalyses ATP hydrolysis to allow RNA polymerase to transition to an open promoter complex(Martinez-Argudo, Little et al. 2004). The lack of identity between the GAF domains suggests that the lack of functionality in this complementation may be due to a lack of regulatory function.

The taurine-inducible vector pLMB509 has been well characterised for expression in *R. leguminosarum* (Tett, Rudder et al. 2012), but has not been well tested in *A. caulinodans*. Growth with 10 mM taurine under free-living conditions showed that the taurine promoter was effective in *A. caulinodans.* Growth of complemented mutants under taurine control led to toxic effects; cultures did not grow in microaerobic conditions when 10 mM taurine was added to the media. This suggests that expression of Rlv3841 *fixAB* is toxic only under nitrogen-fixing conditions; the foreign FixAB proteins may interact unfavourably with other machinery expressed during nitrogen fixation.

Future experiments would need to confirm complementation of these mutants with the native ORS571 *fixAB* under its own promoter. Additionally, the Rlv3841 FixAB proteins could be tested under the ORS571 native promoter to confirm the phenotype seen with the taurine-inducible strains. This would rule out the possibility that overexpression of the Rlv3841 *fixAB* is not the cause for toxicity. Additionally, complementation experiments could be tested in symbiosis with the *A. caulinodans* legume partner, *S. rostrata.*

As well as being a subunit of ETF-QO, FixX is also a ferredoxin like protein. Ferredoxins are iron-sulphur proteins involved in electron transfer in many processes in cells, as well as contributing to catalytic function as part of larger complexes (Bruschi and Guerlesquin

1988). Herrmann, Jayamani et al. (2008) hypothesised that a second ferredoxin is involved in nitrogen fixation, playing the role of carrying electrons to nitrogenase whilst FixX acts in a heterodimer with FixC as an ETF-QO. In consideration of this *R. leguminosarum* bv. *viciae* 3841 was investigated to determine if a second ferredoxin may be involved in this system. There is now a wealth of genetic information available for this strain, including microarray and RNAseq transcriptomic data and a fully annotated genome, allowing for efficient searching of potential ferredoxins for this role. Thirteen ferredoxins were identified in Rlv3841, and of these one was selected due to its high expression in bacteroids relative to the free-living state. A stable mutant was constructed in *fdxB1* and its phenotype characterised. The *fdxB1* mutant OPS0500 had wild-type nitrogen fixation levels and so FdxB1 is not essential for nitrogen fixation. This result does not eliminate the possibility of a second ferredoxin involved in nitrogen fixation, but more genetic information may be required before this ferredoxin, if it exists, is identified, or more biochemical characterisation of the role of FixX may answer this question.

Another potential player in biological nitrogen fixation in this system is menaquinone. It is generally assumed that ubiquinone plays the role of electron transport quinone in bacteroids. A recent transcriptomic study in *M. haukuii* identified that a menaquinone biosynthesis gene was required for nitrogen fixation in association with its legume host *Astralagus sinicus* (Xie, Cheng et al. 2011). In order to build a more complete model of the electron transfer machinery required in *R. leguminosarum* bv. *viciae,* a homologue of the menaquinone methyltransferase was identified and a mutant constructed. The homologue was identified using BLAST analysis combined with microarray data, genome annotation and comparison to homologues identified by Xie, Cheng et al. (2011). Initial attempts to construct a stable mutant were unsuccessful, though an interposon mutant was constructed using pK19mob. An acetylene reduction assay on this interposon mutant showed no symbiotic phenotype; the menaquinone dimethyltransferase encoded by RL3587 is not essential for nitrogen fixation. This suggests that menaquinone is not the sole quinone involved in symbiotic electron transfer.

CHAPTER 5

Developmental and redox effects of *fixAB* mutation

5.1 Introduction

Mutation of the *fixABCX* genes disrupts symbiotic interactions with plants, resulting in a Fix$^-$ phenotype. Nitrogen fixation is an energy-intensive process, requiring sixteen molecules of ATP per molecule of nitrogen fixed. Additionally, this process requires eight electrons per molecule of nitrogen fixed. It could be hypothesised that interruption of this process could lead to massive changes in redox and energy status in the cell.

5.2 Microscopic phenotype of *fixAB* mutants

In order to fully understand the phenotype of mutants LMB771 (*fixAB*::Ωspc) and LMB777 (Δ*fixAB*), nodule sections were fixed for both light and transmission electron microscopy. In nodules hosting wild-type bacteroids most plant cells are uniformly and densely occupied by bacteroids. In nodules hosting the *fixAB* mutants fewer plant cells are occupied by bacteroids, with large variations in the density of occupation by bacteroids (Figure 5.1).

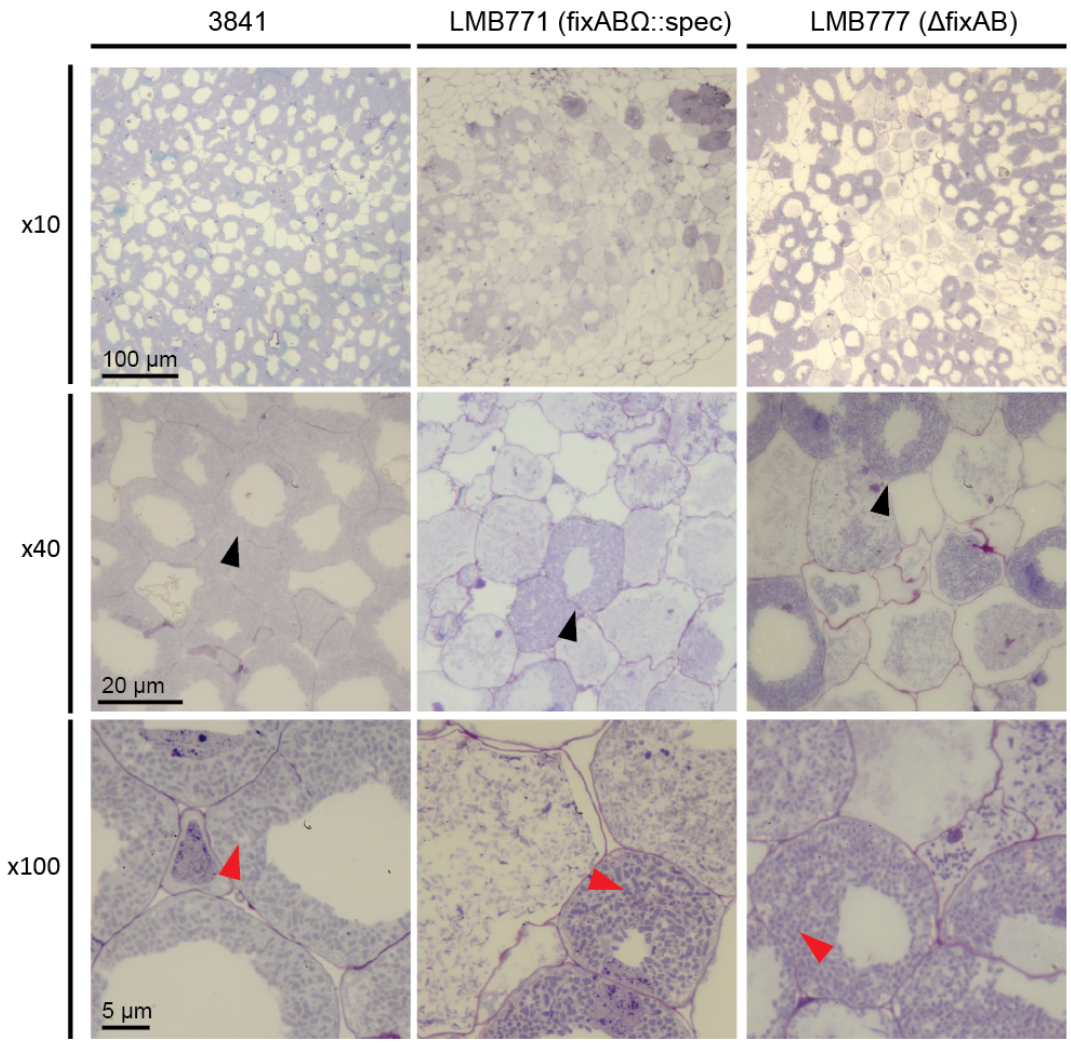


Figure 5.1 Sections of nodules taken from P. sativum inoculated with Rlv3841, LMB771 (*fixAB*::Ωspc) and LMB777 (Δ*fixAB*) at 28 dpi. Black arrows indicate plant cells; Red arrows indicate individual bacteroids. Sections stained with toluidine blue. Visualised by light microscopy at magnification x10, x40 and x100.

Transmission electron microscopy (TEM) was used to determine the morphological changes of bacteroids seen in Fix$^-$ mutants. TEM allows analysis of ultrastructural elements of the bacterial up to the single-cell level. TEM of wild-type nodule sections allow visualisation of the typical Y-shaped bacteroids as well as infection threads carrying undifferentiated bacteria into the infection thread (Lodwig, Leonard et al. 2005). PHB is more prominent in growing infection threads, which also contain small electron-dense granules, hypothesised to be organic polyphosphate (also known as volutin) (Jensen 1968, Craig and Williamson 1972,

Bode, Mauch et al. 1993, Sr, Knebel et al. 2005). A typical section of a mature wild-type nodule is shown in Figure 5.2.

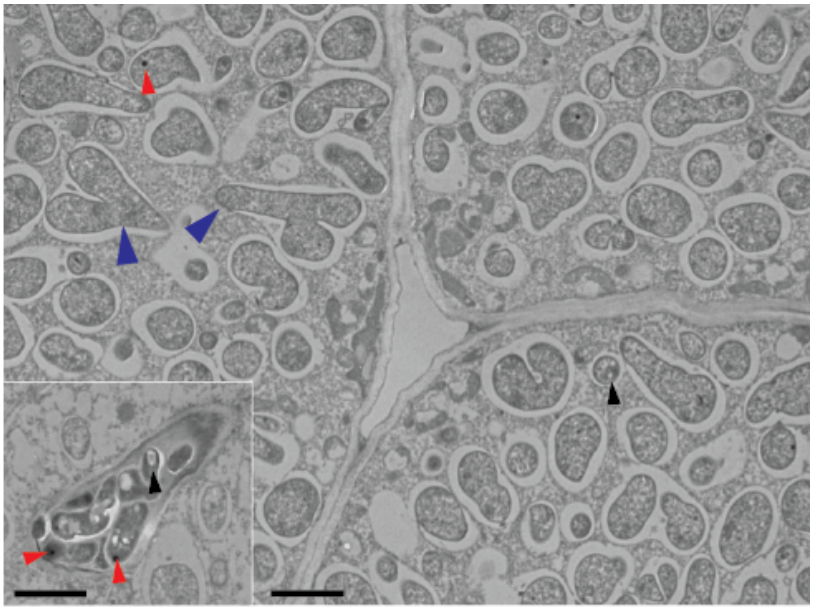


Figure 5.2 Sections of nodules taken from *P. sativum* inoculated with Rlv3841 at 28 dpi; Insert shows an infection thread. Blue arrows indicate typical Y-shaped bacteroid morphology; black arrows indicate PHB granules; red arrows indicate polyphosphate granules. Scale bars 2 µm. Visualised by TEM at x1700. EM slides were prepared and imaged by E. Barclay of JIC Bioimaging.

Low magnification TEM images of nodules show the level of nodule occupancy in the wild-type vs. a *fixAB* mutant (Figure 5.3). As seen in the light micrographs, fewer plant cells are occupied by bacteroids in the *fixAB* mutants.

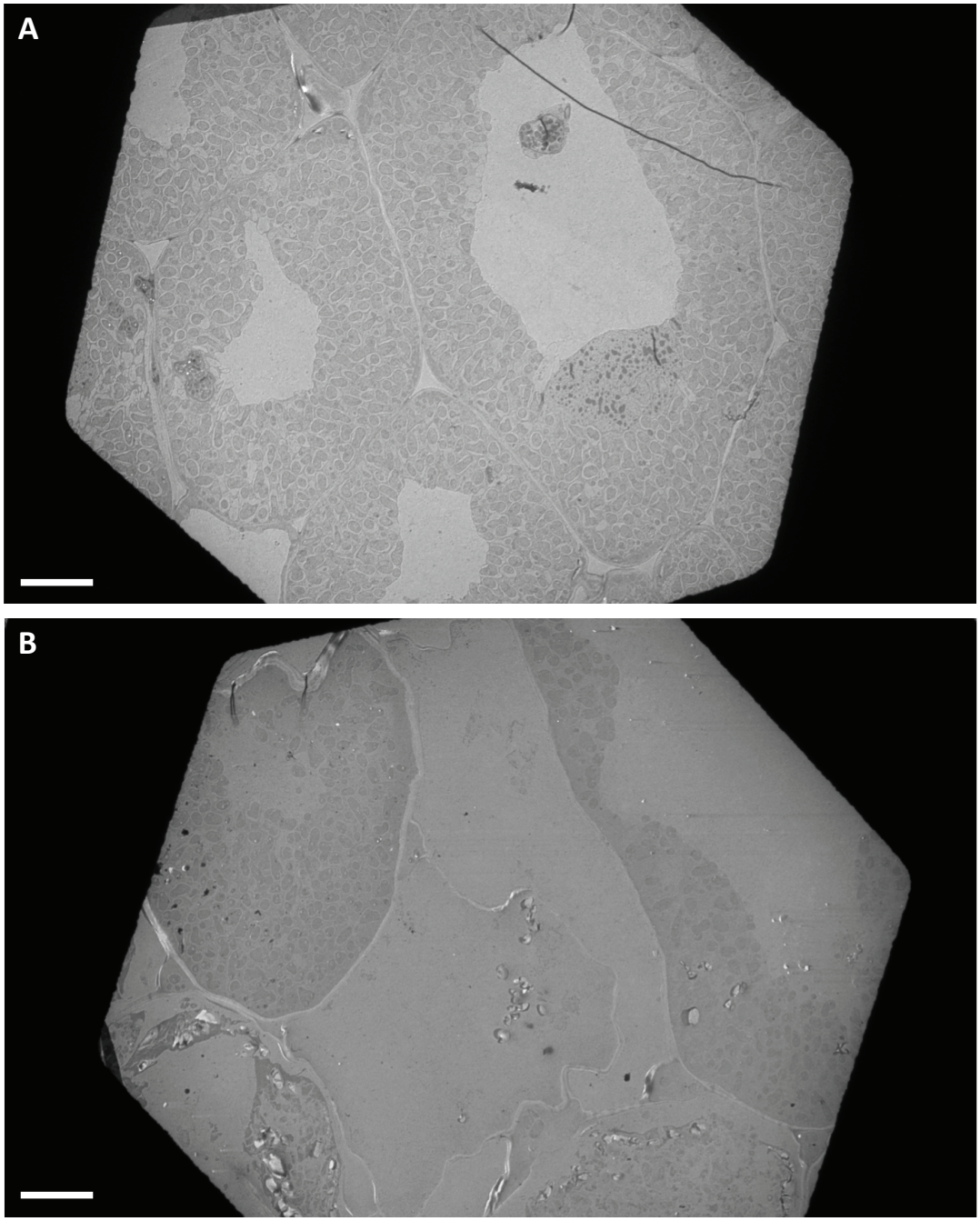


Figure 5.3 Sections of nodules taken from *P. sativum* inoculated with Rlv3841 (A) and LMB771 (B) at 26 dpi. Scale bar 10 μm. Visualised by TEM at x330. EM slides were prepared and imaged by E. Barclay of JIC Bioimaging.

5.2.1 Morphology of *fixAB* mutant bacteroids.

Ultrastructural analysis of *fixAB* mutants reveals drastic morphological changes in bacteroids. In mature nodules (26-28 dpi) two distinct morphologies can be found, indicating two separate fates for cells that cannot fix nitrogen. The first morphology is of PHB-filled bacteroids, which tended to be swollen and larger than those in the wild-type, often losing their Y-shaped form in favour of more irregular shapes (Figure 5.4 C, D, G, H). The second morphology is of smaller bacteroids with electron-dense regions thought to be polyphosphate (Figure 5.4 E, F, I, J). These polyphosphate (polyP) inclusions appear to be positioned at the poles of the bacteroid, in a clearly defined spherical granule (Figure 5.5). High levels of PHB or polyP tend to be associated with the infection thread in the wild-type. Both phenotypes can be seen within the same nodule and neighbouring plant cells can be filled with bacteroids of these different morphologies (Figure 5.6). Typically within a mutant nodule more PHB-phenotype plant cells can be found than polyP-phenotype plant cells. Analysis of cell sizes using ImageJ shows statistically different sizes between these cell types. Comparison with wild-type shows that the smaller, polyp-filled cells are statistically the same size as the wild-type infection thread cells (Figure 5.7).

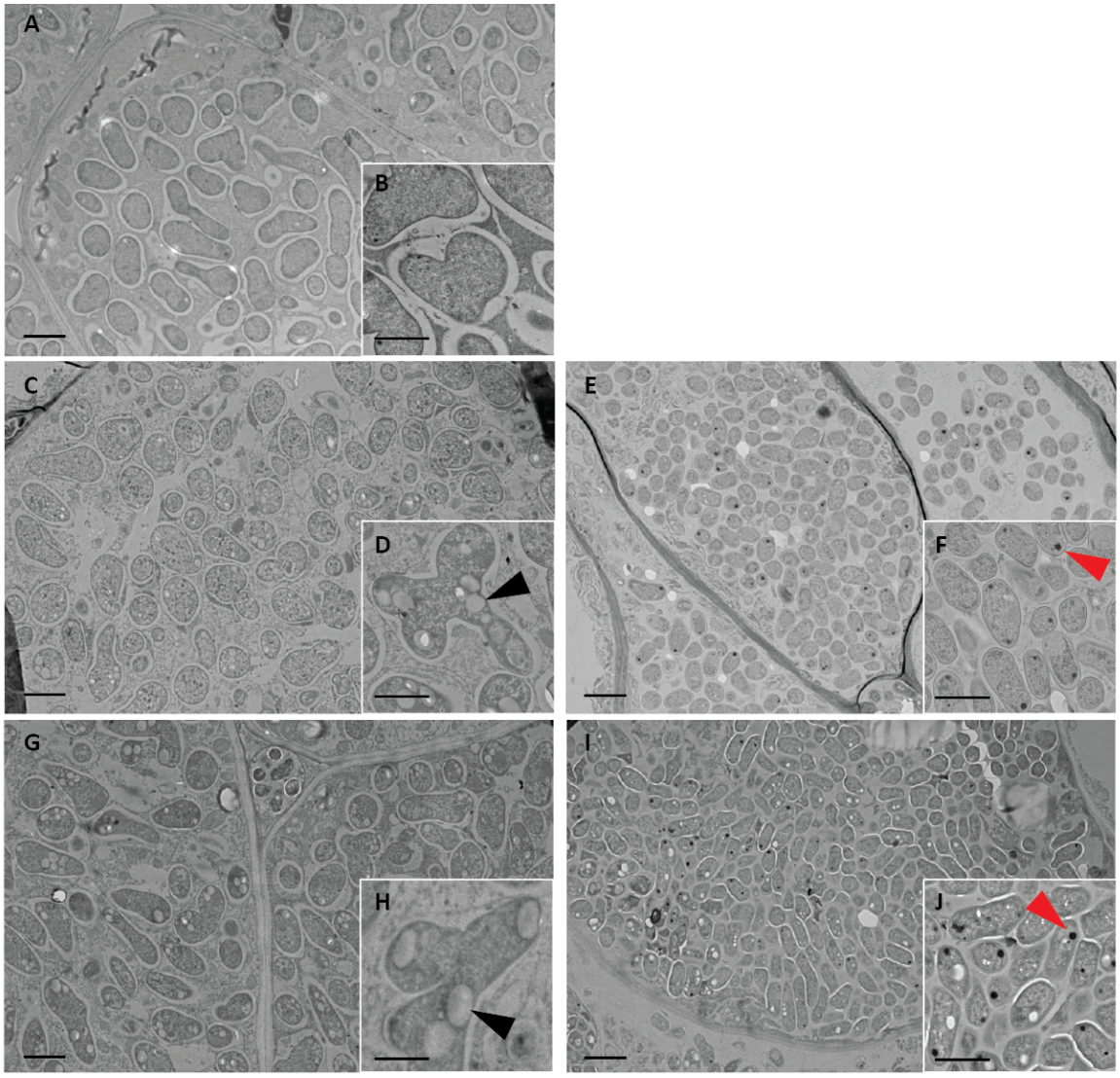


Figure 5.4 Sections of nodules taken from *P. sativum* inoculated with Rlv3841 (A,B); LMB771 (C, D, E, F); LMB777 (G, H, I, J) at 28 dpi. Black arrows indicate PHB granules; Red arrows indicate polyphosphate granules. Scale bars 2 µm (A, C, E, G, I) and 1 µm (B, D, F, H, J). Visualised by TEM at x1700 (A, C, E, G, I) and x3500 (B, D, F, H, J). EM slides were prepared and imaged by E. Barclay of JIC Bioimaging.

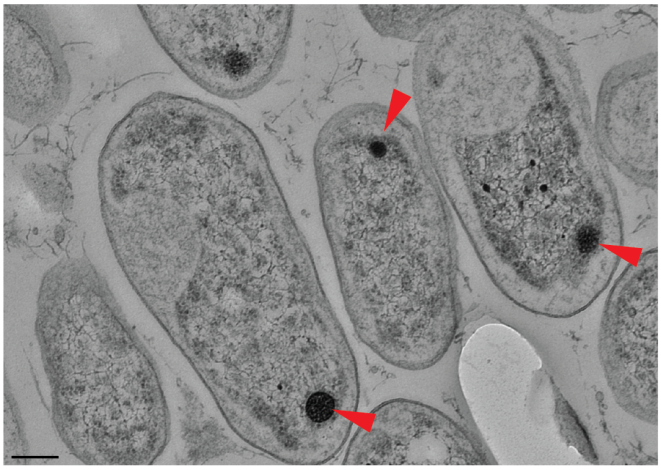


Figure 5.5 Section of a nodule taken from *P. sativum* inoculated with LMB771 (*fixAB*::Ωspc) at 28 dpi. Red arrows indicate polar polyphosphate granules. Scale bar 200 nm. Visualised by TEM at x14500. EM slides were prepared and imaged by E. Barclay of JIC Bioimaging.

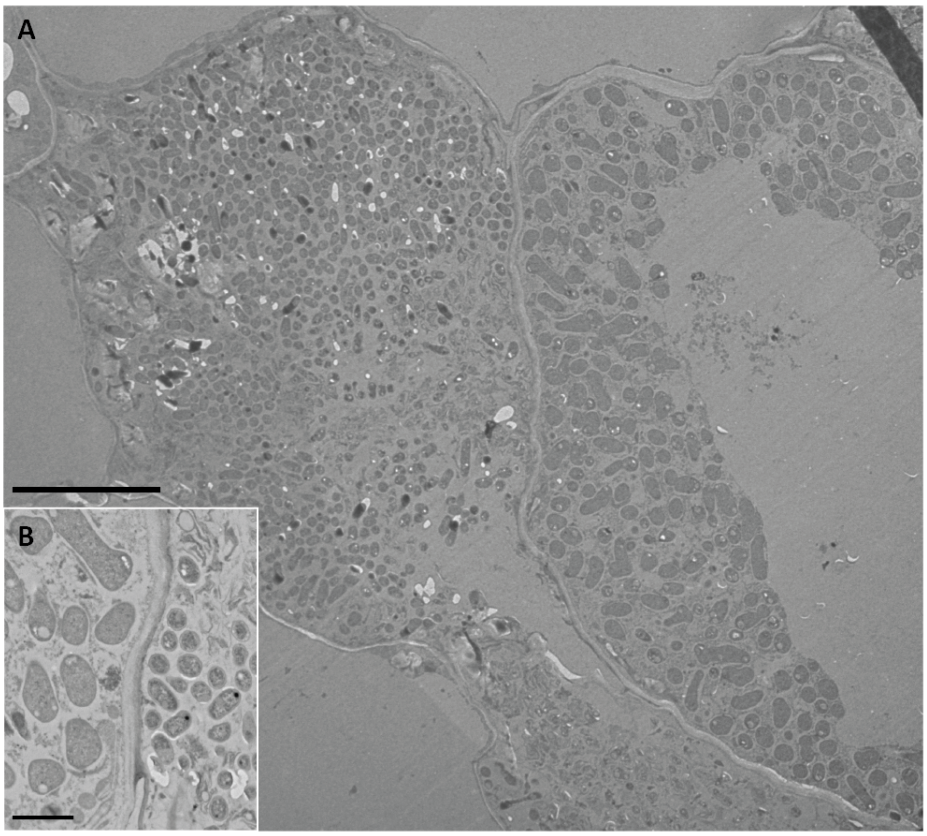


Figure 5.6 Section of nodules taken from *P. sativum* inoculated with LMB771 (*fixAB*::Ωspc) at 28 dpi. Scale bars 10 µm (A) and 2 µm (B). Visualised by TEM at x420 (A) and x1700 (B). EM slides were prepared and imaged by E. Barclay of JIC Bioimaging.

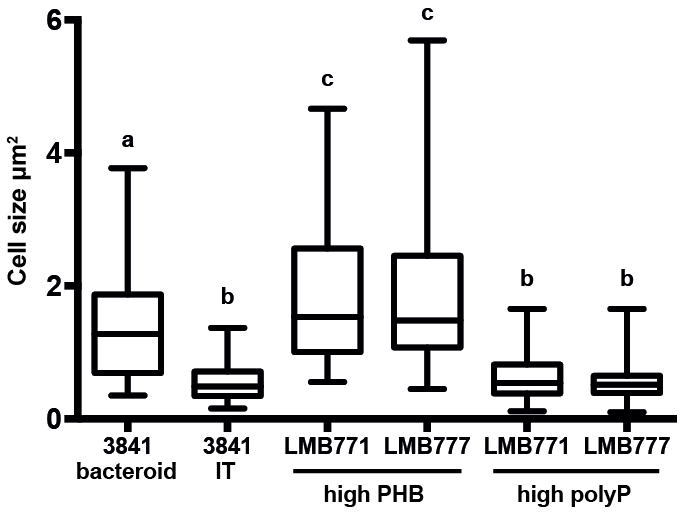


Figure 5.7 Cell sizes determined from TEM images of nodule sections taken from *P. sativum* at 28 dpi. Size in μm^2. n=100; a, b, c represent statistically distinct groups as determined by one-way ANOVA and post-hoc Tukey test. Cells were measured using ImageJ.

5.2.2 Developmental phenotype of *fixAB* mutant bacteroids

Since the polyP-type bacteroids appear much smaller than those filled with PHB it was hypothesised that they may be younger, still-developing bacteroids. Analysis was carried out on immature nodules to investigate this hypothesis. Typically nodules begin appearing on roots from seven dpi onwards. Bacteroid sections were taken of nodules at 14, 21 and 26 dpi, which aimed to capture a time course of bacteroids in developing nodules. At 26 dpi, the same PHB and polyP phenotypes can be seen in electron micrographs (not shown). At fourteen days (Figure 5.9) the high-polyP phenotype cannot be found. In all strains several symbiosomes can be found with sparse, smaller bacteroids. These may be characteristic of maturing nodules whose bacteroids have not yet finished differentiating. This is seen in the wild-type and in the *fix* mutants, and so may be a developmental phenotype rather than a mutant phenotype. Figure 5.8 shows this sparse phenotype found next to an infection thread which appears to have just released its cells. In other wild-type sections, the bacteroids resemble the mature bacteroids seen at 28 dpi. In the mutants, these developed bacteroids tend to have much higher levels of PHB. At 21 dpi (Figure 5.10) the wild-type phenotype resembles that at 28 dpi, with small amounts of both PHB and polyP inclusions seen within bacteroids. In the *fix* mutants the PHB-filled phenotype is becoming more prominent, with entire plant cells filled with bacteroids full of PHB. A second phenotype is seen of cells with

more electron-dense regions at the poles of bacteroids, with a few polyP granules appearing. Visually, the sizes of these two phenotypes are beginning to differ. The two distinct bacteroid phenotypes appears to be a developmental phenotype, appearing later in nodule development.

Quantification of bacteroid sizes at 14 dpi and 21 dpi was made comparing the 'sparse', undeveloped phenotype and the 'normal', developed bacteroid phenotype. Micrographs taken at 14 dpi (Figure 5.11 A) show a clear difference in bacteroid size in mutant and wild-type cells that have and have not developed. No difference is between between strains. At 21 dpi (Figure 5.11 B) the wild-type cells show statistically similar size to the mutant PHB-containing phenotype. No difference is seen between the sparser, developing cells.

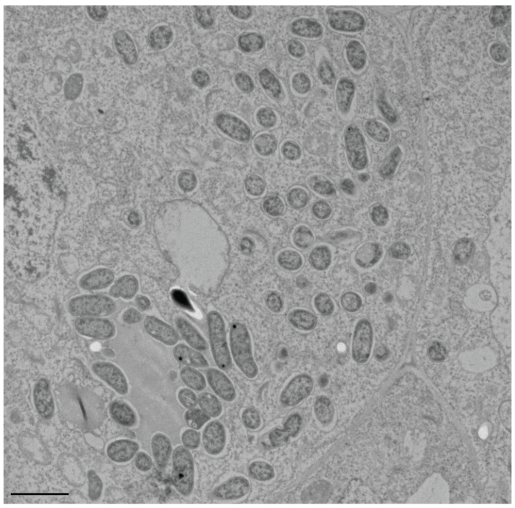


Figure 5.8 Sections of wild-type nodule taken from *P. sativum* at 21 dpi. Scale bar 2 μm. Visualised by TEM at x1700. EM slides were prepared and imaged by E. Barclay of JIC Bioimaging.

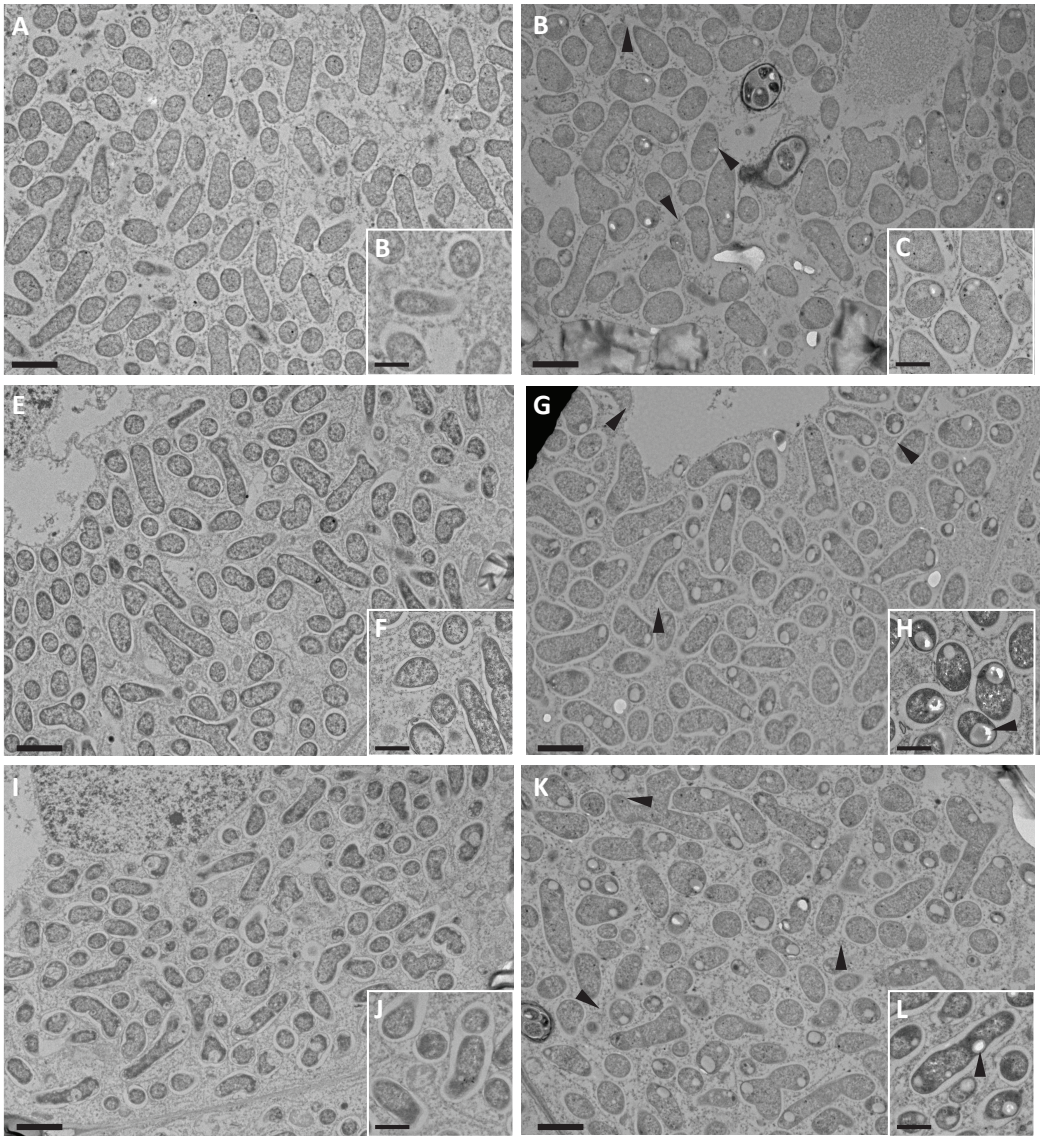


Figure 5.9 Sections of nodules taken from *P. sativum* at 14 dpi; Inoculated with 3841 (A-D); LMB771 (E-H); LMB777 (I-L). Black arrows show PHB inclusions. Scale bars 2 μm (A, C, E, G, I, K) and 1 μm (B, D, F, H, J, L). Visualised by TEM at x1700 (A, C, E, G, I, K) and x3500 (B, D, F, H, J, L). EM slides were prepared and imaged by E. Barclay of JIC Bioimaging.

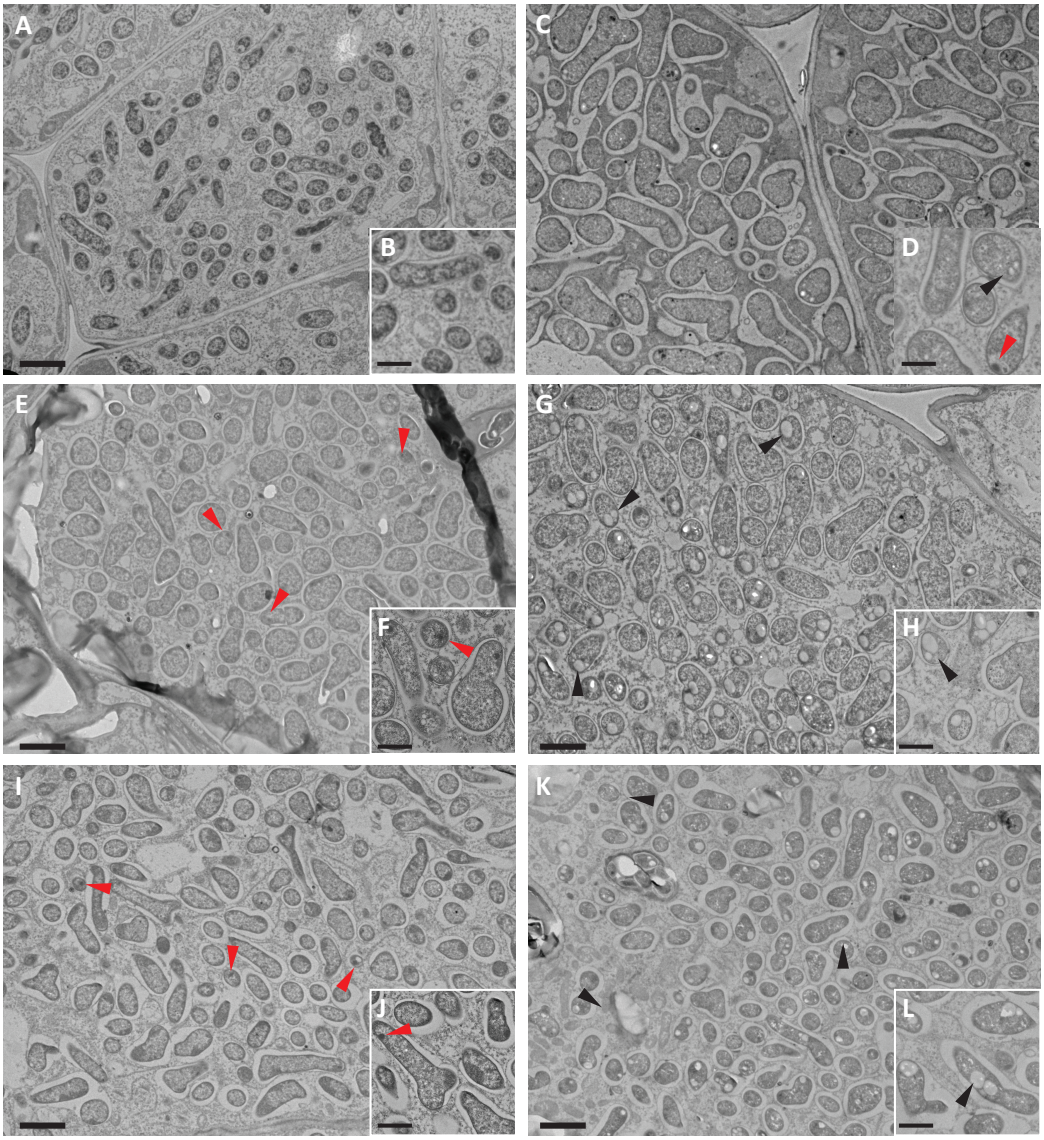


Figure 5.10 Sections of nodules taken from *P. sativum* at 21 dpi; Inoculated with 3841 (A,B); LMB771 (C-F); LMB777 (G-J); Black arrows show PHB inclusions; Red arrows show polyP granules. Scale bars 2 µm (A, C, E, G, I) and 1 µm (B, D, F, H, J). Visualised by TEM at x1700 (A, C, E, G, I) and x3500 (B, D, F, H, J). EM slides were prepared and imaged by E. Barclay of JIC Bioimaging.

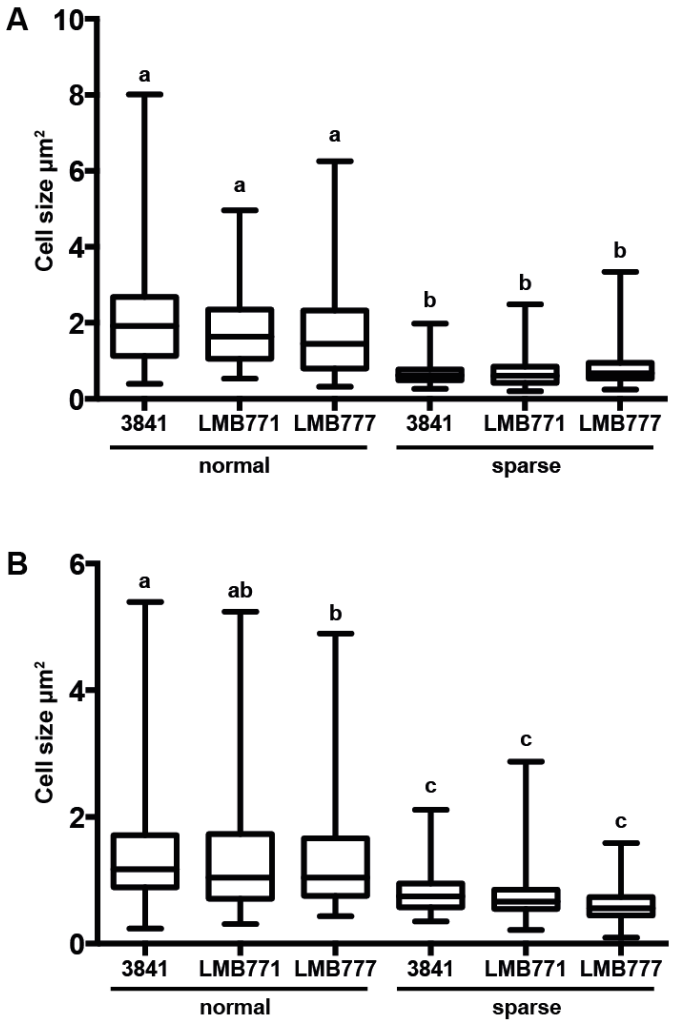


Figure 5.11 Cell sizes determined from TEM images of nodule sections taken from *P. sativum* (pea) inoculated with Rlv3841, LMB771 and LMB777 at 14 dpi (A) and 21 dpi (B). Size in μm^2. n=100; a, b, c represent statistically distinct groups as determined by one-way ANOVA and post-hoc Tukey test. Cells were measured using ImageJ.

5.2.3 Identifying the electron-dense inclusions found in infection threads and mutant bacteroids

Whilst the black polar inclusions have been assumed to be polyphosphate, it should be taken into consideration that lipids may also be seen as electron-dense in electron microscopy, especially if osmium is used during the staining process (Casley-Smith 1967, Craig and Williamson 1972, Wigglesworth 1975, Angermuller and Fahimi 1982). In order to determine if the black inclusions were lipid or PHB, nodule sections were also prepared for electron

microscopy without the osmium-staining step (Figure 5.12). The dark inclusions were still visible in sample prepared without osmium; this suggests that the inclusions are not lipids.

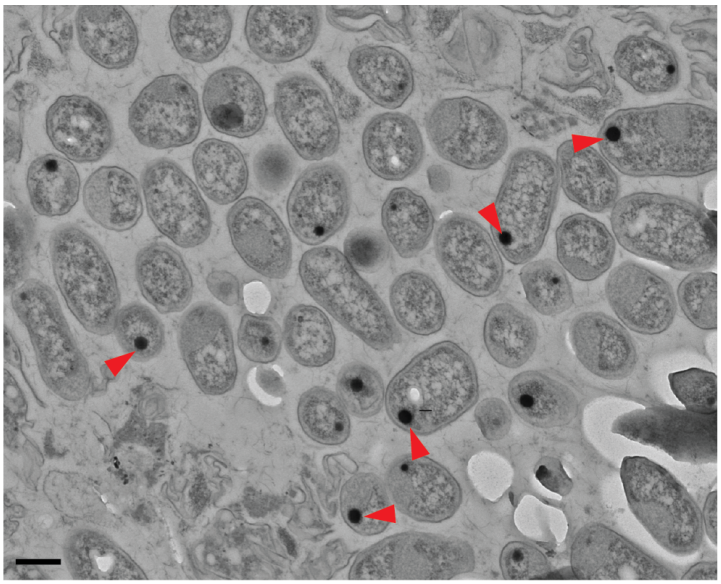


Figure 5.12 Section of a nodule taken from *P. sativum* inoculated with LMB771 at 28 dpi. Osmium was omitted from all fixatives and preparation. Red arrows indicate polar polyphosphate granules. Scale bar 500 nm. Visualised by TEM at x5000. EM slides were prepared and imaged by E. Barclay of JIC Bioimaging.

A second approach to investigating lipid identity is to treat cells with hexane prior to TEM imaging (Figure 5.13). Hexane can be used as a solvent for lipid extraction from tissues (Hara and Radin 1978). If the electron-dense inclusions were lipids, hexane-treated nodules would lack any black inclusions in TEMs. Treatment with hexane resulted bacteroids that retained black dots, supporting the osmium data that the black inclusions are not lipids. The hexane-treated bacteroids had large holes in the images, thought to be due to the effect of hexane causing lysis of PHB bodies.

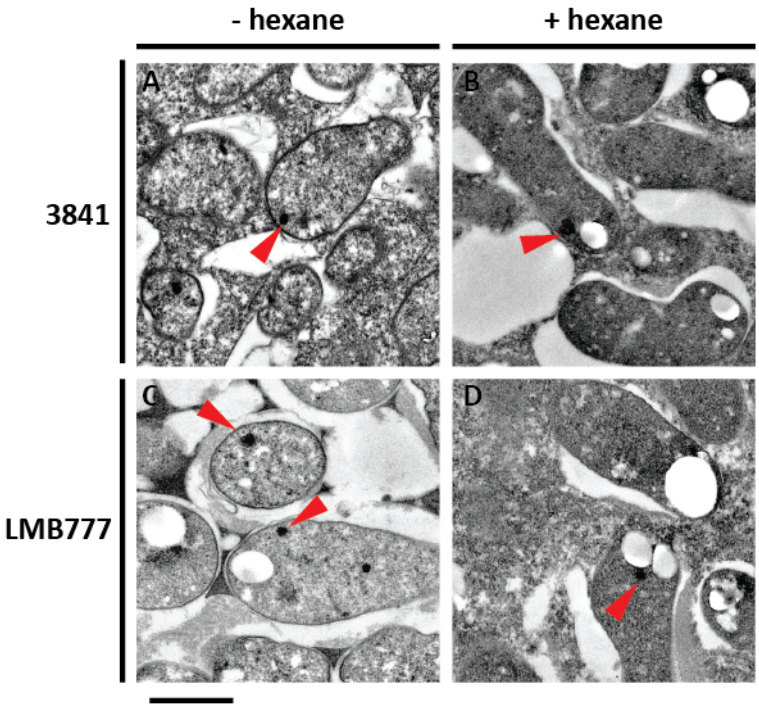


Figure 5.13 Section of a nodule taken from *P. sativum* inoculated with LMB777 at 28 dpi. Red arrows indicate polar polyphosphate granules. Scale bar 1 µm. Visualised by TEM. EM slides were prepared and imaged by E. James of the James Hutton Institute.

To confirm if the inclusions are polyphosphate, a polyphosphate mutant was constructed. Polyphosphate synthesis in bacteria is typically by polyphosphate kinase, *ppk*. Some bacterial species have two Ppk proteins PPK1 and PPK2, though *R. leguminosarum* only has the PPK2 form, which is able to reversibly catalyse formation of polyphosphate from nucleoside triphosphates (Rao, Gomez-Garcia et al. 2009). A search of the annotated Rlv3841 genome resulted in identification of a putative polyphosphate kinase annotated as *ppk* (RL1599). An omega-transposon mutant was constructed in the *ppk* gene. The mutant was constructed using an omega-tetracycline cassette to allow for construction of double and triple mutants with other genes of interest. Plasmids and strains used in mutant construction are described in Chapter 2.2.

Polyphosphate was quantified in the *ppk* mutant (OPS0706), as well as in a FixAB mutant (LMB771) (Figure 5.14). Polyphosphate was significantly lower in both mutants; the *ppk* mutant cannot make WT levels of polyphosphate. Additionally, the FixAB mutant had the same low levels of polyphosphate. Polyphosphate can also be detected using monochromatic blue stains, such as Toluidene Blue. This explains why the bacteroids in the light

micrographs (Figure 5.1) showed such variation in intensity of blue colour; polyphosphate shows up as a deep purple under this staining.

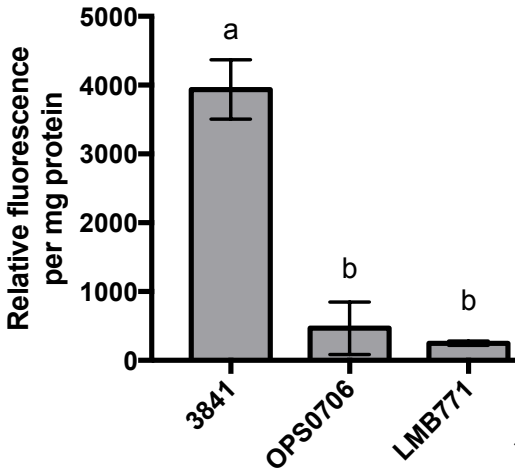


Figure 5.14 Quantification of polyphosphate in wild-type (3841), polyphosphate synthase mutant (OPS0706) and *fixAB*::Ωspc (LMB771). n≥3; a,b,c represent statistically distinct groups as determined by one-way ANOVA and post-hoc Tukey test.

5.3 Investigation of mutants in storage polymer biosynthesis

Several double and double mutants were constructed involved in polyphosphate and PHB synthesis and the FixAB genes. Rlv3841 has two copies of the polyhydroxybutyrate synthase gene *phaC*, one on the chromosome (*phaC1*) and one on the symbiotic plasmid pRL10 (*phaC2*). A double mutant (*phaC1/phaC2*), LMB816 has been published by the Poole lab (Terpolilli, Masakapallic et al. 2016). Phage transduction was used to conjugate *ppk::*Ωtc into LMB777 (Δ*fixAB*) and LMB816 (*phaC1*::Tn5 *phaC2*::Ωspc) to give strains OPS0707 and OPS0744 respectively. All strains showed wild-type growth on rich media.

Strain	Description	Doubling time /hours
3841	Wild-type	4.76 ± 0.14
LMB816	*phaC1*::Tn5 *phaC2*::Ωspc	4.09 ± 0.15
OPS0706	*ppk*::Ωtc	4.60 ± 0.12
LMB777	*ΔfixAB*	4.37 ± 0.39
OPS0707	*ΔfixAB ppk*::Ωtc	4.45 ± 0.16
OPS0739	*ΔfixAB phaC2*::Ωspc	4.59 ± 0.24
OPS0744	*ppk*::Ωtc *phaC1*::Tn5 *phaC2*::Ωspc	4.46 ± 0.06

Table 5.1 Free-living growth phenotype of storage mutant strains grown on rich media in a BMG FluoStar Omega plate reader; 28°C; 500 rpm orbital shaking.

All storage mutant strains were inoculated on plants in order to determine if lack of storage pathways affected nitrogen fixation ability and plant phenotype. Mutants lacking FixAB were unable to fix nitrogen, but all other strains were able to fix at wild-type levels. The plant phenotypes of these storage mutants are shown in Figure 5.15. Plants inoculated with strains which contain a disruption of *fixAB* have small, white nodules and stunted, yellowing plants, typical of non-fixing nodules. All other strains resulted in plants with resembling plants inoculated with wild-type, with no obvious phenotypic effect.

Strain	Description	Symbiotic phenotype	Nodule colour	Acetylene reduction µmol ethylene plant^{-1} hr^{-1}
3841	Wild-type	Nod$^+$ Fix$^+$	Pink	1.99 ± 0.66^a
LMB816	*phaC1*::Tn5 *phaC2*::Ωspc	Nod$^+$ Fix$^+$	Pink	2.21 ± 0.22^a
OPS0706	*ppk*::Ωtc	Nod$^+$ Fix$^+$	Pink	2.65 ± 0.31^a
OPS0707	*ΔfixAB ppk*::Ωtc	Nod$^+$ Fix$^-$	White	0.24 ± 0.16^b
OPS0739	*ΔfixAB phaC2*::Ωspc	Nod$^+$ Fix$^-$	White	0.06 ± 0.06^b
OPS0744	*ppk*::Ωtc *phaC1*::Tn5 *phaC2*::Ωspc	Nod$^+$ Fix$^+$	Pink	3.21 ± 0.34^a
Uninoculated control	n/a	Nod$^-$ Fix$^-$	n/a	0.02 ± 0.03^b

Table 5.2 Symbiotic phenotype of storage mutant strains on pea (*P. sativum*). Acetylene reduction assay was carried out 28 days post inoculation ± SEM; n ≥ 3; a, b represent statistically distinct groups as determined by one-way ANOVA and post-hoc Tukey test.

PHB and polyphosphate are both energy storage molecules. In consideration of this, strain RU1448 (*glgA*::Tn5 in *R. leguminosarum* bv. *viciae* A34 background) (Lodwig, Leonard et al. 2005) was also used in plant assays, in order to determine the effect of glycogen mutation compared to mutations that influence storage molecule biosynthesis. Phage transduction using RL38 was carried out to transfer the *glgA* mutation across species. The *glgA* mutation was created in the A34 strain background, and the recipients are in the 3841 background. There is a 96% similarity between the *glgA* gene region of A34 and 3841, allowing successful transduction between background strains. The *glgA* mutation was transduced into LMB816 (*phaC1C2*) and OPS0744 (*phaC1C2 ppk)* to give strains OPS0770 and OPS0771 respectively. Plants assays using these strains showed wild-type levels of fixation in RU1448 (*glgA*) and OPS0770 (*glgA phaC1C2*). Strain OPS0771, defective in glycogen, PHB and polyphosphate showed significantly higher rates of nitrogen fixation, however this should be repeated, and dry weight assays carried out in order to confirm this higher rate.

Strain	Description	Symbiotic phenotype	Nodule colour	Acetylene reduction μmol ethylene plant^{-1} hr^{-1}
3841	Wild-type	Nod$^+$ Fix$^+$	Pink	2.13 ± 0.20^a
RU1448	*glgA::*Tn*B60*	Nod$^+$ Fix$^+$	Pink	2.14 ± 0.06^a
OPS0770	*phaC1*::Tn5 *phaC2*::Ωspc *glgA::*Tn*B60*	Nod$^+$ Fix$^+$	Pink	2.21 ± 0.15^a
OPS0771	*phaC1*::Tn5 *phaC2*::Ωspc *glgA::*Tn*B60* *ppk*::Ωtc	Nod$^+$ Fix$^+$	Pink	2.99 ± 0.23^b
Uninoculated control	n/a	Nod$^-$ Fix$^-$	n/a	0.02 ± 0.01^c

Table 5.3 Symbiotic phenotype of glycogen mutant strains on pea (*P. sativum*). Acetylene reduction assay was carried out 28 days post inoculation ± SEM; n ≥ 3; a, b represent statistically distinct groups as determined by one-way ANOVA and post-hoc Tukey test.

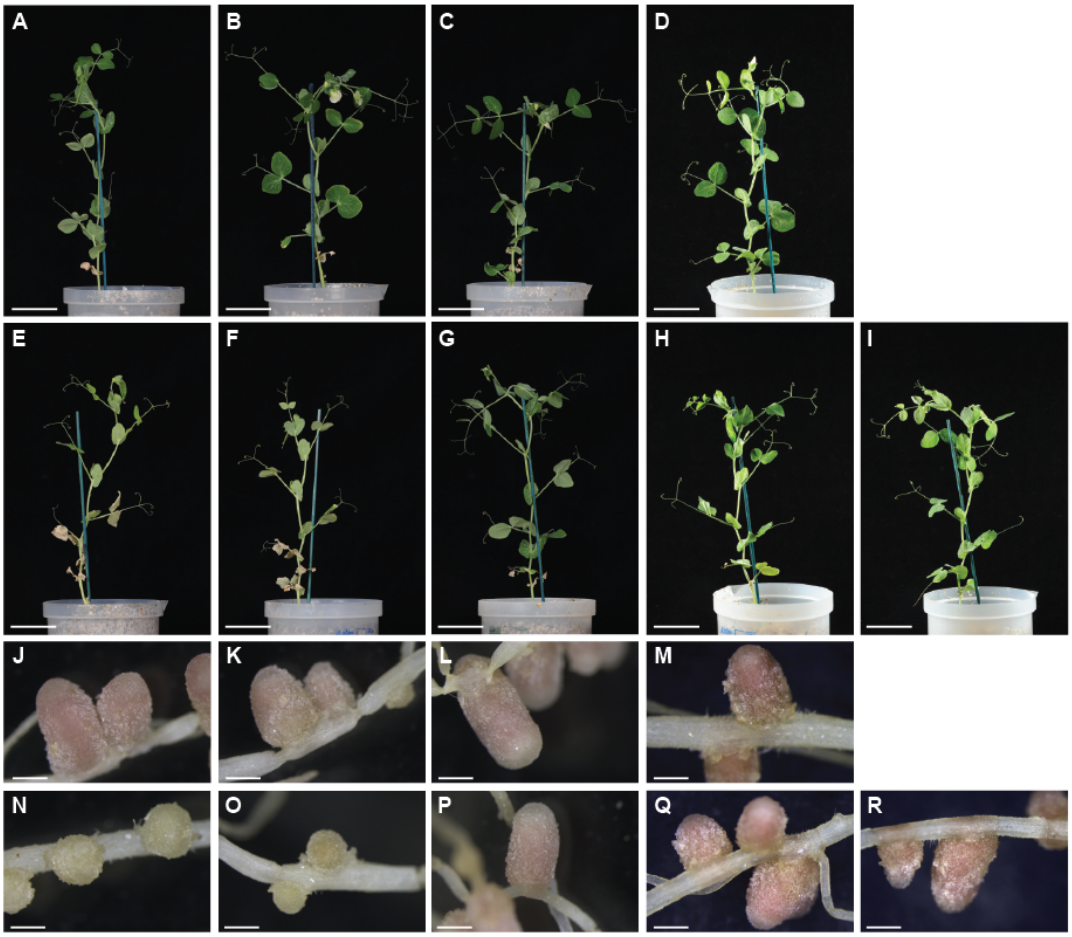


Figure 5.15 Plant phenotype of plants inoculated with storage mutants: A,J: Rlv3841; B,K: LMB816 (*phaC1*::Tn5 *phaC2*::Ωspc); C,L: OPS0706 (*ppk*::Ωtc); D,M: RU1448 (*glgA*::*TnB60*); E,N: OPS0707 (*ΔfixAB ppk*::Ωtc); F,O; OPS0739 (*ΔfixAB phaC1*::*Tn5 phaC2*::Ωspc); G,P: OPS0744 (*ppk*::Ωtc *phaC1*::Tn5 *phaC2*::Ωspc); H,Q: OPS0770 (*phaC1*::Tn5 *phaC2*::Ωspc *glgA*::*TnB60*); I,R: OPS0771 (*ppk*::Ωtc *phaC1*::Tn5 *phaC2*::Ωspc *glgA*::*TnB60*). Plants are representative of four replicates. Photographs taken at 28 days post-inoculation. Scale bars A-I: 5 cm; J-R: 1 mm.

5.4 Using Raman microscopy to phenotype mutants at a single-cell level

Single-cell Raman microscopy (SCRM) is a spectroscopic technique that combines a confocal microscope with Raman spectroscopy, which uses molecular vibrations to assign unique spectra to a single cell. This allows creation of a trace with identifiable peaks corresponding to macromolecules (Huang, Griffiths et al. 2004).

5.4.1 Optimising Raman microscopy for use in *R. leguminosarum* bv. *viciae*

Raman microscopy was used as a strategy to determine the biochemical phenotype of *fix* mutants at a single-cell level. In order to test the effectiveness of this technique on *R. leguminosarum*, an initial test was used to determine if the Raman microscope could determine the difference between the wild-type in free-living culture vs. terminally differentiated bacteroids (Figure 5.16), known to be highly different from transcriptomic data (Karunakaran, Ramachandran et al. 2009). All samples were standardised against a phenyalanine peak at ~1120 cm^{-1}. Several peaks can be identified as different between the two states, including four cytochrome peaks corresponding to the terminal cbb_3-oxidase FixNOQP. Principal component analysis (PCA) confirms the distinct difference between the two samples.

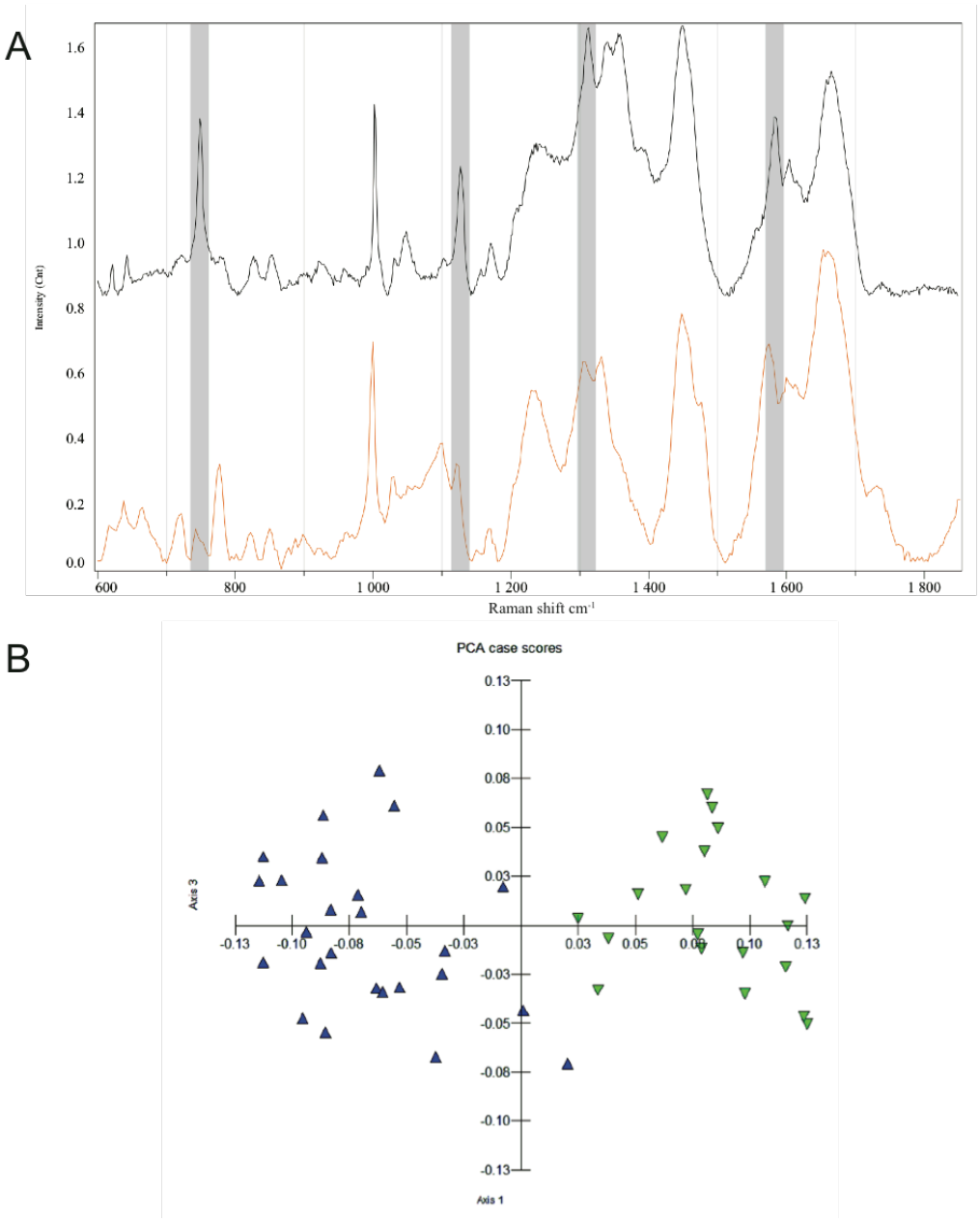


Figure 5.16 A. Raman spectra for Rlv3841 in the free-living (orange) and bacteroid (black) state. Grey bars highlight cytochrome peaks; B. PCA analysis of spectra for Rlv3841 free-living culture (green) versus bacteroids (blue). Raman measurements were carried out by J. Xu.

5.4.2 Analysis of bacteroid populations in *fixAB* mutants

Raman microscopy allows distinction between phenotypes in heterogeneous cell populations (Hermelink, Brauer et al. 2009) and so can be used to analyse bacteroids extracted from 28 dpi nodules of the *fixAB* mutant. These nodules contain mutant bacteroids with different phenotypes. Randomly chosen cells (n=75) were measured, and spectra were initially separated by presence or absence of PHB, with 56 spectra having larger than wild-type PHB peaks, and 19 spectra lacking a clear PHB peak (Figure 5.17). The average size of cells with large PHB peaks was 3.12 μm^2 and cells without a PHB peak averaged 1.39 μm^2. This agrees with observations made in the electron micrographs. NifH mutant RU3940 was also analysed using Raman spectroscopy. The same heterogeneity was seen in this sample, with spectra able to be split into high and low PHB phenotypes (data not shown).

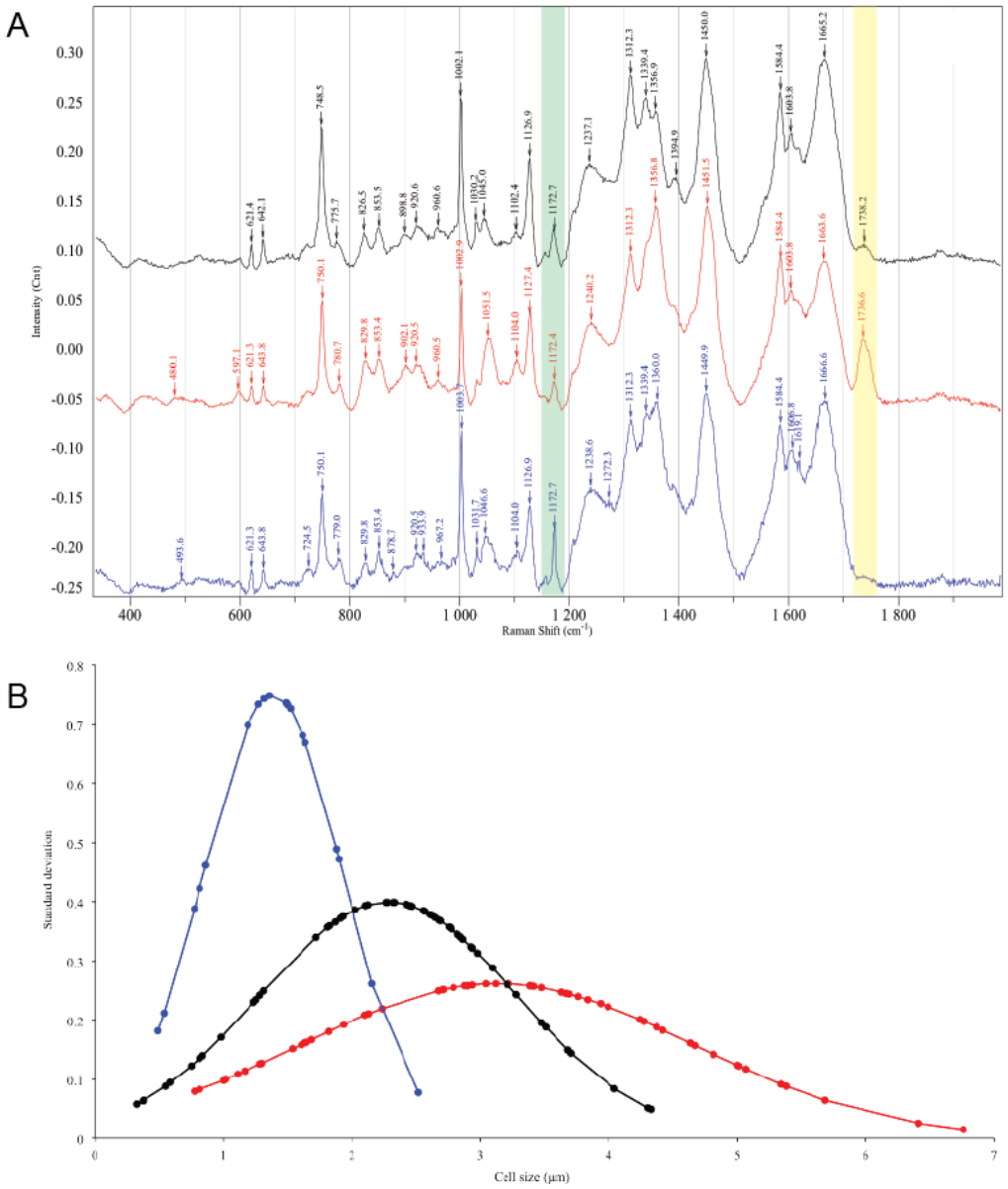


Figure 5.17 Raman spectra of bacteroid samples of Rlv3841 (wild-type) and LMB777 (Δ*fixAB*). Individual LMB777 spectra were separated by presence or absence of a PHB peak. Spectra were then averaged within groups; Black: wild-type; Red: LMB777 with a PHB peak; Blue: LMB777 spectra lacking a PHB peak. B. Cell sizes were quantified for each group, and plotted against standard deviation to show the variation between groups. Raman microscope measurements were carried out by J. Xu.

Mutant samples were reanalysed by size, and split as larger than or smaller than the average for the wild-type. Raman is semi-quantitative (Majed and Gu 2010) and so these samples were quantified for PHB and polyphosphate (Figure 5.18). A huge increase in PHB is seen in the larger cells of the *fix*AB mutants, correlating with the phenotype seen in electron micrographs. The small infection-thread-like cells have similar composition to the wild-type; since these cells are much smaller the abundance of polyphosphate may be more prominent in electron micrographs.

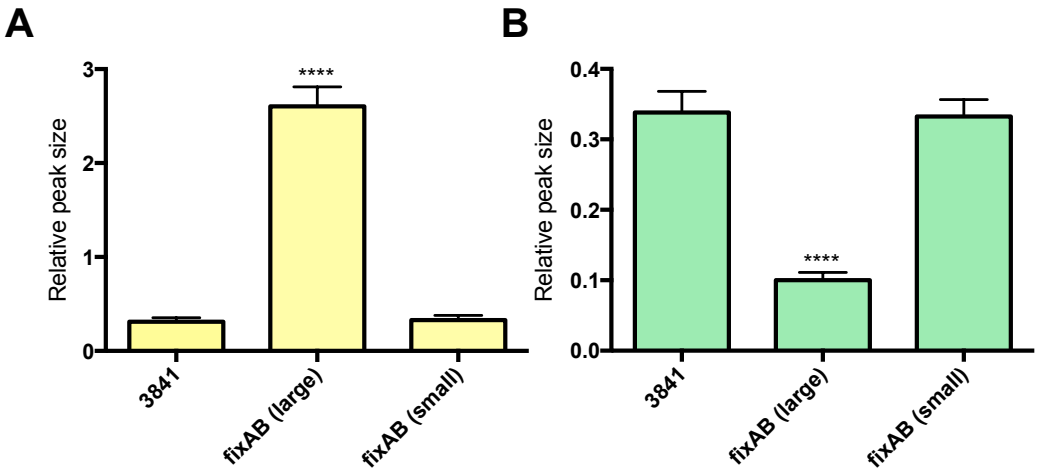


Figure 5.18. Relative peak size of A: PHB and B: polyphosphate in Raman spectra of 28-day bacteroids. Peak size was normalised against the phenylalanine peak; n=30; **** indicates significant difference from the wild-type ($p \leq 0.0001$).

5.4.3 Characterising a novel *ppk* mutant

A polyphosphate mutant was also analysed to confirm the identity of the mutant against a known polyphosphate peak. In both the *ppk* mutant (OPS0706) and *ppk fixAB* double mutant (OPS0707) a smaller peak is seen for polyphosphate than in the wild-type (Figures 5.19 and 5.21 A).

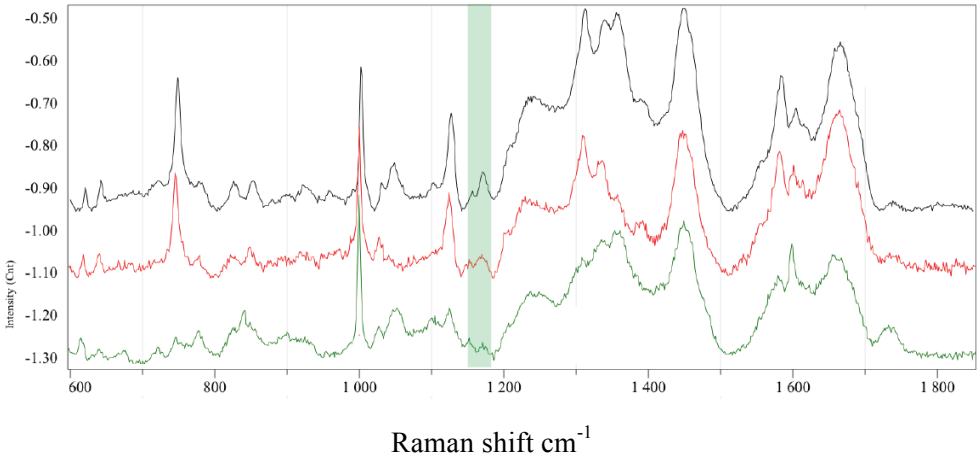


Figure 5.19 Raman spectra of bacteroid samples of Rlv3841 (black); OPS0706 (*ppk*) (red); OPS0707 (*fixAB ppk*) (green). Green bar indicates peak for polyphosphate. Raman measurements were carried out by J. Xu.

5.4.4 Confirming peak identity of PHB and glycogen using characterised mutants

Known mutants were used to confirm the identity of hypothesised PHB and glycogen peaks. As expected, there was no PHB peak in a PHB mutant and no glycogen peak in a glycogen mutant (Figure 5.20).

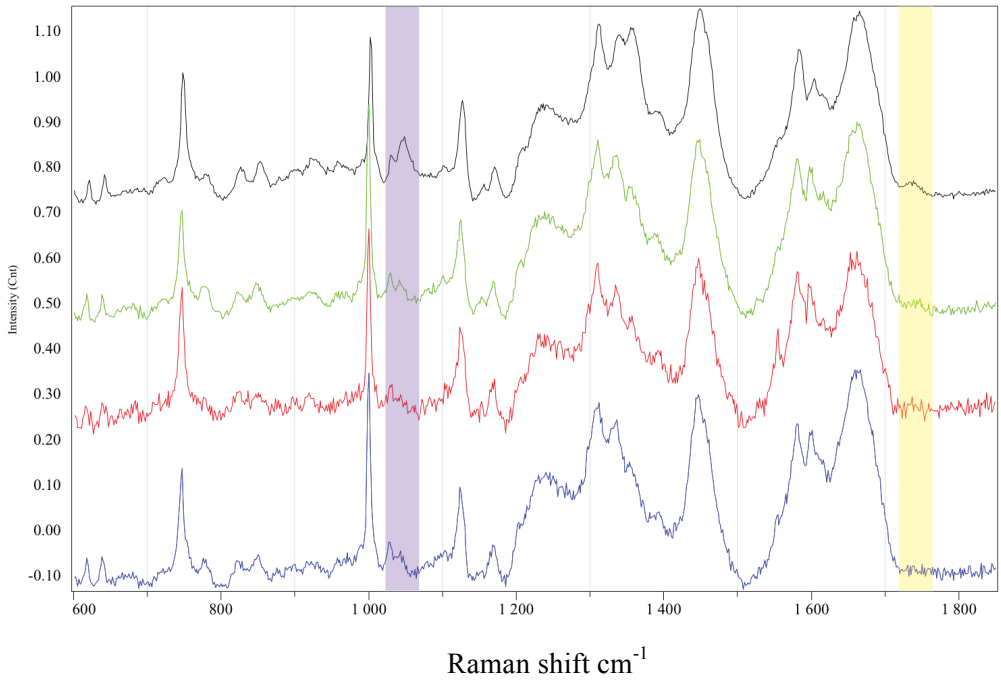


Figure 5.20 Raman spectra of bacteroid samples of Rlv3841 (black); LMB816 (*phaC12*) (green); RU1448 (*glgA*) (red); RU1478 (*glgA phaC*) (blue). Bars indicate peaks for glycogen (purple) and PHB (yellow). Raman measurements were carried out by J. Xu.

PHB, glycogen and polyphosphate peak sizes can be quantified in all mutants analysed by Raman. Phenotyped mutants can be used to confirm the identity of peaks, and in a complementary manner Raman can be used to confirm the identity of mutants. Polyphosphate kinase (*ppk,* RL1599) mutants show a decrease in polyphosphate compared to the wild-type. Considering the reciprocal peaks, a glycogen mutant has low PHB, and a PHB mutant has glycogen levels similar to that of a glycogen mutant. This coupled PHB-glycogen phenotype was recognised by Lodwig (2001) in *R. leguminosarum* bv. *viciae* A34, where a *phaC* mutant had 4.5-fold less glycogen than the wild-type. Further evidence for this comes from the *fixAB* mutants, where the large, PHB-filled cells have higher glycogen levels than the wild-type. Glycogen and PHB quantification was attempted in these strains, but insufficient nodule material was available and results were below the limit of the assay.

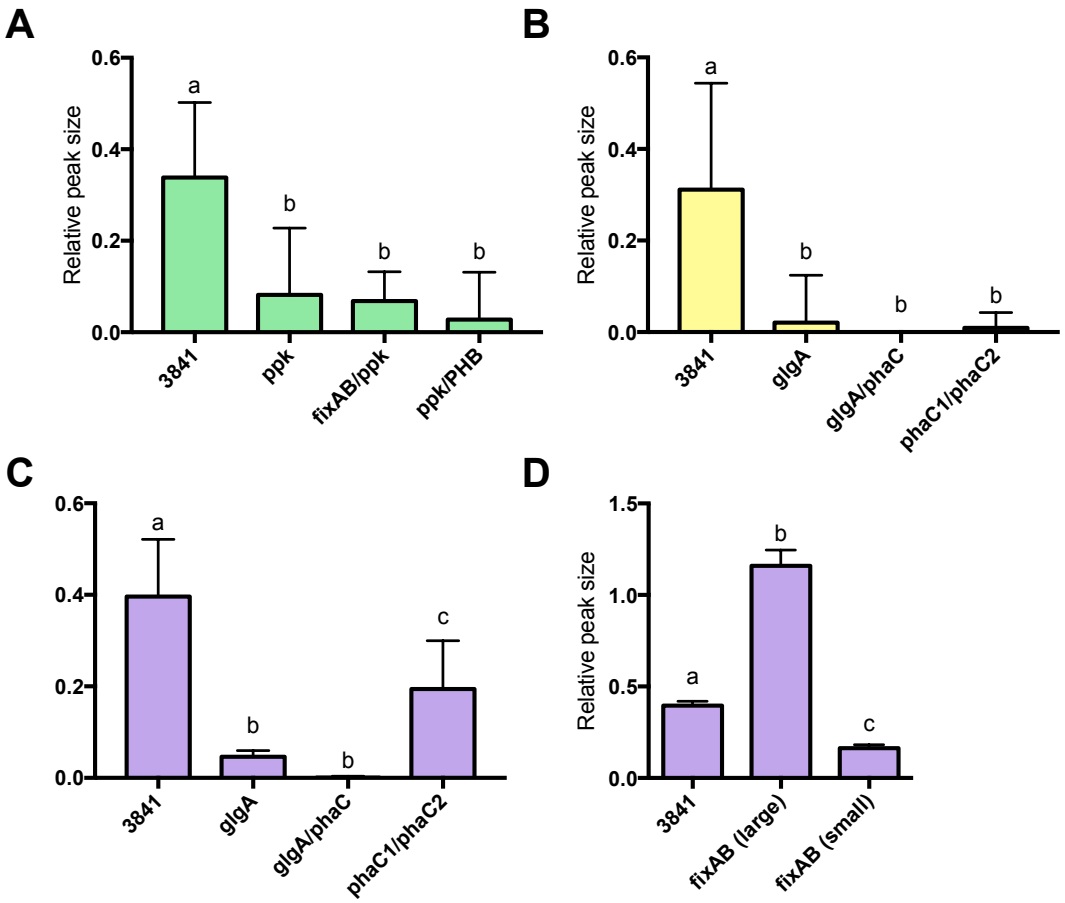


Figure 5.21 A: Relative polyphosphate peak size in *ppk* mutant bacteroids isolated 28-days post inoculation; B: Relative PHB peak size in *phaC* and *glgA* mutant bacteroids isolated 28-days post inoculation; C: Relative glycogen peak size in *phaC* and *glgA* mutant bacteroids isolated 28-days post inoculation; D: Relative glycogen peak size *fixAB* mutant bacteroids isolated 28-days post inoculation; peak size normalised against phenylalanine peak. n=30; a,b,c indicates statistically separate groups as determined by ANOVA and post-hoc Tukey test ($p \leq 0.0001$).

5.4.5 Analysis of double and triple mutants involved in polymer biosynthesis and electron transport

Double and triple mutants of the strains investigated in this chapter were also sent for Raman microscopy (Figure 5.22). Combination of mutation in *fixAB* and synthetic pathways appears to have an effect on the polymers found. A *fixAB ppk* double mutant appears to make much higher levels of glycogen and PHB. A *fixAB phaC2* double mutant makes lower amounts of polyphosphate than the wild-type. The *fixAB phaC2* mutant has wild-type levels of PHB; this

strain still has the chromosome-encoded *phaC1*, and so may be retaining PHB produced in the infection thread.

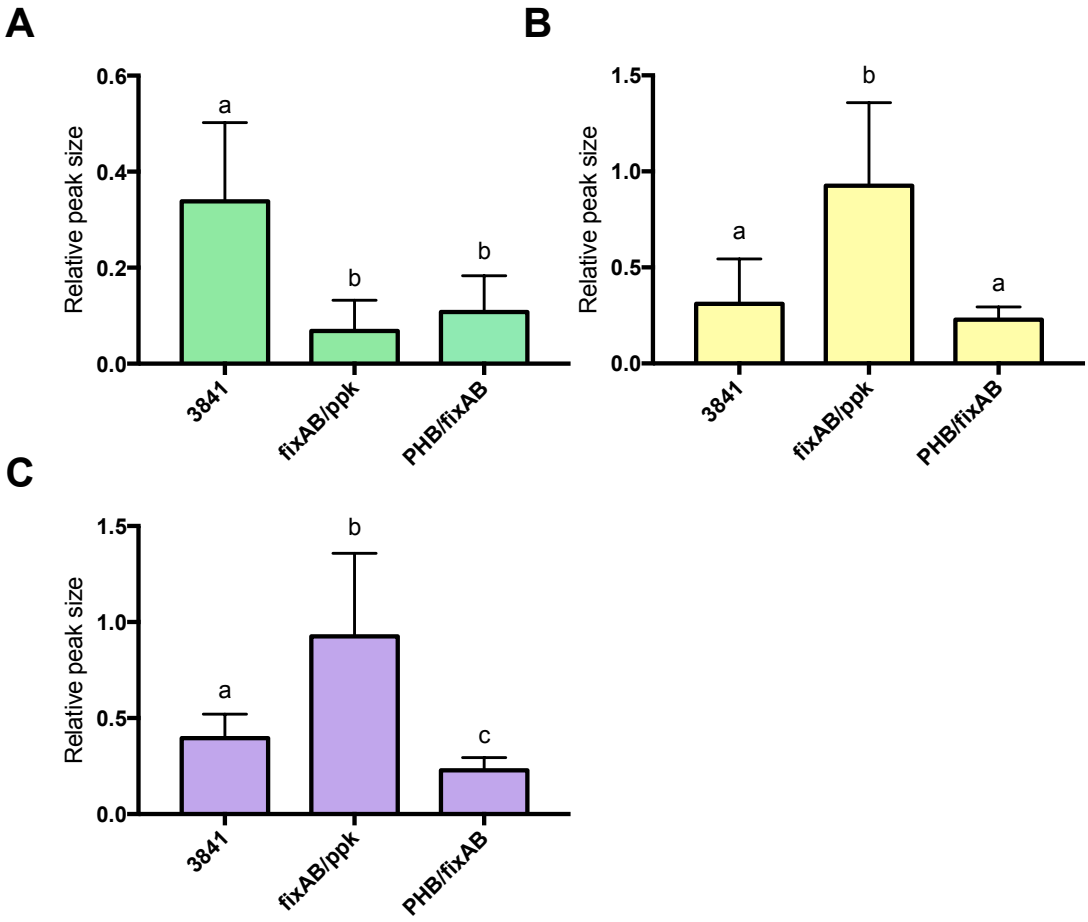


Figure 5.22 A: Relative polyphosphate peak size in *fixAB* double mutant bacteroids isolated 28-days post inoculation; B: Relative PHB peak size in in *fixAB* double mutant bacteroids isolated 28-days post inoculation; C: Relative glycogen peak size in in *fixAB* double mutant bacteroids isolated 28-days post inoculation; peak size normalised against phenylalanine peak; n=30; a,b,c indicates statistically separate groups as determined by ANOVA and post-hoc Tukey test ($p \leq 0.0001$).

5.4 Discussion

The FixAB mutants characterised in this work show hugely different bacteroid morphology to the wild-type Rlv3841 when inoculated on *P. sativum*. Plant cell occupancy is much lower than in the wild-type, as seen in both light micrographs (Chapter 4) and electron micrographs. This, along with the smaller, white nodules suggests that the plant ceases providing an optimal environment for bacteroids when it recognises that the strain is not providing fixed nitrogen in return for the plant carbon input.

In young nodules two phenotypes can be seen; differentiating bacteria just emerged from the infection thread, and developed bacteroids. The smaller, developing bacteroids are found near infection threads and are seen in both the wild-type and in the FixAB mutants. These tend to be found more sparsely within the plant cell. No difference is seen in these developing bacteroids between the strains; these strains may not have finished differentiating, and so *fix* genes may not yet be expressed.

The larger, mature bacteroids show a morphological difference between strains. In the FixAB mutants, these developed bacteroids contain high levels of polyhydroxybutyrate (PHB), a polymer produced from acetyl-CoA. Previous work in *R. leguminosarum* bv. *viciae* has suggested a role for PHB as a carbon storage compound which accumulates in the infection thread and can be broken down in to fuel bacteroid differentiation (Lodwig, Leonard et al. 2005). *R. leguminosarum* has two PHB synthases. PhaC1 is encoded on the chromosome and is a type I PHB synthase. There is a type III PHB synthase encoded by *phaEphaC2* (pRL100104-5), found on the symbiotic plasmid pRL10 under the control of NifA. The *phaE* gene is upregulated 40-fold in 21-day wild-type bacteroids (Karunakaran, Ramachandran et al. 2009), and so is likely already being used as a storage pathway during nitrogen fixation. Investigation of *phaC1* and *phaC2* mutants showed that PHB produced in the bacteroids is by the PhaC2 enzyme, and PHB produced in infection threads is by the PhaC1 enzyme (Terpolilli, Masakapallic et al. 2016). This suggests that PHB found in wild-type bacteroids is not the result of accumulation in the infection thread; bacteroids produce PHB during symbiosis.

In *Azotobacter spp.* PHB has been proposed to play a role in protecting cells from excess reductant under microaerobic conditions (Senior, Beech et al. 1972, Page and Knosp 1989). Excess NAD(P)H will inhibit pyruvate dehydrogenase, a key enzyme in symbiotic nitrogen fixation as well as TCA cycle enzymes such as isocitrate dehydrogenase and α-ketoglutarate dehydrogenase (Berg 2002). Mutants in PHB synthesis in both *R. etli* and *A. caulinodans* have raised levels of reduced nucleotides, confirming a role for PHB as a reductant sink

(Cevallos, Encarnacion et al. 1996, Mandon, Michel-Reydellet et al. 1998). Strains defective in FixAB appear to accumulate this storage polymer rather than breaking it down, suggesting reductant accumulation in the non-fixing bacteroids. This accumulation is expected since the nitrogenase reaction requires eight electrons per molecule of dinitrogen fixed. If electrons cannot reach nitrogenase, a large amount of reductant will build up in bacteroids that cannot be oxidised due to the low TCA cycle turnover under microaerobic conditions. PHB production would be one solution for removal of excess reductant. At 28 days post-inoculation this high PHB phenotype becomes more pronounced with huge quantities of PHB accumulating, and the cells become larger and much more misshapen as more reductant needs to be sequestered in order to prevent redox stresses.

At 26-28 days post-inoculation an additional phenotype appears within the FixAB mutant nodules. Developing bacteroids can no longer be detected within the plant cell. Instead, there are significantly smaller bacteria, filled with electron-dense granules initially thought to be either polyphosphate or lipid (Craig and Williamson 1972). These cells resemble undifferentiated cells found in infection threads. This phenotype appears late during nodule development, and is not seen in earlier stage bacteroids.

Two strategies were used to determine the identity of the electron-dense inclusions seen in TEM images. The first was to section a new batch of nodules with omission of the osmium staining step, as osmium is known to stain lipids (Angermuller and Fahimi 1982). The second strategy was hexane treatment of nodules, expected to remove lipids. TEM images of hexane-treated samples retained the dark inclusions; these are most likely not lipids. A second identity proposed for these inclusions is polyphosphate. A putative polyphosphate kinase (*ppk*) was identified in Rlv3841 and an omega-tetracycline mutant was constructed. A polyphosphate assay was used to confirm identity of this mutant.

Polyphosphate production is seen across the Kindoms of Life (Kornberg, Rao et al. 1999, Rao, Gomez-Garcia et al. 2009), and is thought to have been part of early biotic evolution (Achbergerova and Nahalka 2011). Bacteria such as *E. coli* and *Lactobacillus* have been shown to accumulate polyphosphate under conditions of stress including osmotic stress, oxidative stress and nutrient starvation (Ault-Riche, Fraley et al. 1998, Alcantara, Blasco et al. 2014). In *E. coli* inorganic phosphate accumulates under stress, and plays a role in expression of RpoS; the stationary-phase σ-factor. A similar phenomenon has been demonstrated in *Pseudomonas putida* (Nikel, Chavarria et al. 2013), *Salmonella typhimurium* (Cheng and Sun 2009) and *Mycobacterium tuberculosis* (Sureka, Dey et al. 2007, Sanyal, Banerjee et al. 2013). A direct role in stress responses has also been suggested; under oxidative stress polyphosphate can act as a chemical chaperone, promoting protein

stability and proteolysis of degraded and aggregated proteins, and it can also be converted to ATP to power production of protein chaperones (Price-Carter, Fazzio et al. 2005, Gray, Wholey et al. 2014). This stress-response role for polyphosphate appears relatively conserved across bacteria. *R. leguminosarum* does not have a direct homologue of RpoS, however it does have a two copies of heat shock sigma factor RpoH (RpoH1, RpoH2). In *R. etli* RpoH1 has been shown to be involved in response to oxidative stress (Martinez-Salazar, Sandoval-Calderon et al. 2009).

Both polyphosphate and PHB are produced under conditions of oxidative stress. It could be hypothesised that PHB production in infection threads is occurring as a response to the environment, rather than a strategy to fuel bacteroid differentiation.

In the case of the FixAB mutants, the high polyphosphate phenotype does not arise until later in nodule development. The smaller, polyphosphate-rich cells resemble those seen in the infection thread in electron micrographs, both in appearance and size. These polyphosphate-rich 'bacteroids' may be infection thread bacteria that never differentiated. The plant ceases provision of carbon and other signals, such as NCR peptides, removing the signals required for bacteroid development. A lack of input from the plant is seen in the nodule phenotype; the mutant nodules are white, lacking in plant leghaemaglobin. It would be interesting to analyse temporal and spatial NCR peptide expression in plants inoculated with these nitrogen fixation mutants. The polyphosphate may have been synthesised as a stress response in these cells, though we cannot deduce whether this was produced within the infection thread or subsequent to release from the infection thread. The infection thread may already be under oxidative stress; reactive oxygen species (ROS) are thought to be involved in infection thread formation (Jones, Kobayashi et al. 2007, Chang, Damiani et al. 2009). The polyphosphate-rich 'bacteroids' contain low levels of PHB; they may be using up the existing stores produced by PhaC1 to fuel the TCA cycle in the absence of plant-derived carbon.

Figure 5.23 shows a diagram representing the changes occurring in fixing versus non-fixing bacteroids.

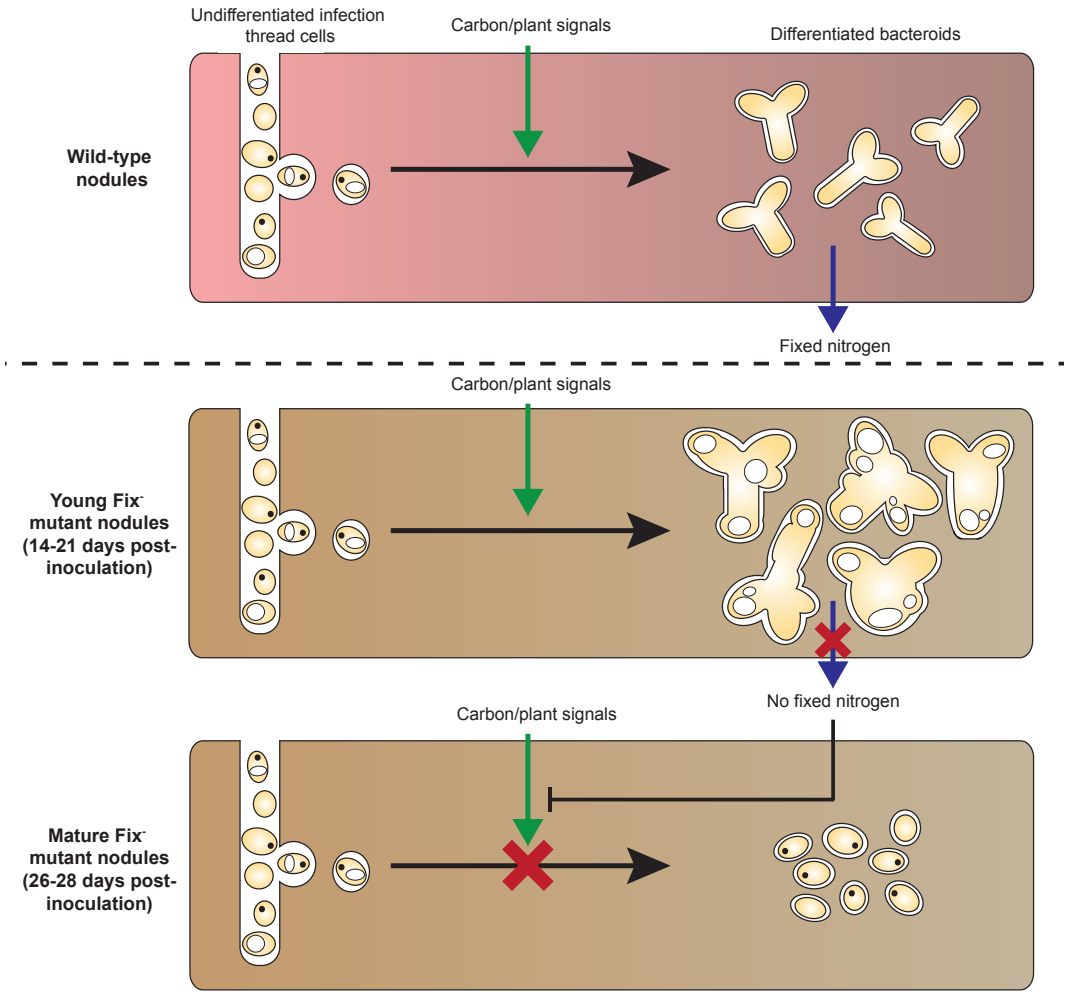


Figure 5.23 Schematic of bacteroid differentiation in A. wild-type *R. leguminosarum* and B. a non-fixing mutant. Bacteria released from the infection thread receive signals from the plant, including C_4-dicarboxylates, which lead to differentiation into Y-shaped bacteroids. Bacteria within the infection thread show signs of nutrient stress, including PHB granules (white) and polyphosphate (black). Differentiated bacteroids release fixed nitrogen, in the form of ammonia. In the absence of nitrogen fixation, reductant is removed via sinks such as PHB. These cells do not release fixed nitrogen. At some point ~21-26 days post-inoculation the plant responds by ceasing to send carbon/other signals. Bacteria no longer differentiate into bacteroids and begin to use existing carbon stores such as PHB for fuel.

Analysis of bacteroid morphology in TEM images has shown a tight relationship between nitrogen fixation and storage polymers. In addition to PHB, other storage polymers have been identified in nitrogen fixation in *R. leguminosarum* bv. *viciae*, including glycogen (Lodwig, Leonard et al. 2005) and lipids (Terpolilli, Masakapallic et al. 2016). A careful balance is required to ensure that the bacteroid does not undergo detrimental redox stress during nitrogen fixation. Multiple routes for relief of redox stress means that disruption of one route is not detrimental to survival. This is seen in single, double and triple mutants of PHB, glycogen and polyphosphate synthesis. Figure 5.24 shows a scheme for potential electron sinks and storage routes within a nitrogen-fixing bacteroid and a non-fixing mutant. Input into electron sinks is higher in a non-fixing mutant as it directs reductant down alternative pathways. Biochemical assays on these mutants are influenced by the cell heterogeneity; a polyphosphate assay yielded relative fluorescence matching that of a polyphosphate mutant, despite polyphosphate being detectable in TEM images. Glycogen and PHB assays were inconclusive due to a lack of nodule material, however this heterogeneity may have led to the same masked phenotypes. Cell sorting may enable biochemical assays for polyphosphate, PHB, lipids and glycogen in these strains.

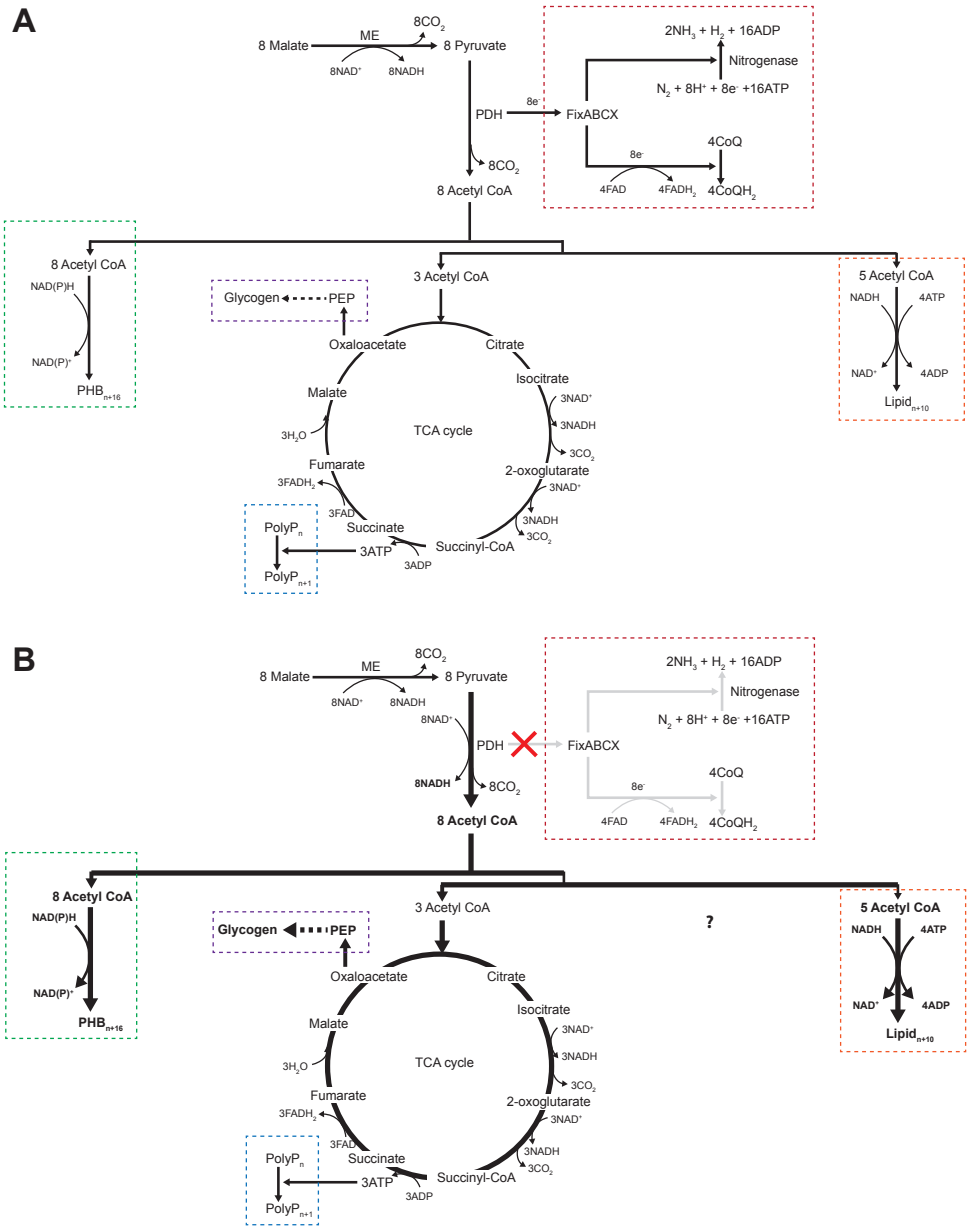


Figure 5.24. Scheme showing routes for electrons and redox energy within a A. nitrogen-fixing bacteroid and B. non-fixing mutant bacteroid. Nitrogen fixation (red); PHB (green); glycogen (purple); polyphosphate (blue); lipid (orange). ME: malic enzyme; PDH: pyruvate dehydrogenase; PEP: phosphoenolpyruvic acid.

Biochemical investigation of *fixAB* mutants is difficult due to the heterogeneity seen in nodules; there are multiple bacteroid phenotypes seen in mature bacteroids. Raman microscopy provides a strategy to investigate phenotypes within heterogenous microbial populations (Hermelink, Brauer et al. 2009). This technique is able to investigate the biochemical phenotype of a culture at the single-cell level. Bacteroid preparations were analysed using a Raman microscope in order to investigate the phenotype of the *fixAB* mutant. Comparison between free-living and bacteroid cells of Rlv3841 served to demonstrate the effectiveness of Raman microscopy in our strain, clearly showing differences between these populations.

The *fixAB* cultures were separated manually, initially depending on size of PHB peak and latterly based on size. Clear differences could be seen between the PHB peaks in oversized cells, confirming the character of the electron-transparent inclusions of the swollen, misshapen bacteroids seen in electron microscopy. The polyphosphate peak in these larger cells was much lower, whilst the polyphosphate peak in the smaller infection-thread-like cells resembled that of wild-type. Polyphosphate can be detected in wild-type bacteroids in electron micrographs, but it is only prevalent in infection threads. This data suggests that polyphosphate levels are as high in wild-type bacteroids as in infection threads; they may be obscured in electron micrographs due to the comparative cell size. When polyphosphate was quantified in the *fixAB* mutant the levels were at almost zero. This suggests that presence of the larger PHB-filled cells may have been masking the phenotype of the other, smaller cells.

Double mutants of *fixAB* and *ppk* appear to make higher than normal levels of both PHB and glycogen, resembling the phenotype of the larger *fixAB* mutant phenotype, though at about half the relative peak size. This may be a result of averaging between the two cell sizes.A double mutant in *fixAB* and *phaC2* (the bacteroid-specific PHB synthase) is reduced in polyphosphate production to a level similar to a polyphosphate mutant. Coupling of this data to transmission electron microscopy data may allow for more interpretation, as presence or absence can be confirmed, or it may be that these strains should be separated by cell size as with the *fixAB* single mutants.

In order to confirm the identity of PHB and glycogen mutants, the *glgA* mutant characterised by Lodwig, Leonard et al. (2005) and the *phaC1phaC2* mutant characterised by Terpolilli, Masakapallic et al. (2016) were inoculated on plants and bacteroids analysed by Raman. The data showed a coupling between PHB and glycogen production. This was supported by an increase in glycogen in the *fixAB* mutant bacteroids with high PHB. In *R. etli* the opposite phenotype has been suggested; a PHB mutant makes higher levels of glycogen than the wild-type (Cevallos, Encarnacion et al. 1996). Lodwig (2001) hypothesised that either reserves are

used up faster when only one polymer is found, or that synthesis occurs at a lower rate in the respective mutants. Biochemical data is available for glycogen quantification in a PHB mutant; Lodwig (2001) showed a lack of glycogen production in a PHB mutant of *R leguminosarum* bv. *viciae* A34. The reciprocal data of PHB content in a glycogen synthase mutant was not available. Not enough nodule material was available to carry out a PHB assay during this study, or to repeat glycogen quantification on the newer PHB synthase double mutant. Growth of these strains in larger volumes will yield more material in order to carry out these tests. This data may provide insights into control of carbon metabolism more generally in bacteroids. Assays on free-living cultures will reveal if this is a symbiosis-specific phenotype.

Raman microscopy has great potential for research in this field. Raman peaks are shifted by differences in molecular weight, and so radiolabelled substrates can be followed, either fed to the plant in substrates or added to the atmosphere surrounding the inoculated plant. Radiolabelling will allow the fate of carbon and nitrogen to be followed in the wild-type, as well as in mutants of interest. There is also potential for this technique in microbial ecology, as strains will provide unique fingerprints; this style of fingerprinting has already been shown using MALDI-TOF mass spectrometry (Ziegler, Mariotti et al. 2012). Outside of the symbiosis field, this technique may be useful for single-cell phenotyping of difficult-to-culture bacteria such as from deep ocean or soil samples.

Raman and transmission electron microscopy have proven to be complementary in this study, allowing much better biochemical characterisation of nitrogen fixation mutants. These techniques have revealed a temporal phenotype for bacteroids that are unable to fix nitrogen. Initially the bacteroid diverts electrons and reductant into alternative pathways such as PHB and glycogen synthesis. As nodules mature and the plant detects a lack of input from the bacterial symbiont, it ceases providing carbon and other signals to the nodule, and bacteria no longer differentiate upon release from the infection thread.

CHAPTER 6

Regulation of symbiotic nitrogen fixation

6.1 Introduction

The *fixABCX* operon comprises some of the most upregulated genes in bacteroids (Karunakaran, Ramachandran et al. 2009). Control of *fixABCX* has been shown in many strains to be under the control of NifA, the general transcriptional activator of nitrogen fixation (Martinez, Palacios et al. 2004, Salazar, Diaz-Mejia et al. 2010). In *Rhizobium leguminosarum* bv. *viciae* 3841, the *fixA* gene is preceded by a 1070 bp intergenic region (IGR). Initially I set out to determine which regulatory elements could be found within this IGR, as well as determining whether low oxygen conditions were sufficient to activate *fix* gene expression. This led to further investigation into the expression of the *nifA* gene, and demonstration of NifA autoregulation in strain Rlv3841 by both promoter analysis and complementation studies.

6.2 Analysis of NifA-controlled genes in *R. leguminosarum* bv. *viciae* 3841

Consensus binding sequences can be found for the predicted regulatory elements of *fixA*. A NifA consensus sequence was defined in *R. etli* by Salazar, Diaz-Mejia et al. (2010) as a palindromic sequence, 5'-TGT-N_{10}-ACA-3'. The same study described a consensus sequence for RpoN as 5'-TGGCACG-N_4-TTGCW-3'. An RpoN site was identified in *R. leguminosarum* bv. *viciae* UPM791 as 5'-TGGCAC-N_6-TGCT-3' (Martinez, Palacios et al. 2004).

Several operons controlled by *nifA* promoters in Rlv3841 were identified by Karunakaran, Ramachandran et al. (2009). In addition, several other genes on pRL10 are preceded by the NifA consensus sequence, as determined by a manual search with sequence 5'-TCT-N_{10}-ACA-3'. Genes with a NifA-binding site identified within 500 bp of the ATG codon were selected as putative NifA-controlled genes. These are described in Table 6.1. A similar search was carried out using the RpoN consensus sequence identified by (Martinez, Palacios et al. 2004). This sequence could identified upstream of *fixA*. This sequence could be found three more times on pRL10, each time within a gene, and not immediately upstream as would be expected for sigma-54 binding. These genes were pRL100062, pRL100345 pRL100428. None of these sites appeared near identified NifA consensus binding sites. It may be that the RpoN consensus for Rlv3841 differs from that identified for RlvUPM791. In order to identify the Rlv3841 RpoN consensus sequence, promoter regions from putative NifA-controlled genes were extracted. A MEME search of these upstream regions updid not

yield any potential new consensus sequences (Bailey, Boden et al. 2009). The majority of NifA-controlled operons found on pRL10 are highly upregulated in bacteroids, though are not necessarily essential for nitrogen fixation. Those not upregulated are generally psuedogenes or encode hypothetical proteins. These psuedogenes genes and their binding sites may be a result of duplications within the genome. This may be the case for pRL100100, a short 147 bp ORF that shows homology to the 5' end of *nifH*.

Gene	Fold elevation of expression of gene[a]	Operon[b]	Gene mutated	Description	Phenotype of mutants	Source
pRL100026	0.7	pRL100026	n/a	conserved hypothetical	n/a	This work
pRL100072	0.62	pRL100072	n/a	putative SBP of ABC transporter PAAT	n/a	This work
pRL100100	65	pRL100100-097	pRL100100	putative nitrogenase iron protein, pseudogene	Fix^+	Karunakaran, Ramachandran et al. (2009)
			pRL100099	conserved hypothetical	Fix^+	
			pRL100097	conserved hypothetical	n/a	
pRL100101	105	pRL100101-106	pRL100101	pseudogene, acetolactate synthase	Fix^+	Karunakaran, Ramachandran et al. (2009)
			pRL100103	alcohol dehydrogenase	Fix^+	
			pRL100104	conserved hypothetical	Fix^+	
			pRL100105	putative polyhydroxyalkanoate synthase subunit C	Fix^+	
			pRL100106	hypothetical	Fix^+	
pRL100115	2.11	unknown	n/a	putative acetolactate synthase, pseudogene	n/a	This work
pRL100119	34	pRL100119-122	pRL100119	propionate CoA-transferase	Fix^+	Karunakaran, Ramachandran et al. (2009)
pRL100126	1.14	pRL100126-124	n/a	putative transposase	n/a	This work
pRL100144	0.98	pRL100144-145	n/a	conserved hypothetical	n/a	This work

pRL100154	5	pRL100154	n/a	hypothetical	n/a	Karunakaran, Ramachandran et al. (2009)
pRL100155	1.65	pRL100155	n/a	putative transmembrane protein	n/a	This work
fdxB1 (pRL100156)	9.46	pRL100156	*fdxB1*	putative ferredoxin	Fix^+	This work
nifH (pRL100162)	60	pRL100162-156	*nifH*	nitrogenase iron protein	Fix^-	Karunakaran, Ramachandran et al. (2009)
fixA (pRL100200)	120.42	pRL100200-197	*fixAB*	electron transfer protein	Fix^-	This work
			fixB	electron transfer protein	Fix^-	
			fixC	electron transfer protein	Fix^-	
			fixX	electron transfer protein	Fix^-	
pRL100201	26	pRL100201-202	pRL100201	conserved hypothetical	Fix^+	Karunakaran, Ramachandran et al. (2009)
pRL100396	0.66	pRL100396	n/a	putative SBP of ABC transporter CUT2	n/a	This work

Table 6.1 Genes either preceded by a NifA-binding sequence as identified in this work by manual search of the NifA-binding TGT-N_{10}-ACA consensus sequence (Martinez, Palacios et al. 2004) or; identified by Karunakaran, Ramachandran et al. (2009) as having an upstream copy of the *nifH* promoter. a: compared to 28 dpi bacteroids; b: putative operons for this work were classified using the BioCyc.org database; genes >2fold upregulated are highlighted in green. n/a: no mutation created in this gene.

6.3 Regulation of the *fixABCX* operon

6.3.1 Identifying regulatory elements in the *fixA* promoter

Analysis of the *fixABCX* upstream promoter region reveals the presence of three potential NifA-binding regions, as well as a single RpoN binding region at the -68 position (with respect to the *fixA* ATG start codon). Transcriptional 5'-profiling of the Rlv3841 genome has been carried out in the Poole lab. (Green and Poole, unpublished). This involved growing Rlv3841 under multiple conditions in order to express the majority of genes, before RNAseq profiling to identify putative transcriptional start sites. Similar work has been carried out in *S. meliloti* (Schluter, Reinkensmeier et al. 2013). The Poole lab data set was used to determine potential transcriptional start sites within the *fixA* promoter region. A putative start site was found at the -17 position (with respect to the *fixA* ATG start codon), found downstream of the RpoN site and the proximal NifA site. An additional putative start site can be found at the -679 position, but this is unlikely to be the functional start site due to its distance from the *fixA* start codon. The distal NifA-binding site may act as part of the regulation for divergent gene pRL100201, which is also upregulated in Rlv3841 bacteroids. These regulatory elements are annotated on the IGR in Figure 6.1.

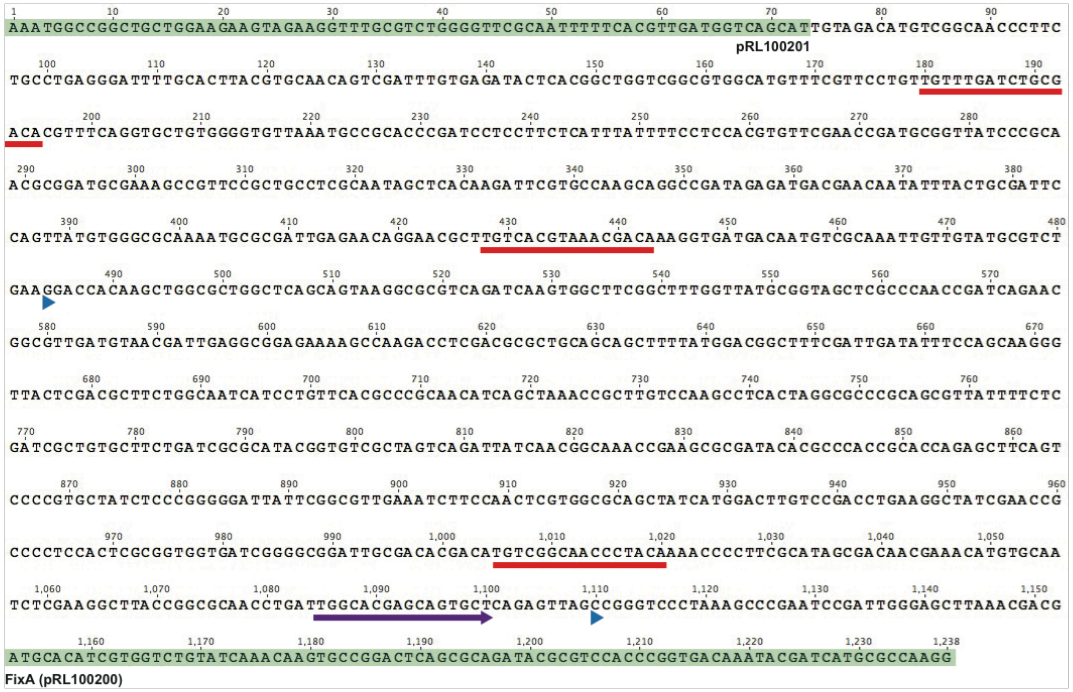


Figure 6.1 IGR upstream of *fixA* from beginning of divergent gene pRL100201 to beginning of *fixA*; annotations show genes (green), NifA binding sites (red), RpoN binding site (purple) and potential transcription start sites (blue) as determined by 5' transcriptional start site mapping (Green and Poole, unpublished).

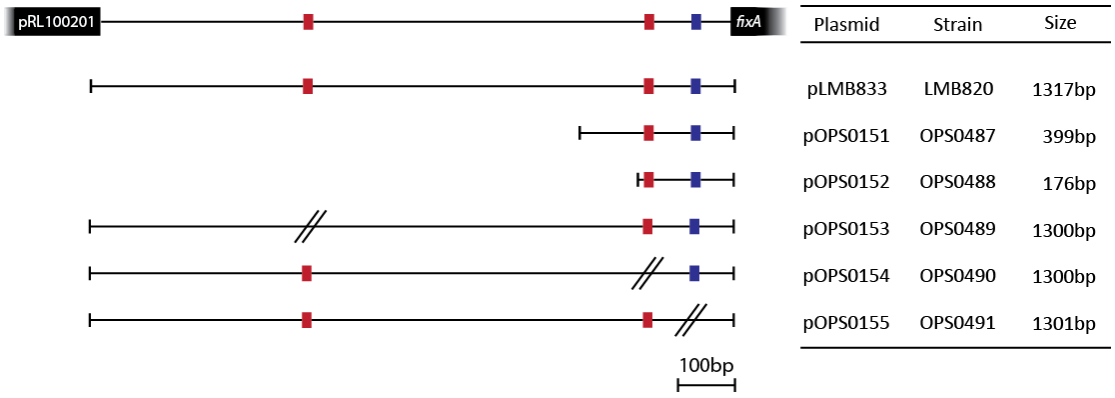
6.3.2 Elucidating essential regulatory elements in the *fixA* promoter using a *lux*-based reporter system

A light (*lux*)-based reporter system was used to test which of the identified regulatory sequences are involved in *fixA* expression. The *lux* system, described by Frederix, Edwards et al. (2014), utilises a reporter plasmid, pIJ11268. pIJ11268 contains the luciferase genes *luxCDABE* constructed in the backbone of pJP2, a stable vector designed for use in rhizobia (Prell, Boesten et al. 2002). A promoter region of interest can be cloned into the multiple cloning site (MCS) preceding the *lux* genes, and luminescence can then be detected by a photon counting camera in order to determine promoter activity semi-quantitatively.

A series of promoter fusions were designed in the pIJ11268 backbone in order to investigate which regions of the IGR are required for *fix* gene expression. Two strategies were taken. The first was to truncate the promoter region down from ~1 kb to find a minimal region for expression. The entire IGR was amplified using primers pr1642 and pr1643, and the resulting fragment was cloned into pIJ11268 using *KpnI/BamHI* to give plasmid pLMB833.

Sense primers oxp0374 and oxp0375 were used with pr1643 to give pOPS0151 and pOPS0152 respectively.

A second strategy used overlap PCR to delete the regulatory elements within the promoter region. To remove NifA-binding sites the entire 5'TGT-N_{10}-ACA-3' consensus sequence was deleted (annotated in Figure 6.1 in red). To remove RpoN-binding site the entire 5'-TGGCAC-N_6-TGCT-3' (annotated in Figure 6.1 in purple) consensus sequence was deleted. This resulted in pOPS0153, lacking the distal NifA site (deleted using oxp0494/oxp0495); pOPS0154, lacking the proximal NifA site (deleted using oxp0496/oxp0497); and pOPS0154 (deleted using oxp0498/oxp0499), lacking the RpoN site. The promoter regions, plasmids and strains used are illustrated in Figure 6.2.



Plasmid	Strain	Size
pLMB833	LMB820	1317bp
pOPS0151	OPS0487	399bp
pOPS0152	OPS0488	176bp
pOPS0153	OPS0489	1300bp
pOPS0154	OPS0490	1300bp
pOPS0155	OPS0491	1301bp

Figure 6.2 Promoter regions used in *lux*-based reporter for analysis of the *fixA* promoter. Annotations show the NifA (red) and RpoN (blue) binding sites. The NifA-binding site is based on consensus TGT-N_{10}-ACA (Martinez, Palacios et al. 2004). The RpoN-binding site is based on consensus TGGCAC-N_6-TGCT-3' (Benhassine, Fauvart et al. 2007).

Luminescence was initially tested in free-living culture, in order to confirm data seen by microarrays that the *fix* operon is not expressed under free-living conditions. The strains containing the *fix* reporter plasmids were grown on rich media agar and imaged using a NightOwl photon counting camera. No luminescence was seen from any of the strains (Figure 6.3). The empty pIJ11268 in WT background, LMB542, was used as a negative control. Plasmid pIJ11282, the *lux* plasmid containing a constitutive promoter was used as a positive control, conjugated into the WT background to give strain D5250 (Frederix, Edwards et al. 2014). This result was confirmed by growth of the strains in TY media in a FluoStar Omega plate reader with measurement of both optical density and luminescence (measured in relative light units, RLU). This allowed for determination of maximum

luminescence during exponential growth (Table 6.2). Only the positive control was able to express a significant amount of luminescence under free-living growth conditions.

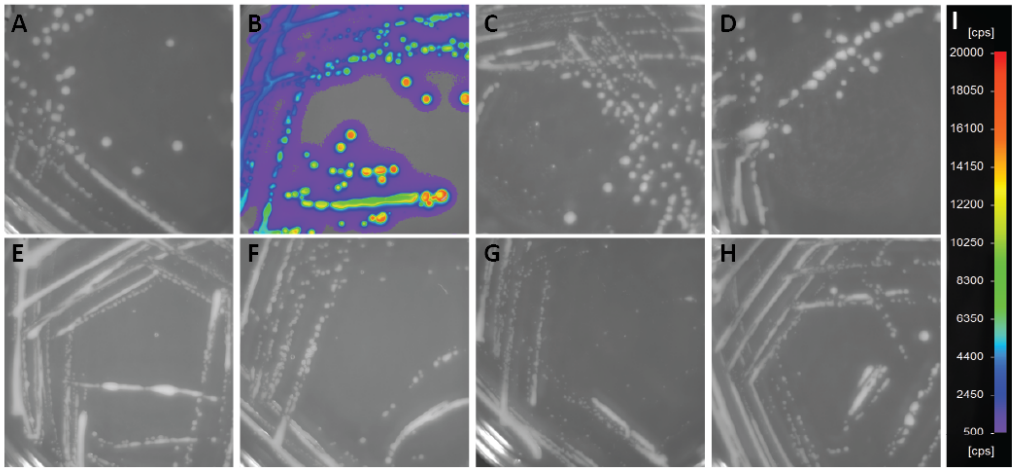


Figure 6.3 Plate luminescence phenotype of *fixA* promoter reporter strains. A: LMB542 (negative control); B: D5250 (positive control); C: LMB820; D: OPS0487; E: OPS0488; F: OPS0489; G: OPS0490; H: OPS0491; I: Scale bar from 500 (purple) to 25000 (red) counts per second (cps).

Plasmid	Strain	Description	Maximum luminescence RLU per OD_{600}
pIJ11268	LMB542	Empty vector	2.9 ± 2.9[a]
pIJ11282	D5250	pIJ11268 containing constitutive neomycin promoter	409921 ± 3407[b]
pLMB833	LMB820	pIJ11268 containing full *fixA* promoter	1044 ± 12[a]
pOPS0151	OPS0487	pIJ11268 containing truncated 400 bp *fixA* promoter	489 ± 52[a]
pOPS0152	OPS0488	pIJ11268 containing truncated 176 bp *fixA* promoter	138 ± 19[a]
pOPS0153	OPS0489	pIJ11268 containing full *fixA* promoter lacking distal NifA binding site	1028 ± 90[a]
pOPS0154	OPS0490	pIJ11268 containing full *fixA* promoter lacking proximal NifA binding site	836 ± 95[a]
pOPS0155	OPS0491	pIJ11268 containing full *fixA* promoter lacking RpoN binding site	6.8 ± 6.8[a]

Table 6.2 Maximum luminescence per OD_{600} for p*fixA* reporter constructs quantified using FluoStar Omega plate reader. RLU: relative light units. a, b: statistically distinct groups as determined by an ANOVA and post-hoc Tukey test.

In order to confirm the functionality of the *fixA* reporter fusions, the strains were grown on plants to induce expression under N_2-fixing conditions. Pea (*P. sativum*) seeds were planted in sterile vermiculite and inoculated with $1x10^7$ colony-forming units (cfu) of the strain of interest. Each plant was grown for 21 days and then imaged using the NightOwl camera. The results of this are shown in Figure 6.4. From this initial qualitative view of *lux* expression it can be concluded that both the proximal NifA-binding site and the RpoN-binding site are required for expression of *fixA*. Since strain OPS0488 was able to express the luciferase genes, it can be concluded that the 176 bp region containing both these sites is the minimal required promoter region for expression of these genes.

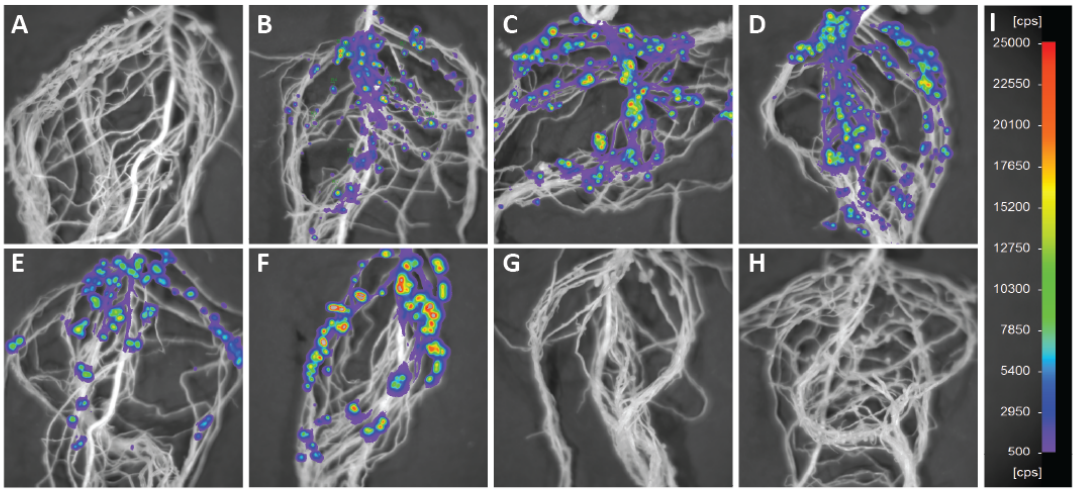


Figure 6.4 Plant luminescence phenotype of *fixA* promoter reporter strains 21 dpi. A: LMB542; B: D5250; C: LMB820; D: OPS0487; E: OPS0488; F: OPS0489; G: OPS0490; H: OPS0491; I: Scale bar from 500 (purple) to 25000 (red) counts per second (cps).

Luminescence was quantified by calculating the average luminescence per nodule for a sample set of 10-20 nodules per plant. Quantified luminescence values can be seen in Figure 6.5. Quantitation of the *lux* expression in these clones allows for better analysis of the promoter constructs. The *fix* promoter is one of the most over-expressed in bacteroids, as highlighted by the fact it is up to four-fold more expressed than the constitutive neomycin promoter (D5250). A two to three-fold increase in expression can be seen in constructs lacking the distal NifA binding site (OPS0487 and OPS0489). This effect was not seen in OPS0488, however, suggesting possible additional regulatory elements in the 223 bp region included in the OPS0487 promoter. The higher expression from promoters lacking the second NifA site may be due to competitive effects between the multiple NifA promoter-binding sites found in the IGR.

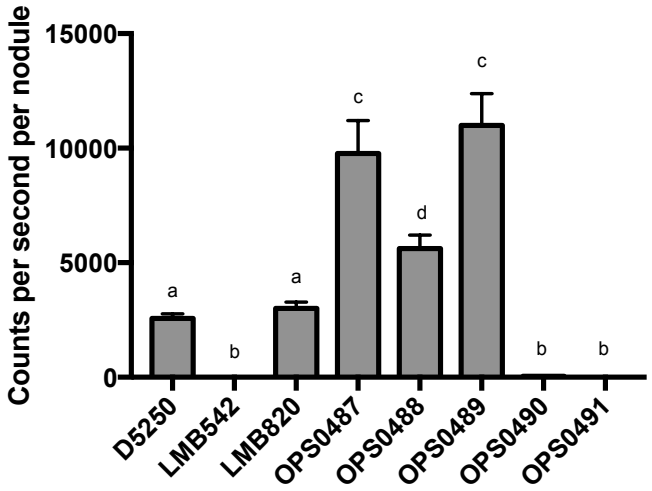


Figure 6.5 Quantified luminescence values for nodules 21 dpi. Values presented as counts per second per nodule; n= 30-70; a, b, c, d represent statistically distinct groups as determined by an ANOVA and post-hoc Tukey test.

6.3.3 Expression from the *fixA* promoter under microaerobic conditions

The *fixABCX* operon is under the control of NifA, the general transcriptional regulator of nitrogen fixation, whose expression is thought to be mainly under nitrogen and oxygen control. To determine if NifA-dependent *fix* gene expression is solely oxygen-dependent, a luciferase assay was carried out in the FluoStar Omega plate reader under 1% oxygen conditions in rich media. Expression of the full p*fix* promoter was analysed under ambient and low oxygen conditions. D5250 (constitutive *lux* expression) was included to account for changes in luciferase expression due to altered metabolic activity under microaerobic conditions. A two-way ANOVA determined that a significant change can be seen between 21% and 1% O_2 in strain D5250, but not in LMB542 or LMB820; expression of the *fixA* promoter has not been affected by microaerobic conditions (Figure 6.6). Other factors within the nodule must be involved in expression of nitrogen fixation genes. This may be another environmental condition such as low nitrogen, pH, or provision of unknown plant factors. To test the effect of plant factors, a luciferase assay was carried out with nodule extract in the media. Rich media was supplemented with 10% nodule extract, and cultures grown at 1% O_2 to recreate the nodule environment. Addition of nodule extract did not significantly increase the level of expression from the *pfixA* promoter.

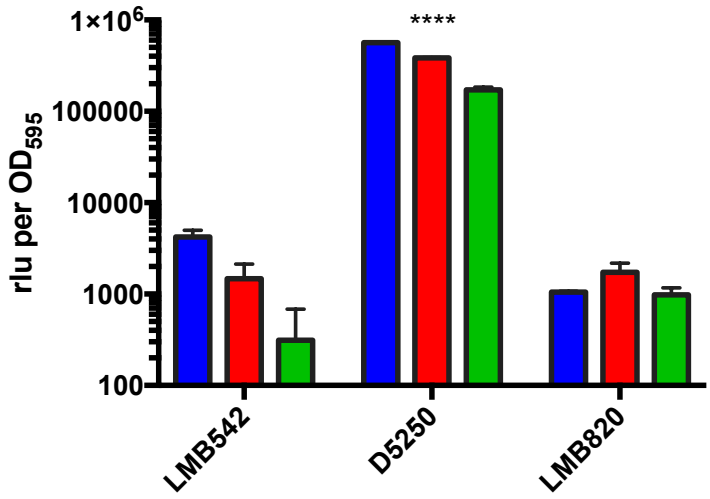


Figure 6.6 Maximum luminescence per OD_{600} for p*fixA* reporter construct quantified using FluoStar Omega plate reader under 21% O_2 (blue), 1% O_2 (red) and 1% O_2 plus nodule extract (green). LMB542: empty vector. D5250: constitutive neomycin promoter. LMB820: *pfixA* promoter. *Lux* activity expressed in relative light units per OD_{595}. ****: statistical significance between groups determined by an ANOVA and post-hoc Sidak's test.

6.3.4 Confirmation of the minimal promoter region for *fixA*

The luciferase system has shown a minimal region of 176 bp is required for expression of the *fixABCX* operon. In order to confirm this, a promoter (PU) module was designed for use in Golden Gate cloning. This allows generation of a construct in which the p*fixA* promoter drives alternative reporter molecules, such as sfGFP, which have been developed for construction using Golden Gate. Primers oxp0822 and oxp0823 were used to amplify the 176 bp promoter region. The resulting PCR product was cloned into pL0V-PU using *BpiI* to form pOGG095. A Golden Gate cloning reaction was set up using pOGG095 (pL0M-PU-pfix), pOGG037 (pL0M-sfGFP), pOGG003 (pharmacia terminator pL0M-T-pharma) and pOGG026 (pJP2 homolog, Geddes et al., in production) to construct plasmid pOPS0325. Plasmid pOPS0325 was conjugated into Rlv3841 to give strain OPS0749. This strain was grown on peas (*P. sativum*) and imaged at 28 dpi. Plants were imaged using both a stereo microscope and the NightOwl camera, both with GFP filters (excitation wavelength 450-490 nm, emission wavelength 500-550 nm). Plants inoculated with wild-type 3841 show background autofluorescence levels expected from roots and nodules. GFP expression can be clearly seen in mature nodules of OPS0749. Using GFP as a reporter system provides far better spatial resolution than the luciferase-based assay and shows localisation of *fix* gene expression to the nitrogen-fixing zone of the nodules, whilst the nodule meristem shows no

fluorescence. Whole-plant fluorescence imaging using the NightOwl system shows expression across all nodules, showing good stability of this plasmid. This acts as a proof of concept for use of the *fixA* promoter module in future Golden Gate constructs for expression in the nitrogen-fixing zone of the nodules. Additionally, this acts proof of concept for stability of GFP expression from the pOGG026 vector in mature bacteroids, something previously not demonstrated using this new vector.

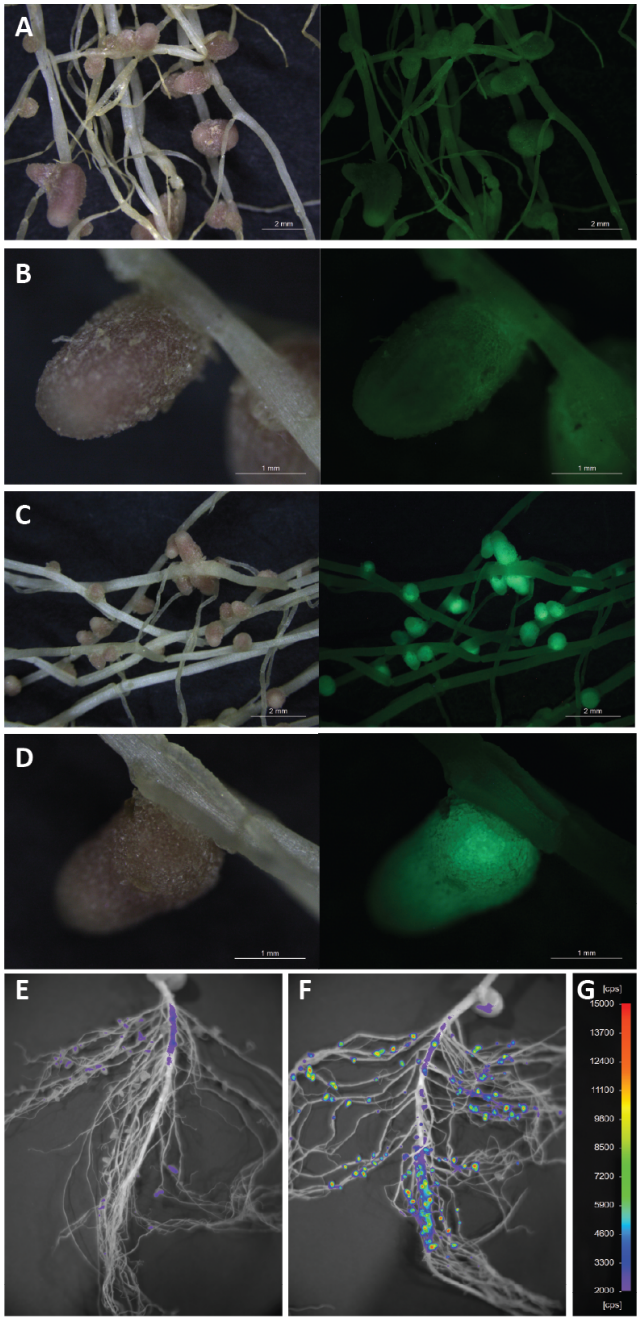


Figure 6.7 Root GFP phenotype of *fixA* promoter reporter strain at 28 dpi. A, B, E: 3841; C, D, F: OPS0749; A, B, C, D: Images taken using Leica using both a visible light filter (left) and GFP filter (right); E, F: Images taken using NightOwl camera. G: Scale bar from 2000 (purple) to 15000 (red) counts per second (cps).

6.4 Elucidating a *nifA* promoter in *R. leguminosarum* bv. *viciae* 3841

Activity of the NifA protein has been shown to be affected by the presence of oxygen (Huala and Ausubel 1989, Morett, Fischer et al. 1991). In some diazotrophs *nifA* gene expression is regulated in response to both oxygen and nitrogen status (Dixon and Kahn 2004). To determine if the lack of free-living expression of *fixA* is due to lack of expression of the *nifA* gene or lack of activity of the NifA protein, the promoter region of *nifA* was also characterised using the *lux* system. Previous studies have suggested that NifA autoregulates its own expression from an upstream site with low-level expression from a promoter immediately upstream of *nifA*. This has been demonstrated in *R. leguminosarum* bv. *viciae* UPM791 (Martinez, Palacios et al. 2004) as well as in *S. meliloti* (Buikema, Szeto et al. 1985). In UPM791 expression is from a promoter within an open reading frame, *orf71*, found within a gene cluster *orf71 orf79 fixW orf5 fixA fixB fixC fixX nifA nifB*. In this strain of *R. leguminosarum* there is no intergenic region upstream of *fixA*. In *S. meliloti*, NifA autoregulation is thought to be from the *fixA* promoter. In both species, a second promoter is predicted just upstream of *nifA*. The *nifA* promoter identified in *R. leguminosarum* strain UPM791 showed no expression in the free-living state (Martinez, Palacios et al. 2004) although expression was seen from the *S. meliloti* p*nifA* in both symbiotic and free-living conditions (Kim, Helinski et al. 1986, Benhassine, Fauvart et al. 2007). In *R. etli* the *nifA* promoters were predicted to be upstream of the 3' end of *fixX,* rather than in the *fixX-nifA* IGR. In UPM791 it was hypothesised that this low-level promoter was found upstream of *fixA*, though this could not be proven experimentally (Martinez, Palacios et al. 2004).

6.4.1 Comparison of the *fixABCXnifA* region across rhizobia

The BioCyc database was used to investigate synteny of *nifA* within multiple rhizobial species (Figure 6.8). Fast-growing rhizobia (members of the *Rhizobiacae*) are syntenous with *R. leguminosarum* bv *viciae* 3841, where the *nifA* gene is found immediately downstream of the *fixABCX* operon with a 100-300 bp intergenic region between *fixX* and *nifA*. This organisation does not occur in in free-living diazotrophs such as *B. japonicum* or *A. caulinodans*, though they do possess *nifA* and *fixABCX* genes.

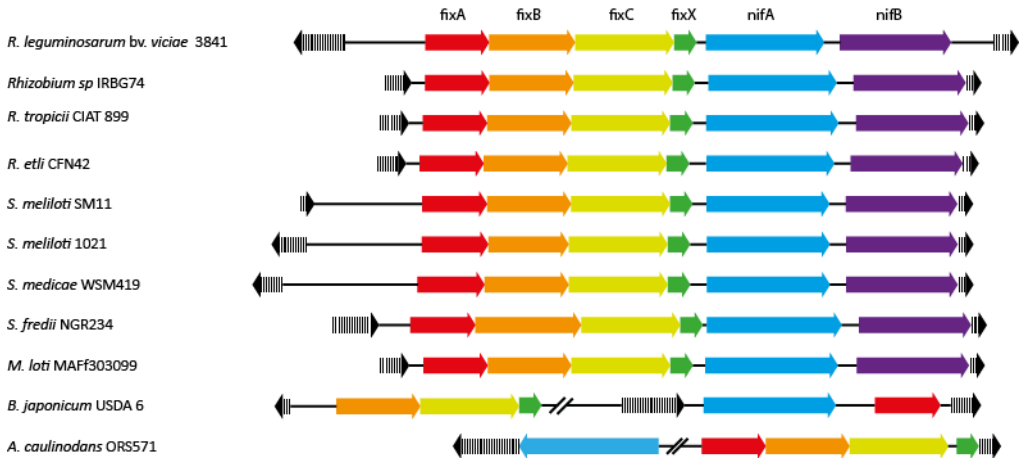


Figure 6.8 Alignment of the *fixABCXnifA* gene cluster of several sequenced rhizobia. *fixABCXnifAB* gene cluster annotated in red (*fixA*), orange (*fixB*), yellow (*fixC*), green (*fixX*), blue (*nifA*) and purple (*nifB*). Flanking genes are shown in hashed black arrows. Sequences obtained from BioCyc.org.

6.4.2 Construction of *nifA* promoter fusions

Less data is available for prediction of the *nifA* promoter sequence. A FixJ-binding region for has been described within the *nifA* promoter of *S. meliloti* (TGAMGTAG) (Ferrieres and Kahn 2002), however Rlv3841 lacks any likely orthologs. This binding sequence cannot be found upstream of the Rlv3841 *nifA* gene. Bioinformatics was used to predict putative promoter elements within the region upstream of *nifA*. PePPER promoter prediction software (de Jong, Pietersma et al. 2012) was used to search for a *nifA* promoter sequence, which searches for classical prokaryotic -10 and -35 promoter motifs. A putative -10/-35 motif was predicted at the -74 position within the *fixX-nifA* IGR. 5'-transcription start site mapping data (Green and Poole, unpublished) was also used to predict the start site for *nifA* expression. Four putative start sites were identified, all of which can be found within the *fixC* ORF. ARNold software (Gautheret and Lambert 2001) was used to search for transcriptional terminators within *fixABCXnifA* in order to determine if these genes formed a single transcriptional unit. No terminators could be found within this region.

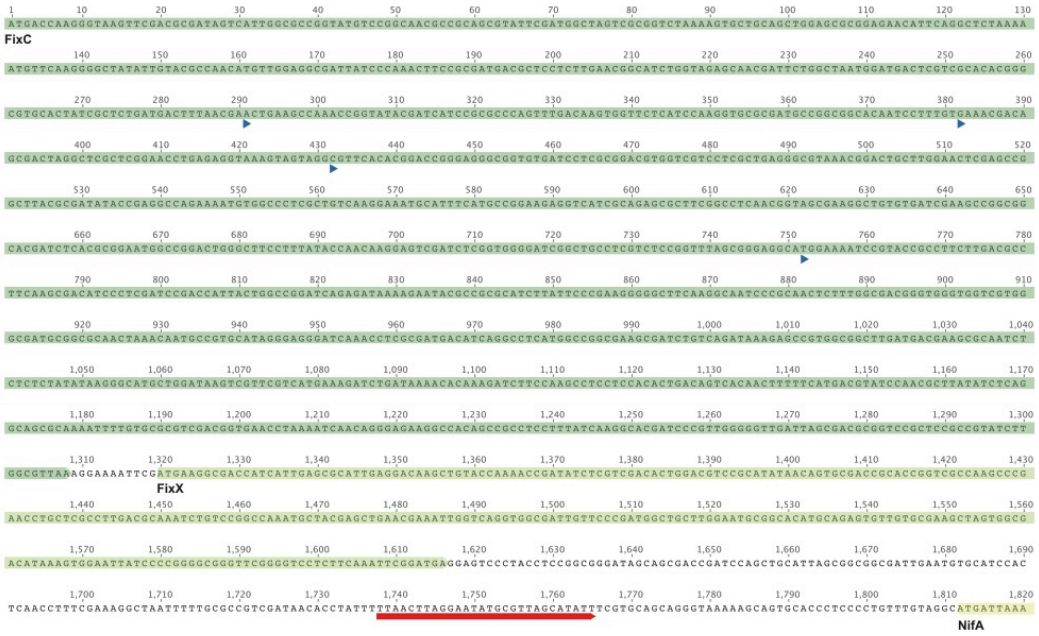


Figure 6.9 Region upstream of *nifA* annotated with putative promoter and transcription start sites; annotations show genes (green/yellow); predicted promoter (red) as predicted by PePPER; and potential transcription start sites (blue) as determined by 5'-mapping (Green and Poole, unpublished).

Initially, three constructs were designed to try and elucidate the region of the plasmid containing the *nifA* promoter. The first comprised the entire *fixCX-IGR* region, amplified using oxp0382 and oxp385, resulting in plasmid pOPS0156. The second lacked *fixC* (oxp0383/oxp0385; pOPS0157) and the third consisted of just the *fixX-nifA* IGR (oxp0384/oxp0385; pOPS0158). Preliminary data led to the construction of three further constructs, truncating *fixX* into two increasingly shorter regions (using forward primers oxp0791 and oxp792 to give pOPS0279 and pOPS0300 respectively), as well as a construct amplified from the *fixX* deletion mutant (OPS0469) using oxp0382 and oxp0385, so containing just *fixC* and the IGR (pOPS0301). The promoters used are outlined in Figure 6.10.i along with the resulting plasmids and strain in the WT background.

Expression was seen from *nifA* constructs in free-living colonies grown on rich media (Figure 6.10.iv A-F); *nifA* may be under control of a basal constitutive promoter. To elucidate if this basal promoter is active in plants, each strain was grown on peas for 21 days before imaging in the NightOwl camera (Figure 6.10.iv). Again, nodule expression was quantified using a sample of 10-20 nodules per plant. Expression from the *nifA* promoter is <30% of the expression from the *fixA* promoter constructs (Figure 6.10.ii). This higher expression is seen in 21 day bacteroid microarray data, where *fixA* is expressed 132-fold

relative to free-living culture, compared to 29-fold for *nifA* (Karunakaran, Ramachandran et al. 2009). Only one promoter was unable to express the luciferase genes. This was the promoter containing only the *fixX-nifA* intergenic region, correlating with data seen in previous studies that expression occurs upstream of the 3' end of *fixX*. However, a region as short as 100 bp of *fixX* is sufficient to allow expression, suggesting that a basal promoter is present within the final 100 bp of *fixX*. However, expression from a promoter region lacking *fixX* suggests that there may be additional basal promoters within the *fixC* open reading frame. This suggests the presence of multiple basal promoters controlling *nifA* gene expression. It should be considered that the cloning procedure may have led to introduction of new promoter regions or terminators ahead of the *lux* genes. Confirmation using a new system such as the Golden Gate-based GFP system may help to support or refute the data seen from these lux constructions.

Quantitative values for these constructs show an incremental decrease in expression with the length of the promoter. This phenomenon has been shown to occur previously by Benhassine, Fauvart et al. (2007), who saw decreases in expression with progressive deletions from the 5' end of the *nifA* promoter. Removal of the *fixX* region does not lead to a decrease in expression, but removal of the *fixC* region does, suggesting that the majority of expression comes from within *fixC*. Truncation of *fixX* leads to decreases in expression; there are multiple promoters within *fixX*. Expression from the *nifA* promoters *in planta* is ~10% of the expression seen for the *fixA* promoter.

6.4.3 Expression of *nifA* under free-living and microaerobic conditions

Previous work in *R. leguminosarum* (Martinez, Palacios et al. 2004) suggested that expression of *nifA* only occurs within the nodule, and that there is no basal expression in the free-living state. The *nifA* reporter fusions generated in this work luminesce on agar plates. Free-living expression was further tested by growing all *nifA* promoter fusions in the FluoStar Omega plate reader under ambient oxygen in rich media. Expression was seen from all the functional promoter regions. Though not as pronounced as in the plant assay, a decrease in expression can be seen in the progressive deletions, with the longest promoter region showing higher expression than the shortest functional promoter region (Figure 6.10.iii). Additionally, the same experiment was run under microaerobic conditions to determine if *nifA* basal expression was higher under 1% oxygen. A two-way ANOVA and post-hoc test was used to determine if a significant difference in expression was seen between 21% and 1% oxygen. A significant decrease was seen for strains OPS0492 and OPS0493 though this could be due to the metabolic effects of growth at low oxygen rather

than changes in gene expression as seen previously with constitutively expressed strain D5250 (Figure 6.6). No significant difference was detected between any other strains. This suggests that *nifA* basal expression is constitutive under the free-living conditions, and not induced by a microaerobic environment.

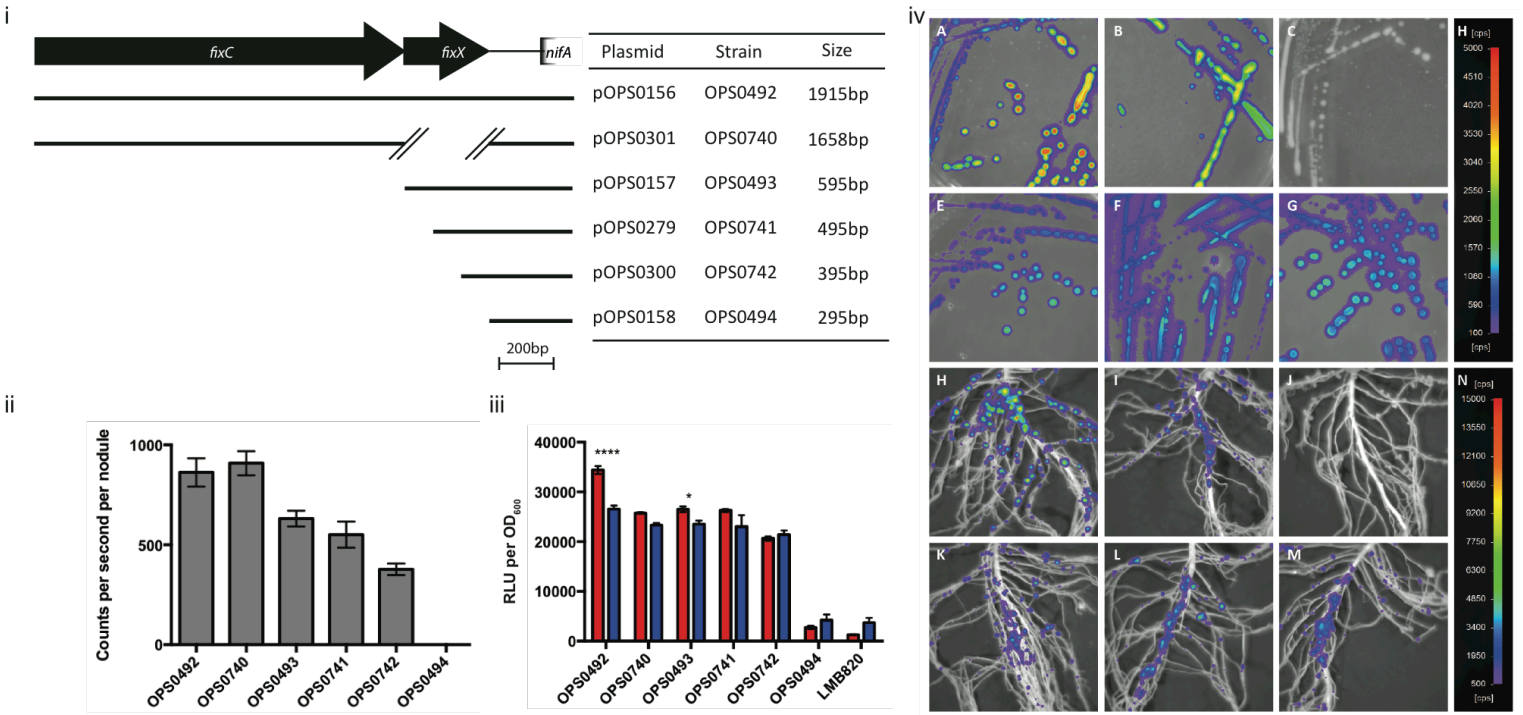


Figure 6.10 i: Promoter regions used in *lux* system for analysis of the *nifA* promoter; ii: Quantified luminescence values for nodules 21 dpi. Values presented as counts per second per nodule; n = 60; iii: Maximum luminescence per OD_{600} for p*nifA* reporter constructs quantified using FluoStar Omega plate reader under 21% (red) and 1% (blue) oxygen conditions. Expression measured in relative light units per OD_{595}. Asterisks show statistically significant distances between groups as determined by an ANOVA and post-hoc Sidak's test; * = $p \leq 0.05$, **** = $p \leq 0.0001$; iv: Plate and plant luminescence phenotype of *nifA* promoter reporter strains 21 dpi. A,H: OPS0491; B,I: OPS0493; C,J: OPS0494; D,K: OPS0740; E,L: OPS0741; F,M: OPS0742; G: Scale bar for plate images from 100 (purple) to 5000 (red) counts per second (cps); N: Scale bar for plate images from 500 (purple) to 15000 (red) counts per second (cps).

6.5 Confirmation of a *nifA* basal promoter and symbiotic *nifA* autoregulation using mutant complementation

6.5.1 Insights from polar *fixABCX* mutants

In all *fix* genes mutated, the complemented polar mutants fix nitrogen at a rate lower than that of the wild-type, whilst an in frame mutant can be returned to wild-type fixation levels by complementation. This could be explained by *nifA* autoregulation from the *fixA* promoter. Whilst the basal promoter is able to express enough NifA to switch on the nitrogen-fixation machinery, the autoregulation would increase expression of *nifA* and increase levels of the regulator in the cell. A polar mutant in the *fix* genes would prevent this autoregulation, and so less NifA is available to switch on nitrogen fixation genes. This indicates that post-transcriptional control of NifA determines expression of nitrogen fixation activity.

As previously discussed in Chapter 4, a polar *fixX* mutant cannot be complemented by provision of *fixABCX*. This would agree with the presence of a basal promoter within *fixX* required for *nifA* expression. The presence of multiple promoters within both *fixC* and *fixX* would explain why both an in frame *fixC* mutant and an in frame *fixX* mutant can be functionally complemented with *fixABCX*. In order to investigate this further, a complementing plasmid expressing the *nifA* gene is required.

6.5.2 Determining the transcriptional unit of *nifA*

Before designing any constructs for complementation of functional *nifA* mutants, we first had to consider whether or not *nifA* was co-transcribed with *nifB,* found immediately downstream on the chromosome. The work by Martinez, Palacios et al. (2004) suggested that *nifB* may be part of the same transcriptional unit as *nifA*, but had no direct evidence for this. Work in *S. meliloti* (Klipp, Reilander et al. 1989) showed *nifA* and *nifB* to be cotranscribed, though a low-level NifA-dependent promoter was also found in front of *nifB*. A Biocyc.org search showed that the *fdxN* gene is found downstream of *nifAB* in several rhizobial species including *R. etli* CIAT 652, *Sinorhizobium fredii* NGR234 and *Rhizobium tropici* CIAT 899. The FdxN ferredoxin has been shown to be required for nitrogen fixation in *Bradyrhizobium* and *Sinorhizobium* (Klipp, Reilander et al. 1989, Hauser, Pessi et al. 2007) and is highly upregulated in *A. caulinodans* bacteroids (Tsukada, Aono et al. 2009). If Rlv3841 also has a downstream *fdxN* gene then this needs to be considered as part of the operon. In 3841 the product of the gene downstream from *nifB*, pRL100194, is annotated as a hypothetical protein, and BLAST analysis of the pRL100194 gene product shows no

identity with ferredoxin FdxN. An additional BLAST search suggests that no FdxN ortholog can be found in Rlv3841. A search for the Rlv3841 *nifA* on Biocyc.org suggests that *nifA* alone forms a transcriptional unit, and *nifB* is expressed separately. Analysis of Rlv3841 microarray data (Karunakaran, Ramachandran et al. 2009) indicates higher expression of *nifB* than *nifA* in bacteroids (up to 4-fold more expression) whilst *nifA* tended to be slightly more upregulated under free-living conditions, such as in the rhizosphere (Figure 6.11). In mnay diazotrophs, such as *R. etli*, the *nifB* gene is part of the NifA-RpoN regulon, under the control of a sigma-54 promoter (Salazar, Diaz-Mejia et al. 2010). This differing control would corroborate the hypothesis that they are under separate control. ARNold (Gautheret and Lambert 2001) was used to search for transcriptional terminators within the *nifAB* genetic region in order to confirm that they form separate units, but no transcriptional terminators could be found for either gene. Additionally, no -10/-35 putative promoters could be identified upstream of *nifB* using PePPER (de Jong, Pietersma et al. 2012).

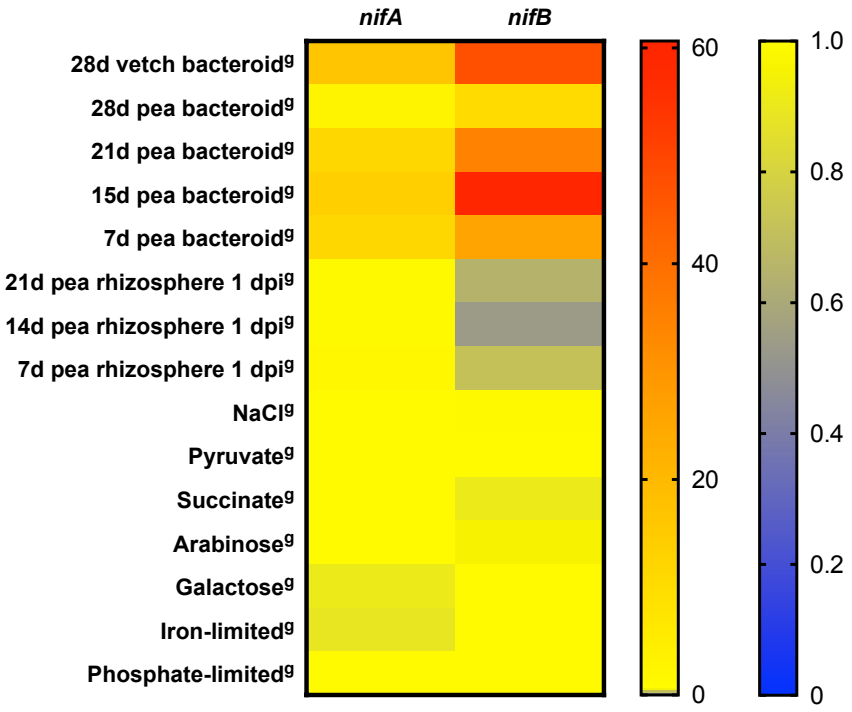


Figure 6.11 Changes in expression of the *nifA* and *nifB* (red) genes under different growth conditions. Scale bars: downregulated genes from 0-fold (blue) to 1-fold (yellow); upregulated genes from 1-fold (yellow) to 60-fold (red) All values are relative to free-living 3841 grown with either glucose (g) or pyruvate (p). dpi: days post inoculation.

6.5.3 Complementation of polar *fixX* mutant

Since there is no conclusive evidence for whether *nifAB* is transcribed as a single operon or not, two constructs were designed to complement the *fixX* mutants. The first contained only the *fixX-nifA* region. The second contained the *fixX-nifAB* region. Both were designed for use in a Golden Gate cloning reaction. Conventional cloning proved difficult since the *nifAB* gene region contains recognition sites for the majority of commonly used restriction enzymes found in the pJP2 multiple cloning site. Sense primer oxp0819 bound to the 5'-end of *fixX*, whilst antisense primers oxp0820 and oxp0821 bound the 3'-end of *nifA* and *nifB* respectively. The *nifAB* region contained several *BsaI* sites, required for Golden Gate cloning assembly. To overcome this, the *fixXnifAB* region was domesticated using Gibson assembly. Primers oxp0801-oxp0808 were used to amplify four fragments of *fixXnifAB*, with primers designed to alter a single base pair in order to remove the *BsaI* recognition site whilst maintaining translational codons. Overlap regions were designed on the ends of each fragment to allow correct Gibson assembly. The resulting DNA region was cloned into pJET1.2/blunt to give plasmid pOPS0302. Amplification using primers oxp0819, oxp0820 and oxp0821 was used to generate fragments for a Golden Gate reaction. Golden Gate assembly was carried out to assemble several constructs, shown in Table 6.3. Initially, a constitutive pNeo promoter was used. No correctly assembled plasmids could be obtained from the reaction. This led to the conclusion that constitutive expression of one of these genes was fatal in *E. coli*, and so colonies containing a correctly assembled construct could not be obtained. Instead, the bacteroid-inducible p*fixA* promoter confirmed using GFP was used (Section 6.2.4). Although this promoter requires NifA for expression, leaky expression from the plasmid may provide enough basal expression to start autoregulation within the construct. Use of the inducible promoter allowed correct assembly of the *fixXnifA* construct. However, the construct containing *nifB* would still not assemble correctly, suggesting that *R. leguminosarum* bv. *viciae* NifB is toxic in *E. coli* even in very low levels caused by leaky plasmid induction.

Promoter	Gene	Terminator	Vector backbone	Correct colonies obtained?
pNeo	*fixXnifA*	Pharmacia	pOGG026	No
pNeo	*fixXnifAB*	Pharmacia	pOGG026	No
p*fixA*	*fixXnifA*	Pharmacia	pOGG026	Yes
p*fixA*	*fixXnifAB*	Pharmacia	pOGG026	No

Table 6.3 Combinations of parts used in Golden Gate reactions for construction of a *nifA* complementing plasmid.

The successful *fixXnifA* plasmid was designated pOPS0303. pOPS0303 was conjugated into 3841, LMB829 (*fixX*::Ωspc) and OPS0469 (*ΔfixX*) to give strains OPS0746, OPS0747 and OPS0748 respectively. These strains were inoculated on pea (*P. sativum*), and these plants harvested at 28 days for acetylene reduction assays (Figure 6.12). Plasmid pOPS0325 was able to complement the in-frame *fixX* mutant, restoring nodule phenotype and nitrogen-fixation ability. This confirms that this plasmid functions correctly in the bacteroid. The polar mutant, however, could not be complemented by this plasmid. This suggests that *nifB* may be co-transcribed with *nifA*, and so a larger genetic region was required on the complementing plasmid.

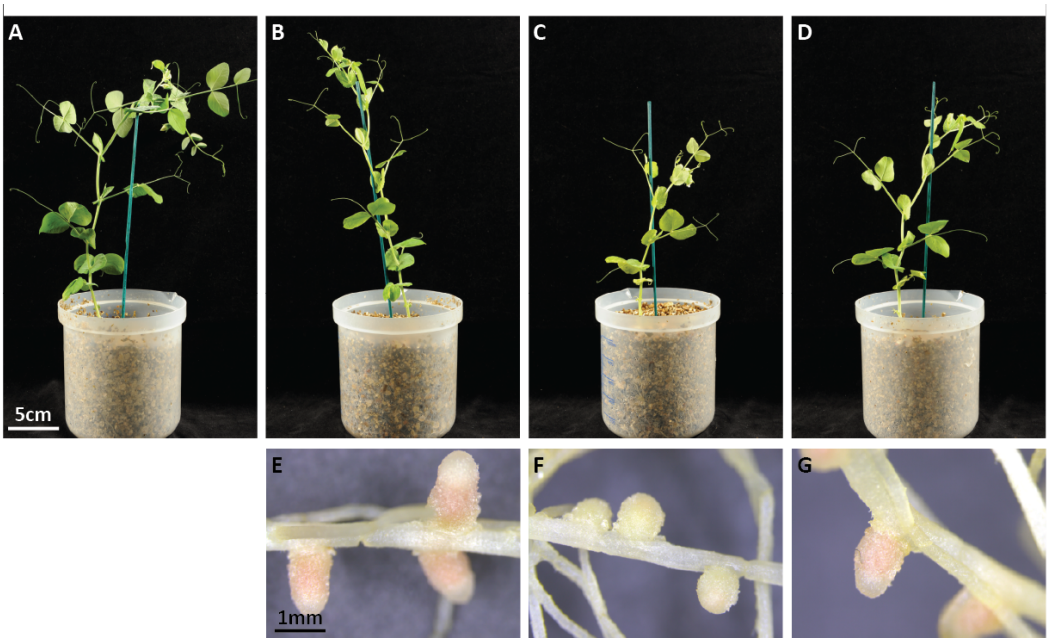


Figure 6.12 Plant phenotype inoculated with *fixX* complements at 28dpi: A: Rlv3841; B,E: OPS0746 (3841:pOPS0325); C,F: OPS0747 (*fixX*::Ωspc:pOPS0325); D,G: OPS0748 (Δ*fixX*:pOPS0325).

Strain	**Description**	**Symbiotic phenotype**	**Nodule colour**	**Acetylene reduction** µmol ethylene plant^{-1} hr^{-1}
3841	Wild-type	Nod^{+} Fix^{+}	Pink	2.6 ± 0.2^{a}
OPS0746	3841:pOPS0325	Nod^{+} Fix^{+}	Pink	2.9 ± 0.1^{a}
OPS0747	*fixX*::Ωspc:pOPS0325	Nod^{+} Fix^{-}	White	0.1 ± 0.0^{b}
OPS0748	Δ*fixX*:pOPS0325	Nod^{+} Fix^{+}	Pink	2.4 ± 0.1^{a}
Uninoculated control	n/a	Nod^{-} Fix^{-}	n/a	0.1 ± 0.0^{b}

Table 6.4 Symbiotic phenotype of complemented *fixX* mutant strains on pea (*P. sativum*). Acetylene reduction assay was carried out 28 days post inoculation, with values expressed as µmol ethylene plant^{-1} hr^{-1}; ± SEM, n ≥ 3. a, b represent statistically distinct groups ($p \leq 0.05$) determined using one-way ANOVA and Tukey's multiple comparisons test.

6.5 Discussion

Several rhizobial genes are thought to be under the control of NifA, This includes key nitrogenase machinery, but also genes not apparently involved in nitrogen fixation, such as an alcohol dehydrogenase (pRL100103). This can be seen in other published NifA-RpoN regulons; the *R. etli* regulon includes sugar transporters and lantibiotic secretion systems (Salazar, Diaz-Mejia et al. 2010) and the *M. loti* regulon includes genes for terpenoid synthesis (Sullivan, Brown et al. 2013). Despite being a regulator of nitrogen fixation, many NifA-regulated genes are not essential for nitrogen fixation, such as *fdxB1*, a stable mutant of which has been produced during this work. The roles of many of these genes in symbiosis is yet to be elucidated, such as pRL100201, the divergent gene from *fixA*, a hypothetical protein with no recognised functional domains. A mutant of pRL100201 has a Fix^+ phenotype, despite it being 26-fold upregulated in 28 day bacteroids (Karunakaran, Ramachandran et al. 2009).

Published NifA and RpoN DNA-binding sequences allowed better characterisation of the *fixA* promoter. These binding sites were combined with 5' transcriptional start site data (Green and Poole, unpublished) to predict which sites are required for *fixA* expression. A site upstream of the CDS and downstream of the sigma factor binding site suggests successful mapping of the 5' transcriptional start site of *fixA*. Bioinformatic characterisation of the *fix* promoter directed design of promoter regions for the *lux*-based reporter system (Frederix, Edwards et al. 2014). The published use of the pIJ11268 plasmid was as a bioreporter to determine bacterial attachment to roots. This work has adapted the system to act as a non-invasive, bacteroid-stable reporter fusion for promoter mining. Combination of bioinformatics and this *lux*-based technique determined that a single NifA-binding site at the -132 position and an RpoN-binding site at the -52 position are essential for expression of the *fixA* gene.

Previous literature suggests a role for the *fixA* promoter in autoregulation of *nifA* (Buikema, Szeto et al. 1985, Martinez, Palacios et al. 2004). Autoregulation of *nifA* was demonstrated in *R. leguminosarum* bv. *viciae* UPM791, from a site predicted to be within a gene cluster including the *fixABCX* operon. The *fixABCX* gene region of strain UPM791 is shown in Figure 6.13. Though expression from a *fixA* promoter was considered, this was not thought to be the case due to the lack of an intergenic region between *fixA* and upstream gene *orf5*.

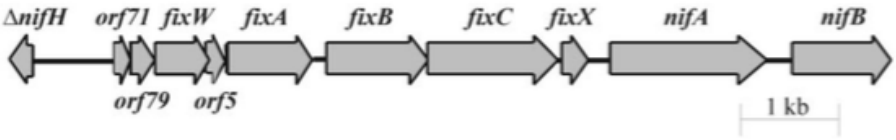


Figure 6.13. *fixABCXnifA* gene cluster in *R. leguminosarum* bv. *viciae* UPM791.

Alignment of several other *Rhizobiacae* gene clusters shows conservation of an intergenic region upstream of *fixA,* suggesting that strain UPM791 used in previous autoregulation studies may be an exception to a conserved autoregulation system from the *fixA* promoter. Additionally, the smallest IGR found upstream of *fixA* was 175 bp; the minimal region required for *fixA* expression deduced in this work was 176 bp. It could be hypothesised that *nifA* autoregulation from *pfixA* is a common strategy across the *Rhizobiacae*.

In addition to autoregulated expression, the *nifA* gene also has been shown to be under a low-level basal promoter, thought to be found upstream of the 3'-end of *fixX* in *R. leguminosarum* bv. *viciae UPM791* (Martinez, Palacios et al. 2004). In *S. meliloti* the NifJ protein is thought to activate *nifA* expression via a binding site found at the -55 position (relative to the ATG start site) (Hertig, Li et al. 1989), within the IGR. *Lux*-based promoter mining has indicated the presence of multiple promoters upstream of *nifA*, found in both the *fixC* and *fixX* ORF, but none within the *fixX-nifA* IGR. Expression from these constructs *in planta* is approximately 10% of the expression seen *in planta* from the *fixA* promoter constructs; *fixA* is one of the most upregulated genes in Rlv3841. At least one promoter can be found in *fixX*, expressed from within the last 100 bp of the ORF. Progressive deletions of the promoter region led to incremental decreases in *lux* expression in both plant and free-living assays. The same pattern was seen with truncated *nifA* promoters in *R. etli* (Benhassine, Fauvart et al. 2007). This incremental decrease in expression suggests the presence of multiple low-level promoters, a phenomenon common to constitutive promoters (Juven-Gershon and Kadonaga 2010). No promoter activity was seen from the IGR alone.

Reports in the literature are at odds as to whether this basal *nifA* expression is bacteroid-specific or constitutive under free-living conditions (Kim, Helinski et al. 1986, Martinez, Palacios et al. 2004, Benhassine, Fauvart et al. 2007). A *lux* promoter fusion with a bacteroid-specific promoter such as *fixA* (this work) or *nifH (*Pini et al., in production) only luminesces in bacteroids and not in the free-living state. A promoter fusion with a *nifA* promoter region is able to luminesce in both the free-living state and in bacteroids. This indicates that the basal promoters of *nifA* are constitutively expressed under aerobic, free-living growth.

Transcriptional regulation by NifA is affected by the oxygen status of the cell (Dixon and Kahn 2004). This is thought to be due to the effect of oxygen on both gene expression and activity of the NifA protein. Growth of *nifA* reporter strains under microaerobic conditions showed no difference in gene expression compared to aerobic conditions. Growth of a *fixA* reporter strain under the same conditions also showed no effect on *fixA* expression in the free-living state. Other factors must be involved in regulation of the nitrogen fixation machinery in Rlv3841. This factor may be nitrogen limitation, another known player in NifA control (Jack, De Zamaroczy et al. 1999, Martinez-Argudo, Little et al. 2004). Other factors could be pH, since the nodule environment is at low pH (Bhat and Carlson 1992, Perez-Galdona and Kahn 1994, Pierre, Engler et al. 2013), or provision of unknown plant factors. The identity of these plant factors may be revealed as more understanding is gained of the hundreds of NCR peptides provided by the host plant (Silverstein, Moskal et al. 2007). The luciferase fusions constructed in this work, as well as a GFP fusion with higher resolution will allow future investigation of factors required for bacteroid gene expression.

Complementation of mutants of the *fixABCX* operon supports the existence of multiple *nifA* promoters; a basal promoter found immediately upstream of *nifA* (‘*pnifA*’), and a bacteroid-specific promoter found upstream of *fixA* (‘*pfixA*’). Reduced nitrogenase activity in polar *fix* mutants can be explained by the lack of activity from p*fixA* leading to lower levels of NifA in the cell, and so lower expression from NifA-dependent promoters such as that driving *nifHDK* expression. A polar *fixX* mutant lacks the basal promoters found within the *fixX* gene, and also is blocked in expression from the *fixA* promoter and basal promoters found within *fixC,* resulting in a mutant unable to produce NifA, and therefore a Fix^- phenotype. qRT-PCR or microarray analysis of NifA-dependent promoters could be used to confirm lower or abolished activity in these polar mutant strains.

An in-frame *fixX* mutant can be complemented to wild-type fixation levels by a plasmid expressing *fixX-nifA* under a *fixA* promoter, showing that expression from the *nifA* promoters within *fixC* are sufficient to result in autoregulation from the *fixA* promoter. A polar *fixX* mutant cannot be complemented with *fixX-nifA* alone, however. This is most likely due to the possibility that *nifB* is found in the same transcriptional unit as *nifA*. The *nifB* gene product is part of the molecular machinery required for assembly of a functional nitrogenase MoFe protein (Figure 6.14). Bioinformatic analysis of the *nifAB* gene region yields little information regarding the transcriptional unit, and genetic regulation of the region is not well enough understood to predict what elements are involved in *nifB* control. The *nifB* gene is found immediately downstream of *nifA* in all strains aligned in this work, which may suggest a conserved transcriptional unit.

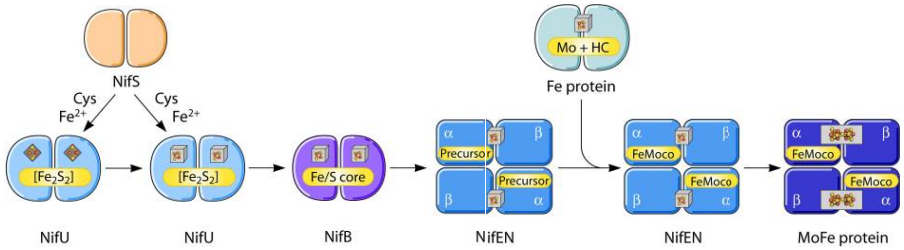


Figure 6.14 Biosynthesis of the MoFe protein of nitrogenase in *A. vinelandii*. NifB provides a scaffold for assembly of the Fe/S core, which is then passed to the NifEN complex. Adapted from (Hu and Ribbe 2011).

Construction of a complementing plasmid containing *nifB* was not successful using the Golden Gate assembly method. Transformations with the result of a Golden Gate reaction involving the *nifB* gene only yielded mis-assembled plasmids. (determined by blue-white selection). This was the case under both constitutive and inducible promoters, suggesting that leaky expression from this system is sufficient cause toxic effects in *E. coli*. Functional complementation of the polar *fixX* mutant is still required. However, complementation of an in-frame *fixX* is sufficient to confirm that *fixX* is required for nitrogen fixation.

The autoregulation of NifA can be added to the oxygen regulation scheme proposed by Hood (2013) for *R. leguminosarum* bv. *viciae* 3841. Although there is still work to be done to confirm this scheme, further information has been yielded about the mechanism of regulation in this strain. McMurty et al. (undergraduate thesis) used the *lux* system to demonstrate induction of expression of *fixN1, fixN2* and *fnrN* under 1% oxygen conditions. The same study used FixL mutants to determine that *fixN2* expression is dependent on FixLc but not FixLp. The role of FnrN has not yet been characterised in strain 3841. In *R. leguminosarum* bv. *viciae* UPM791 FnrN is hypothesised to control expression of FixNOQP and FixGHIS, whilst NifA controls other *nif* and *fix* genes (Ruiz-Argueso, Palacios et al. 2001). A *fnrN*::Ωtet mutant has been constructed in Rlv3841, but the *lux* system has not been used in this mutant due to the tetracycline resistance gene being present in both strain and plasmid. Plasmid pOPS0325 (p*fixA*-GFP) has kanamycin resistance and so could be used to investigate whether FnrN affects *fix* gene expression in Rlv3841. Integration of current data for Rlv3841 combined with data available for other *Rhizobium spp.* allows construction of an updated scheme for regulation of nitrogen fixation, as shown in Figure 6.15. This scheme is far from complete, and requires more investigation.

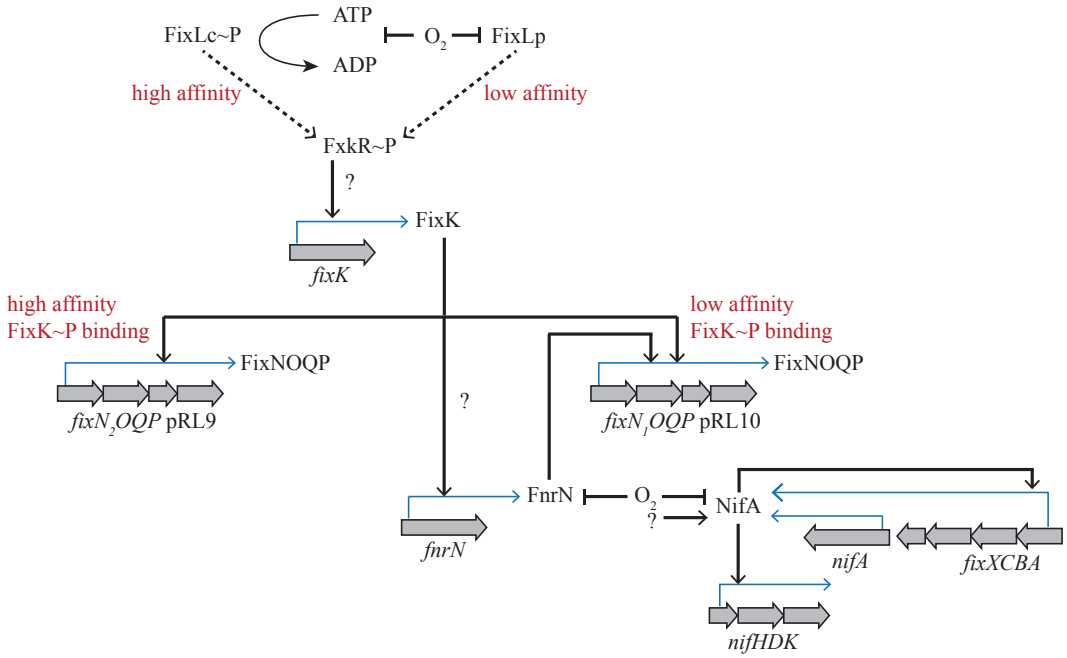


Figure 6.15 Proposed scheme for oxygen regulation of nitrogen fixation in *R. leguminosarum* bv. *viciae* 3841. Dashed lines indicate phosphorylation steps. FixLc has high affinity for oxygen, and controls expression of *fixNOQP* (FixN1) in a high-affinity system. The second *fixNOQP* operon (FixN1) is under the control of both FixLc and FixLp, in a low affinity system. This model is not yet complete, ? indicates unproven elements – there is no evidence yet of FxkR in *R. leguminosarum* bv. *viciae* 3841, nor is there direct evidence for FixK activating *fixNOQP* or *fnrN* transcription. The downstream processes from FixK need to be further elucidated, as does the regulation of the basal *nifA* promoter.

CHAPTER 7

Protein-protein interactions of FixAB

7.1 Introduction

The FixABCX proteins are hypothesised to work as electron transfer proteins carrying the electrons required for nitrogen fixation to the nitrogenase enzyme. Mechanisms have been proposed for the role of FixABCX in free-living diazotrophs (Scott and Ludwig 2004, Edgren and Nordlund 2006), but these proteins have not been fully characterised in the symbiotic state. ETF proteins are defined by (Buckel and Thauer 2013) as "heterodimeric FAD containing proteins mediating electron transport between a dehydrogenase and an ETF-quinone oxidoreductase". FixAB is an ETF and FixCX is an ETF-quinone oxidoreductase (ETF-QO), so most likely FixAB is passing electrons to FixCX. In *A. caulinodans* there is evidence suggesting that FixABCX interacts with pyruvate dehydrogenase under microaerobic nitrogen-fixing conditions, as shown by genetic interaction between *fixB* and *lpdH* (encoding dihydrolipoamide dehydrogenase, the E3 subunit of pyruvate dehydrogenase) though this is unsupported by biochemical evidence. This interaction with pyruvate dehydrogenase may also occur in the symbiotic system. In mammals the ETF-QO (FixCX) enzyme passes electrons to the quinone pool (Watmough and Frerman 2010). It is not known whether FixABCX sends electrons exclusively to nitrogenase, or whether it also passes electrons to the quinone pool and onto terminal electron acceptors. In Figure 7.1 we propose a mechanism for energisation of nitrogen fixation in the *Rhizobium*-legume symbiosis, in which pyruvate dehydrogenase interacts with FixABCX in order to pass electrons onto nitrogenase and the quinone pool.

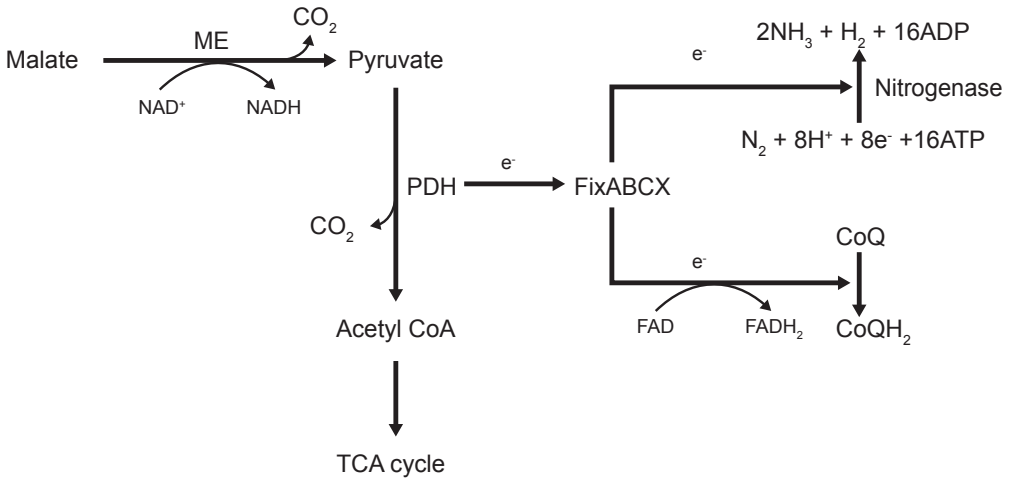


Figure 7.1 Scheme showing proposed role for FixABCX in *R. leguminosarum* nitrogen fixation. ME: malic enzyme; PDH: pyruvate dehydrogenase.

In order to better characterise FixAB within this symbiotic system we have looked at the protein-protein interactions occurring within *R. leguminosarum*, in both the free-living and symbiotic states. Polyclonal antibodies against *A. caulindans* ORS571 FixAB were provided by N. Watmough, UEA (Ruston 2003). FixA is 30,083 Da and FixB is 40,055 Da.

7.2 Expression of FixAB for use in assays

In order to investigate FixAB biochemically, large amounts of proteins are required. Addition of protein tags will allow for purification and interaction experiments using commercial systems. Several strategies were tried; expression from *E. coli*, expression from bacteroids under native conditions and inducible expression from free-living *R. leguminosarum* cultures.

7.2.1 Expressing FixAB from a T7-based expression system

Previous attempts at purifying FixAB from *E. coli* resulted in inclusion bodies (Watmough, N., pers. comm.). The pOPIN system is a T7-based vector recommended as a high-throughput expression system in a set of versatile vectors for expression in *E. coli* (Berrow, Alderton et al. 2007). The *fixAB* genes were cloned into pOPINF and pOPINS3C. pOPINF is one of the basic pOPIN expression vectors, whilst pOPINS3C includes a SUMO tag, which aids protein stability. A variant was constructed in pOPINF with both an N-terminal triple-FLAG tag and a C-terminal Twin-Strep tag. The vectors were expressed using a T7-expression protocol. In other successful pOPIN expressions in the Poole lab (IolG from pOPINF, Geddes, B., pers. comm.), the expressed proteins were clearly visible in a Coomassie-stained SDS-PAGE gel. Expression of FixAB from the pOPIN system did not demonstrate this clear overexpression (Figure 7.2). A Western blot was used to test expression of the tagged version of FixAB using both anti-FLAG and anti-Strep antibodies. The resultant Western blots showed a successful expression of tagged FixA, but no expression of a Strep-tagged FixB. This may mean that FixB has not been expressed properly in this system. The pOPIN system was not pursued further due to the lack of FixB expression.

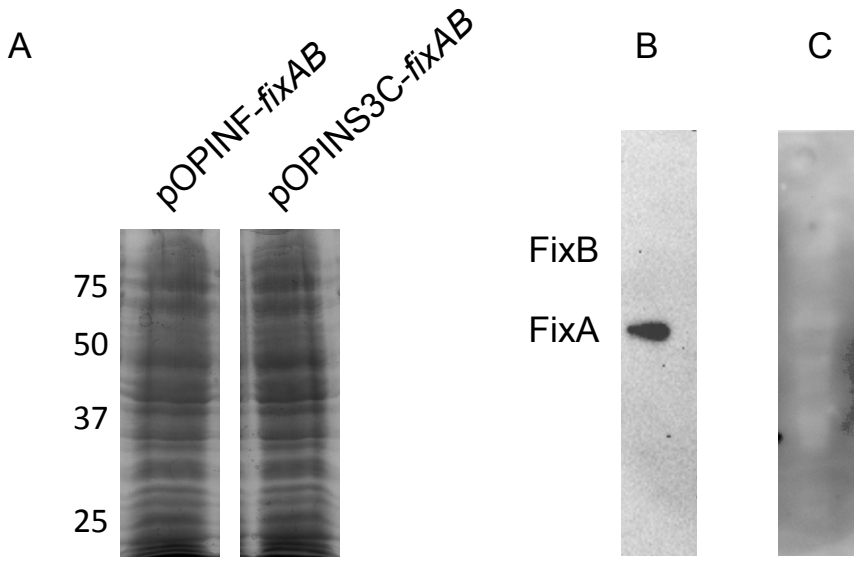


Figure 7.2 A. Coomassie Blue-stained SDS-PAGE gel of induced pOPINF-*fixAB* and pOPINS3C-*fixAB E. coli* cultures; B. Western blot against induced pOPINF-3xFLAG-*fixAB*-2xStrep using anti-FLAG M2 antibody; C. Western blot against induced pOPINF-3xFLAG-*fixAB*-2xStrep using anti-Strep antibody.

7.2.2 Expression of tagged FixAB under native conditions

In order to investigate FixAB in its native environment (expression within bacteroids) the *fixAB* genes were cloned into pJP2 under their native promoter; in this case the entire upstream intergenic region. Three different protein tags were initially cloned onto the C-terminus of FixB: His_6-, FLAG- and Strep-tag. Sequences of these tags are outlined in Chapter 2. These clones were constructed in pJP2, a stable vector for expression *in planta*. Additionally a control was constructed with no tag.

An initial test was carried out to ensure that presence of the tag did not influence nitrogen fixation ability. The constructed pJP2 clones were conjugated into 3841, LMB771 (*fixAB*::Ωspc) and LMB777 (Δ*fixAB*). As shown in Chapter 4, the omega insertion strain LMB771 cannot be complemented by FixAB alone. Additionally, fixation in the in-frame deletion LMB777 cannot be restored to wild-type fixation levels by FixAB, instead showing a Fix^{red} phenotype. Some variation is seen between the tagged strains, in particular the single Strep-tag appears to reduce nitrogen fixation ability. Presence of the FLAG- and His_6-tag had no significant effect on nitrogen fixation phenotype in either the wild-type or mutant background. Nodule count and weight were also measured in these complemented strains, and no clear effect was seen in the tagged strains (Figure 7.3). These strains were taken ahead for use in protein biochemistry experiments.

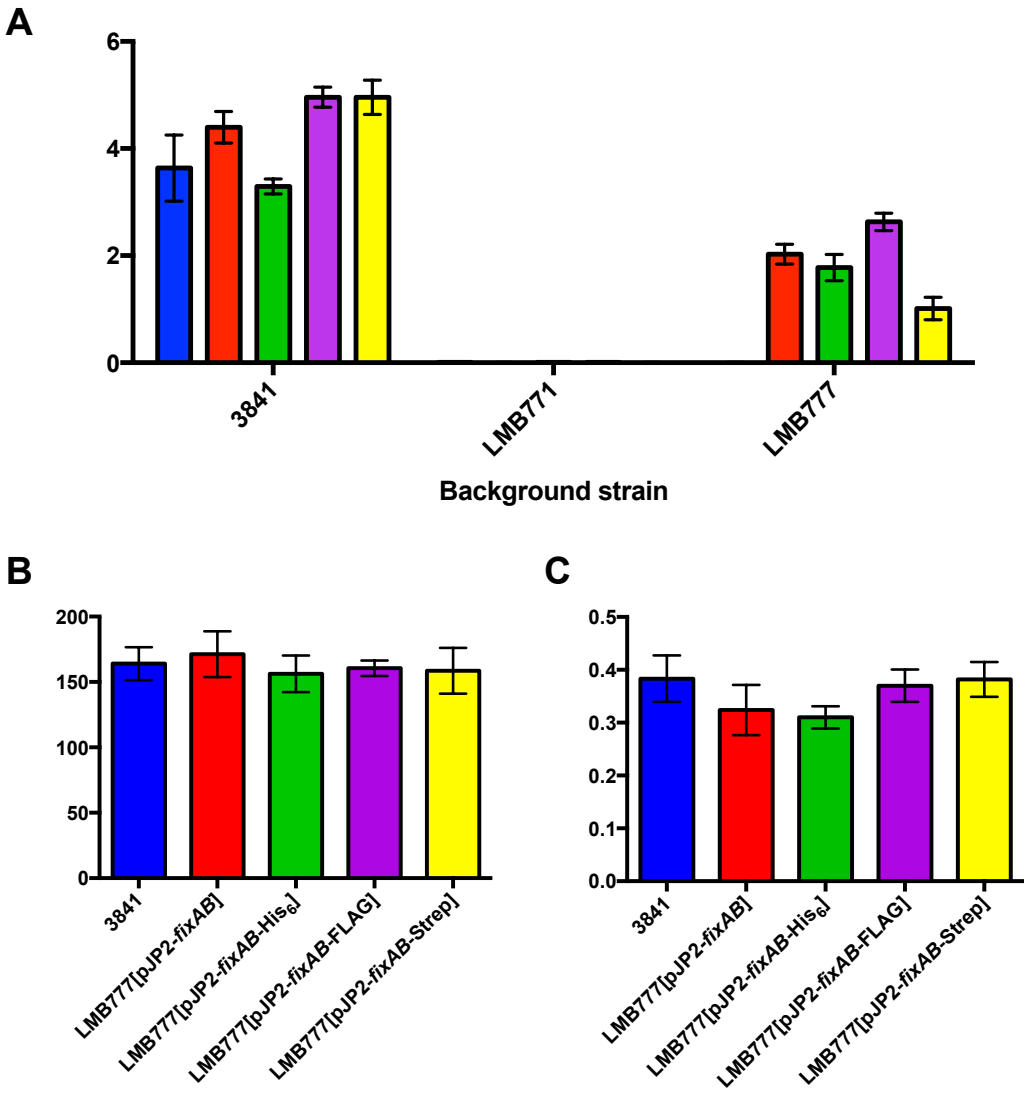


Figure 7.3 A. Acetylene reduction assay on complemented *fixAB* mutants with C-terminal tags inoculated on *P. sativum* at 28 dpi; No vector (blue); pJP2:*fixAB* (red); pJP2:*fixAB*-His_6 (green); pJP2:*fixAB*-FLAG (purple); pJP2:*fixAB*-Strep (yellow). Nitrogen fixation expressed as µmol ethylene per plant per hour; n=5. B. Nodule counts on complemented *fixAB* mutants; values expressed as nodules per plant; n=5. C. Nodule weight of complemented *fixAB* mutants; values expressed as nodules per plant $g^{-1;}$ n=5.

An anti-FixA Western blot was used to confirm expression from the pJP2 vector, as shown in Figure 7.4.

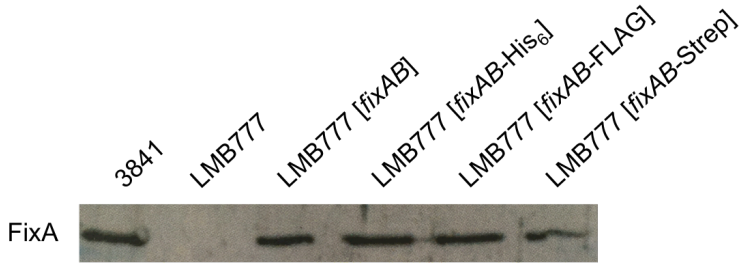


Figure 7.4 Western blot of tagged complementing strains expressed from pJP2 in bacteroids harvested 28 dpi; anti-FixA antibody.

The FLAG-tag antibody was used to check expression of a functional FLAG-tag on FixB from the pJP2 clone as shown in Figure 7.5. FLAG-tagged FixB protein could be detected by Western blot. A FixB antibody was used to confirm complementation of the *fixAB* mutant with both the tagged and untagged versions of *fixAB*.

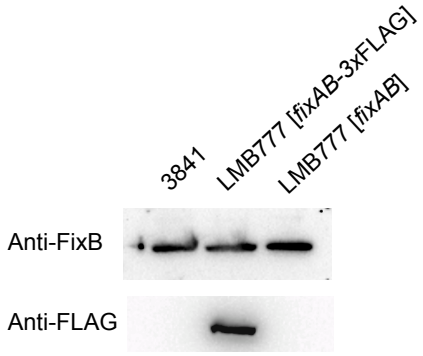


Figure 7.5 Western blot of FLAG-tagged complementing strain expressed from pJP2 in bacteroids harvested 28 dpi; anti-FixB and anti-FLAG M2 antibodies.

In addition to FLAG-tag, a triple(3x)FLAG-tag has been designed (Sigma Aldrich), which is reported to be more efficient for protein purification. A 3xFLAG-tagged version of FixAB in the pJP2 vector was constructed subsequently. Presence of a 3xFLAG-tag had no effect on nitrogen fixation ability, so tagged FixB is still functional. A Western blot was used to confirm the 3xFLAG-tag on FixB (Figure 7.6). 3xFLAG-tag was used for future experiments

due to higher efficiency. An acetylene reduction assay was used to confirm that the 3xFLAG-tag did not affect nitrogen fixation ability.

Strain	**Description**	**Acetylene reduction** μmol ethylene plant^{-1} hr^{-1}
3841	Wild-type	5.2 ± 0.3
LMB777	Δ*fixAB*	0.0 ± 0.0
OPS0328	LMB777 [*fixAB*-3xFLAG]	4.2 ± 0.7
Uninoculated control	n/a	0.0 ± 0.0

Table 7.1 Symbiotic phenotype of triple-FLAG complemented FixAB mutant strains on pea (*P. sativum*). Acetylene reduction assay was carried out 28 dpi; ± SEM, n ≥ 6.

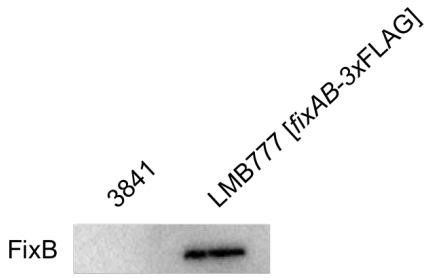


Figure 7.6 Western blot of triple-FLAG-tagged complementing strain expressed from pJP2 in bacteroids harvested 28 dpi; anti-FLAG M2 antibody.

The anti-Strep antibody was used to check for the presence of a correct Strep-tag from the pJP2 clone. No Strep-tagged protein could be detected in these Western blots (Figure 7.7). This also acted to show that there were no contaminating proteins within a bacteroid sample. A purified Strep-tagged cytoglobin (Watmough, N., pers. comm.) was used as a positive control for the Strep-tag antibody.

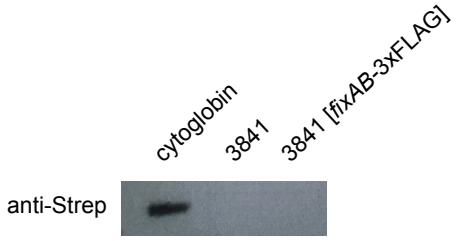


Figure 7.7 Western blots of tagged complementing strains expressed from pJP2 in bacteroids harvested 28 dpi; anti-Strep antibody; Strep-tagged cytoglobin was used as a positive control.

7.2.3 Expression of FixAB under inducible conditions

In order to obtain large amounts of FixAB for protein experiments, a taurine-inducible expression vector was used to express the FixAB proteins in free-living cultures. The *fixAB* genes were cloned into pLMB509 (Tett, Rudder et al. 2012). The pLMB509 vector has a C-terminal His_6 tag. Constructs were designed with and without stop codons in order to keep or remove the His_6-tag. Constructs were initially designed with a C-terminal single FLAG and Strep tag, with and without His_6. All plasmids constructed under the taurine promoter are described in Table 7.2.

Plasmid	Strain in 3841 background	N-terminal tag	C-terminal tag
pLMB779	LMB779	None	None
pLMB780	LMB780	None	Strep His_6
pLMB781	LMB781	None	Strep
pLMB782	LMB782	None	FLAG His_6
pLMB783	LMB783	None	FLAG
pOPS129	OPS0287	Twin-Strep	3xFLAG
pOPS0130	OPS0288	Twin-Strep	3xFLAG His_6
pOPS0131	OPS0289	3xFLAG	Twin-Strep
pOPS0132	OPS0290	3xFLAG	Twin-Strep His_6

Table 7.2 Inducible plasmids designed to express FixAB in free-living culture.

An initial test was carried out using pLMB779 at a range (0-30 mM) of taurine concentrations, in order to determine the optimum amount of taurine required for induction of FixAB (Figure 7.8). As expected no FixA expression was seen in the uninduced free-living culture. Induction was seen with all of the taurine concentrations used. In order to minimise effect of taurine on other cellular processes, 10 mM was chosen for use in future induction assays.

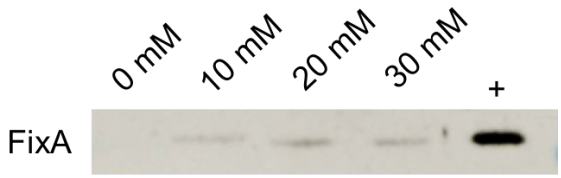


Figure 7.8 Western blots of strain LMB779 grown under increasing taurine concentrations; anti-FixA antibody; 3841 bacteroid sample used as positive control for FixA.

Strains were grown under 10 mM taurine to test for expression of tagged FixAB. Initial strains with C-terminal His_6, single FLAG and single Strep tags were tested (Figure 7.9). The His-tag could be detected by Western blot in both tagged strains. The FLAG-tagged FixB could be detected in LMB780, but a successful Western blot showing FLAG tag in LMB781 was not obtained. A Strep-tagged FixB was never detected in these initial expression plasmids.

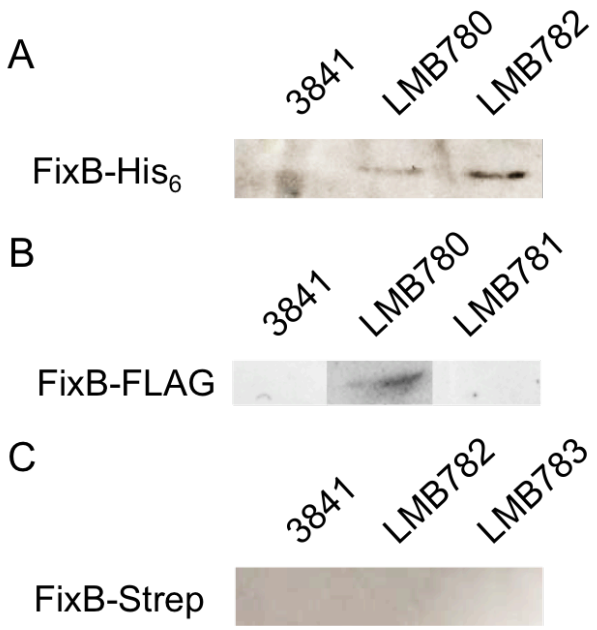


Figure 7.9 Western blots of tagged FixAB strains grown for 24 hours in UMS minimal media supplemented with 10 mM taurine; A: anti-His antibody; B: Anti-FLAG antibody; C: Anti-Strep antibody; 3841 was used as a negative control; All tags are on the C-terminus of FixB.

Subsequent to the lack of detection with single-Strep tag, primers were designed to add both Twin-Strep and 3x-FLAG tags on FixAB, including on the N-terminus. These multiple tags are expected to have higher efficiency in pull-down assays. A new set of plasmids was designed, with both N-terminal tags on FixA and C-terminal tags on FixB. These new plasmids were conjugated into 3841 and tested using 10 mM taurine in minimal media. Expression of correctly tagged proteins was seen in all strains (Figure 7.10). Expression from OPS0288 was slightly lower than in other strains, and no FixB could be detected. Strains OPS0287 and OPS0289 shower best expression and so were carried forward for future work.

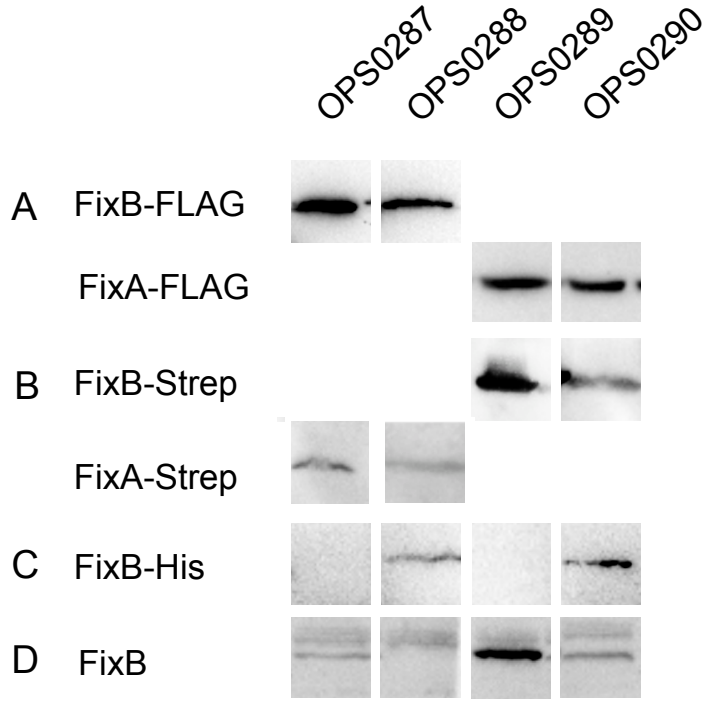


Figure 7.10 Western blots of tagged FixAB strains grown for 24 hours in UMS minimal media supplemented with 10 mM taurine; A: anti-FLAG antibody; B: Anti-Strep antibody; C: anti-His antibody; D: anti-FixB antibody (top bands are non-specific binding, lower band is FixB protein); All tags on FixA are on the N-terminus. All tags on FixB are on the C-terminus.

7.3 Purification of tagged FixAB using gravity-flow columns

Gravity-flow columns were used to purify tagged FixA and FixB proteins. FLAG-tag purification was carried out using FLAG M2 affinity gel, consisting of beads coated with M2 anti-FLAG antibody (Sigma Aldrich). When sample is washed over the gel FLAG-tagged proteins stick to the beads. The bound protein can be eluted by a solution containing pure FLAG M2 peptide. The Strep-tag® purification uses high selective binding of Strep-Tactin®, an engineered streptavidin, to Strep-tagged proteins. Samples containing tagged proteins were washed over the Strep-Tactin® Sepharose column and tagged proteins are bound to the streptavidin. An elution buffer containing desthiobiotin is then used to elute the bound Strep-tagged proteins. In order to check the success of the pull-downs, samples were collected before the column protocol, after flow through the column, from the wash steps and from the eluates.

Initial purifications were carried out with crudely lysed cells from taurine-induced cultures. These cultures are easier to obtain than bacteroid samples and faster to prepare than gently lysed cells, and so were used for column optimisation. Crude lysis is expected to interrupt protein-protein interactions, so this protocol cannot be used to investigate interactions, only purify the protein of interest. These samples were used to troubleshoot the protocol. Western blots were used to confirm the success of the immunoprecipitations (Figure 7.11). Purifications of tagged proteins were successful with both the Strep-tag and FLAG-tag columns. Tagged-protein can be seen in the sample, flow-through and the early wash stages, although in later washes tagged protein stops being detected. Appearance of tagged protein in the flow-through suggests saturation of the column. Enough protein is eluted that a faint band corresponding to FLAG-tagged FixB can be detected in a Coomassie gel of a FLAG-tag pull down on bacteroid sample OPS0328 (Figure 7.12). Tagged protein is then detected in the eluates. In these purifications the FLAG-tagged samples gave a much higher yield of protein than Strep-tagged samples. The FLAG-tag system was chosen for further work.

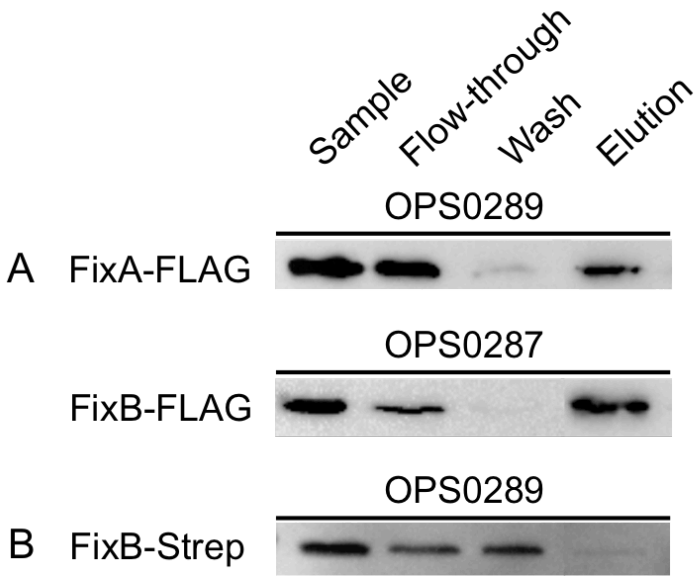


Figure 7.11 Western blots of initial column purifications using taurine-induced strains grown in minimal media supplemented with 10 mM taurine. Samples were prepared using a crude lysis. A: FLAG-tag pull downs carried out on OPS0289 and OPS0287 respectively using anti-FLAG M2 antibody. B: Strep-tag pull down carried out on OPS0289 using anti-Strep antibody.

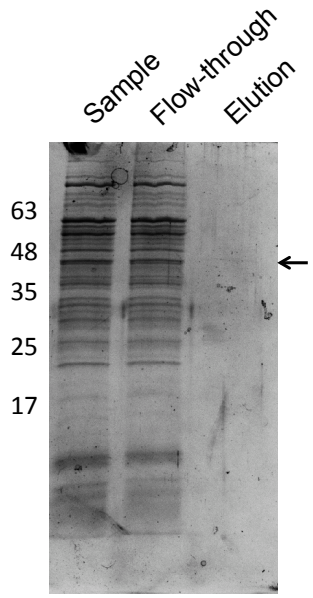


Figure 7.12 Western blots of initial column purifications using taurine-induced strains grown in minimal media supplemented with 10 mM taurine. Samples were prepared using a crude lysis. A: FLAG-tag pull downs carried out on OPS0289 and OPS0287 respectively using anti-FLAG M2 antibody. B: Strep-tag pull down carried out on OPS0289 using anti-Strep antibody.

7.4 Pull-down of FLAG-tagged FixAB using gently lysed cells

Once immunoprecipitation protocols had been optimised, gently lysed samples were used for pull-downs. These cells have been treated with a gentle lysis buffer with low ionic strength, containing non-ionic detergents. This buffer should maintain stable protein-protein interactions. Figure 7.12 shows Coomassie gels of FLAG-tag pull downs using gently lysed cells. Several protein bands can be identified, suggesting maintenance of interactions between the tagged protein and its interacting partners. The buffer and protocol for this lysis is outlined in Chapter 2.7.9.

Pull-downs on gently lysed samples were optimised using induced, free-living strains. These were tested using anti-FLAG M2 antibody to check presence of FLAG-tagged proteins in the eluate fraction (Figure 7.13). Coomassie Blue was used to stain SDS-PAGE gels to check quality of samples. In the tagged samples the associating proteins can be detected in the eluate fraction in a Coomassie gel, suggesting successful immunoprecipiation. In samples with a high protein concentration bands at can be seen at ~30 kDa (FixA) and ~40 kDa

(FixB), as well as a strong band at ~60 kDa. As discussed later in this chapter, this 60 kDa protein is thought to be GroEL, a chaperonin.

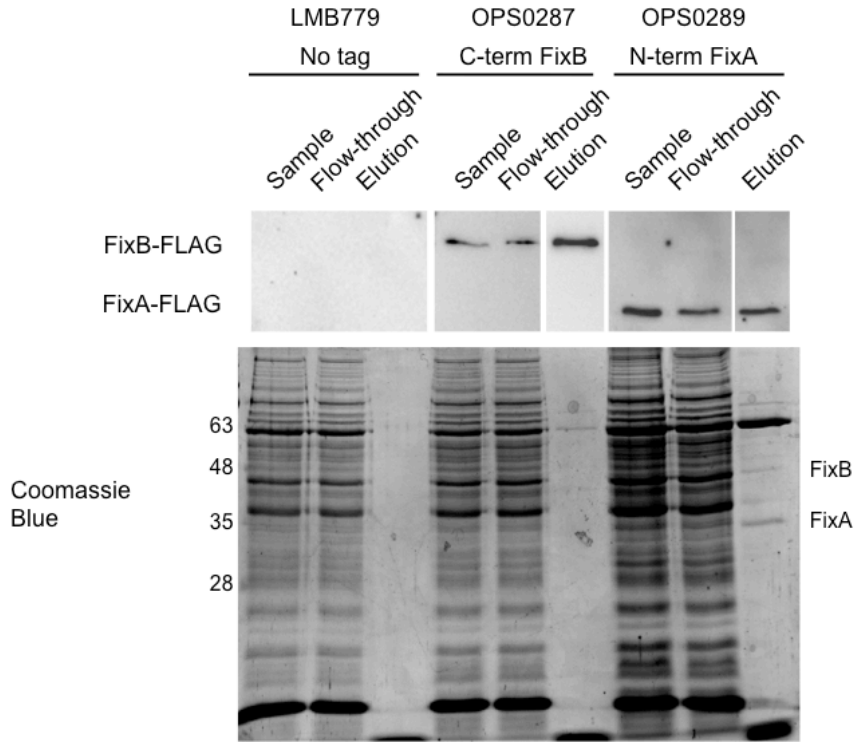


Figure 7.13 Western blots and Coomassie stained SDS-PAGE gel of pull-down experiments using taurine-induced strains grown in minimal media supplemented with 10 mM taurine. Samples were prepared using a gentle lysis to retain protein-protein interactions. Western blots were carried out with an anti-FLAG antibody.

The optimised pull-down procedure was then used on bacteroid samples. Bacteroids were harvested from *P. sativum* nodules at 28 dpi. Again, Western blots and Coomassie gels were used to check the quality of the pull-downs. An anti-FixA antibody was used to check the negative control strain to ensure that untagged proteins were not being retained on the column. Interacting proteins can be seen in the Coomassie gel, including at 60 kDa, as seen in the free-living Coomassie gels (Figure 7.14).

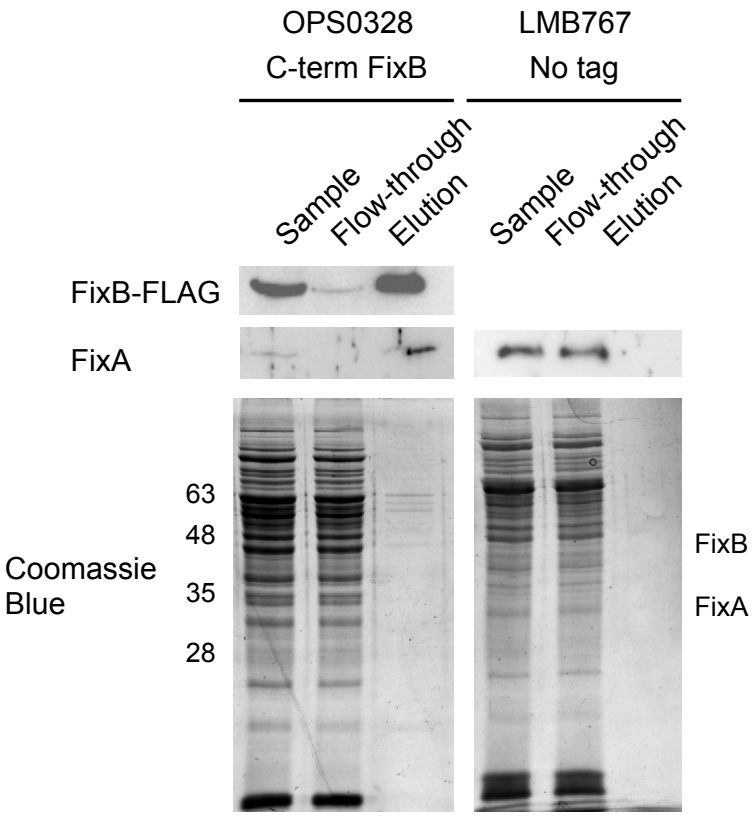


Figure 7.14 Western blots and Coomassie stained SDS-PAGE gel of pull-down experiments using bacteroids extracted from nodules at 28 dpi. Samples were prepared using a gentle lysis to retain protein-protein interactions. Western blots were carried out using an anti-FLAG and anti-FixA antibody.

Pull-downs were attempted with gently lysed samples containing Strep-tagged proteins, but no successful pull-down was ever obtained. The FLAG system was optimised to take further.

7.5 Identifying interactions between FixAB and other proteins

7.5.1 Prediction of interacting partners from a proposed model

The model proposed at the start of this chapter suggests that FixAB interacts with pyruvate dehydrogenase, and proceeds to carry electrons to nitrogenase. In view of this, antibodies were raised against NifH (nitrogenase iron subunit), PdhA (pyruvate dehydrogenase subunit E1) and LpdH (pyruvate dehydrogenase subunit E3, dihydrolipolyl dehydrogenase). A an interaction has previously been demonstrated between *fixB* and *lpdH* in *A. caulinodans*, in which the an *lpdH* mutation could rescue the Fix$^-$ phenotype of a *fixB* missense mutation (Scott and Ludwig 2004). Antibodies were obtained from GenScript (NJ, USA).

These antibodies were obtained in order to detect the presence of their protein targets in elutions from a gentle lysis pull-down with FLAG-tagged FixB. If present in Western blots, this would provide evidence for an interaction between FixB and the protein of interest. Elutions from a bacteroid sample with 3xFLAG-tag FixB were tested using Western blots with antibodies against these three proteins. All three proteins could be detected in the lysed cell sample, but none could be seen in the Western blots, shown in Figure 7.15.

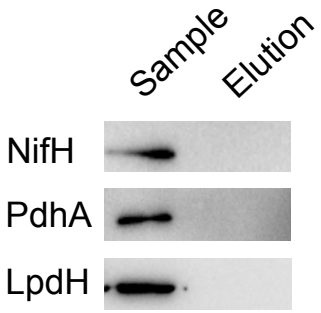


Figure 7.15 Western blots of pull-down experiments using bacteroids extracted from nodules at 28 dpi. Samples were prepared using a gentle lysis to retain protein-protein interactions. Western blots were carried out using anti-NifH, anti-PdhA and anti-LpdH antibodies.

7.5.2 Bioinformatic prediction of FixAB interacting partners

STRING 10.0 (Szklarczyk, Franceschini et al. 2015) was used to predict interacting partners for FixAB. STRING is a database of protein-protein interactions, derived from genomic predictions, experimental data and previous knowledge in other databases. Figure 7.15 shows STRING prediction networks for FixA and FixB. Neither network is supported by any direct biochemical evidence; both are entirely constructed from genomic context and computational prediction. FixA and FixB are both predicted to interact with FixC and FixX. Additionally, interaction was predicted with NifA and NifB. Since predictions are made based on genetic context, this may be due to the conserved genetic region of *fixABCXnifAB* (see Chapter 7). pRL120522 and pRL120521 encode EtfA2 and EtfB2, and RL4319 and RL4320 encode EtfA1 and EtfB1 respectively. These interactions are most likely predicted due to FixAB being an electron transfer flavoprotein. IvdH is a putative isovaleryl-CoA dehydrogenase and is predicted to interact due to concurrence across genomes. Both pRL100302 and pRL100376 are putative acyl-CoA dehydrogenases and are also predicted as interacting partners due to concurrence between genomes.

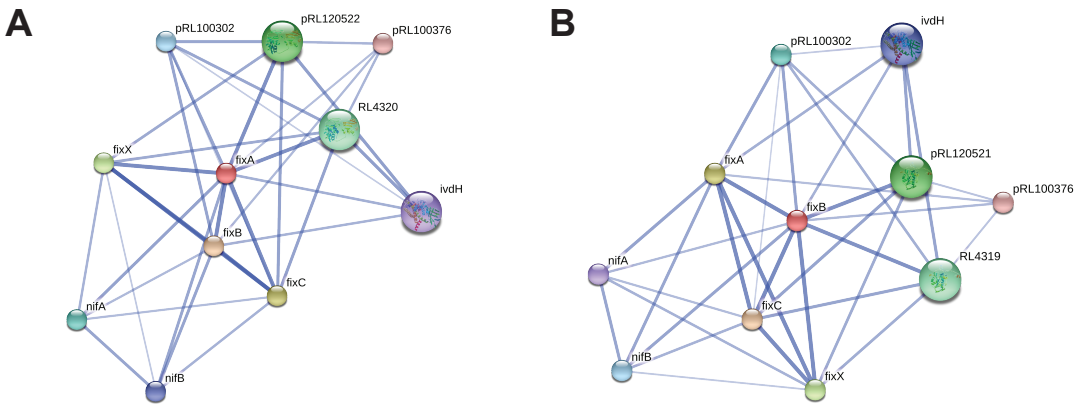


Figure 7.16 Predicted protein-protein interaction networks constructed for A: FixA and B: FixB. Predictions carried out by STRING 9.0.

7.5.3 Determining protein-protein interactions using protein mass spectrometry

In order to gain a comprehensive list of interacting proteins, the eluate fraction from the pull-down experiments was sent for tandem mass spectrometry (MS/MS) (Brymora, Valova et al. 2004). Negative controls were included containing untagged FixAB on the respective plasmids. Samples were washed and concentrated before a digestion with Trypsin to create peptides which were then run through a mass spectrometer. After a first stage of mass spectrometry (MS), ions are separated by mass-to-charge ratio (m/z). Fragmentation then occurs by collision-induced dissociation. A second MS run then determines a sequence which can be combined with peptide mass, allowing identification of the protein (Aebersold and Mann 2003, Han, Aslanian et al. 2008).

Initial tests of protein MS/MS yielded peptides from 1263 proteins for FixB in the bacteroid fraction, far too high to make meaningful conclusions about strongly interacting proteins. The pull-down protocol was refined to reduce the level of background in the eluate fraction. The original protocol used 15 washes of TBS between the sample and elution steps. A sample was tested with 40 washes of TBS. Additionally, Triton X100 was added in small percentages as detergent to break apart non-specific interactions. Finally, a high-salt wash was tested. TBS is 150 mM NaCl, so a higher concentration of 250 mM was tested. SYPRO Ruby was used to stain SDS-PAGE gels to allow higher resolution of proteins, and Western blots were used to confirm that FLAG-tagged protein was retained in the eluate despite treatments (Figure 7.17). Samples with fewer bands were chosen, as these hopefully retained real interaction whilst losing non-specific background. A combination of 40 washes plus 0.25% Triton was selected, and pull-downs repeated for MS.

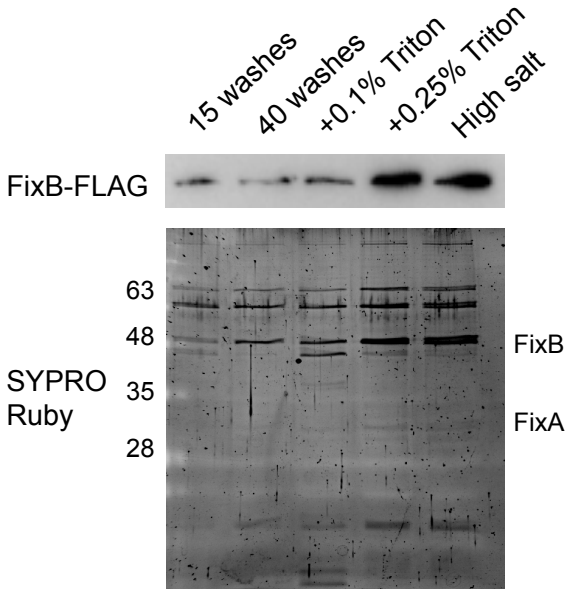


Figure 7.17 Western blots and SYPRO Ruby stained SDS-PAGE gel of pull-down optimisation. Eluate fraction of each test was run on the gel. Experiments were carried out using taurine-induced OPS0287 grown in minimal media supplemented with 10 mM taurine. Samples were prepared using a gentle lysis to retain protein-protein interactions. Western blots were carried out using an anti-FLAG antibody.

After optimisation of the pull-down protocol, fresh samples were sent for MS/MS. Samples sent for MS/MS are described in Table 7.3. Three replicates of each strain were analysed. Tagged bacteroid samples will give an indication of interactions occurring in the native state, whilst induced free-living samples will seek to supplement the bacteroid data. This run resulted in peptides from 590 proteins in the bacteroid FixB sample, a ~50% reduction from the test runs.

Strain	Description	Tagged protein	Type of sample
LMB779	3841 [pLMB509-*fixAB*]	Negative control	Free-living culture induced with 10 mM taurine
OPS0287	3841 [pLMB509-twin-Strep-*fixAB*-3xFLAG]	FixB	Free-living culture induced with 10 mM taurine
OPS0289	3841 [pLMB509-3xFLAG-*fixAB*-twin-Strep]	FixA	Free-living culture induced with 10 mM taurine
LMB767	3841 [pJP2-*fixAB*]	Negative control	Bacteroids harvested from nodules 28 dpi
OPS0328	3841 [pJP2-*fixAB*-3xFLAG]	FixB	Bacteroids harvested from nodules 28 dpi

Table 7.3 Strains used for pull-downs and subsequent protein MS/MS.

Identified peptides from MS/MS were mapped to the Rlv3841 proteome (obtained from UniProt) and matched to a locus tag from the Rlv3841 complete gene annotation. Figures 7.18 and 7.19 show hierarchical clusterings of the pull-down MS data.

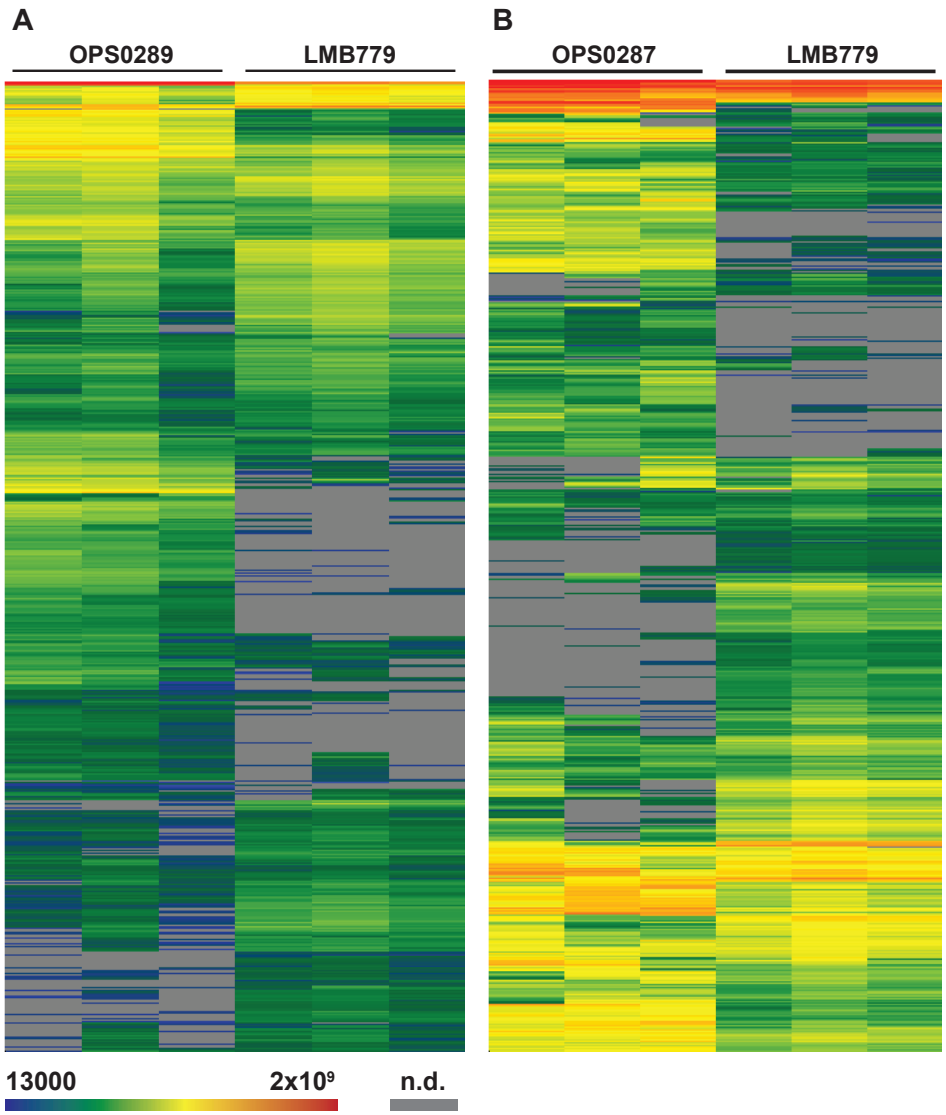


Figure 7.18 Hierarchical clustering of samples from MS/MS. A: Free-living tagged FixA (OPS0289) versus free-living negative control (LMB779); B: Free-living tagged FixB (OPS0287) versus free-living negative control (LMB779). Colours show protein intensities (abundance) per sample on a logarithmic scale. n.d.: not detected in sample.

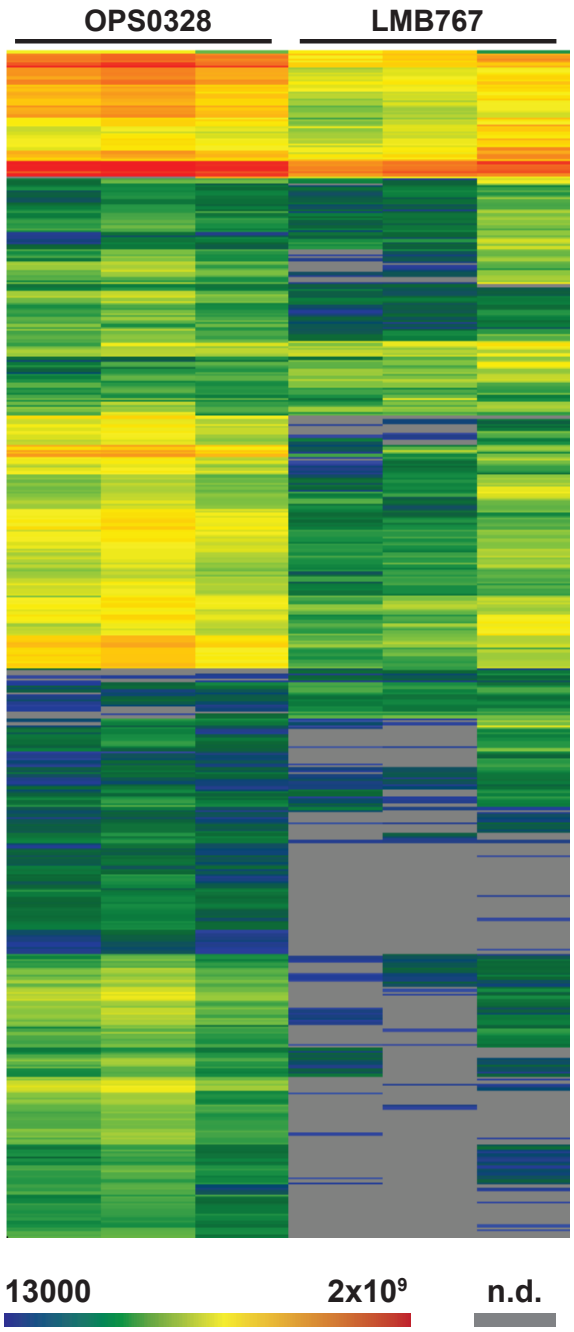


Figure 7.19 Hierarchical clustering of samples from MS/MS. Tagged FixB in bacteroids (OPS0328) versus bacteroid negative control (LMB767). Colours show protein intensities (abundance) per sample on a logarithmic scale. n.d.: not detected in sample.

Comparisons were made between tagged proteins and their respective negative controls. Any proteins that were detected more highly in the negative control than the tagged samples were discarded. A t-test was used to determine proteins with a significantly higher intensity in tagged samples than in the negative control, with a p-value of 0.05 (-log[p]=1.3). Some peptides matched to multiple proteins; usually where multiple homologues exist within the genome. In these cases, the majority protein was annotated.

In order to test whether MS protein intensity was as a result of abundance in the cell rather than interaction, comparison was made with microarray data. The putative interacting proteins encoded on the chromosome were plotted against their relative expression in bacteroids compared to free-living 3841 (Figure 7.20). Highly upregulated proteins involved in nitrogen fixation are also expected to interact with FixB, so the chromosomal proteins were chosen as a subset to remove this bias. There was no obvious relationship between the expression level measured by a microarray and the amount of protein pulled down by FixB-3xFLAG.

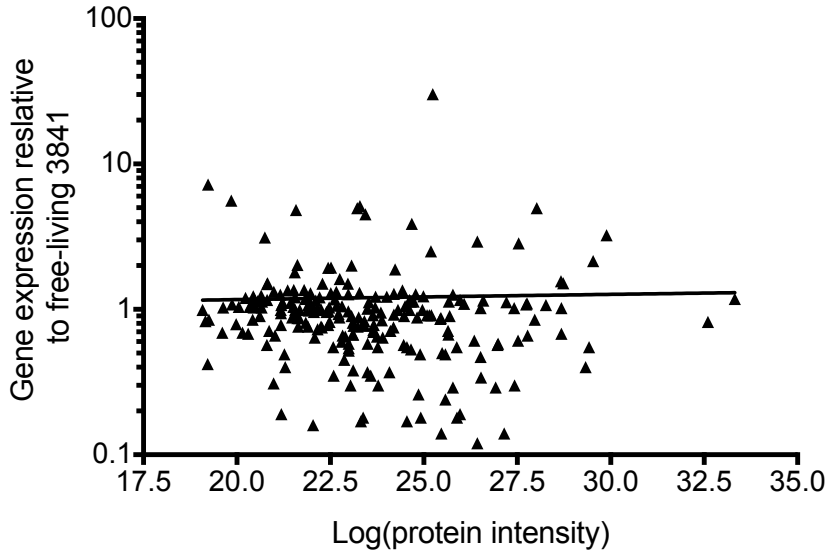


Figure 7.20 Scatter plot comparing abundance (log_2[intensity]) of chromosomally-encoded proteins pulled down by FixB-3xFLAG in bacteroid MS sample to relative gene expression under free-living conditions grown on succinate as determined by microarray (Karunakaran, Ramachandran et al. 2009). Linear regression fitted with R value of $1.39x10^{-4}$.

An initial analysis of the associating proteins showed the locations of the genes encoding these proteins. When sequenced in 2006 the 3841 genome was determined to have 7263 genes (Young, Crossman et al. 2006); this value has since risen to 7286 due to newly

discovered ORFs. The genome comprises a chromosome plus six large plasmids, including pRL10, the 'sym' plasmid. The majority of genes involved in symbiosis and nitrogen fixation are found on pRL10, including *fixNOQP*, *nifHDK* and *fixABCX*. Proteins encoded on pRL10 are mostly only expressed under these symbiotic conditions and so are expected to be pulled-down in the bacteroid sample, but not in the free-living samples. More proteins are also expected to come from pRL9 in the bacteroid sample compared to free-living, as pRL9 encodes many copies of genes from pRL10, such as *fixNOQP*.

Sample	**Chromosome**	**pRL7**	**pRL8**	**pRL9**	**pRL10**	**pRL11**	**pRL12**	**Total**
FixA free-living	196	1	3	3	9	5	5	222
FixB free-living	188	1	2	2	4	3	13	213
FixB bacteroid	230	1	2	15	28	3	4	282
Total in 3841	4739	188	142	313	471	644	789	7286

Table 7.4 Gene locations for proteins significantly pulled down by FLAG-tagged FixA and FixB, as determined by t-test; n=3.

Proteins were classified according to Riley codes. Broader groupings were then designated to allow functional classifications of all interacting proteins. The functional classifications for each Riley code are shown in the supplementary material. The functional classifications for interacting proteins in each strain is shown in Figure 7.21.

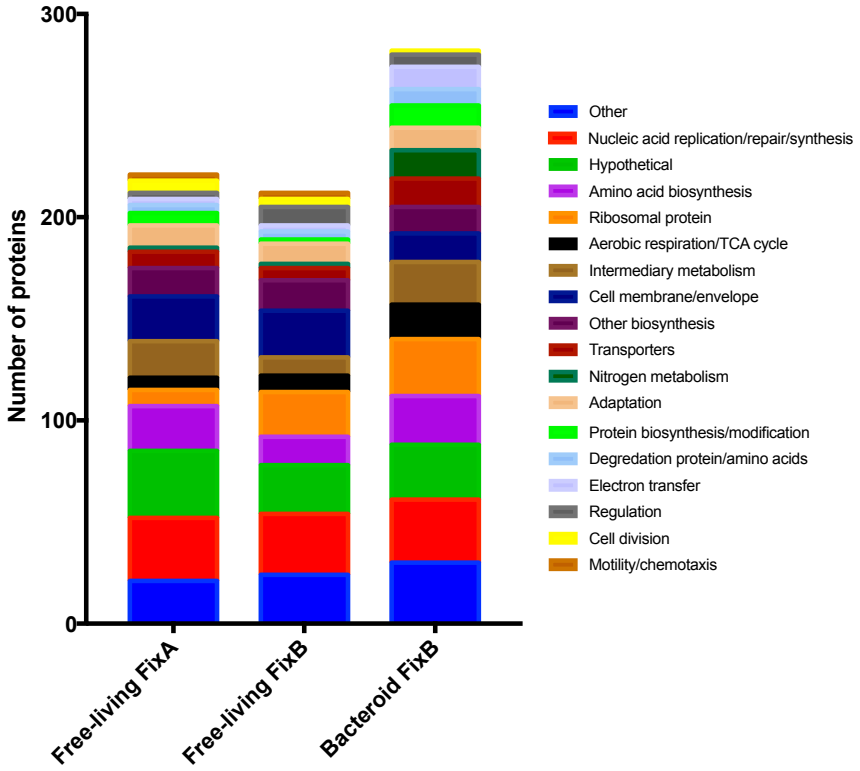


Figure 7.21 Functional classifications of all proteins associating with FixA/B in protein MS. An initial comparison was made between the three samples to identify any proteins coming down in all samples. A list of 53 common proteins was obtained. No nitrogen-fixation proteins are expected in this list, as none would be expressed in the free-living state other than FixAB. FixA and FixB are in this common list, confirming that they interact; as two subunits of a heterodimer this is expected. These 53 proteins were grouped according to functional classification, and the ratios of these can be seen in Figure 7.22. Many of these appear to be housekeeping genes such as heat-shock proteins, chaperonins and ribosomal proteins. Three additional proteins stand out from this list of 53. The first two are PdhA and PdhB, which are subunit A and B respectively of the heterodimeric E1 protein of the pyruvate dehydrogenase complex. The third protein is SucA, annotated as a putative E1 subunit of 2-oxoglutarate dehydrogenase.

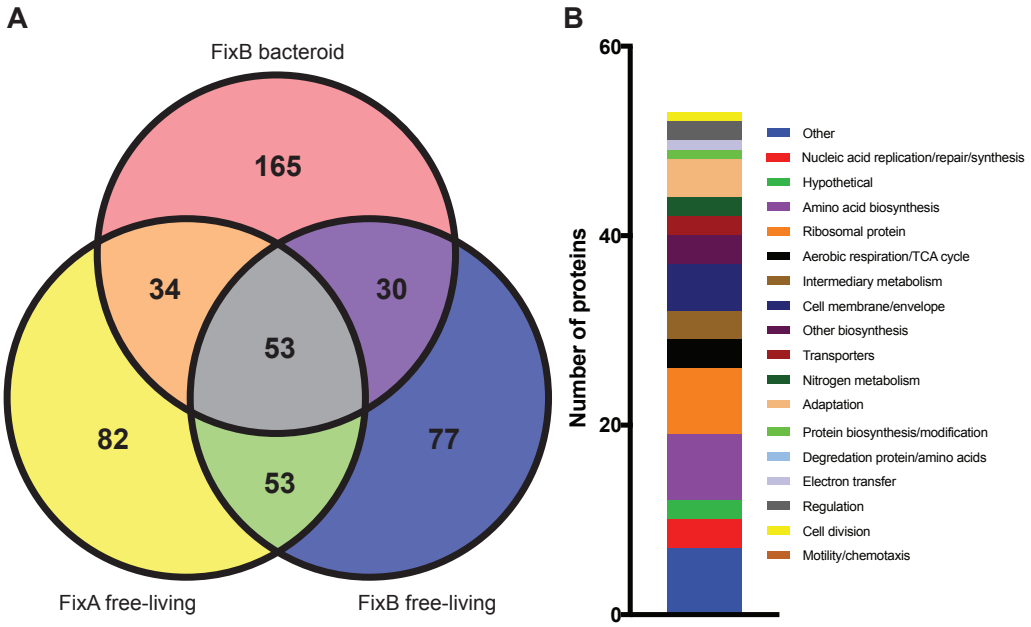


Figure 7.22 A. Venn diagram showing significant associated proteins common to sample groups. B. Functional classifications of the 53 common associated proteins.

Bioinformatic prediction using STRING yielded several potential interacting partners. This included EtfAB1 and EtfAB2 and several dehydrogenases. None of these bioinformatically predicted proteins were found to be significantly interacting with FixAB under any conditions. Interaction with FixCX was seen in the bacteroid sample; FixCX are bacteroid-specific so would not be detected in free-living samples. Interaction with NifB was seen in the bacteroid sample but NifA was not significantly pulled-down.

In order to investigate the individual pull-downs, samples were ranked by intensity in the positive sample. Only proteins that were significantly pulled down were considered. The functional groups of these rankings are shown in Figure 7.23. The top 20 ranked proteins across all samples are shown in Tables 7.5 and 7.6. There was an abundance of housekeeping proteins; chaperonins and ribosomal subunits. Proteins falling into these functional categories were removed from the list before ranking.

Five proteins appeared in both the free-living FixA and free-living FixB top twenty. These include FixAB. The other proteins were RpoB (a putative DNA-directed RNA polymerase), SecA (a preprotein translocase) and RkpK, a putative UDP-glucose 6-dehydrogenase.

PdhAB appeared in the free-living FixB sample as highly pulled down, and although they were not in the top twenty list for FixA, they also had high intensity in this sample. Some electron-transfer proteins were detected from the free-living pull-downs. FixB interacted

with PetA, a putative ubiquinol-cytochrome c reductase iron-sulfur subunit. FixA interacted with RL2369, a putative NADPH:ferredoxin reductase.

The tagged bacteroid sample provides a more representative dataset of the native environment for FixAB. The Fix and Nif proteins involved in nitrogen fixation are expressed in bacteroids. In the top twenty list for bacteroids, half of the proteins are Nif or Fix proteins. This includes FixABC, although FixX is not seen in the top ranked list. FixOP, are also detected. There are multiple FixNOQP complexes in *R. leguminosarum* bv. *viciae* 3841, one encoded on pRL10 (FixN1, FixO1, FixQ1, FixP1), one on pRL9 (FixN2, FixO2, FixQ2, FixP2) and one on the chromosome (FixN3, FixO3, FixQ3, FixP3). These make up the high-affinity cytochrome cbb_3 oxidase required under microaerobic conditions. FixNOQP1 and FixNOQP2 show 96% similarity, whilst FixNOQP3 shows 60% similarity with both the N1 and N2 complexes. All peptides detected for FixO are found in both FixO1 and FixO2, and so no conclusions can be made about which FixO is interacting. Out of 22 unique peptides identified for FixP, three were unique to FixP1 and three were unique to FixP2. The others peptides were conserved across both FixP1 and FixP2. The other subunits of the cbb_3 oxidase, FixN and FixQ were not found to be significantly associated with FixB. NifE and NifB, required for nitrogenase assembly were all present in the sample. Assembly factor NifN was not found in the top ranking list, however the protein was found to be significantly interacting with FixB.

Several proteins were identified in the bacteroid sample that are involved in the TCA cycle and upstream carbon metabolism. PdhAB were identified in the top twenty. The E2 and E3 subunits of pyruvate dehydrogenase, PdhC and LpdH were not found in this top list. SucA and SucB were also found, which encode the E1 and E2 subunits of the 2-oxoglutarate dehydrogenase complex. The 2-oxoglutarate dehydrogenase complex is in the same family as the pyruvate dehydrogenase complex, and though they do not share high sequence similarity, they do share functional domains. SucA and PdhA share a common domain, the dehydrogenase E1 domain, and SucA and PdhB share a transketolase pyrimidine binding domain. SucB and PdhC, encoding the E2 subunit also share conserved domains; the 2-oxoacid dehydrogenases acyltransferase catalytic domain, lipolyl domain and E3 binding domain. RkpK was detected in this list, which also appeared in both free-living samples. As a putative UDP-glucose 6-dehydrogenase, this protein is also involved in carbon metabolism.

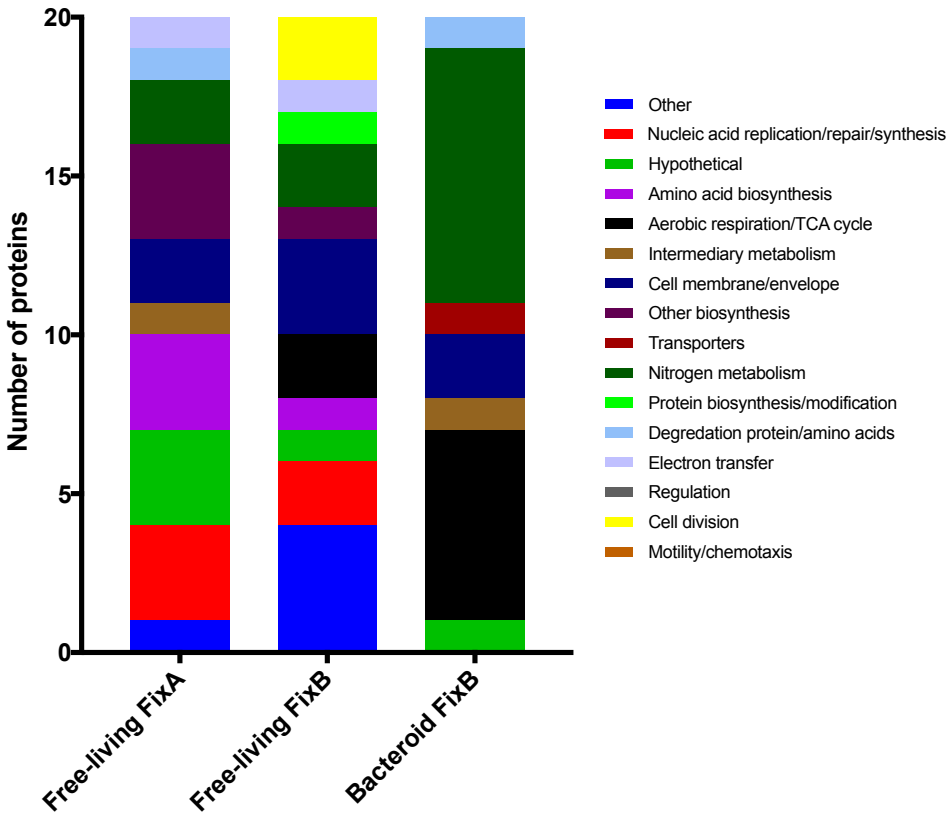


Figure 7.23 Functional classification of top 20 proteins in pull-down ranked by intensity

Gene	Protein	Description	Average log(Intensity)	
			FixA	FixB
clpX	ClpX	putative ATP-dependent Clp protease ATP-binding subunit ClpX	26.69	
cobS	CobS	putative aerobic cobaltochelatase	27.98	25.78
cysD	CysD	putative sulfate adenylyltransferase subunit 2	26.47	22.04
dkgA	DkgA	putative 2,5-diketo-D-gluconic acid reductase A	26.08	
fcl	Fcl	putative GDP-L-fucose synthetase	27.23	24.85
fixA	FixA	electron transfer protein FixA	34.61	29.15
fixB	FixB	electron transfer protein FixB	27.38	32.69
frcA	FrcA	putative ATP-binding component of ABC transporter CUT2 fructose transporter from homology with S.mel SMc02169	26.78	23.72
ftsZ2	FtsZ2	putative cell division protein FtsZ	25.84	26.97
glnII	GlnII	putative glutamine synthetase II	26.78	25.19
ilvI	IlvI	putative acetolactate synthase isozyme III large subunit	27.3	
metK	MetK	putative S-adenosylmethionine synthetase	25.28	26.65
*pdhA**	PdhA*	putative pyruvate dehydrogenase E1 subunit A	24.91	26.28
*pdhB**	PdhB*	putative pyruvate dehydrogenase E1 subunit B	25.82	26.88
petA	PetA	putative ubiquinol-cytochrome c reductase iron-sulfur subunit		27.60
prs	Prs	putative ribose-phosphate pyrophosphokinase	27.99	26.12
*rkpK**	RkpK*	putative UDP-glucose 6-dehydrogenase	26.83	27.12
ropA2	RopA2	putative outer membrane porin protein RopA		26.34
rpoA	RpoA	putative DNA-directed RNA polymerase alpha chain		26.22
rpoB	RpoB	putative DNA-directed RNA polymerase beta chain	26.93	28.35
secA	SecA	putative preprotein translocase subunit	28.89	26.41

		SecA		
typA	TypA	putative GTP-binding protein TypA/BipA (tyrosine phosphorylated protein A)		26.88
vbsS	VbsS	putative non-ribosomal peptide synthetase		30.73
RL0161		putative cell division DNA translocase protein		27.06
RL0717		conserved hypothetical protein (TPR repeat family)	25.15	27.19
RL0814		conserved hypothetical protein	26.86	23.50
RL0822		putative DegT family aminotransferase	26.81	25.48
RL1945		putative vitamin B12-dependent ribonucleotide reductase	26.89	25.59
RL2369		putative NADPH:ferredoxin reductase	27.48	
RL2700		conserved hypothetical exported protein	26.88	25.76
RL3210		putative glucosyl transferase		30.36
RL3217		putative MarR family transcriptional regulator	22.21	29.26
RL3626		putative ATP-binding:ATP-binding (ABC:ABC) component of ABC transporter CUT2		26.84
RL4716		conserved hypothetical exported protein	26.13	
pRL70107		putative zeta toxin	24.38	26.54

Table 7.5. Top ranked interacting proteins in free-living FLAG-tag pull-down eluates as determined by protein MS. Results from pull-downs with both tagged FixA and tagged FixB. Top twenty proteins according to average abundance (intensity) and significance compared to a negative control. Grey shading indicates values falling outside of the top twenty for that protein species. Black shading indicates that the relevant protein species did not significantly interact with the tagged protein. Chaperones and ribosomal proteins were excluded from this list. * indicates proteins which also appeared in the bacteroid top twenty.

Gene	Protein	Description	Average log(Intensity)
atpA	AtpA	putative ATP synthase alpha chain	29.42
fixA	FixA	electron transfer protein FixA	32.89
fixB	FixB	electron transfer protein FixB	30.89
fixC	FixC	nitrogen fixation protein FixC	29.35
fixO	FixO	putative cytochrome oxidase subunit	28.67
fixP	FixP	putative cytochrome oxidase subunit	28.46
lon	Lon	putative ATP-dependent protease	28.73
nifB	NifB	FeMo cofactor biosynthesis protein NifB	28.06
nifD	NifD	nitrogenase molybdenum-iron protein alpha chain NifD	33.32
nifE	NifE	putative nitrogenase iron-molybdenum cofactor biosynthesis protein NifE	27.85
nifH	NifH	nitrogenase iron protein NifH	32.91
nifK	NifK	nitrogenase molybdenum-iron protein beta chain NifK	32.84
*pdhA**	PdhA	putative pyruvate dehydrogenase E1 subunit A	28.68
*pdhB**	PdhB	putative pyruvate dehydrogenase E1 subunit B	29.33
*rkpK**	RkpK	putative UDP-glucose 6-dehydrogenase	28.68
sucA	SucA	putative 2-oxoglutarate dehydrogenase E1 component	29.53
sucB[a]	SucB[a]	putative dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase	29.89
RL1878		putative peptidoglycan binding protein	28.02
pRL100098		conserved hypothetical protein	28.04
pRL100103		putative hydroxyvalerate/ alcohol dehydrogenase	28.32

Table 7.6. Top ranked interacting proteins in bacteroid FLAG-tag pull-down eluates as determined by protein MS. Results are from C-terminal triple FLAG tag on FixB. Top twenty proteins according to average abundance (intensity) and significance compared to a negative control. Chaperones and ribosomal proteins were excluded from this list. a. *sucB* is annotated as *citM* and *odhA* by some sources. * indicates proteins which appeared in the free-living FixAB top twenties.

7.6 Validation of protein-protein interactions using a Bacterial Two-Hybrid assay

This study suggests interaction between FixAB and several subunits of pyruvate dehydrogenase. The PDH (pyruvate dehydrogenase complex) is made up of three subunits. Subunit E1, encoded by *pdhAB*, is a pyruvate dehydrogenase. Subunit E2, encoded by *pdhC,* is a dihydrolipoyl transacetylase. Finally subunit E3, encoded by RL2246 (now annotated as *lpdH*), is a dihydrolipolyl dehydrogenase. The two subunits encoding E1 were both strongly predicted in protein MS to interact with FixB. A Bacterial Two-Hybrid assay (BTH) (Karimova, Pidoux et al. 1998) was used to confirm these interactions.

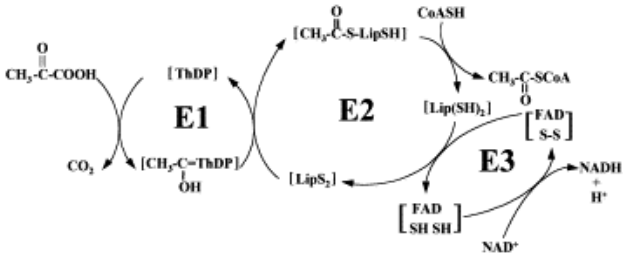


Figure 7.24 Reactions carried out by the pyruvate dehydrogenase complex in Gram negative bacteria (de Kok, Hengeveld et al. 1998). Subunit E1 comprises PdhAB, subunit E2 is PdhC and subunit E3 is LpdH.

Two cloning vectors are used in the bacterial two-hybrid system, each encoding part of an adenylate cyclase (T18 and T25). These adenylate cyclase fragments are fused to proteins of interest, and if the proteins interact they work together to express reporter genes such as *lacZ*. *E. coli* strains containing interacting plasmids can be tested for *lacZ* expression using a quantitative β-galactosidase assay.

Constructs were designed for FixA, FixB, PdhA, PdhB, PdhC and LpdH in both the T18 and T25 vector. All variations of these plasmids were tested, as well as the empty vectors as negative control. FixAB and the PDH complex together will act as positive controls for all plasmids. Testing reciprocal combinations of T18-T25 allows for confirmation of interactions.

None of the combinations involving FixA and FixB showed any significant interactions. Crucially FixA did not interact with FixB even though these proteins must form a complex in the cell to be active.

PdhA/B (E1 subunits), PdhB/C (E1-E2) and PdhC/LpdH (E2-E3) all show positive interactions, so the BTH system works with all PDH complex subunits. Figure 7.25 shows the activity as determined by qualitative β-galactosidase assays.

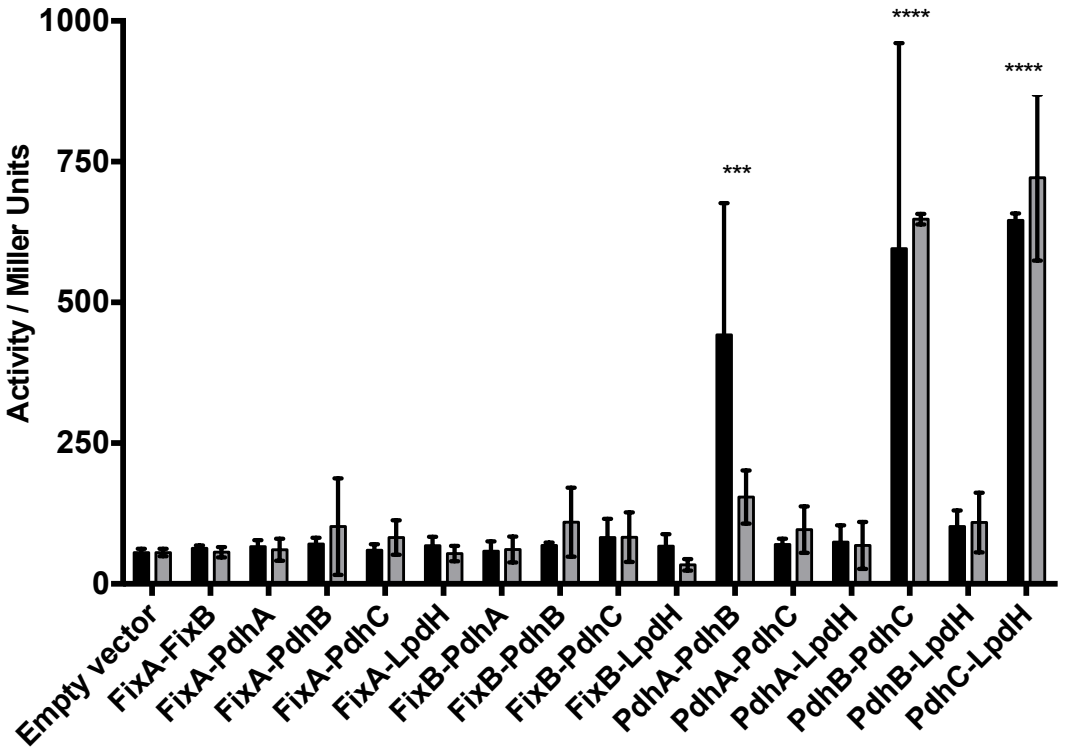


Figure 7.25 β-galactosidase assay to determine interaction in the bacterial-two hybrid system. Values expressed in Miller units. Black bars show T18-T25 arrangement, and grey bars show the reciprocal T25-T18 arrangement. n=3, error bars show SEM. Statistically significant combinations were tested using a two-way ANOVA with row mean comparisons; *** $p<0.005$, **** $p<0.0005$.

7.7 Investigating FixCX

7.7.1 Constructing tagged versions of FixC and FixX

A tagged version of FixC was constructed using the pJP2 backbone. In order to keep *fixC* under its native control, the entire upstream p*fix* region plus the *fixABC* genes were amplified, and a 3xFLAG tag was included on the C-terminal of *fixC*. This plasmid was conjugated into the in-frame *fixC* mutant and grown on *P. sativum*. The tagged FixC was unable to complement the deletion mutant, suggesting that the C-terminal tag is interfering

with the activity of FixC. An anti-FLAG Western blot was used to confirm the presence of tagged FixC in these complemented bacteroids (data not shown).

Strain	**Description**	**Symbiotic phenotype**	**Nodule colour**	**Acetylene reduction** μmol ethylene $plant^{-1}$ hr^{-1}
3841	Wild-type	Nod^+ Fix^+	Pink	2.7 ± 0.3
OPS0280	Δ*fixC*	Nod^+ Fix^-	White	0.0 ± 0.0
OPS0357	Δ*fixC:*pJP2-*fixABC*-3xFLAG	Nod^+ Fix^-	White	0.0 ± 0.0
Uninoculated control	n/a	Nod^- Fix^-	n/a	0.0 ± 0.0

Table 7.7 Symbiotic phenotype of complemented FixC mutant strains on pea (*P. sativum*). Acetylene reduction assay was carried out 28 days post inoculation; ± SEM, n ≥ 6.

In order to confirm the lack of activity with a C-terminal FLAG-tag, Golden Gate cloning was used to create tagged and untagged versions of FixC under the constitutive *nptII* promoter. Constructs were also made for FixX. These clones were constructed in a domesticated version of the pJP2 vector designated pOGG041 (Geddes et al., in production). This vector had not yet been tested *in planta*. Untagged versions and versions with a C-terminal FLAG tag were cloned and conjugated into Rlv3841 and the respective in-frame deletion mutants.

Acetylene reductions were carried out on the Golden Gate complement mutants (Figure 7.26). None of the mutant strains could be complemented to the wild-type nitrogen fixation levels. A small amount of fixation could be seen in the FLAG-tag *fixC* complemented strain, however this wasn't seen in the untagged version. The FLAG-tag plasmid also appeared to result in slightly higher fixation rates in the wild-type background.

Twelve nodules from each strain were crushed and streaked on TY and individual colonies subsequently patched onto TY plus tetracycline in order to determine whether the plasmid had been retained. All strains had 100% tetracycline resistance, suggesting that the plasmid is present, but the construct is not expressing correctly.

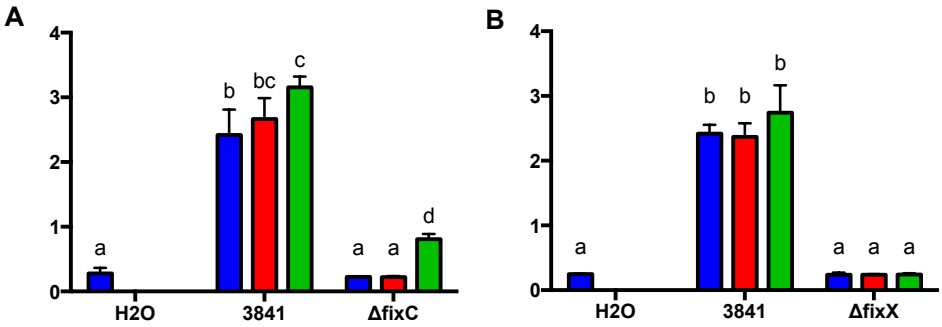


Figure 7.26 Symbiotic phenotype of complemented Δ*fixC* (A) and Δ*fixX* (B) mutant strains on pea (*P. sativum*). Blue bars are the background strain lacking any plasmid. Red bars show strains complemented with pJP2-*fixC/*X. Green bars show strains complemented with pJP2-*fixC/X*-3xFLAG. Acetylene reduction assay was carried out 28 days post inoculation, with values expressed as μmol ethylene plant^{-1} hr^{-1}; ± SEM, n ≥ 4. a, b, c, d represent statistically distinct groups (p≤0.05) determined using one-way ANOVA and Tukey's multiple comparisons test.

7.7.2 Cellular localisation of FixC protein

The mammalian ETF-QO is membrane localised (Watmough and Frerman 2010). A FLAG-tagged FixC enables detection of FixC localisation in *R. leguminosarum*. An initial bioinformatic analysis can be used to predict the localisation of the FixC protein, as well as the FixCX complex. There are many prediction programmes available which search for subcellular localisation signal sequences, amino acid sequence, features such as transmembrane domains and homology data (Cserzo, Wallin et al. 1997, Gardy and Brinkman 2006). The majority of prediction softwares used (pSORT (Yu, Wagner et al. 2010), SubLoc v1.0 (Hua and Sun 2001) and CELLO2GO (Yu, Cheng et al. 2014)) predicted FixC and FixCX to be cytoplasmic proteins with high probability (>90%). Transmembrane domains were only detected when FixC was analysed through TopPred (Claros and Vonheijne 1994), which is a software designed to detect transmembrane helices in membrane proteins. TopPred predicts two highly possible transmembrane helices at amino acids 6-26 and 234-254, as well as a putative transmembrane domain at 317-337.

FixC localisation was tested using a membrane fractionation protocol. A cell extract from bacteroids containing the pJP2-*fixC*-3xFLAG construct (OPS0357) was disrupted in a ribolyser in order to break apart the cells. The resulting extract was centrifuged to pellet the cytoplasmic fraction and large cell fragments. The remaining supernatant should consist of small membrane fragments. The supernatant was then spun at 83850 G for two hours to

pellet the membrane fraction. The remaining supernatant was retained. All fractions were tested in an anti-FLAG Western blot. Tagged FixC could be detected in the cell extract, and then was retained in the cytoplasmic fraction and could not be detected in the membrane fraction (Figure 7.27). A FLAG-tagged AapJ protein (Poole lab, unpublished) was used as a positive control. AapJ is a periplasmic protein involved in amino-acid transport in *R. leguminosarum* (Walshaw and Poole 1996). Prediction software pSORT agrees that AapJ is located in the periplasm. A differential fractionation of a strain containing FLAG-tagged AapJ showed protein present in the supernatant of the ultracentrifugation step. This demonstrates that the ultracentrifugation spin is able to separate proteins by localisation.

A second positive control was desired; a known-membrane associated protein such as AapP. This was not successfully obtained within the time frame of this study. A FLAG-tagged AapP is available in the Poole lab, in plasmid pRU1138 (Prell, Mulley et al. 2012). Attempts to express tagged AapP from strain LMB738 (wild-type background) were unsuccessful. New conjugation or differed growth conditions may be required. Without this second control it is unknown whether the membrane has been broken into sufficiently small fractions. Larger membrane fractions may be retained in the cytoplasmic fraction; FixC may have been retained as a result of this if it is indeed membrane associated as predicted by homology with mammalian proteins.

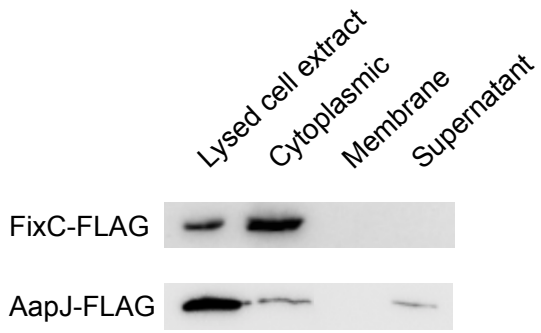


Figure 7.27 Anti-FLAG Western blot to detect FLAG-tagged proteins in soluble and membrane cell fractions. FixC-FLAG sample obtained from bacteroids harvested from *P. sativum* inoculated with OPS0357 at 28 dpi. AapJ-FLAG from strain RU1893 grown in rich media overnight at 200 rpm.

7.8 Discussion

The precise role of FixABCX in symbiotic nitrogen fixation is currently unclear, and there is little biochemical data available for the *Rhizobium*-legume symbioses. In free-living symbiotic species such as *R. rubrum* these proteins are proposed to carry electrons to nitrogenase (Edgren and Nordlund 2006). There is also genetic evidence of interaction between *fixB* and *lpdH*, the dihydrolipoyl dehydrogenase subunit of pyruvate dehydrogenase (Scott and Ludwig 2004). Additionally, ETF complexes are defined by Buckel and Thauer (2013) as interacting with dehydrogenases. This led to a proposed mechanism in which FixAB interacting with pyruvate dehydrogenase, and carries electrons downstream to nitrogenase. This hypothesis has been tested using protein pull-downs to determine interacting partners for FixAB in *R. leguminosarum*.

Several strategies were used to enable purification and pull-down of *R. leguminosarum* bv. *viciae* 3841 FixAB. Attempts to express FixAB in *E. coli* showed some evidence of expression, but not in high amounts. Successful expression of other proteins with the pOPIN system carried out within the Poole lab showed clear evidence of overexpression in a Coomassie gel. Overexpression of foreign electron transfer proteins in *E. coli* may lead to poorly folded proteins or formation of inclusion bodies. Native rhizobial expression was therefore chosen for future work.

As shown in Chapters 3 and 6, FixAB are only expressed in bacteroids, and are not detected in the free-living state. The taurine-inducible vector pLMB509 (Tett, Rudder et al. 2012) was used to express FixAB with various N- and C-terminal tags to allow for purification and pull-down. Expression could be detected under taurine induction, allowing for large volumes of culture to be produced if required. The *fixAB* genes were also cloned under their native promoter with C-terminal tags to allow expression from in bacteroids; allowing investigation of the proteins in a biologically relevant system.

His_6, single-FLAG, triple-FLAG, single-Strep and twin-Strep tags were tested. Single-Strep tags could not be detected in initial constructs, and so twin-Strep was used. Both twin-Strep and 3xFLAG have higher affinity in column purification according to manufacturers (Domanski, Molloy et al. 2012, Schmidt, Batz et al. 2013). These improved tags showed better detection in the tagged FixAB samples and so were taken forward for further experiments. Strains with an added His_6-tag appeared to have slightly lower expression in Western blots, so the no-His_6-tag strains were used for pull-down assays.

In this work we also began constructing vectors to express tagged FixC and FixX. A 3xFLAG-tagged FixC was constructed in the stable vector pJP2, and grown on plants. The

presence of a FLAG-tag interrupts FixC function, and FixC-3xFLAG is unable to complement a FixC deletion mutant. An untagged version of FixC was able to complement an in-frame *fixC* mutant. In addition to conventional cloning, a version of FixC was constructed using Golden Gate cloning during the process of optimising the Golden Gate system in the Poole lab. Tagged FixX constructs were also designed. This acted as a test for the functionality of a domesticated (removal of Golden Gate-compatible restriction sites) version of pJP2. The vector was not functional, and so functional tagged plasmids for FixC and FixX were never generated. A minimal version of pJP2 was subsequently constructed and confirmed as functional in bacteroids (see Chapter 6 for characterisation). This confirmed vector, along with the bacteroid-specific p*fixA* promoter (available for Golden Gate, see chapter 6) can be used as a high-throughput method for assembly of constructs expressing tagged proteins in the future.

The tagged FixC construct was also used to detect the subcellular localisation of FixC. In mammals the ETF-QO is membrane localised. A differential spin was used for membrane fractionation, and tagged FixC appeared to be retained in the cytoplasmic fraction. This agrees with bioinformatic analysis carried out on the FixC amino acid sequence. The tagged version of FixC was unable to complement a *fixC* mutant; lack of functionality in tagged FixC may mean it is not found in its native cell location. However, a known membrane protein was never used to confirm the protocol due to expression problems with AapP-FLAG. Additional controls are also required to confirm the membrane fractionation protocol and so localisation of FixC.

Crude lysis was used to optimise the use of gravity columns prior to pull-down assays. In order to carry out pull-downs, tagged samples were treated with a gentle lysis protocol to retain protein-protein interactions. Lysed samples were then passed over columns and the same protocol followed. A successful pull-down was carried out using the FLAG system. A Strep-tagged protein could be purified using a Strep column from a crudely lysed sample, but much less protein was seen in the eluate fraction compared to a FLAG-tagged purification. A successful Strep-tag immunoprecipitation was never obtained using a gently lysed sample and so no interactions were investigated using this system. The FLAG system was chosen to investigate FixAB protein-protein interactions. Samples from bacteroids with a C-terminal FixB FLAG tag allowed determination of interactions in the native system. In the free-living system tags were put on the N-terminal of FixA and the C-terminal of FixB. Coomassie gels showed the presence of multiple proteins in addition to those tags, suggesting a successful pull-down.

Once the pull-down protocol was optimised for FLAG-tag samples, Western blots were used as an initial test for interactions. FixA and FixB were shown to interact, as expected (they form subunits of a heterodimer). Monoclonal antibodies were designed against proposed interacting partners NifH, PdhA and LpdH. Western blots of pull-down eluates with these custom antibodies didn't show detectable levels of any of these proteins. The eluates of the pull-downs have low protein concentrations, so very high resolution may be needed to detect interacting proteins; this resolution may not be reached with custom antibodies. Tandem protein mass spectrometry (MS/MS) allows much higher resolution for detection of these partners.

MS/MS was used to determine the identity of interacting partners. An initial test run resulted in >1000 potential interacting partners; far too high to make any strong conclusions. A high-resolution protein stain, SYPRO Ruby was used to optimise wash conditions, including use of detergents, high salt and additional wash steps. Further MS showed a significant (~50%) reduction in potential interacting partners; detergent and additional wash steps removed non-specific binding. Comparison to other pull-down assays carried out in the Poole lab suggested that the tagged FixAB proteins had a particularly high affinity for non-specific binding. For example a pull-down using the same system on a C-terminal 3xFLAG-tagged PtsN yielded only 47 significant interacting proteins when put through the same data analysis pipeline (Sanchez-Canizares, C. pers. comm.).

The results of protein MS resulted in ≤282 significantly pulled-down proteins for each sample (Free-living FixA, free-living FixB and bacteroid FixB). Comparison of this proteomic data to microarray analysis of bacteroids did not support the hypothesis that protein abundance in the cell is the main factor influencing the proteins that are pulled-down by FixB-3xFLAG.

The list of putative interacting proteins includes a large number of chaperones and ribosomal proteins. This includes a GroEL protein, which is the 60 kDa band detected in Coomassie Blue-stained gels. Since these proteins form large complexes within the cell, their presence may be an artefact of the experimental procedure. Alternatively, these chaperones may be a result of overexpression of a large electron-transfer complex under free-living conditions. Pull-downs were optimised for this project, and so there was little previous data or relevant controls for comparison. Since the untagged control did not yield high levels of these housekeeping proteins, the presence of a tagged protein bound to the beads appears to cause this result. Additional pull-down experiments using other FLAG-tagged proteins in the bacteroid will provide important control data. This includes experiments using highly

upregulated bacteroid proteins which did not appear in any of the pull-downs carried out using FixAB, as well as a pull-down with a tagged GroEL chaperone.

A few hundred proteins were found to be significantly interacting with FixA/B under the various pull-down conditions. In order to make some biological conclusions from the data housekeeping proteins were removed from the lists; any proteins identified as heat-shock, chaperones or ribosomal were discarded. Proteins were then ranked by intensity, a value indicating the abundance of peptides picked up by the mass spectrometer that related to each protein. A cut-off of the top twenty according to intensity was selected in order to assess the strongest interactions.

The lists from the free-living samples contained few proteins that were predicted in our model, or that stood out as biologically important. As mentioned above, some housekeeping proteins were removed; and only further pull-downs with other controls in this system will indicated which proteins are artefacts or true interacting partners. The expression of FixA and FixB in free-living samples is not representative of what happens in nature, as they are not usually expressed under these conditions. As part of the nitrogen fixation process, it could be predicted that their true interacting partners may only be revealed in experiments in bacteroids. In both free-living cases both FixA and FixB were pulled down, showing an interaction between FixA and FixB, as expected since they are subunits of a heterodimer. Subunit E1 (PdhAB) of the pyruvate dehydrogenase complex, which is predicted to interact with FixAB in our model, was detected in the free-living tagged samples. This strengthens our hypothesis that PDH interacts with FixAB.

The bacteroid sample is more biologically relevant, as FixB is expressed as part of its native system, and all its natural interacting partners will be present in the cell. The data from the tagged FixB bacteroid sample provides some exciting potential interacting partners. Half of the top proteins are bacteroid-specific proteins involved in nitrogen fixation. FixC appears in the top twenty but FixX does not; however FixX is still found to be relatively abundant in the sample. All three subunits of nitrogenase, NifHDK appeared in the top twenty. Since NifHDK forms a single large complex, interaction of FixAB with one of these subunits may result in the entire complex being pulled down. Some proteins involved in nitrogenase biosynthesis and maturation, NifE and NifB, were also found in this list; again these may have been pulled down due to close interactions with NifHDK. NifN was also found to be significantly enriched in the tagged FixB pull-down. Two subunits of the cytochrome cbb_3-oxidase FixNOQP were detected. There are multiple copies of this operon encoded in Rlv3841, and a distinction could not be made between FixO. Peptides unique to FixP1 and FixP2 were detected, suggesting that both the pRL9- and pRL10-encoded complexes are

interacting. Interaction with FixNOQP adds support to the hypothesis that FixABCX may be passing electrons to both nitrogenase and to terminal electron acceptors.

PdhA and PdhB, which form subunits of pyruvate dehydrogenase subunit E1 were also highly abundant in the sample. These subunits were also detected in the free-living state, and it seems likely that FixAB and PDH subunit E1 are interacting within the cell. The E1 and E2 subunit of 2-oxoglutarate dehydrogenase were also pulled down in this sample. Figure 7.28 shows a model of bacteroid metabolism annotated with the putative interacting partners of FixAB.

Pyruvate dehydrogenase has been proposed to be a key enzyme in bacteroid carbon metabolism, required for production of acetyl-CoA (Driscoll and Finan 1993). The role of the PDH complex is to oxidise pyruvate, resulting in the production of acetyl-CoA that is fed into the TCA cycle. NAD^+ is the normal electron acceptor for pyruvate dehydrogenase, but evidence here suggests that FixAB may be acting as an electron acceptor for this complex. Biochemical interaction of pyruvate dehydrogenase with FixAB has not previously been demonstrated in the literature. Work in *A. caulinodans* has demonstrated a genetic interaction between *fixB* and *lpdH* (Scott and Ludwig 2004). A strong interaction with LpdH (subunit E3) was not detected in protein mass spectrometry, but interaction with the PdhAB, the PDH E1 subunit was observed. RT-PCR of *S. meliloti* has shown higher transcription of the genes encoding PDH during symbiosis compared to the free-living state (Cabanes, Boistard et al. 2000), although this is not the case in *R. leguminosarum* bv. *viciae* 3841 (Karunakaran, Ramachandran et al. 2009). In *A. caulinodans* ORS571 pyruvate dehydrogenase is required for nitrogen fixation in symbiosis with *S. rostrata*. In contrast, a *pdhB* mutant in *A. caulinodans* is able to fix nitrogen in free-living conditions at 10 μM dissolved oxygen tension (partially aerobic), though at a lower rate than the wild-type, and with poorer growth with succinate as a carbon source. Under the microaerobic conditions of the nodule, PDH becomes essential for growth (Pauling, Lapointe et al. 2001). Scott and Ludwig (2004) hypothesise that the similarity between the LpdH subunits of the 2-oxoacid dehydrogenase complexes of *A. caulinodans* means that FixAB may interact with any of these dehydrogenases. This may provide answers to how a PDH mutant is able to fix nitrogen; FixAB may be acting with a different dehydrogenase. In this work we have shown a potential interaction between FixAB and SucAB, the E1 subunit of the 2-oxoglutarate dehydrogenase (ODH) complex. SucB is the most upregulated TCA cycle enzyme in mature bacteroids (21 and 28 days post-inoculation) compared to the free-living state, with expression 2.5- to 3-fold higher in mature bacteroids (Karunakaran, Ramachandran et al. 2009). FixAB may be able to interact with SucAB in a bifurcating system in order to

energise nitrogen fixation, either in place of PDH, or in a parallel reaction. This may explain why an *A. caulinodans pdhB* mutant is able to fix nitrogen under partially microaerobic conditions; FixAB is interacting with ODH as it functions within the TCA cycle. In a nitrogen fixing bacteroid under very low oxygen tensions, flux through the TCA may be too low for ODH to compensate for a lack of PDH. This alternative pathway could be investigated by determining the nitrogen fixation phenotype of an *A. caulinodans* ODH mutant in free-living and symbiotic nitrogen fixation, as well as an ODH-PDH double mutant.

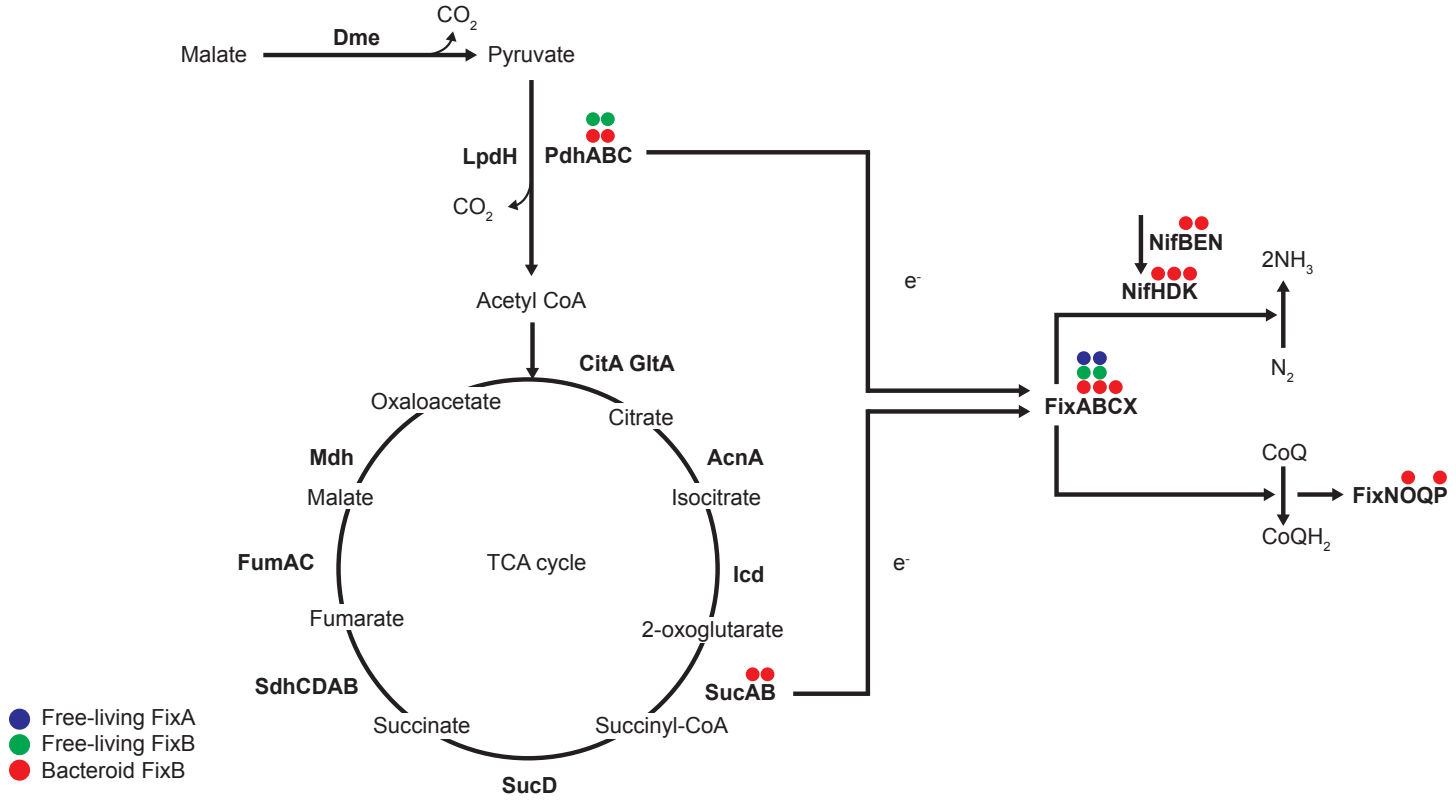


Figure 7.28 Model of FixABCX function within the cell with annotations displaying the strongly interacting proteins as determined by pull-down and protein MS. Dots indicate that a protein was seen in the top twenty most abundant interacting partners for FLAG-tagged FixB in bacteroids (red); FLAG-tagged FixA in the free-living state (blue); FLAG-tagged FixB in the free-living state (green).

In order to confirm the interactions seen in these pull-down assays, several more experiments are required. Reciprocal tagging of FLAG-tags on proposed partners such as PdhA and NifH would act as a confirmation of their interactions. Additionally, optimisation of the Strep-tag protocol would allow a repeat of the FixAB experiments in a different system, and may reinforce or disprove data seen here. Other pull-down systems are available, such as GFP-beads and His_6-tag. In this study pull-down experiments were carried out by washing samples containing FLAG-tagged proteins over gravity-flow columns containing anti-FLAG M2 beads. An alternative strategy is to incubate the samples in microfuge tubes with anti-FLAG M2 beads, and carry out subsequent washes within the tube. A further strategy would be to purify the tagged protein and bind this to the antibody beads. A lysed cell extract could then be added to the beads, and any interacting partners would be expected to bind to the beads. Optimisation of these strategies combined with further protein MS will provide much more evidence to support or refute suggestions made here. As previously mentioned, an important negative control will be the results of pull-downs with unrelated proteins that did not show interaction with FixAB. This will help to understand the presence of housekeeping proteins in the samples.

The bacterial two-hybrid system was used to test the interactions seen between FixAB and PdhAB. All subunits of pyruvate dehydrogenase were tested against FixA and FixB, with reciprocal tagging to ensure complete coverage of the interactions. The pyruvate dehydrogenase subunits clearly interacted with each other, providing a positive control for the two-hybrid protocol. No interaction could be seen between FixA and FixB, or between these proteins and the pyruvate dehydrogenase subunits, suggesting that the plasmids encoding *fixA* and *fixB* are not expressing functional proteins. Previous attempts to express FixAB in *E. coli* have resulted in inclusion bodies (Watmough, N., pers. comm.). These proteins may not be expressing properly from this system. In order to further investigate this new constructs could be designed containing smaller domains of FixA/B rather than the entire protein. Bioinformatic predictions could be used to suggest the most likely amino acid regions for interaction.

Co-immunoprecipitation experiments could be used to test interactions *in vitro*. Tags are already available for FixAB, which could be purified in large volumes from taurine-induced cultures. The PDH subunits express well in *E. coli* in the bacterial two-hybrid system, and so are good candidates for T7-based expression in the POPIN system, which produces large amounts of protein material. Optimisation of expression from Golden Gate vectors has allowed a high-throughput method of producing tagged proteins for expression in rhizobium; this could be used to produce tagged nitrogenase subunits, or other units of interest. The

strength of interactions between purified proteins could then be tested using Biacore (GE Healthcare).

The mass spectrometry data provides a wealth of potential interacting partners for the FixAB ETF complex. The most interesting of these are pyruvate dehydrogenase and the nitrogenase complex. It may be that these three large complexes are interact together in the cell in one larger supercomplex. There may also be a role for 2-oxyglutarate dehydrogenase in place of pyruvate dehydrogenase in this system. This could explain why so many putative interacting proteins were seen in FixAB pull-down experiments; secondary interactions with either pyruvate dehydrogenase or nitrogenase may also be pulled-down. This is already seen with NifB, NifE and NifN, involved in nitrogenase maturation. Additionally, a very large complex may trap otherwise non-interacting proteins and block their passage through the gravity column during wash steps. Comparison of reciprocal experiments with PHD subunits and nitrogenase would provide data to support or refute existence of a supercomplex.

CHAPTER 8

General discussion and future perspectives

The mechanism of electron transfer to nitrogenase has not been characterised in the *Rhizobium*-legume symbioses. FixABCX have been identified as playing a possible role in electron transfer in several free-living rhizobial species. FixA and FixB show homology with electron transfer flavoproteins, and FixCX shows homology to the ETF cognate electron acceptor ETF-quinone oxidase. ETFs are conserved across the Kingdoms of Life, playing roles in electron transfer in many different processes. In this work I show that the *fixABCX* gene products are essential for nitrogen fixation in *R. leguminosarum* bv. *viciae* 3841 in symbiosis with *P. sativum* (Chapter 3).

Mutation and complementation of *fixABCX* mutants: the *fix* genes are essential for symbiotic nitrogen fixation

Complementation of polar *fixAB* mutants with the entire *fixABCX* operon showed that *fixABCX* almost certainly form an operon. However, an odd result was obtained with a polar *fixC* mutant (LMB827) which appeared able to fix nitrogen, albeit at a reduced rate. However, this mutant showed incorrect PCR mapping, suggesting a complex insertion. Therefore, an in-frame *fixC* mutant, and an interposon-insertion mutant in *fixC* were isolated and shown to be unable to fix nitrogen. While LMB827 was not futher studied as part of this project, the strain is being sequenced to deduce the nature of the insertion. This is because it is a useful strain due to its stability and reduced rate of nitrogen fixation. This mutant is now being used within the Poole lab to investigate competition and efficiency within the rhizosphere.

Complementation of a *fixX* mutant suggests autoregulation of *nifA*

A polar *fixX* mutant could not be complemented with *fixABCX*, whilst other polar *fix* mutants could only be returned to approximately 50% of wild-type levels of nitrogen fixation. Investigation of the regulatory network controlling nitrogen fixation provided an explanation for this lack of fixation. The *nifA* gene is found immediately downstream of *fixX*. This gene encodes the transcription factor controlling many nitrogen fixation genes including *fixABCX* and *nifHDK* (encoding nitrogenase). A luciferase-based system was adopted to determine the region of pRL10 involved in expression of the *nifA* gene, and whether this would explain the lack of complementation. A series of basal promoters were detected within *fixC* and *fixX*. These promoters are expressed under ambient oxygen, contrasting to some rhizobia that require low oxygen for *nifA* expression (Dixon and Kahn 2004). Genes under NifA control

are not expressed under ambient oxygen, suggesting that regulation of nitrogen fixation occurs at the post-transcriptional level. Under ambient oxygen expression of *nifA* occurs from these basal promoters, but the NifA protein is not functional. Under bacteroid conditions a functional NifA is able to activate the *fixA* promoter. The NifA-dependent transcription from the *fixA* promoter is then 10-fold higher than the basal *nifA* promoter region *in planta*. The *fixA* promoter results in transcription of downstream genes, including *nifA* in an autoregulatory manner. This autoregulation explains why polar mutants in *fix* genes cannot be complemented to wild-type nitrogen fixation levels; *nifA* expression is limited to its low, basal expression level in these mutants, and subsequently expression of *fixA* and *nifH* is likely lower.

While a non-polar fixX mutant could be complemented with *fixXnifA*, a polar *fixX* mutant (which disrupts transcription from both the basal nifA and inducible fixA promoters) could not, suggesting *nifA* may be part of a larger operon consisting of at least *nifAB*. Cloning of constructs containing *nifB* was unsuccessful using the Golden Gate system. This is typically characteristic of constructs that are toxic in *E. coli*. Complementation of a polar *fixX* mutant is therefore yet to be demonstrated. Genetic techniques such as qRT-PCR may aid in construct design by identifying the full length of the operon.

A *lux*-based system can be used to investigate promoter structure non-invasively

The luciferase-based reporter system was used to investigate the upstream promoter region of *fixA*. In Rlv3841 there is over 1 kb of intergenic region upstream of *fixA*, flanked by a divergent NifA-controlled gene designated pRL100201. Whilst the role of pRL100201 is yet to be identified (annotated as a hypothetical protein), previous work in Rlv3841 has shown that it is not essential for nitrogen fixation (Karunakaran, Ramachandran et al. 2009). Whilst the *fixABCX* operon is conserved across the *Rhizobiacae,* this large intergenic region is not. The regulatory elements found in this region were investigated, and it was found that a minimal region of 176 bp was sufficient for *fixA* expression. This region comprised a consensus-binding site for the NifA transcription factor as well as a consensus-binding site for RpoN (σ^{54}), the cognate sigma-factor to NifA. These regulatory sites were found upstream of a putative transcription start site identified by 5'-mapping. Much of the 5'-mapping data obtained by Green and Poole (unpublished) has not been validated. This lux-based system provides a method for validating these start sites, and could be combined with bioinformatic promoter prediction techniques and consensus-sequence based searching to investigate regulatory networks throughout Rlv3841.

Expression from the *fixA* promoter was only seen within bacteroids. Expression assays were carried out under 1% oxygen, and in the presence of filtered nodule extract. Neither condition was sufficient to cause expression from the *fixA* promoter. This lack of expression demonstrates that post-translational control of NifA is not due to low oxygen alone. Further assays are required to determine what other factors are involved in NifA control. These factors may include nitrogen limitation, pH effects or presence of unknown plant-derived peptides or metabolites. There may also be additional regulatory factors expressed by the bacteroid. Further reporter assays under multiple conditions, as well as assays carried out in mutant backgrounds may provide more insight into post-translational NifA control. Further investigation of NifA control as well as control of other oxygen-regulated genes will allow creation of a more detailed regulatory network in Rlv3841.

Nitrogen fixation mutants have developmental phenotypes as a response to redox status and changes in plant input

In addition to genetic effects caused by mutation of *fix* genes the absence of nitrogen fixation in *R. leguminosarum* is accompanied by a series of developmental metabolic effects. Electron microscopy of *P. sativum* nodule sections was used to investigate morphological phenotypes of bacteroids, with wild-type displaying a typical Y-shape. Infection threads can also be seen, filled with smaller, round cells rich in polyphosphate and polyhydroxybutryate (PHB), both of which are produced under stress. Electron microscopy of mature nodules of *P. sativum* inoculated with *R. leguminosarum* bv. *viciae* strain LMB771 (*fixAB*::Ωspc) or LMB777 (Δ*fixAB*) showed a dual phenotype which appears to be under temporal control. Nodules at all stages contain differentiated bacteroids filled with large PHB granules, a sign of redox stress. In mature nodules (seen from 26 days post-inoculation) a second bacterial morphotype can be seen occupying plant cells. This second morphotype resembles that of undifferentiated cells seen in infection threads. Here we propose a feedback loop between the host and bacterioids. Before infection with *R. leguminosarum* bv. *viciae* 3841 and derived strains, *P. sativum* is unable to choose its symbiont partner by its nitrogen fixation ability (Westhoek, pers. comm.). The plant will act as in a functional symbiosis, providing carbon and other signals that lead to differentiation into bacteroids. Only once the bacteria have differentiated will the absence of nitrogen provision be detected. This eventually leads to a negative feedback, where the plant ceases providing the signals required for bacteria to differentiate after release from later infection threads. These undifferentiated cells remain rich in polyphosphate, but appear depleted in PHB compared to the infection thread. PHB may be being utilised as a carbon source in the absence of plant-derived C_4-decarboxylates.

Production of PHB in earlier-differentiated bacteroids of *fix* mutants is as a response to redox stress caused by the lack of nitrogen fixation. Nitrogen fixation requires eight electrons per molecule of nitrogen fixed, and so if this process is blocked reductant will build up. Microaerobic conditions limit the use of the TCA cycle to remove reductant, and so alternative storage strategies such as PHB must be co-opted to reduce stress on the cells. PHB accumulates in bacteroids, which grow and begin losing their classic Y-shaped morphology as they accommodate large storage molecules.

Single-cell microscopy techniques allowed biochemical phenotyping of bacteroids

Single-cell Raman microscopy (SCRM) provided a novel solution to phenotyping the rhizobia within these mutant nodules. SCRM, combined with electron micrographs, allowed separation and subsequent phenotyping of the dual phenotype in a relatively simple way. An alternative strategy would be to use a cell sorter to separate cells by size, with subsequent biochemical assays to determine PHB, glycogen and polyphosphates. This would require a huge volume of bacteroid material. Raman microscopy allowed simultaneous assaying of several macromolecules, and with a small volume of cells. Raman microscopy has high potential for use in this field as it can distinguish between isotopically labelled samples in a quantitative manner, and so labelled nitrogen or carbon could be followed through the symbiosis, in both the plant and bacterial symbiont. During optimisation of Raman for this work we identified coupled biosynthesis of PHB and glycogen, in both overexpressing and mutant strains. This strategy could be used to investigate other biosynthetic mutants, especially in complex mutants such as those regulating carbon metabolism. Being able to separate phenotypes means that comparisons could be made between infection thread bacteria and differentiated bacteroids. This could help understand gene duplication and redundancy, for example PHB synthase *phaC1,* expressed in the infection thread versus *phaC2* expressed under NifA control in bacteroids. It could also allow investigation of multiple occupancy in nodules. Raman microscopy may have a similar potential to techniques such as MALDI-TOF mass spectrometry for fingerprinting and identification of rhizobial strains (Ziegler, Mariotti et al. 2012).

Determining the interacting partners of FixAB: The Fix proteins may form a supercomplex with pyruvate dehydrogenase and nitrogenase

The FixABCX proteins have been hypothesised to play a role in electron transfer to nitrogenase. This function would explain their essential role in symbiotic nitrogen fixation. No biochemical evidence is available for this role in the legume-rhizobia symbioses. I therefore carried out protein pull-down assays using FLAG-tagged FixA and FixB in both free-living and bacteroid conditions in order to determine interacting partners of the FixAB proteins. The FixAB proteins had a high number of interacting partners; nearly 300 significantly associating with the tagged protein in some cases. This included a large number of housekeeping proteins; in the free-living case this may be expected due to expression of electron-transfer proteins outside of their native environment. Upon the removal of housekeeping genes a list of putative interacting partners was created. In all samples, FixA and FixB interacted, which is expected since these form the beta and alpha subunits of a heterodimer. In the bacteroid sample a tagged FixB was able to pull-down both FixC and FixX as well as all three units of nitrogenase (NifHDK) and several other proteins involved in nitrogen fixation; FixOP (part of the cytochrome cbb_3-oxidase required for respiration under microaerobic conditions), and NifB and NifE, involved in nitrogenase maturation. NifBE could be appearing due to interaction with the NifHDK proteins.

In addition to bacteroid-specific proteins, the E1 subunit of both 2-oxoglutarate dehydrogenase (SucAB) and pyruvate dehydrogenase (PdhAB) were identified as putative interacting proteins. The E3 subunit of pyruvate dehydrogenase has previously been suggested as interacting with FixAB as a result of genetic data (Scott and Ludwig 2004), but this is the first instance of biochemical data to support the interaction between pyruvate dehydrogenase and FixAB. We proposed a model by which pyruvate dehydrogenase or 2-oxoglutarate dehydrogenase is able to act with FixAB in a complex in order to pass electrons onwards to nitrogen fixation. This may be by electron bifurcation, for which growing evidence has accumulated over the past decade as a mechanism of interaction between electron transfer flavoproteins and dehydrogenases. It has been proposed that a bifurcating system may occur in the case of FixAB (Herrmann, Jayamani et al. 2008) though no dehydrogenase partner/s have been suggested previously. Bifurcation would overcome the high-energy requirement for reduction of ferredoxin. FixX is a ferredoxin, however it is hypothesised to be working with FixC as an integral part of part of the ETF-QO. Across the Kingdoms of Life ETF-QO proteins are encoded by a single ORF, suggesting that the ferredoxin (iron-sulphur cluster-binding) is essential for ETF-QO function. A highly expressed ferredoxin under NifA control, FdxB1 was identified in this work, however it was shown to be non-essential for nitrogen fixation although this may be due to redundancy.

FixCX also interacts with FixAB, whilst FdxB1 was not one of the highest-interacting proteins seen in pull-down data. In the case of *R. leguminosarum* bv. *viciae* 3841 this FixAB-dehydrogenase pair may be occurring in a larger supercomplex with nitrogenase and FixCX. The existence of a supercomplex may explain the volume of proteins being captured during protein pull-down assays. FixABCX may be passing electrons directly to nitrogenase, with no additional ferredoxin required. FixCX may also be passing electrons to the quinone pool, which would carry them onwards to respiratory complexes. Interaction studies or purification and biochemical/electrochemical assays would provide experimental evidence for this. Figure 8.1 illustrates the proposed scheme for FixABCX function in symbiotic nitrogen fixation.

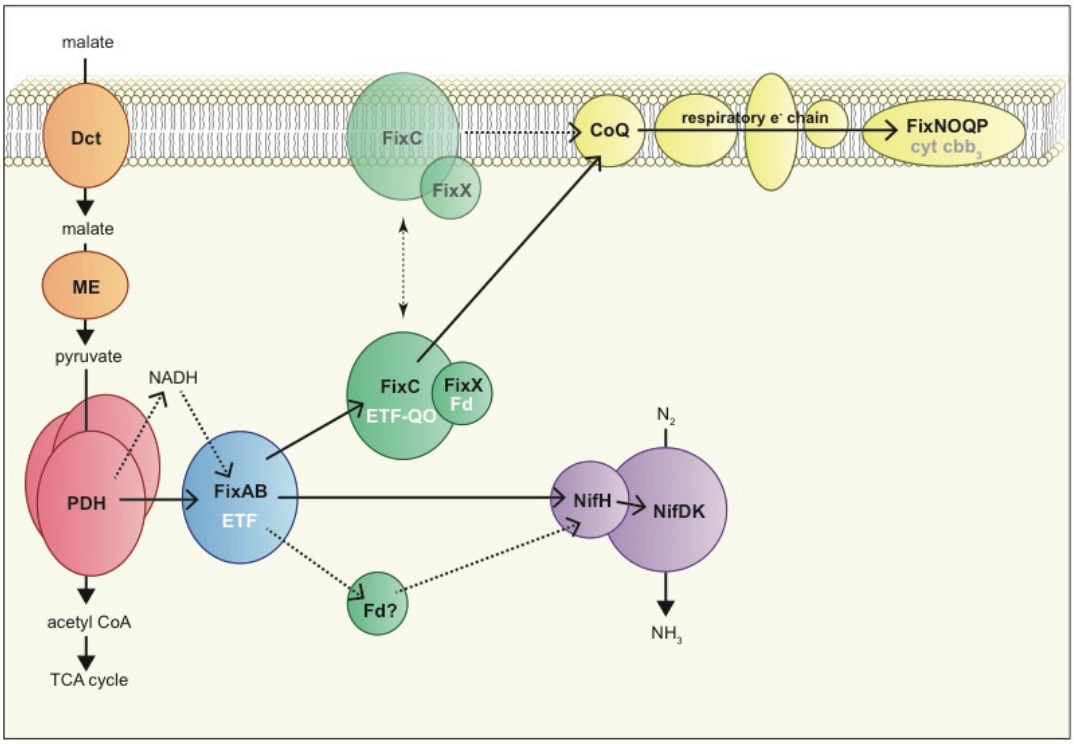


Figure 8.1 Proposed scheme for the role of FixABCX in symbiotic nitrogen fixation. Dashed lines represent hypothetical alternatives. C_4-dicarboxylates such as malate are transported into the bacteroid, where they are converted to pyruvate via malic enzyme (ME). The pyruvate dehydrogenase (PDH) complex then converts pyruvate to acetyl CoA which enters the TCA cycle. Electrons produced by PDH are then passed to FixAB, either directly or via NADH. Electrons are then passed from FixAB to FixCX. FixAB may be coupling the exergonic reduction of PDH to the endergonic reduction of a ferredoxin. FixCX may be located in the membrane as determined by homology with other ETF-QOs. Electrons from FixCX are then split between two routes. The first route is the quinone pool, where they are then passed along the respiratory electron transport chain to the terminal cbb_3 oxidase FixNOQP. The second route is towards nitrogenase. FixX may play the role of ferredoxin in both routes, or a second, currently unidentified ferredoxin, may be involved in this electron transfer. Not shown: the 2-oxoglutarate dehydrogenase complex may play the same role as pyruvate dehydrogenase in passing electrons to FixAB.

There are many important future experiments required to confirm these interactions. The first is to carry out reciprocal experiments using tagged Nif and Fix proteins and tagged PDH/ODH proteins. Unrelated proteins that showed no interaction with FixAB will provide a useful negative control to ensure that any identified proteins are not interacting with the FLAG-peptide rather than the tagged protein. Co-immunoprecipitation using complementary tags, or split-YFP experiments will provide *in vitro* data to support results of protein mass spectrometry. These experiments were carried out as a result of optimisation of the FLAG-tag system, however there are many other purification systems available, including His_6-tag, Twin-Strep® and GFP; complementary pull-downs using an alternative system will provide confidence in the data seen using FLAG-tag.

This FixAB-pyruvate/2-oxoglutarate dehydrogenase interaction may extend beyond *R. leguminosarum* and may be conserved across the rhizobia. Free-living diazotroph *A. caulinodans* requires pyruvate dehydrogenase activity for symbiotic nitrogen fixation, but not for free-living nitrogen fixation under partially-microaerobic conditions. In bacteroids flux through the TCA cycle is low (Borah, pers. comm.), and so 2-oxoglutarate dehydrogenase activity may be too low to compensate for a lack of pyruvate dehydrogenase. Under the free-living condition the flux through 2-oxoglutarate dehydrogenase may provide enough activity to energise nitrogen fixation via FixAB.

Investigating the *fix* genes in a free-living diazotroph: *A. caulinodans*

In addition to investigation of the Rlv3841 *fix* genes, FixAB mutants were also generated in the *A. caulinodans* ORS571 background. *A. caulinodans* ORS571 is able to fix nitrogen both in symbiosis with *S. rostrata* and in free-living microaerobic conditions. This makes ORS571 another useful model organism, as it can be investigated in multiple conditions, and work can be done with free-living culture rather than bacteroids. Mutants in *A. caulinodans* ORS571 *fixAB* were also unable to fix nitrogen under nitrogen- and oxygen-limited free-living conditions (Chapter 4), confirming their essential role in nitrogen fixation. The *A. caulinodans* mutants should be grown in symbiosis with its legume partner *S. rostrata* in order to fully confirm their requirement for nitrogen fixation. Attempts to complement an ORS571 mutant with *fixAB* from Rlv3841 failed possibly because they appeared toxic to ORS571 under nitrogen-fixing conditions. This expression was under a taurine-inducible promoter; a construct should be generated under the native ORS571 *fixA* promoter in order to determine whether overexpression effects led to toxicity. Furthermore, the native ORS571 *fixABCX* genes should be checked for complementation of the *A. caulinodans* mutants. Work in *A. caulinodans* was preliminary, and the mutants constructed will now be available for

complete complementation and phenotyping *in planta*. Pull-down experiments could also be carried out using ORS571; free-living culture will provide a much larger volume of material for assays, and data can be compared to the existing knowledge about pyruvate dehydrogenase functioning within nitrogen fixation.

Characterising the role of menaquinone biosynthesis in symbiotic nitrogen fixation

Work in *M. huakuii* suggested a role for menaquinone in bacteroid electron transfer, which may act in place of quinone as an acceptor from ETF-QO (FixCX). A putative menaquinone dimethyltransferase was identified by homology with a characterised protein from *M. huakuii*. Mutation of *dmtH* showed no symbiotic phenotype; suggesting menaquinone may not be involved in symbiotic electron transfer in pea bacteroids. Additional quinone and menaquinone biosynthesis genes could be investigated in order to confirm that menaquinone is not involved in this process. Investigation of bacteroids using newly emerging high-throughput mutant screening techniques, such as insertion sequencing (InSeq) (Perry and Yost 2014) may reveal more answers about essential electron transfer genes involved in nitrogen fixation.

Concluding remarks

In this work we have characterised the role of FixAB(CX) in symbiotic nitrogen fixation. The *fixABCX* genes are required for nitrogen fixation in both symbiotic and free-living diazotrophs. These genes form a larger operon consisting of *nifA,* the general transcriptional regulator of nitrogen fixation, and so their expression is tightly-linked to expression of nitrogen fixation. Removal of the *fixAB* gene products from a functioning symbiosis leads to huge changes in the redox level of the cell, requiring use of redox sinks such as PHB in order to maintain bacteroid function. Lack of nitrogen fixation ability leads to plant responses that feed back into the nodule to prevent further bacteroid differentiation; the plant won't provide input if it receives nothing in return. Finally, we propose a hypothesis for a supercomplex required to energise nitrogen fixation; FixAB interacts with key dehydrogenases in order to pass electrons to nitrogenase.

BIBLIOGRAPHY

Achbergerova, L. and J. Nahalka (2011). "Polyphosphate - an ancient energy source and active metabolic regulator." Microbial Cell Factories **10**.

Aebersold, R. and M. Mann (2003). "Mass spectrometry-based proteomics." Nature **422**(6928): 198-207.

Alcantara, C., A. Blasco, M. Zuniga and V. Monedero (2014). "Accumulation of polyphosphate in *Lactobacillus spp*. and Its involvement in stress resistance." Applied and Environmental Microbiology **80**(5): 1650-1659.

Allaway, D., E. M. Lodwig, L. A. Crompton, M. Wood, R. Parsons, T. R. Wheeler and P. S. Poole (2000). "Identification of alanine dehydrogenase and its role in mixed secretion of ammonium and alanine by pea bacteroids." Molecular Microbiology **36**(2): 508-515.

Alunni, B., Z. Kevei, M. Redondo-Nieto, A. Kondorosi, P. Mergaert and E. Kondorosi (2007). "Genomic organization and evolutionary insights on *GRP* and *NCR* genes, two large nodule-specific gene families in *Medicago truncatula*." Molecular Plant-Microbe Interactions **20**(9): 1138-1148.

Ampe, F., E. Kiss, F. Sabourdy and J. Batut (2003). "Transcriptome analysis of *Sinorhizobium meliloti* during symbiosis." Genome Biology **4**(2).

Anderson, A. J. and E. A. Dawes (1990). "Occurrence, metabolism, metabolic role, and industrial uses of bacterial polyhydroxyalkanoates." Microbiological Reviews **54**(4): 450-472.

Andrio, E., D. Marino, A. Marmeys, M. D. de Segonzac, I. Damiani, A. Genre, S. Huguet, P. Frendo, A. Puppo and N. Pauly (2013). "Hydrogen peroxide-regulated genes in the *Medicago truncatula*-*Sinorhizobium meliloti* symbiosis." New Phytologist **198**(1): 179-189.

Angermuller, S. and H. D. Fahimi (1982). "Imidazole-buffered osmium tetroxide: an excellent stain for visualization of lipids in transmission electron microscopy." Histochemical Journal **14**(5): 823-835.

Appleby, C. A. (1984). "Leghemoglobin and *Rhizobium* Respiration." Annual Reviews of Plant Physiology **35**: 443-478.

Ardissone, S., H. Kobayashi, K. Kambara, C. Rummel, K. D. Noel, G. C. Walker, W. J. Broughton and W. J. Deakin (2011). "Role of BacA in Lipopolysaccharide Synthesis, Peptide Transport, and Nodulation by *Rhizobium* sp Strain NGR234." Journal of Bacteriology **193**(9): 2218-2228.

Arigoni, F., P. A. Kaminski, H. Hennecke and C. Elmerich (1991). "Nucleotide sequence of the fixABC region of Azorhizobium caulinodans ORS571: similarity of the fixB product with

eukaryotic flavoproteins, characterization of fixX, and identification of nifW." Molecular & General Genetics **225**(3): 514-520.

Arigoni, F., P. A. Kaminski, H. Hennecke and C. Elmerich (1991). "Nucleotide-sequence of the *fixABC* region of *Azorhizobium caulinodans* ORS571 - similarity of the *fixB* product with eukaryotic flavoproteins, characterization of *fixX* and identification of *nifW*." Molecular & General Genetics **225**(3): 514-520.

Aslam, S. N., M. A. Newman, G. Erbs, K. L. Morrissey, D. Chinchilla, T. Boller, T. T. Jensen, C. De Castro, T. Ierano, A. Molinaro, R. W. Jackson, M. R. Knight and R. M. Cooper (2008). "Bacterial polysaccharides suppress induced innate immunity by calcium chelation." Current Biology **18**(14): 1078-1083.

Ault-Riche, D., C. D. Fraley, C. M. Tzeng and A. Kornberg (1998). "Novel assay reveals multiple pathways regulating stress-induced accumulations of inorganic polyphosphate in Escherichia coli." Journal of Bacteriology **180**(7): 1841-1847.

Bailey, T. L., M. Boden, F. A. Buske, M. Frith, C. E. Grant, L. Clementi, J. Y. Ren, W. W. Li and W. S. Noble (2009). "MEME SUITE: tools for motif discovery and searching." Nucleic Acids Research **37**: W202-W208.

Barney, B. M., H. I. Lee, P. C. Dos Santos, B. M. Hoffmann, D. R. Dean and L. C. Seefeldt (2006). "Breaking the N(2) triple bond: insights into the nitrogenase mechanism." Dalton Transactions(19): 2277-2284.

Bartlett, K. and S. Eaton (2004). "Mitochondrial beta-oxidation." European Journal of Biochemistry **271**(3): 462-469.

Bartlett, K. and M. Pourfarzam (1998). "Recent developments in the detection of inherited disorders of mitochondrial beta-oxidation." Biochemical Society Transactions **26**(2): 145-152.

Batut, J., M. L. Daveranmingot, M. D. J. Jacobs, A. M. Garnerone and D. Kahn (1989). "*fixK*, a gene homologous with *fnr* and *crp* from *Escherichia coli*, regulates nitrogen-fixation genes both positively and negatively in *Rhizobium meliloti*." EMBO Journal **8**(4): 1279-1286.

Batut, J., B. Terzaghi, M. Gherardi, M. Huguet, E. Terzaghi, A. M. Garnerone, P. Boistard and T. Huguet (1985). "Localization of a symbiotic fix region on Rhizobium meliloti pSYM megaplasmid more than 200 kilobases from the nod-nif region." Molecular & General Genetics **199**(2): 232-239.

Bauer, E., T. Kaspar, H. M. Fischer and H. Hennecke (1998). "Expression of the *fixR-nifA* operon in Bradyrhizobium japonicum depends on a new response regulator, RegR." Journal of Bacteriology **180**(15): 3853-3863.

Becana, M. and R. V. Klucas (1992). "Oxidation and Reduction of Leghemoglobin in Root-Nodules of Leguminous Plants." Plant Physiology **98**(4): 1217-1221.

Becker, A., H. Berges, E. Krol, C. Bruand, S. Ruberg, D. Capela, E. Lauber, E. Meilhoc, F. Ampe, F. J. de Bruijn, J. Fourment, A. Francez-Charlot, D. Kahn, H. Kuster, C. Liebe, A. Puhler,

S. Weidner and J. Batut (2004). "Global changes in gene expression in *Sinorhizobium meliloti* 1021 under microoxic and symbiotic conditions." Molecular Plant-Microbe Interactions **17**(3): 292-303.

Benedito, V. A., I. Torres-Jerez, J. D. Murray, A. Andriankaja, S. Allen, K. Kakar, M. Wandrey, J. Verdier, H. Zuber, T. Ott, S. Moreau, A. Niebel, T. Frickey, G. Weiller, J. He, X. Dai, P. X. Zhao, Y. Tang and M. K. Udvardi (2008). "A gene expression atlas of the model legume *Medicago truncatula*." Plant Journal **55**(3): 504-513.

Benhassine, T., M. Fauvart, J. Vanderleyden and J. Michiels (2007). "Interaction of an IHF-like protein with the Rhizobium etli nifA promoter." FEMS Microbiology Letters **271**(1): 20-26.

Berg, J. M. T., J.L.; Stryer L. (2002). Biochemistry. New York, W. H. Freeman.

Bergersen, F. J. and G. L. Turner (1990). "Bacteroids from soybean root-nodules - accumulation of poly-beta-hydroxybutyrate during supply of malate and succinate in relation to N2-fixation in flow-chamber reactions." Proceedings of the Royal Society of London B -Biological Sciences **240**(1297): 39-59.

Beringer, J. E. (1974). "R factor transfer in Rhizobium-leguminosarum." Journal of General Microbiology **84**(Sep): 188-198.

Beringer, J. E., A. W. B. Johnston and B. Wells (1977). "Isolation of conditional ineffective mutants of Rhizobium-leguminosarum." Journal of General Microbiology **98**(Feb): 339-343.

Berrow, N. S., D. Alderton, S. Sainsbury, J. Nettleship, R. Assenberg, N. Rahman, D. I. Stuart and R. J. Owens (2007). "A versatile ligation-independent cloning method suitable for high-throughput expression screening applications." Nucleic Acids Research **35**(6): e45.

Bertsch, J., A. Parthasarathy, W. Buckel and V. Muller (2013). "An Electron-bifurcating Caffeyl-CoA Reductase." Journal of Biological Chemistry **288**(16): 11304-11311.

Beynon, J., M. Cannon, V. Buchanan-Wollaston, A. Ally, R. Setterquist, D. Dean and F. Cannon (1988). "The nucleotide sequence of the *nifT, nifY, nifX* and *nifW* genes of *K. pneumoniae*." Nucleic Acids Research **16**(20): 9860.

Beynon, J. L., M. K. Williams and F. C. Cannon (1988). "Expression and functional analysis of the *Rhizobium meliloti nifA* gene." EMBO Journal **7**(1): 7-14.

Bhandari, B. and D. J. Nicholas (1985). "Proton motive force in washed cells of *Rhizobium japonicum* and bacteroids from *Glycine max*." Journal of Bacteriology **164**(3): 1383-1385.

Bhat, U. R. and R. W. Carlson (1992). "Chemical characterization of pH-dependent structural epitopes of lipopolysaccharides from *Rhizobium leguminosarum* biovar *phaseoli*." Journal of Bacteriology **174**(7): 2230-2235.

Bird, L. E. (2011). "High throughput construction and small scale expression screening of multi-tag vectors in Escherichia coli." Methods **55**(1): 29-37.

Bode, G., F. Mauch, H. Ditschuneit and P. Malfertheiner (1993). "Identification of structures containing polyphosphate in *Helicobacter pylori*." Journal of General Microbiology **139**: 3029-3033.

Boesten, B. and U. B. Priefer (2004). "The C-terminal receiver domain of the *Rhizobium leguminosarum* bv. *viciae* FixL protein is required for free-living microaerobic induction of the *fnrN* promoter." Microbiology **150**: 3703-3713.

Bonfante, P. and A. Genre (2010). "Mechanisms underlying beneficial plant-fungus interactions in mycorrhizal symbiosis." Nature Communications **1**: 48.

Bornhorst, J. A. and J. J. Falke (2000). "Purification of proteins using polyhistidine affinity tags." Methods in Enzymology **326**: 245-254.

Brandt, U. (1996). "Bifurcated ubihydroquinone oxidation in the cytochrome bc1 complex by proton-gated charge transfer." FEBS Letters **387**(1): 1-6.

Breakspear, A., C. Liu, S. Roy, N. Stacey, C. Rogers, M. Trick, G. Morieri, K. S. Mysore, J. Wen, G. E. Oldroyd, J. A. Downie and J. D. Murray (2014). "The root hair "infectome" of *Medicago truncatula* uncovers changes in cell cycle genes and reveals a requirement for Auxin signaling in rhizobial infection." Plant Cell **26**(12): 4680-4701.

Brewin, N. J. (2004). "Plant cell wall remodelling in the *Rhizobium*-legume symbiosis." Critical Reviews in Plant Sciences **23**(4): 293-316.

Brito, B., M. Martinez, D. Fernandez, L. Rey, E. Cabrera, J. M. Palacios, J. Imperial and T. RuizArgueso (1997). "Hydrogenase genes from *Rhizobium leguminosarum* bv *viciae* are controlled by the nitrogen fixation regulatory protein NifA." Proceedings of the National Academy of Sciences of the United States of America **94**(12): 6019-6024.

Brophy, J. A. and C. A. Voigt (2014). "Principles of genetic circuit design." Nature Methods **11**(5): 508-520.

Bruschi, M. and F. Guerlesquin (1988). "Structure, Function and Evolution of Bacterial Ferredoxins." Fems Microbiology Letters **54**(2): 155-175.

Brymora, A., V. A. Valova and P. J. Robinson (2004). "Protein-protein interactions identified by pull-down experiments and mass spectrometry." Curr Protoc Cell Biol **Chapter 17**: Unit 17 15.

Buchanan-Wollaston, V. (1979). "Generalized transduction in *Rhizobium leguminosarum*." Journal of General Microbiology **112**(May): 135-142.

Buckel, W. and R. K. Thauer (2013). "Energy conservation via electron bifurcating ferredoxin reduction and proton/Na(+) translocating ferredoxin oxidation." Biochimica et Biophysica Acta **1827**(2): 94-113.

Buikema, W. J., W. W. Szeto, P. V. Lemley, W. H. Orme-Johnson and F. M. Ausubel (1985). "Nitrogen fixation specific regulatory genes of Klebsiella pneumoniae and Rhizobium meliloti share homology with the general nitrogen regulatory gene ntrC of K. pneumoniae." Nucleic Acids Research **13**(12): 4539-4555.

Burgess, B. K. and D. J. Lowe (1996). "Mechanism of molybdenum nitrogenase." Chemical Reviews **96**(7): 2983-3012.

Cabanes, D., P. Boistard and J. Batut (2000). "Symbiotic induction of pyruvate dehydrogenase genes from Sinorhizobium meliloti." Molecular Plant-Microbe Interactions **13**(5): 483-493.

Cardenas, L., J. Dominguez, O. Santana and C. Quinto (1996). "The role of the *nodI* and *nodJ* genes in the transport of Nod metabolites in *Rhizobium etli*." Gene **173**(2): 183-187.

Casley-Smith, J. R. (1967). "Some observations on the electron microscopy of lipids." Journal of the Royal Microscopical Society **87**(3): 463-473.

Cebolla, A., J. M. Vinardell, E. Kiss, B. Olah, F. Roudier, A. Kondorosi and E. Kondorosi (1999). "The mitotic inhibitor *ccs52* is required for endoreduplication and ploidy-dependent cell enlargement in plants." EMBO Journal **18**(16): 4476-4484.

Cevallos, M. A., S. Encarnacion, A. Leija, Y. Mora and J. Mora (1996). "Genetic and physiological characterization of a *Rhizobium etli* mutant strain unable to synthesize poly-beta-hydroxybutyrate." Journal of Bacteriology **178**(6): 1646-1654.

Chan, J. M., J. Christiansen, D. R. Dean and L. C. Seefeldt (1999). "Spectroscopic evidence for changes in the redox state of the nitrogenase P-cluster during turnover." Biochemistry **38**(18): 5779-5785.

Chang, C., I. Damiani, A. Puppo and P. Frendo (2009). "Redox changes during the legume-Rhizobium symbiosis." Molecular Plant **2**(3): 370-377.

Chang, W. S., W. L. Franck, E. Cytryn, S. Jeong, T. Joshi, D. W. Emerich, M. J. Sadowsky, D. Xu and G. Stacey (2007). "An oligonucleotide microarray resource for transcriptional profiling of *Bradyrhizobium japonicum*." Molecular Plant-Microbe Interactions **20**(10): 1298-1307.

Charpentier, M. and G. E. Oldroyd (2013). "Nuclear calcium signaling in plants." Plant Physiology **163**(2): 496-503.

Cheng, Y. Y. and B. L. Sun (2009). "Polyphosphate kinase affects oxidative stress response by modulating cAMP receptor protein and rpoS expression in Salmonella typhimurium." Journal of Microbiology and Biotechnology **19**(12): 1527-1535.

Chowdhury, N. P., A. M. Mowafy, J. K. Demmer, V. Upadhyay, S. Koelzer, E. Jayamani, J. Kahnt, M. Hornung, U. Demmer, U. Ermler and W. Buckel (2014). "Studies on the mechanism of electron bifurcation catalyzed by electron transferring flavoprotein (Etf) and butyryl- CoA dehydrogenase (Bcd) of *Acidaminococcus fermentans*." Journal of Biological Chemistry **289**(8): 5145-5157.

Chun, Y. and Z. D. Yin (1998). "Glycogen assay for diagnosis of female genital Chlamydia trachomatis infection." Journal of Clinical Microbiology **36**(4): 1081-1082.

Claros, M. G. and G. Vonheijne (1994). "Toppred-Ii - an improved software for membrane-protein structure predictions." Computer Applications in the Biosciences **10**(6): 685-686.

Colebatch, G., G. Desbrosses, T. Ott, L. Krusell, O. Montanari, S. Kloska, J. Kopka and M. K. Udvardi (2004). "Global changes in transcription orchestrate metabolic differentiation during symbiotic nitrogen fixation in *Lotus japonicus*." Plant Journal **39**(4): 487-512.

Colombo, M. V., D. Gutierrez, J. M. Palacios, J. Imperial and T. Ruiz-Argueso (2000). "A novel autoregulation mechanism of *fnrN* expression in *Rhizobium leguminosarum* bv *viciae*." Molecular Microbiology **36**(2): 477-486.

Colonna-Romano, S., W. Arnold, A. Schluter, P. Boistard, A. Puhler and U. B. Priefer (1990). "An Fnr-like protein encoded in *Rhizobium leguminosarum* biovar *viciae* shows structural and functional homology to *Rhizobium meliloti* FixK." Molecular & General Genetics **223**(1): 138-147.

Craig, A. S. and K. I. Williamson (1972). "Three inclusions of rhizobial bacteroids and their cytochemical character." Archiv Fur Mikrobiologie **87**(2): 165-172.

Crane, F. L. and H. Beinert (1956). "Mechanism of dehydrogenation of fatty acyl derivatives of coenzyme-A .2. Electron-transferring flavoprotein." Journal of Biological Chemistry **218**(2): 717-731.

Crane, F. L. and H. Beinert (1956). "On the mechanism of dehydrogenation of fatty acyl derivatives of coenzyme A. II. The electron-transferring flavoprotein." Journal of Biological Chemistry **218**(2): 717-731.

Cserzo, M., E. Wallin, I. Simon, G. vonHeijne and A. Elofsson (1997). "Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method." Protein Engineering **10**(6): 673-676.

Curatti, L. and L. M. Rubio (2014). "Challenges to develop nitrogen-fixing cereals by direct nif-gene transfer." Plant Science **225**: 130-137.

Czernic, P., D. Gully, F. Cartieaux, L. Moulin, I. Guefrachi, D. Patrel, O. Pierre, J. Fardoux, C. Chaintreuil, P. Nguyen, F. Gressent, C. Da Silva, J. Poulain, P. Wincker, V. Rofidal, S. Hem, Q. Barriere, J. F. Arrighi, P. Mergaert and E. Giraud (2015). "Convergent evolution of endosymbiont differentiation in dalbergioid and inverted repeat-lacking clade legumes mediated by nodule-specific cysteine-rich peptides." Plant Physiology **169**(2): 1254-1265.

Day, D. A. and L. Copeland (1991). "Carbon metabolism and compartmentation in nitrogen-fixing legume nodules." Plant Physiology and Biochemistry **29**(2): 185-201.

de Jong, A., H. Pietersma, M. Cordes, O. P. Kuipers and J. Kok (2012). "PePPER: a webserver for prediction of prokaryote promoter elements and regulons." BMC Genomics **13**: 299.

de Kok, A., A. F. Hengeveld, A. Martin and A. H. Westphal (1998). "The pyruvate dehydrogenase multi-enzyme complex from Gram-negative bacteria." Biochimica Et Biophysica Acta **1385**(2): 353-366.

de Souza, E. M., C. E. Granada and R. A. Sperotto (2016). "Plant pathogens affecting the establishment of plant-symbiont interaction." Frontiers in Plant Science **7**.

Deistung, J., F. C. Cannon, M. C. Cannon, S. Hill and R. N. Thorneley (1985). "Electron transfer to nitrogenase in *Klebsiella pneumoniae. nifF* gene cloned and the gene product, a flavodoxin, purified." Biochemical Journal **231**(3): 743-753.

Delgado, M. J., E. J. Bedmar and J. A. Downie (1998). "Genes involved in the formation and assembly of rhizobial cytochromes and their role in symbiotic nitrogen fixation." Advances in Microbial Physiology **40**: 191-231.

Delmotte, N., S. Mondy, B. Alunni, J. Fardoux, C. Chaintreuil, J. A. Vorholt, E. Giraud and B. Gourion (2014). "A proteomic approach of *Bradyrhizobium*/*Aeschynomene* root and stem symbioses reveals the importance of the *fixA* locus for symbiosis." International Journal of Molecular Sciences **15**(3): 3660-3670.

Denison, R. F. (2000). "Legume sanctions and the evolution of symbiotic cooperation by rhizobia." American Naturalist **156**(6): 567-576.

Ditta, G., S. Stanfield, D. Corbin and D. R. Helinski (1980). "Broad host range DNA cloning system for gram-negative bacteria - construction of a gene bank of *Rhizobium meliloti*." Proceedings of the National Academy of Sciences of the United States of America **77**(12): 7347-7351.

Dixon, R. (1998). "The oxygen-responsive NifL-NifA complex: a novel two-component regulatory system controlling nitrogenase synthesis in gamma-proteobacteria." Archives of Microbiology **169**(5): 371-380.

Dixon, R., R. R. Eady, G. Espin, S. Hill, M. Iaccarino, D. Kahn and M. Merrick (1980). "Analysis of regulation of *Klebsiella pneumoniae* nitrogen fixation (*nif*) gene cluster with gene fusions." Nature **286**(5769): 128-132.

Dixon, R. and D. Kahn (2004). "Genetic regulation of biological nitrogen fixation." Nature Reviews Microbiology **2**(8): 621-631.

Dixon, R. O. (1967). "The origin of the membrane envelope surrounding the bacteria and bacteroids and the presence of glycogen in clover root nodules." Archiv Fur Mikrobiologie **56**(2): 156-&.

Djordjevic, M. A. (2004). "*Sinorhizobium meliloti* metabolism in the root nodule: a proteomic perspective." Proteomics **4**(7): 1859-1872.

Domanski, M., K. Molloy, H. Jiang, B. T. Chait, M. P. Rout, T. H. Jensen and J. LaCava (2012). "Improved methodology for the affinity isolation of human protein complexes expressed at near endogenous levels." Biotechniques **0**(0): 1-6.

Downie, J. A. (2005). "Legume haemoglobins: symbiotic nitrogen fixation needs bloody nodules." Current Biology **15**(6): R196-198.

Dreyfus, B. L. and Y. R. Dommergues (1981). "Nitrogen-fixing nodules induced by Rhizobium on the stem of the tropical legume Sesbania-rostrata." FEMS Microbiology Letters **10**(4): 313-317.

Dreyfus, B. L., C. Elmerich and Y. R. Dommergues (1983). "Free-living *Rhizobium* strain able to grow on N-2 as the sole nitrogen-source." Applied and Environmental Microbiology **45**(2): 711-713.

Driscoll, B. T. and T. M. Finan (1993). "Nad+-dependent malic enzyme of Rhizobium meliloti is required for symbiotic nitrogen-fixation." Molecular Microbiology **7**(6): 865-873.

Driscoll, B. T. and T. M. Finan (1997). "Properties of NAD(+)- and NADP(+)-dependent malic enzymes of *Rhizobium* (*Sinorhizobium*) *meliloti* and differential expression of their genes in nitrogen-fixing bacteroids." Microbiology **143**: 489-498.

Dusha, I., S. Kovalenko, Z. Banfalvi and A. Kondorosi (1987). "Rhizobium meliloti insertion element ISRm2 and its ise for identification of the fixX gene." Journal of Bacteriology **169**(4): 1403-1409.

Earl, C. D., C. W. Ronson and F. M. Ausubel (1987). "Genetic and structural analysis of the *Rhizobium meliloti fixA, fixB, fixC* and *fixX* genes." Journal of Bacteriology **169**(3): 1127-1136.

Edgren, T. and S. Nordlund (2004). "The *fixABCX* genes in *Rhodospirillum rubrum* encode a putative membrane complex participating in electron transfer to nitrogenase." Journal of Bacteriology **186**(7): 2052-2060.

Edgren, T. and S. Nordlund (2005). "Electron transport to nitrogenase in *Rhodospirillum rubrum*: Identification of a new *fdxN* gene encoding the primary electron donor to nitrogenase." FEMS Microbiology Letters **245**(2): 345-351.

Edgren, T. and S. Nordlund (2006). "Two pathways of electron transport to nitrogenase in Rhodospirillum rubrum: the major pathway is dependent on the fix gene products." FEMS Microbiology Letters **260**(1): 30-35.

Eichler, K., A. Buchet, F. Bourgis, H. P. Kleber and M. A. Mandrandberthelot (1995). "The *fix Escherichia coli* region contains four genes related to carnitine metabolism." Journal of Basic Microbiology **35**(4): 217-227.

Elmerich, C., B. Dreyfus and J. P. Aubert (1983). "Nicotinic acid requirement and degradation by Sesbania-Rhizobium strain ORS571." FEMS Microbiology Letters **19**(2-3): 281-284.

Esseling, J. J., F. G. Lhuissier and A. M. Emons (2003). "Nod factor-induced root hair curling: continuous polar growth towards the point of nod factor application." Plant Physiology **132**(4): 1982-1988.

Fellay, R., J. Frey and H. Krisch (1987). "Interposon mutagenesis of soil and water bacteria - a family of DNA fagments designed for in vitro insertional mutagenesis of gram-negative bacteria." Gene **52**(2-3): 147-154.

Fernandez-Lopez, M., W. D'Haeze, P. Mergaert, C. Verplancke, J. C. Prome, M. Van Montagu and M. Holsters (1996). "Role of *nodl* and *nodJ* in lipo-chitooligosaccharide secretion in *Azorhizobium caulinodans* and *Escherichia coli*." Molecular Microbiology **20**(5): 993-1000.

Ferrieres, L. and D. Kahn (2002). "Two distinct classes of FixJ binding sites defined by in vitro selection." FEBS Letters **517**(1-3): 185-189.

Finan, T., A. Hirsch, J. Leigh, E. Johansen, G. Kuldeu, S. Deegan, G. Walker and E. Signer (1985). "Symbiotic mutants of *Rhizobium meliloti* that uncouple plant from bacterial differentiation." Cell **40**(4): 869-877.

Finan, T. M., E. Mcwhinnie, B. Driscoll and R. J. Watson (1991). "Complex symbiotic phenotypes result from gluconeogenic mutations in *Rhizobium meliloti*." Molecular Plant-Microbe Interactions **4**(4): 386-392.

Fischer, H. M. (1994). "Genetic regulation of nitrogen fixation in rhizobia." Microbiological Reviews **58**(3): 352-386.

Fischer, H. M., A. Alvarezmorales and H. Hennecke (1986). "The pleiotropic nature of symbiotic regulatory mutants - *Bradyrhizobium japonicum nifA* gene is involved in control of *nif* gene expression and formation of determinate symbiosis." EMBO Journal **5**(6): 1165-1173.

Fleischman, D. and D. Kramer (1998). "Photosynthetic rhizobia." Biochimica Et Biophysica Acta - Bioenergetics **1364**(1): 17-36.

Fournier, J., A. C. Timmers, B. J. Sieberer, A. Jauneau, M. Chabaud and D. G. Barker (2008). "Mechanism of infection thread elongation in root hairs of *Medicago truncatula* and dynamic interplay with associated rhizobial colonization." Plant Physiology **148**(4): 1985-1995.

Frederix, M., A. Edwards, A. Swiderska, A. Stanger, R. Karunakaran, A. Williams, P. Abbruscato, M. Sanchez-Contreras, P. S. Poole and J. A. Downie (2014). "Mutation of praR in Rhizobium leguminosarum enhances root biofilms, improving nodulation competitiveness by increased expression of attachment proteins." Molecular Microbiology **93**(3): 464-478.

Frugier, F., S. Kosuta, J. D. Murray, M. Crespi and K. Szczyglowski (2008). "Cytokinin: secret agent of symbiosis." Trends in Plant Science **13**(3): 115-120.

Fu, W. G., R. F. Jack, T. V. Morgan, D. R. Dean and M. K. Johnson (1994). "NifU gene product from *Azotobacter vinelandii* Is a homodimer that contains 2 identical [2fe-2s] clusters." Biochemistry **33**(45): 13455-13463.

Gage, D. J. (2002). "Analysis of infection thread development using Gfp- and DsRed-expressing *Sinorhizobium meliloti*." Journal of Bacteriology **184**(24): 7042-7046.

Gage, D. J. (2004). "Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes." Microbiology and Molecular Biology Reviews **68**(2): 280-300.

Gardy, J. L. and F. S. L. Brinkman (2006). "Methods for predicting bacterial protein subcellular localization." Nature Reviews Microbiology **4**(10): 741-751.

Gautheret, D. and A. Lambert (2001). "Direct RNA motif definition and identification from multiple sequence alignments using secondary structure profiles." Journal of Molecular Biology **313**(5): 1003-1011.

Gerace, E. and D. Moazed (2015). Chapter Seven - Affinity Pull-Down of Proteins Using Anti-FLAG M2 Agarose Beads. Methods in Enzymology. R. L. Jon, Academic Press. **Volume 559:** 99-110.

Gerson, T., J. J. Patel and M. N. Wong (1978). "Effects of age, darkness and nitrate on poly-beta-hydroxybutyrate levels and nitrogen-fixing ability of *Rhizobium* in *Lupinus angustifolius*." Physiologia Plantarum **42**(4): 420-424.

Geurts, R., E. Fedorova and T. Bisseling (2005). "Nod factor signaling genes and their function in the early stages of *Rhizobium* infection." Current Opinion in Plant Biology **8**(4): 346-352.

Glazebrook, J., A. Ichige and G. C. Walker (1993). "A *Rhizobium meliloti* homolog of the *Escherichia coli* peptide-antibiotic transport protein SbmA is essential for bacteroid development." Genes & Development **7**(8): 1485-1497.

Glenn, A. R., P. S. Poole and J. F. Hudman (1980). "Succinate uptake by free-living and bacteroid forms of *Rhizobium leguminosarum*." Journal of General Microbiology **119**(Jul): 267-271.

Gordon, G. B., L. R. Miller and K. G. Bensch (1963). "Fixation of tissue culture cells for ultrastructural cytochemistry." Experimental Cell Research **31**(2): 440-&.

Gourion, B., F. Berrabah, P. Ratet and G. Stacey (2015). "*Rhizobium*-legume symbioses: the crucial role of plant immunity." Trends in Plant Science **20**(3): 186-194.

Graham, M. A., K. A. Silverstein, S. B. Cannon and K. A. VandenBosch (2004). "Computational identification and characterization of novel genes from legumes." Plant Physiology **135**(3): 1179-1197.

Granqvist, E., J. Sun, R. Op den Camp, P. Pujic, L. Hill, P. Normand, R. J. Morris, J. A. Downie, R. Geurts and G. E. Oldroyd (2015). "Bacterial-induced calcium oscillations are common to nitrogen-fixing associations of nodulating legumes and non-legumes." New Phytologist **207**(3): 551-558.

Gray, M. J., W. Y. Wholey, N. O. Wagner, C. M. Cremers, A. Mueller-Schickert, N. T. Hock, A. G. Krieger, E. M. Smith, R. A. Bender, J. C. A. Bardwell and U. Jakob (2014). "Polyphosphate Is a primordial chaperone." Molecular Cell **53**(5): 689-699.

Gronger, P., S. S. Manian, H. Reilander, M. O'Connell, U. B. Priefer and A. Puhler (1987). "Organization and partial sequence of a DNA region of the Rhizobium leguminosarum symbiotic plasmid pRL6JI containing the genes fixABC, nifA, nifB and a novel open reading frame." Nucleic Acids Research **15**(1): 31-49.

Gruber, N. and J. N. Galloway (2008). "An Earth-system perspective of the global nitrogen cycle." Nature **451**(7176): 293-296.

Gubler, M. and H. Hennecke (1986). "*fixA, B* and *C* genes are essential for symbiotic and free-living, microaerobic nitrogen-fixation." FEBS Letters **200**(1): 186-192.

Gubler, M., T. Zurcher and H. Hennecke (1989). "The *Bradyrhizobium japonicum fixBCX* operon - identification of *fixX* and of a 5' messenger RNA region affecting the level of the *fixBCX* transcript." Molecular Microbiology **3**(2): 141-148.

Guefrachi, I., M. Nagymihaly, C. I. Pislariu, W. Van de Velde, P. Ratet, M. Mars, M. K. Udvardi, E. Kondorosi, P. Mergaert and B. Alunni (2014). "Extreme specificity of NCR gene expression in *Medicago truncatula*." BMC Genomics **15**: 712.

Gutierrez, D., Y. Hernando, J. M. Palacios, J. Imperial and T. Ruiz-Argueso (1997). "FnrN controls symbiotic nitrogen fixation and hydrogenase activities in *Rhizobium leguminosarum* biovar *viciae* UPM791." Journal of Bacteriology **179**(17): 5264-5270.

Haag, A. F., M. Baloban, M. Sani, B. Kerscher, O. Pierre, A. Farkas, R. Longhi, E. Boncompagni, D. Herouart, S. Dall'angelo, E. Kondorosi, M. Zanda, P. Mergaert and G. P. Ferguson (2011). "Protection of *Sinorhizobium* against host cysteine-rich antimicrobial peptides is critical for symbiosis." PLoS Biology **9**(10): e1001169.

Hageman, R. and R. H. Burris (1978). "Electron allocation to alternative substrates of *Azotobacter* nitrogenase is controlled by the electron flux through dinitrogenase." Biochimica et Biophysica Acta **37**(6): 1420-1420.

Hakoyama, T., K. Niimi, H. Watanabe, R. Tabata, J. Matsubara, S. Sato, Y. Nakamura, S. Tabata, J. C. Li, T. Matsumoto, K. Tatsumi, M. Nomura, S. Tajima, M. Ishizaka, K. Yano, H. Imaizumi-Anraku, M. Kawaguchi, H. Kouchi and N. Suganuma (2009). "Host plant genome overcomes the lack of a bacterial gene for symbiotic nitrogen fixation." Nature **462**(7272): 514-U210.

Han, X. M., A. Aslanian and J. R. Yates (2008). "Mass spectrometry for proteomics." Current Opinion in Chemical Biology **12**(5): 483-490.

Hara, A. and N. S. Radin (1978). "Lipid extraction of tissues with a low-toxicity solvent." Analytical Biochemistry **90**(1): 420-426.

Hartmann, A. and R. H. Burris (1987). "Regulation of nitrogenase activity by oxygen in *Azospirillum brasilense* and *Azospirillum lipoferum*." Journal of Bacteriology **169**(3): 944-948.

Hauser, F., G. Pessi, M. Friberg, C. Weber, N. Rusca, A. Lindemann, H. M. Fischer and H. Hennecke (2007). "Dissection of the *Bradyrhizobium japonicum* NifA+sigma(54) regulon, and identification of a ferredoxin gene (*fdxN*) for symbiotic nitrogen fixation." Molecular Genetics and Genomics **278**(3): 255-271.

Heidstra, R., W. C. Yang, Y. Yalcin, S. Peck, A. M. Emons, A. vanKammen and T. Bisseling (1997). "Ethylene provides positional information on cortical cell division but is not involved in Nod factor-induced root hair tip growth in *Rhizobium*-legume interaction." Development **124**(9): 1781-1787.

Heinrich, K., M. H. Ryder and P. J. Murphy (2001). "Early production of rhizopine in nodules induced by Sinorhizobium meliloti strain L5-30." Canadian Journal of Microbiology **47**(2): 165-171.

Hermelink, A., A. Brauer, P. Lasch and D. Naumann (2009). "Phenotypic heterogeneity within microbial populations at the single-cell level investigated by confocal Raman microspectroscopy." Analyst **134**(6): 1149-1153.
Herridge, D. F., M. B. Peoples and R. M. Boddey (2008). "Global inputs of biological nitrogen fixation in agricultural systems." Plant Soil **311**(1-2): 1-18.
Herrmann, G., E. Jayamani, G. Mai and W. Buckel (2008). "Energy conservation via electron-transferring flavoprotein in anaerobic bacteria." Journal of Bacteriology **190**(3): 784-791.
Hertig, C., R. Y. Li, A. M. Louarn, A. M. Garnerone, M. David, J. Batut, D. Kahn and P. Boistard (1989). "Rhizobium meliloti regulatory gene FixJ activates transcription of R. meliloti nifA and fixK genes in Escherichia coli." Journal of Bacteriology **171**(3): 1736-1738.
Hill, S. and E. P. Kavanagh (1980). "Roles of *nifF* and *nifJ* gene products in electron-transport to nitrogenase in *Klebsiella pneumoniae*." Journal of Bacteriology **141**(2): 470-475.
Hinnemann, B. and J. K. Norskov (2004). "Structure of the FeFe-cofactor of the iron-only nitrogenase and possible mechanism for dinitrogen reduction." Physical Chemistry Chemical Physics **6**(4): 843-853.
Hirsch, A. M., M. Bang and F. M. Ausubel (1983). "Ultrastructural analysis of ineffective alfalfa nodules formed by *ni*::Tn5 mutants of *Rhizobium meliloti*." Journal of Bacteriology **155**(1): 367-380.
Hirsch, A. M. and C. A. Smith (1987). "Effects of *Rhizobium meliloti nif* and *fix* mutants on alfalfa root nodule development." Journal of Bacteriology **169**(3): 1137-1146.
Hoffman, B. M., D. R. Dean and L. C. Seefeldt (2009). "Climbing nitrogenase: toward a mechanism of enzymatic nitrogen fixation." Accounts of Chemical Research **42**(5): 609-619.
Hood, G. A. (2013). Physiological response of *Rhizobium leguminosarum* during bacteroid development, University of East Anglia.
Hoover, T. R., J. Imperial, P. W. Ludden and V. K. Shah (1988). "Homocitrate cures the NifV phenotype in *Klebsiella pneumoniae*." Journal of Bacteriology **170**(4): 1978-1979.
Hoover, T. R., J. Imperial, P. W. Ludden and V. K. Shah (1989). "Homocitrate is a component of the iron molybdenum cofactor of nitrogenase." Biochemistry **28**(7): 2768-2771.
Hosie, A. H., D. Allaway, C. S. Galloway, H. A. Dunsby and P. S. Poole (2002). "*Rhizobium leguminosarum* has a second general amino acid permease with unusually broad substrate specificity and high similarity to branched-chain amino acid transporters (Bra/LIV) of the ABC family." Journal of Bacteriology **184**(15): 4071-4080.
Howard, J. B. and D. C. Rees (1994). "Nitrogenase: A nucleotide-dependent molecular switch." Annual Review of Biochemistry **63**: 235-264.
Hu, Y., A. W. Fay, C. C. Lee and M. W. Ribbe (2007). "P-cluster maturation on nitrogenase MoFe protein." Proceedings of the National Academy of Sciences of the United States of America **104**(25): 10424-10429.

Hu, Y., C. C. Lee and M. W. Ribbe (2012). "Vanadium nitrogenase: a two-hit wonder?" Dalton Transactions **41**(4): 1118-1127.

Hu, Y. L. and M. W. Ribbe (2011). "Biosynthesis of the metalloclusters of molybdenum nitrogenase." Microbiology and Molecular Biology Reviews **75**(4): 664-+.

Hua, S. J. and Z. R. Sun (2001). "Support vector machine approach for protein subcellular localization prediction." Bioinformatics **17**(8): 721-728.

Huala, E. and F. M. Ausubel (1989). "The central domain of *Rhizobium meliloti* NifA is sufficient to activate transcription from the *R. meliloti* nifH promoter." Journal of Bacteriology **171**(6): 3354-3365.

Huang, W. E., R. I. Griffiths, I. P. Thompson, M. J. Bailey and A. S. Whiteley (2004). "Raman microscopic analysis of single microbial cells." Analytical Chemistry **76**(15): 4452-4458.

Ichige, A. and G. C. Walker (1997). "Genetic analysis of the *Rhizobium meliloti bacA* gene: Functional interchangeability with the *Escherichia coli sbmA* gene and phenotypes of mutants." Journal of Bacteriology **179**(1): 209-216.

Igarashi, R. Y. and L. C. Seefeldt (2003). "Nitrogen fixation: The mechanism of the Mo-dependent nitrogenase." Critical Reviews in Biochemistry and Molecular Biology **38**(4): 351-384.

Iismaa, S. E. and J. M. Watson (1987). "A gene upstream of the *Rhizobium trifolii nifA* gene encodes a ferredoxin-like protein." Nucleic Acids Research **15**(7): 3180-3180.

Jack, R., M. De Zamaroczy and M. Merrick (1999). "The signal transduction protein GlnK is required for NifL-dependent nitrogen control of nif gene expression in Klebsiella pneumoniae." Journal of Bacteriology **181**(4): 1156-1162.

Jamet, A., K. Mandon, A. Puppo and D. Herouart (2007). "H2O2 is required for optimal establishment of the *Medicago sativa*/*Sinorhizobium meliloti* symbiosis." Journal of Bacteriology **189**(23): 8741-8745.

Jensen, T. E. (1968). "Electron microscopy of polyphosphate bodies in a blue-green alga *Nostoc pruniforme*." Archiv Fur Mikrobiologie **62**(2): 144-&.

Jervis, A. J. and J. Green (2007). "In vivo demonstration of FNR dimers in response to lower O-2 availability." Journal of Bacteriology **189**(7): 2930-2932.

Johnston, A. W. B. and J. E. Beringer (1975). "Identification of *Rhizobium* strains in pea root nodules using genetic markers." Journal of General Microbiology **87**(APR): 343-350.

Jones, K. M., H. Kobayashi, B. W. Davies, M. E. Taga and G. C. Walker (2007). "How rhizobial symbionts invade plants: the Sinorhizobium-Medicago model." Nature Reviews Microbiology **5**(8): 619-633.

Jones, K. M. and G. C. Walker (2008). "Responses of the model legume *Medicago truncatula* to the rhizobial exopolysaccharide succinoglycan." Plant Signalling & Behavour **3**(10): 888-890.

Jording, D., P. K. Sharma, R. Schmidt, T. Engelke, C. Uhde and A. Puhler (1993). "Regulatory aspects of the C4-dicarboxylate transport in *Rhizobium meliloti* - transcriptional activation and dependence on effective symbiosis." Journal of Plant Physiology **141**(1): 18-27.

Juven-Gershon, T. and J. T. Kadonaga (2010). "Regulation of gene expression via the core promoter and the basal transcriptional machinery." Developmental Biology **339**(2): 225-229.

Kadouri, D., E. Jurkevitch and Y. Okon (2003). "Poly beta-hydroxybutyrate depolymerase (PhaZ) in *Azospirillum brasilense* and characterization of a *phaZ* mutant." Archives of Microbiology **180**(5): 309-318.

Kahn, D., M. David, O. Domergue, M. L. Daveran, J. Ghai, P. R. Hirsch and J. Batut (1989). "*Rhizobium meliloti fixGHI* sequence predicts involvement of a specific cation pump in symbiotic nitrogen-fixation." Journal of Bacteriology **171**(2): 929-939.

Kaminski, P. A., F. Norel, N. Desnoues, A. Kush, G. Salzano and C. Elmerich (1988). "Chracterization of the *fixABC* region of *Azorhizobium caulinodans* ORS571 and identification of a new nitrogen fixation gene." Molecular & General Genetics **214**(3): 496-502.

Kamst, E., J. Pilling, L. M. Raamsdonk, B. J. Lugtenberg and H. P. Spaink (1997). "*Rhizobium* nodulation protein NodC is an important determinant of chitin oligosaccharide chain length in Nod factor biosynthesis." Journal of Bacteriology **179**(7): 2103-2108.

Karimova, G., J. Pidoux, A. Ullmann and D. Ladant (1998). "A bacterial two-hybrid system based on a reconstituted signal transduction pathway." Proceedings of the National Academy of Sciences of the United States of America **95**(10): 5752-5756.

Karr, D. B., J. K. Waters and D. W. Emerich (1983). "Analysis of poly-beta-hydroxybutyrate in *Rhizobium japonicum* bacteroids by ion-exclusion high-pressure liquid-chromatography and UV detection." Applied and Environmental Microbiology **46**(6): 1339-1344.

Karunakaran, R., A. F. Haag, A. K. East, V. K. Ramachandran, J. Prell, E. K. James, M. Scocchi, G. P. Ferguson and P. S. Poole (2010). "BacA is essential for bacteroid development in nodules of galegoid, but not phaseoloid, legumes." Journal of Bacteriology **192**(11): 2920-2928.

Karunakaran, R., V. K. Ramachandran, J. C. Seaman, A. K. East, B. Mouhsine, T. H. Mauchline, J. Prell, A. Skeffington and P. S. Poole (2009). "Transcriptomic analysis of *Rhizobium leguminosarum* biovar *viciae* in symbiosis with host plants *Pisum sativum* and *Vicia cracca*." Journal of Bacteriology **191**(12): 4002-4014.

Kereszt, A., P. Mergaert and E. Kondorosi (2011). "Bacteroid development in legume nodules: evolution of mutual benefit or of sacrificial victims?" Molecular Plant-Microbe Interactions **24**(11): 1300-1309.

Kim, C. H., D. R. Helinski and G. Ditta (1986). "Overlapping transcription of the *nifA* regulatory gene in *Rhizobium meliloti*." Gene **50**(1-3): 141-148.

Kitts, C. L., J. P. Lapointe, V. T. Lam and R. A. Ludwig (1992). "Elucidation of the complete *Azorhizobium* nicotinate catabolism pathway." Journal of Bacteriology **174**(23): 7791-7797.

Klipp, W., H. Reilander, A. Schluter, R. Krey and A. Puhler (1989). "The *Rhizobium meliloti fdxN* gene encoding a ferredoxin-like protein is necessary for nitrogen fixation and is cotranscribed with *nifA* and *nifB*." Molecular & General Genetics **216**(2-3): 293-302.

Kondorosi, E., F. Roudier and E. Gendreau (2000). "Plant cell-size control: growing by ploidy?" Current Opinion in Plant Biology **3**(6): 488-492.

Kornberg, A., N. N. Rao and D. Ault-Riche (1999). "Inorganic polyphosphate: A molecule of many functions." Annual Review of Biochemistry **68**: 89-125.

Korner, H., H. J. Sofia and W. G. Zumft (2003). "Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: exploiting the metabolic spectrum by controlling alternative gene programs." FEMS Microbiology Reviews **27**(5): 559-592.

Kouchi, H., K. Shimomura, S. Hata, A. Hirota, G. J. Wu, H. Kumagai, S. Tajima, N. Suganuma, A. Suzuki, T. Aoki, M. Hayashi, T. Yokoyama, T. Ohyama, E. Asamizu, C. Kuwata, D. Shibata and S. Tabata (2004). "Large-scale analysis of gene expression profiles during early stages of root nodule formation in a model legume, *Lotus japonicus*." DNA Research **11**(4): 263-274.

Kulakova, A. N., D. Hobbs, M. Smithen, E. Pavlov, J. A. Gilbert, J. P. Quinn and J. W. McGrath (2011). "Direct quantification of inorganic polyphosphate in microbial cells using 4 '-6-diamidino-2-phenylindole (DAPI)." Environmental Science & Technology **45**(18): 7799-7803.

Kuster, H., M. F. Vieweg, K. Manthey, M. C. Baier, N. Hohnjec and A. M. Perlick (2007). "Identification and expression regulation of symbiotically activated legume genes." Phytochemistry **68**(1): 8-18.

Kuzma, M. M., S. Hunt and D. B. Layzell (1993). "Role of oxygen in the limitation and inhibition of nitrogenase activity and respiration rate in individual soybean nodules." Plant Physiology **101**(1): 161-169.

Labes, M., V. Rastogi, R. Watson and T. M. Finan (1993). "Symbiotic nitrogen fixation by a *nifA* deletion mutant of *Rhizobium meliloti* - the role of an unusual *ntrC* allele." Journal of Bacteriology **175**(9): 2662-2673.

Lang, C. and S. R. Long (2015). "Transcriptomic analysis of *Sinorhizobium meliloti* and *Medicago truncatula* symbiosis using nitrogen fixation deficient nodules." Molecular Plant-Microbe Interactions **28**(8): 856-868.

Lee, K. B., P. De Backer, T. Aono, C. T. Liu, S. Suzuki, T. Suzuki, T. Kaneko, M. Yamada, S. Tabata, D. M. Kupfer, F. Z. Najar, G. B. Wiley, B. Roe, T. T. Binnewies, D. W. Ussery, W. D'Haeze, J. D. Herder, D. Gevers, D. Vereecke, M. Holsters and H. Oyaizu (2008). "The genome of the versatile nitrogen fixer *Azorhizobium caulinodans* ORS571." BMC Genomics **9**: 271.

Lhuissier, F. G. P., N. C. A. De Ruijter, B. J. Sieberer, J. J. Esseling and A. M. C. Emons (2001). "Time course of cell biological events evoked in legume root hairs by *Rhizobium* Nod factors: State of the art." Annals of Botany **87**(3): 289-302.

Li, F., J. Hinderberger, H. Seedorf, J. Zhang, W. Buckel and R. K. Thauer (2008). "Coupled ferredoxin and crotonyl coenzyme A (CoA) reduction with NADH catalyzed by the butyryl-CoA dehydrogenase/Etf complex from *Clostridium kluyveri*." Journal of Bacteriology **190**(3): 843-850.

Libault, M., A. Farmer, L. Brechenmacher, J. Drnevich, R. J. Langley, D. D. Bilgin, O. Radwan, D. J. Neece, S. J. Clough, G. D. May and G. Stacey (2010). "Complete transcriptome of the soybean root hair cell, a single-cell model, and its alteration in response to *Bradyrhizobium japonicum* infection." Plant Physiology **152**(2): 541-552.

Limpens, E., C. Franken, P. Smit, J. Willemse, T. Bisseling and R. Geurts (2003). "LysM domain receptor kinases regulating rhizobial Nod factor-induced infection." Science **302**(5645): 630-633.

Lodwig, E. (2001). Regulation of carbon metabolism in Rhizobium leguminosarum, University of Reading.

Lodwig, E. and P. Poole (2003). "Metabolism of Rhizobium bacteroids." Critical Reviews in Plant Sciences **22**(1): 37-78.

Lodwig, E. M., A. H. F. Hosie, A. Bordes, K. Findlay, D. Allaway, R. Karunakaran, J. A. Downie and P. S. Poole (2003). "Amino-acid cycling drives nitrogen fixation in the legume-*Rhizobium* symbiosis." Nature **422**(6933): 722-726.

Lodwig, E. M., M. Leonard, S. Marroqui, T. R. Wheeler, K. Findlay, J. A. Downie and P. S. Poole (2005). "Role of polyhydroxybutyrate and glycogen as carbon storage compounds in pea and bean bacteroids." Molecular Plant-Microbe Interactions **18**(1): 67-74.

Lohar, D. P., N. Sharopova, G. Endre, S. Penuela, D. Samac, C. Town, K. A. Silverstein and K. A. VandenBosch (2006). "Transcript analysis of early nodulation events in *Medicago truncatula*." Plant Physiology **140**(1): 221-234.

Lopez-Gomez, M., N. Sandal, J. Stougaard and T. Boller (2012). "Interplay of flg22-induced defence responses and nodulation in *Lotus japonicus*." Journal of Experimental Botany **63**(1): 393-401.

Lopez-Torrejon, G., E. Jimenez-Vicente, J. M. Buesa, J. A. Hernandez, H. K. Verma and L. M. Rubio (2016). "Expression of a functional oxygen-labile nitrogenase component in the mitochondrial matrix of aerobically grown yeast." Nature Communications **7**: 11426.

Ludden, P. W. (1993). *nif* gene products and their roles in nitrogen fixation. New Horizons in Nitrogen Fixation: Proceedings of the 9th International Congress on Nitrogen Fixation, Cancún, Mexico, December 6–12, 1992. R. Palacios, J. Mora and W. E. Newton. Dordrecht, Springer Netherlands**:** 101-104.

Maathuis, F. J. (2009). "Physiological functions of mineral macronutrients." Current Opinion in Plant Biology **12**(3): 250-258.

Madison, L. L. and G. W. Huisman (1999). "Metabolic engineering of poly(3-hydroxyalkanoates): From DNA to plastic." Microbiology and Molecular Biology Reviews **63**(1): 21-+.

Madsen, E. B., L. H. Madsen, S. Radutoiu, M. Olbryt, M. Rakwalska, K. Szczyglowski, S. Sato, T. Kaneko, S. Tabata, N. Sandal and J. Stougaard (2003). "A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals." Nature **425**(6958): 637-640.

Majed, N. and A. Z. Gu (2010). "Application of Raman microscopy for simultaneous and quantitative evaluation of multiple intracellular polymers dynamics functionally relevant to enhanced biological phosphorus removal processes." Environmental Science & Technology **44**(22): 8601-8608.

Mandon, K., P. A. Kaminski and C. Elmerich (1994). "Functional analysis of the *fixNOQP* region of *Azorhizobium caulinodans*." Journal of Bacteriology **176**(9): 2560-2568.

Mandon, K., N. Michel-Reydellet, S. Encarnacion, P. A. Kaminski, A. Leija, M. A. Cevallos, C. Elmerich and J. Mora (1998). "Poly-beta-hydroxybutyrate turnover in *Azorhizobium caulinodans* is required for growth and affects *nifA* expression." Journal of Bacteriology **180**(19): 5070-5076.

Marchler-Bauer, A., S. N. Lu, J. B. Anderson, F. Chitsaz, M. K. Derbyshire, C. DeWeese-Scott, J. H. Fong, L. Y. Geer, R. C. Geer, N. R. Gonzales, M. Gwadz, D. I. Hurwitz, J. D. Jackson, Z. X. Ke, C. J. Lanczycki, F. Lu, G. H. Marchler, M. Mullokandov, M. V. Omelchenko, C. L. Robertson, J. S. Song, N. Thanki, R. A. Yamashita, D. C. Zhang, N. G. Zhang, C. J. Zheng and S. H. Bryant (2011). "CDD: a Conserved Domain Database for the functional annotation of proteins." Nucleic Acids Research **39**: D225-D229.

Marino, D., E. Andrio, E. G. Danchin, E. Oger, S. Gucciardo, A. Lambert, A. Puppo and N. Pauly (2011). "A Medicago truncatula NADPH oxidase is involved in symbiotic nodule functioning." New Phytologist **189**(2): 580-592.

Maroti, G. and E. Kondorosi (2014). "Nitrogen-fixing *Rhizobium-legume* symbiosis: are polyploidy and host peptide-governed symbiont differentiation general principles of endosymbiosis?" Frontiers in Microbiology **5**: 326.

Marroqui, S., A. Zorreguieta, C. Santamaria, F. Temprano, M. Soberon, M. Megias and J. A. Downie (2001). "Enhanced symbiotic performance by *Rhizobium tropici* glycogen synthase mutants." Journal of Bacteriology **183**(3): 854-864.

Martinez, M., J. M. Palacios, J. Imperial and T. Ruiz-Argueso (2004). "Symbiotic autoregulation of *nifA* expression in *Rhizobium leguminosarum* bv. *viciae*." Journal of Bacteriology **186**(19): 6586-6594.

Martinez-Argudo, I., R. Little and R. Dixon (2004). "Role of the amino-terminal GAF domain of the NifA activator in controlling the response to the antiactivator protein NifL." Molecular Microbiology **52**(6): 1731-1744.

Martinez-Argudo, I., R. Little, N. Shearer, P. Johnson and R. Dixon (2004). "The NifL-NifA system: a multidomain transcriptional regulatory complex that integrates environmental signals." Journal of Bacteriology **186**(3): 601-610.

Martinez-Salazar, J. M., M. Sandoval-Calderon, X. W. Guo, S. Castillo-Ramirez, A. Reyes, M. G. Loza, J. Rivera, X. Alvarado-Affantranger, F. Sanchez, V. Gonzalez, G. Davila and M. A. Ramirez-Romero (2009). "The Rhizobium etli RpoH1 and RpoH2 sigma factors are involved in different stress responses." Microbiology **155**: 386-397.

Maunoury, N., M. Redondo-Nieto, M. Bourcy, W. Van de Velde, B. Alunni, P. Laporte, P. Durand, N. Agier, L. Marisa, D. Vaubert, H. Delacroix, G. Duc, P. Ratet, L. Aggerbeck, E. Kondorosi and P. Mergaert (2010). "Differentiation of symbiotic cells and endosymbionts in *Medicago truncatula* nodulation are coupled to two transcriptome-switches." PLOS One **5**(3): e9519.

Mcdermott, T. R., S. M. Griffith, C. P. Vance and P. H. Graham (1989). "Carbon metabolism in *Bradyrhizobium japonicum* bacteroids." FEMS Microbiology Letters **63**(4): 327-340.

Mckay, I. A., A. R. Glenn and M. J. Dilworth (1985). "Gluconeogenesis in *Rhizobium leguminosarum* Mnf3841." Journal of General Microbiology **131**(Aug): 2067-2073.

Mergaert, P., K. Nikovics, Z. Kelemen, N. Maunoury, D. Vaubert, A. Kondorosi and E. Kondorosi (2003). "A novel family in *Medicago truncatula* consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs." Plant Physiology **132**(1): 161-173.

Mergaert, P., T. Uchiumi, B. Alunni, G. Evanno, A. Cheron, O. Catrice, A. E. Mausset, F. Barloy-Hubler, F. Galibert, A. Kondorosi and E. Kondorosi (2006). "Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*-legume symbiosis." Proceedings of the National Academy of Sciences of the United States of America **103**(13): 5230-5235.

Migita, C. T., K. M. Matera, M. Ikeda-Saito, J. S. Olson, H. Fujii, T. Yoshimura, H. Zhou and T. Yoshida (1998). "The oxygen and carbon monoxide reactions of heme oxygenase." Journal of Biological Chemistry **273**(2): 945-949.

Mitchell, P. (1975). "The protonmotive Q cycle: a general formulation." FEBS Letters **59**(2): 137-139.

Miwa, H., J. Sun, G. E. D. Oldroyd and J. A. Downie (2006). "Analysis of nod-factor-induced calcium signaling in root hairs of symbiotically defective mutants of *Lotus japonicus*." Molecular Plant-Microbe Interactions **19**(8): 914-923.

Monahan-Giovanelli, H., C. A. Pinedo and D. J. Gage (2006). "Architecture of infection thread networks in developing root nodules induced by the symbiotic bacterium *Sinorhizobium meliloti* on *Medicago truncatula*." Plant Physiology **140**(2): 661-670.

Morett, E., H. M. Fischer and H. Hennecke (1991). "Influence of oxygen on DNA binding, positive control, and stability of the Bradyrhizobium japonicum NifA regulatory protein." Journal of Bacteriology **173**(11): 3478-3487.

Mouncey, N. J. and S. Kaplan (1998). "Oxygen regulation of the *ccoN* gene encoding a component of the cbb(3) oxidase in *Rhodobacter sphaeroides* 2.4.1(T): Involvement of the FnrL protein." Journal of Bacteriology **180**(8): 2228-2231.

Muller, J., A. Wiemken and T. Boller (2001). "Redifferentiation of bacteria isolated from *Lotus japonicus* root nodules colonized by *Rhizobium* sp NGR234." Journal of Experimental Botany **52**(364): 2181-2186.

Mulley, G., M. Lopez-Gomez, Y. Zhang, J. Terpolilli, J. Prell, T. Finan and P. Poole (2010). "Pyruvate Is synthesized by two pathways in pea bacteroids with different efficiencies for nitrogen fixation." Journal of Bacteriology **192**(19): 4944-4953.

Mulley, G., J. P. White, R. Karunakaran, J. Prell, A. Bourdes, S. Bunnewell, L. Hill and P. S. Poole (2011). "Mutation of GOGAT prevents pea bacteroid formation and N2 fixation by globally downregulating transport of organic nitrogen sources." Molecular Microbiology **80**(1): 149-167.

Murray, J. D. (2011). "Invasion by invitation: rhizobial infection in legumes." Molecular Plant-Microbe Interactions **24**(6): 631-639.

Mus, F., M. B. Crook, K. Garcia, A. Garcia Costas, B. A. Geddes, E. D. Kouri, P. Paramasivan, M. H. Ryu, G. E. Oldroyd, P. S. Poole, M. K. Udvardi, C. A. Voigt, J. M. Ane and J. W. Peters (2016). "Symbiotic nitrogen fixation and challenges to extending it to non-legumes." Applied and Environmental Microbiology.

Nakabachi, A., S. Koshikawa, T. Miura and S. Miyagishima (2010). "Genome size of *Pachypsylla venusta* (Hemiptera: Psyllidae) and the ploidy of its bacteriocyte, the symbiotic host cell that harbors intracellular mutualistic bacteria with the smallest cellular genome." Bulletin of Entomological Research **100**(1): 27-33.

Nie, S. M. (2001). "Probing single molecules and single nanoparticles by surface-enhanced Raman scattering." Science **275**(5303): 1102-1106.

Nikel, P. I., M. Chavarria, E. Martinez-Garcia, A. C. Taylor and V. de Lorenzo (2013). "Accumulation of inorganic polyphosphate enables stress endurance and catalytic vigour in *Pseudomonas putida* KT2440." Microbial Cell Factories **12**.

Novak, K., L. Lisa and V. Skrdleta (2004). "Rhizobial nod gene-inducing activity in pea nodulation mutants: dissociation of nodulation and flavonoid response." Physiologia Plantarum **120**(4): 546-555.

Oke, V. and S. R. Long (1999). "Bacterial genes induced within the nodule during the *Rhizobium*–legume symbiosis." Molecular Microbiology **32**(4): 837-849.

Oldroyd, G. E. and J. A. Downie (2004). "Calcium, kinases and nodulation signalling in legumes." Nature Reviews Molecular Cell Biology **5**(7): 566-576.

Oldroyd, G. E., J. D. Murray, P. S. Poole and J. A. Downie (2011). "The rules of engagement in the legume-rhizobial symbiosis." Annual Review of Genetics **45**: 119-144.

Oldroyd, G. E. D. and R. Dixon (2014). "Biotechnological solutions to the nitrogen problem." Current Opinion in Biotechnology **26**: 19-24.

Olsen, R. K. J., B. S. Andresen, E. Christensen, P. Bross, F. Skovby and N. Gregersen (2003). "Clear relationship between *ETF/ETFDH* genotype and phenotype in patients with multiple Acyl-CoA dehydrogenation deficiency." Human Mutation **22**(1): 12-23.

Ott, T., J. T. van Dongen, C. Gunther, L. Krusell, G. Desbrosses, H. Vigeolas, V. Bock, T. Czechowski, P. Geigenberger and M. K. Udvardi (2005). "Symbiotic leghemoglobins are crucial for nitrogen fixation in legume root nodules but not for general plant growth and development." Current Biology **15**(6): 531-535.

Paau, A. S., C. B. Bloch and W. J. Brill (1980). "Developmental fate of *Rhizobium meliloti* bacteroids in alfalfa nodules." Journal of Bacteriology **143**(3): 1480-1490.

Page, W. J. and O. Knosp (1989). "Hyperproduction of poly-beta-hydroxybutyrate during exponential growth of *Azotobacter vinelandii* UWD." Applied and Environmental Microbiology **55**(6): 1334-1339.

Parniske, M. (2000). "Intracellular accommodation of microbes by plants: a common developmental program for symbiosis and disease?" Current Opinion in Plant Biology **3**(4): 320-328.

Patschkowski, T., A. Schluter and U. B. Priefer (1996). "*Rhizobium leguminosarum* bv. *viciae* contains a second *fnr/fixK*-like gene and an unusual *fixL* homologue." Molecular Microbiology **21**(2): 267-280.

Pauling, D., J. Lapointe, C. Paris and R. A. Ludwig (2001). "*Azorhizobium caulinodans* pyruvate dehydrogenase activity is dispensable for aerobic but required for microaerobic growth." Microbiology **147**(8): 2233-2245.

Peng, J., B. Hao, L. Liu, S. Wang, B. Ma, Y. Yang, F. Xie and Y. Li (2014). "RNA-Seq and microarrays analyses reveal global differential transcriptomes of *Mesorhizobium huakuii* 7653R between bacteroids and free-living cells." PLOS One **9**(4): e93626.

Perez-Galdona, R. and M. L. Kahn (1994). "Effects of organic acids and low pH on *Rhizobium meliloti* 104A14." Microbiology **140 (Pt 5)**: 1231-1235.

Perry, B. J. and C. K. Yost (2014). "Construction of a *mariner*-based transposon vector for use in insertion sequence mutagenesis in selected members of the *Rhizobiaceae*." BMC Microbiology **14**.

Pierre, O., G. Engler, J. Hopkins, F. Brau, E. Boncompagni and D. Herouart (2013). "Peribacteroid space acidification: a marker of mature bacteroid functioning in Medicago truncatula nodules." Plant, Cell & Environment **36**(11): 2059-2070.

Pini, F., N. J. De Nisco, L. Ferri, J. Penterman, A. Fioravanti, M. Brilli, A. Mengoni, M. Bazzicalupo, P. H. Viollier, G. C. Walker and E. G. Biondi (2015). "Cell cycle control by the master regulator CtrA in *Sinorhizobium meliloti*." PLoS Genetics **11**(5): e1005232.

Pitcher, R. S. and N. J. Watmough (2004). "The bacterial cytochrome cbb(3) oxidases." Biochimica Et Biophysica Acta - Bioenergetics **1655**(1-3): 388-399.

Povolo, S. and S. Casella (2002). Global Perspectives : Proceedings of the 13th International Congress on Nitrogen Fixation.

Preisig, O., R. Zufferey and H. Hennecke (1996). "The *Bradyrhizobium japonicum fixGHIS* genes are required for the formation of the high-affinity cbb(3)-type cytochrome oxidase." Archives of Microbiology **165**(5): 297-305.

Preisig, O., R. Zufferey, L. ThonyMeyer, C. A. Appleby and H. Hennecke (1996). "A high-affinity cbb(3)-type cytochrome oxidase terminates the symbiosis-specific respiratory chain of *Bradyrhizobium japonicum*." Journal of Bacteriology **178**(6): 1532-1538.

Prell, J., B. Boesten, P. Poole and U. B. Priefer (2002). "The Rhizobium leguminosarum bv. viciae VF39 gamma-aminobutyrate (GABA) aminotransferase gene (gabT) is induced by GABA and highly expressed in bacteroids." Microbiology **148**: 615-623.

Prell, J., G. Mulley, F. Haufe, J. P. White, A. Williams, R. Karunakaran, J. A. Downie and P. S. Poole (2012). "The PTS(Ntr) system globally regulates ATP-dependent transporters in Rhizobium leguminosarum." Molecular Microbiology **84**(1): 117-129.

Prell, J. and P. Poole (2006). "Metabolic changes of rhizobia in legume nodules." Trends in Microbiology **14**(4): 161-168.

Prell, J., J. P. White, A. Bourdes, S. Bunnewell, R. J. Bongaerts and P. S. Poole (2009). "Legumes regulate *Rhizobium* bacteroid development and persistence by the supply of branched-chain amino acids." Proceedings of the National Academy of Sciences of the United States of America **106**(30): 12477-12482.

Price-Carter, M., T. G. Fazzio, E. L. Vallbona and J. R. Roth (2005). "Polyphosphate kinase protects Salmonella. enterica from weak organic acid stress." Journal of Bacteriology **187**(9): 3088-3099.

Puppels, G. J., F. F. M. Demul, C. Otto, J. Greve, M. Robertnicoud, D. J. Arndtjovin and T. M. Jovin (1990). "Studying single living cells and chromosomes by confocal Raman microspectroscopy." Nature **347**(6290): 301-303.

Purnick, P. E. and R. Weiss (2009). "The second wave of synthetic biology: from modules to systems." Nature Reviews Molecular Cell Biology **10**(6): 410-422.

Quandt, J. and M. F. Hynes (1993). "Versatile suicide vectors which allow direct selection for gene replacement in gram-negative bacteria." Gene **127**(1): 15-21.

Raitio, M. and M. Wikstrom (1994). "An alternative cytochrome oxidase of *Paracoccus denitrificans* functions as a proton pump." Biochimica Et Biophysica Acta - Bioenergetics **1186**(1-2): 100-106.

Rao, N. N., M. R. Gomez-Garcia and A. Kornberg (2009). "Inorganic polyphosphate: essential for growth and survival." Annual Review of Biochemistry **78**: 605-647.

Ratcliff, W. C., S. V. Kadam and R. F. Denison (2008). "Poly-3-hydroxybutyrate (PHB) supports survival and reproduction in starving rhizobia." FEMS Microbiology Ecology **65**(3): 391-399.

Ratet, P., K. Pawlowski, J. Schell and F. J. Debruijn (1989). "The *Azorhizobium caulinodan*s nitrogen-fixation regulatory gene, *nifA*, is controlled by the cellular nitrogen and oxygen status." Molecular Microbiology **3**(6): 825-838.

Rehm, B. H. and A. Steinbuchel (1999). "Biochemical and genetic analysis of PHA synthases and other proteins required for PHA synthesis." International Journal of Biological Macromolecules **25**(1-3): 3-19.

Riedel, K. U., Y. Jouanneau, B. Masepohl, A. Puhler and W. Klipp (1995). "A *Rhizobium meliloti* ferredoxin (FdxN) purified from *Escherichia coli* donates electrons to *Rhodobacter capsulatus* nitrogenase." European Journal of Biochemistry **231**(3): 742-746.

Roberts, D. L., F. E. Frerman and J. J. Kim (1996). "Three-dimensional structure of human electron transfer flavoprotein to 2.1-A resolution." Proceedings of the National Academy of Sciences of the United States of America **93**(25): 14355-14360.

Roche, P., F. Maillet, C. Plazanet, F. Debelle, M. Ferro, G. Truchet, J. C. Prome and J. Denarie (1996). "The common *nodABC* genes of *Rhizobium meliloti* are host-range determinants." Proceedings of the National Academy of Sciences of the United States of America **93**(26): 15305-15310.

Ronson, C. W., P. M. Astwood, B. T. Nixon and F. M. Ausubel (1987). "Deduced products of C4-dicarboxylate transport regulatory genes of *Rhizobium leguminosarum* are homologous to nitrogen regulatory gene products." Nucleic Acids Research **15**(19): 7921-7934.

Ronson, C. W., P. Littleton and J. G. Robertson (1981). "C4-dicarboxylate transport mutants of *Rhizobium trifolii* form ineffective nodules on *Trifolium repens*." Proceedings of the National Academy of Sciences of the United States of America.

Rosendahl, L., M. J. Dilworth and A. R. Glenn (1992). "Exchange of Metabolites across the Peribacteroid Membrane in Pea Root-Nodules." Journal of Plant Physiology **139**(5): 635-638.

Roth, E., K. Jeon and G. Stacey (1988). Homology in endosymbiotic systems: The term 'symbiosome' St. Paul, MN, American Phytopathological Society Press.

Ruiz-Argueso, T., J. M. Palacios and J. Imperial (2001). "Regulation of the hydrogenase system in Rhizobium leguminosarum." Plant Soil **230**(1): 49-57.

Ruston, E. (2003). An investigation of the FixABCX proteins from Azorhizobium caulinodans, University of East Anglia.

Salazar, E., J. J. Diaz-Mejia, G. Moreno-Hagelsieb, G. Martinez-Batallar, Y. Mora, J. Mora and S. Encarnacion (2010). "Characterization of the NifA-RpoN regulon in *Rhizobium etli* in free life and in symbiosis with *Phaseolus vulgaris*." Applied and Environmental Microbiology **76**(13): 4510-4520.

Santana, M., K. Pihakaski-Maunsback, N. Sandal, K. Marcker and A. Smith (1998). "Evidence that the plant host synthesizes the heme moeiety of leghemoglobin in root nodules." Plant Physiology **116**: 1259-1269.

Sanyal, S., S. K. Banerjee, R. Banerjee, J. Mukhopadhyay and M. Kundu (2013). "Polyphosphate kinase 1, a central node in the stress response network of Mycobacterium tuberculosis, connects the two-component systems MprAB and SenX3-RegX3 and the extracytoplasmic function sigma factor, sigma E." Microbiology **159**: 2074-2086.

Sato, K., Y. Nishina and K. Shiga (1993). "Electron-Transferring Flavoprotein Has an Amp-Binding Site in Addition to the Fad-Binding Site." Journal of Biochemistry **114**(2): 215-222.

Schafer, A., A. Tauch, W. Jager, J. Kalinowski, G. Thierbach and A. Puhler (1994). "Small mobilizable multipurpose cloning vectors derived from the Escherichia coli plasmids pK18 and pK19 - selection of defined deletions in the chromosome of Corynebacterium-glutamicum." Gene **145**(1): 69-73.

Scheibel, M. G. and S. Schneider (2012). "New insights into the biological and synthetic fixation of nitrogen." Angewandte Chemie International Edition **51**(19): 4529-4531.

Scheres, B., F. Vanengelen, E. Vanderknaap, C. Vandewiel, A. Vankammen and T. Bisseling (1990). "Sequential induction of nodulin gene expression in the developing pea nodule." Plant Cell **2**(8): 687-700.

Schlesinger, W. H. (2009). "On the fate of anthropogenic nitrogen." Proceedings of the National Academy of Sciences of the United States of America **106**(1): 203-208.

Schluter, A., T. Patschkowski, J. Quandt, L. B. Selinger, S. Weidner, M. Kramer, L. M. Zhou, M. F. Hynes and U. B. Priefer (1997). "Functional and regulatory analysis of the two copies of the *fixNOQP* operon of *Rhizobium leguminosarum* strain VF39." Molecular Plant-Microbe Interactions **10**(5): 605-616.

Schluter, J. P., J. Reinkensmeier, M. J. Barnett, C. Lang, E. Krol, R. Giegerich, S. R. Long and A. Becker (2013). "Global mapping of transcription start sites and promoter motifs in the symbiotic alpha-proteobacterium Sinorhizobium meliloti 1021." BMC Genomics **14**: 156.

Schmidt, T. G. M., L. Batz, L. Bonet, U. Carl, G. Holzapfel, K. Kiem, K. Matulewicz, D. Niermeier, I. Schuchardt and K. Stanar (2013). "Development of the Twin-Strep-tag and its application for purification of recombinant proteins from cell culture supernatants." Protein Expression and Purification **92**(1): 54-61.

Schumpp, O. and W. J. Deakin (2010). "How inefficient rhizobia prolong their existence within nodules." Trends in Plant Science **15**(4): 189-195.

Sciotti, M. A., A. Chanfon, H. Hennecke and H. M. Fischer (2003). "Disparate oxygen responsiveness of two regulatory cascades that control expression of symbiotic genes in *Bradyrhizobium japonicum*." Journal of Bacteriology **185**(18): 5639-5642.

Scott, J. D. and R. A. Ludwig (2004). "Azorhizobium caulinodans electron-transferring flavoprotein N electrochemically couples pyruvate dehydrogenase complex activity to N2 fixation." Microbiology **150**(1): 117-126.

Seefeldt, L. C. and D. R. Dean (1997). "Role of nucleotides in nitrogenase catalysis." Accounts of Chemical Research **30**(6): 260-266.

Senior, P. J., G. A. Beech, G. A. Ritchie and E. A. Dawes (1972). "The role of oxygen limitation in the formation of poly-beta-hydroxybutyrate during batch and continuous culture of *Azotobacter beijerinckii*." Biochemical Journal **128**(5): 1193-1201.

Shah, V. K., G. Stacey and W. J. Brill (1983). "Electron transport to nitrogenase. Purification and characterization of pyruvate:flavodoxin oxidoreductase. The *nifJ* gene product." Journal of Biological Chemistry **258**(19): 2064-2068.

Shcherbak, I., N. Millar and G. P. Robertson (2014). "Global metaanalysis of the nonlinear response of soil nitrous oxide (N2O) emissions to fertilizer nitrogen." Proceedings of the National Academy of Sciences of the United States of America **111**(25): 9199-9204.

Silverstein, K. A. T., W. A. Moskal, H. C. Wu, B. A. Underwood, M. A. Graham, C. D. Town and K. A. VandenBosch (2007). "Small cysteine-rich peptides resembling antimicrobial peptides have been under-predicted in plants." Plant Journal **51**(2): 262-280.

Smil, V. (2011). "Nitrogen cycle and world food production." World Agriculture **2**: 9-13.

Smit, G., S. Swart, B. J. Lugtenberg and J. W. Kijne (1992). "Molecular mechanisms of attachment of *Rhizobium* bacteria to plant roots." Molecular Microbiology **6**(20): 2897-2903.

Spaink, H. P. (2000). "Root nodulation and infection factors produced by rhizobial bacteria." Annual Review of Microbiology **54**: 257-288.

Sr, P., S. Knebel, T. Polen, P. Klauth, J. Hollender, V. F. Wendisch and S. M. Schoberth (2005). "Formation of volutin granules in *Corynebacterium glutamicum*." FEMS Microbiology Letters **243**(1): 133-140.

Stacey, G., C. B. McAlvin, S. Y. Kim, J. Olivares and M. J. Soto (2006). "Effects of endogenous salicylic acid on nodulation in the model legumes *Lotus japonicus* and *Medicago truncatula*." Plant Physiology **141**(4): 1473-1481.

Streeter, J. G. (1991). "Transport and metabolism of carbon and nitrogen in legume nodules." Advances in Botanical Research **18**: 129-187.

Sullivan, J. T., S. D. Brown and C. W. Ronson (2013). "The NifA-RpoN regulon of *Mesorhizobium loti* strain R7A and its symbiotic activation by a novel LacI/GalR-family regulator." PLOS One **8**(1).

Sureka, K., S. Dey, P. Datta, A. K. Singh, A. Dasgupta, S. Rodrigue, J. Basu and M. Kundu (2007). "Polyphosphate kinase is involved in stress-induced mprAB-sigE-rel signalling in mycobacteria." Molecular Microbiology **65**(2): 261-276.

Suzaki, T., M. Ito, E. Yoro, S. Sato, H. Hirakawa, N. Takeda and M. Kawaguchi (2014). "Endoreduplication-mediated initiation of symbiotic organ development in *Lotus japonicus*." Development **141**(12): 2441-2445.

Swem, D. L. and C. E. Bauer (2002). "Coordination of ubiquinol oxidase and cytochrome cbb(3) oxidase expression by multiple regulators in *Rhodobacter capsulatus*." Journal of Bacteriology **184**(10): 2815-2820.

Szklarczyk, D., A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, M. Simonovic, A. Roth, A. Santos, K. P. Tsafou, M. Kuhn, P. Bork, L. J. Jensen and C. von Mering (2015). "STRING v10: protein-protein interaction networks, integrated over the tree of life." Nucleic Acids Research **43**(D1): D447-D452.

Taiz, L. and E. Zeiger (2010). Plant Physiology, Fifth Edition. Sunderland, Massachusetts, Sinauer Associates Inc.

Taylor, B. L. and I. B. Zhulin (1999). "PAS domains: internal sensors of oxygen, redox potential, and light." Microbiology and Molecular Biology Reviews **63**(2): 479-506.

Temme, K., D. H. Zhao and C. A. Voigt (2012). "Refactoring the nitrogen fixation gene cluster from *Klebsiella oxytoca*." Proceedings of the National Academy of Sciences of the United States of America **109**(18): 7085-7090.

Terpolilli, J. J., G. A. Hood and P. S. Poole (2012). "What determines the efficiency of N(2)-fixing *Rhizobium*-legume symbioses?" Advances in Microbial Physiology **60**: 325-389.

Terpolilli, J. J., S. K. Masakapallic, R. Karunakaran, I. Webb, R. Green, N. Watmough, N. Kruger, G. R. Ratcliffe and P. S. Poole (2016). "Lipogenesis and redox balance in nitrogen-fixing pea bacteroids." Journal of Bacteriology.

Tett, A. J., S. J. Rudder, A. Bourdes, R. Karunakaran and P. S. Poole (2012). "Regulatable vectors for environmental gene expression in Alphaproteobacteria." Applied and Environmental Microbiology **78**(19): 7137-7140.

Tett. A, S. Rudder, A. Bourdes, R. Karunakaran and P. Poole (2012). "Regulatable Vectors for Environmental Gene Expression in Alphaproteobacteria." Applied and Environmental Microbiology **78**(19): 7137-7140.

Timmers, A. C. J., E. Soupene, M. C. Auriac, F. de Billy, J. Vasse, P. Boistard and G. Truchet (2000). "Saprophytic intracellular rhizobia in alfalfa nodules." Molecular Plant-Microbe Interactions **13**(11): 1204-1213.

Tsai, M. H. and M. H. Saier (1995). "Phylogenetic characterization of the ubiquitous electron transfer flavoprotein families ETF-alpha and ETF-beta." Research in Microbiology **146**(5): 397-404.

Tsukada, S., T. Aono, N. Akiba, K. B. Lee, C. T. Liu, H. Toyazaki and H. Oyaizu (2009). "Comparative genome-wide transcriptional profiling of *Azorhizobium caulinodans* ORS571

grown under free-living and symbiotic conditions." Applied and Environmental Microbiology **75**(15): 5037-5046.

Tuckerman, J. R., G. Gonzalez, E. M. Dioum and M. A. Gilles-Gonzalez (2002). "Ligand and oxidation-state specific regulation of the heme-based oxygen sensor FixL from *Sinorhizobium meliloti*." Biochemistry **41**(19): 6170-6177.

Uchiumi, T., T. Ohwada, M. Itakura, H. Mitsui, N. Nukui, P. Dawadi, T. Kaneko, S. Tabata, T. Yokoyama, K. Tejima, K. Saeki, H. Omori, M. Hayashi, T. Maekawa, R. Sriprang, Y. Murooka, S. Tajima, K. Simomura, M. Nomura, A. Suzuki, Y. Shimoda, K. Sioya, M. Abe and K. Minamisawa (2004). "Expression islands clustered on the symbiosis island of the *Mesorhizobium loti* genome." Journal of Bacteriology **186**(8): 2439-2448.

Udvardi, M. and P. S. Poole (2013). "Transport and metabolism in legume-rhizobia symbioses." Annual Review of Plant Biology **64**: 781-805.

Van de Velde, W., G. Zehirov, A. Szatmari, M. Debreczeny, H. Ishihara, Z. Kevei, A. Farkas, K. Mikulass, A. Nagy, H. Tiricz, B. Satiat-Jeunemaitre, B. Alunni, M. Bourge, K. Kucho, M. Abe, A. Kereszt, G. Maroti, T. Uchiumi, E. Kondorosi and P. Mergaert (2010). "Plant peptides govern terminal differentiation of bacteria in symbiosis." Science **327**(5969): 1122-1126.

Vanbatenburg, F. H. D., R. Jonker and J. W. Kijne (1986). "Rhizobium induces marked root hair curling by redirection of tip growth: a computer simulation." Physiologia Plantarum **66**(3): 476-480.

Vasse, J., F. de Billy, S. Camut and G. Truchet (1990). "Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules." Journal of Bacteriology **172**(8): 4295-4306.

Verhagen, M. F., T. O'Rourke and M. W. Adams (1999). "The hyperthermophilic bacterium, *Thermotoga maritima*, contains an unusually complex iron-hydrogenase: amino acid sequence analyses versus biochemical characterization." Biochimica et Biophysica Acta **1412**(3): 212-229.

Walshaw, D. L. and P. S. Poole (1996). "The general L-amino acid permease of *Rhizobium leguminosarum* is an ABC uptake system that also influences efflux of solutes." Molecular Microbiology **21**(6): 1239-1252.

Walt, A. and M. L. Kahn (2002). "The *fixA* and *fixB* genes are necessary for anaerobic carnitine reduction in *Escherichia coli*." Journal of Bacteriology **184**(14): 4044-4047.

Wang, C. X., M. Saldanha, X. Y. Sheng, K. J. Shelswell, K. T. Walsh, B. W. S. Sobral and T. C. Charles (2007). "Roles of poly-3-hydroxybutyrate (PHB) and glycogen in symbiosis of *Sinorhizobium meliloti* with *Medicago* sp." Microbiology **153**: 388-398.

Waters, J. K., B. L. Hughes, 2nd, L. C. Purcell, K. O. Gerhardt, T. P. Mawhinney and D. W. Emerich (1998). "Alanine, not ammonia, is excreted from N2-fixing soybean nodule bacteroids." Proceedings of the National Academy of Sciences of the United States of America **95**(20): 12038-12042.

Watmough, N. J. and F. E. Frerman (2010). "The electron transfer flavoprotein: ubiquinone oxidoreductases." Biochimica et Biophysica Acta **1797**(12): 1910-1916.

Weber, E., C. Engler, R. Gruetzner, S. Werner and S. Marillonnet (2011). "A modular cloning system for standardized assembly of multigene constructs." PLOS One **6**(2).

Werner, G. D. A., W. K. Cornwell, J. I. Sprent, J. Kattge and E. T. Kiers (2014). "A single evolutionary innovation drives the deep evolution of symbiotic N-2-fixation in angiosperms." Nature Communications **5**.

White, J., J. Prell, E. K. James and P. Poole (2007). "Nutrient sharing between symbionts." Plant Physiology **144**(2): 604-614.

Wigglesworth, V. B. (1975). "Lipid staining for the electron microscope: a new method." Journal of Cell Science **19**(3): 425-437.

Wong, P. P. and R. H. Burris (1972). "Nature of oxygen inhibition of nitrogenase from *Azotobacter vinelandii*." Proceedings of the National Academy of Sciences of the United States of America **69**(3): 672-675.

Wong, P. P. and H. J. Evans (1971). "Poly-beta-hydroxybutyrate utilization by soybean (*Glycine max* Merr.) nodules and assessment of its role in maintenance of nitrogenase activity." Plant Physiology **47**(6): 750-&.

Xie, F. L., G. J. Cheng, H. Xu, Z. Wang, L. Lei and Y. G. Li (2011). "Identification of a novel gene for biosynthesis of a bacteroid-specific electron carrier menaquinone." PLOS One **6**(12).

Xu, G., X. Fan and A. J. Miller (2012). "Plant nitrogen assimilation and use efficiency." Annual Review of Plant Biology **63**: 153-182.

Young, J. P. W., L. C. Crossman, A. W. B. Johnston, N. R. Thomson, Z. F. Ghazoui, K. H. Hull, M. Wexler, A. R. J. Curson, J. D. Todd, P. S. Poole, T. H. Mauchline, A. K. East, M. A. Quail, C. Churcher, C. Arrowsmith, I. Cherevach, T. Chillingworth, K. Clarke, A. Cronin, P. Davis, A. Fraser, Z. Hance, H. Hauser, K. Jagels, S. Moule, K. Mungall, H. Norbertczak, E. Rabbinowitsch, M. Sanders, M. Simmonds, S. Whitehead and J. Parkhill (2006). "The genome of *Rhizobium leguminosarum* has recognizable core and accessory components." Genome Biology **7**(4).

Yu, C. S., C. W. Cheng, W. C. Su, K. C. Chang, S. W. Huang, J. K. Hwang and C. H. Lu (2014). "CELLO2GO: A web server for protein subCELlular LOcalization prediction with functional gene ontology annotation." PLOS One **9**(6).

Yu, N. Y., J. R. Wagner, M. R. Laird, G. Melli, S. Rey, R. Lo, P. Dao, S. C. Sahinalp, M. Ester, L. J. Foster and F. S. L. Brinkman (2010). "PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes." Bioinformatics **26**(13): 1608-1615.

Zamorano-Sanchez, D., A. Reyes-Gonzalez, N. Gomez-Hernandez, P. Rivera, D. Georgellis and L. Girard (2012). "FxkR provides the missing link in the *fixL-fixK* signal transduction cascade in *Rhizobium etli* CFN42." Molecular Plant-Microbe Interactions **25**(11): 1506-1517.

Zevenhuizen, L. P. T. M. (1981). "Cellular glycogen, beta-1,2,-glucan, poly beta-hydroxybutyric acid and extracellular polysaccharides in fast-growing species of *Rhizobium*." Antonie Van Leeuwenhoek Journal of Microbiology **47**(6): 481-497.

Zhang, J., F. E. Frerman and J. J. P. Kim (2006). "Structure of electron transfer flavoprotein-ubiquinone oxidoreductase and electron transfer to the mitochondrial ubiquinone pool." Proceedings of the National Academy of Sciences of the United States of America **103**(44): 16212-16217.

Zheng, L. M., R. H. White and D. R. Dean (1997). "Purification of the *Azotobacter vinelandii nifV*-encoded homocitrate synthase." Journal of Bacteriology **179**(18): 5963-5966.

Ziegler, D., A. Mariotti, V. Pfluger, M. Saad, G. Vogel, M. Tonolla and X. Perret (2012). "*In Situ* identification of plant-invasive bacteria with MALDI-TOF mass spectrometry." PLOS One **7**(5).

Zipfel, C. (2014). "Plant pattern-recognition receptors." Trends in Immunology **35**(7): 345-351.

Appendix

Supplementary Table 1: Riley codes for gene classification

Riley codes and corresponding description were previously available in the Poole lab. Functional classifications were created for use in this project.

Code	Description	Functional Classification
0.0.0	Unknown function, no known homologues	Hypothetical
0.0.1	Conserved in E. coli	Hypothetical
0.0.2	Conserved in other organism than E.coli	Hypothetical
1.0.0	Cell processes	Other
1.1.1	Chemotaxis and mobility	Motility/chemotaxis
1.2.1	Chromosome replication	Nucleic acid replication/repair/synthesis
1.3.1	Chaperones	Other
1.4.0	Protection responses	Adaptation
1.4.1	Cell killing	Adaptation
1.4.2	Detoxification	Adaptation
1.4.3	Drug/analogue sensitivity	Adaptation
1.4.4	Radiation sensitivity	Adaptation
1.5.0	Transport/binding proteins	Transporters
1.5.1	Amino acids and amines	Amino acid biosynthesis
1.5.2	Cations	Transporters
1.5.3	Carbohydrates, organic acids and alcohols; PTS	Other biosynthesis
1.5.4	Anions	Transporters
1.5.5	Other	Other
1.6.0	Adaptation	Adaptation

1.6.1	Adaptations, atypical conditions	Adaptation
1.6.2	Osmotic adaptation	Adaptation
1.6.3	Fe storage	Other
1.6.4	Nodulation-related	Other
1.7.1	Cell division	Cell division
1.8.1	Sporulation, differentiation and germination	Other
2.0.0	Macromolecule metabolism	Other biosynthesis
2.1.0	Macromolecule degradation	Other biosynthesis
2.1.1	Degradation of DNA	Nucleic acid replication/repair/synthesis
2.1.2	Degradation of RNA	Nucleic acid replication/repair/synthesis
2.1.3	Degradation of polysaccharides	Degredation protein/amino acids
2.1.4	Degradation of proteins, peptides, glycoproteins	Degredation protein/amino acids
2.2.0	Macromolecule synthesis, modification	Other biosynthesis
2.2.1	Amino acyl tRNA synthesis; tRNA modification	Nucleic acid replication/repair/synthesis
2.2.10	Proteins - translation and modification	Protein biosynthesis/modification
2.2.11	RNA synthesis, modification, DNA transcription	Nucleic acid replication/repair/synthesis
2.2.12	tRNA	Nucleic acid replication/repair/synthesis
2.2.2	Basic proteins - synthesis, modification	Protein biosynthesis/modification
2.2.3	DNA - replication, repair, restriction/modification	Nucleic acid replication/repair/synthesis
2.2.4	Glycoprotein	Cell membrane/envelope
2.2.5	Lipopolysaccharide	Cell membrane/envelope
2.2.6	Lipoprotein	Cell membrane/envelope
2.2.7	Phospholipids	Cell membrane/envelope
2.2.8	Polysaccharides - (cytoplasmic)	Other biosynthesis
2.2.9	Protein modification	Protein biosynthesis/modification
3.0.0	Metabolism of small molecules	Intermediary metabolism

3.1.0	Amino acid biosynthesis	Amino acid biosynthesis
3.1.1	Alanine	Amino acid biosynthesis
3.1.10	Histidine	Amino acid biosynthesis
3.1.11	Isoleucine	Amino acid biosynthesis
3.1.12	Leucine	Amino acid biosynthesis
3.1.13	Lysine	Amino acid biosynthesis
3.1.14	Methionine	Amino acid biosynthesis
3.1.15	Phenylalanine	Amino acid biosynthesis
3.1.16	Proline	Amino acid biosynthesis
3.1.17	Serine	Amino acid biosynthesis
3.1.18	Threonine	Amino acid biosynthesis
3.1.19	Tryptophan	Amino acid biosynthesis
3.1.2	Arginine	Amino acid biosynthesis
3.1.20	Tyrosine	Amino acid biosynthesis
3.1.21	Valine	Amino acid biosynthesis
3.1.3	Asparagine	Amino acid biosynthesis
3.1.4	Aspartate	Amino acid biosynthesis
3.1.5	Chorismate	Amino acid biosynthesis
3.1.6	Cysteine	Amino acid biosynthesis
3.1.7	Glutamate	Amino acid biosynthesis
3.1.8	Glutamine	Amino acid biosynthesis
3.1.9	Glycine	Amino acid biosynthesis
3.2.0	Biosynthesis of cofactors, carriers	Other biosynthesis
3.2.1	Acyl carrier protein (ACP)	Other biosynthesis
3.2.10	Pantothenate, CoA?	Other biosynthesis
3.2.11	Pyridine nucleotide, NAD	Nucleic acid replication/repair/synthesis
3.2.12	Pyridoxine	Nucleic acid replication/repair/synthesis
3.2.13	Riboflavin	Electron transfer
3.2.14	Thiamin	Nucleic acid replication/repair/synthesis
3.2.15	Thioredoxin, glutaredoxin, glutathione	Amino acid biosynthesis

3.2.16	Biotin carboxyl carrier protein (BCCP)	Other
3.2.17	Ferredoxin	Electron transfer
3.2.2	Biotin	Other biosynthesis
3.2.3	Cobalamin	Other biosynthesis
3.2.4	Enterochelin	Other biosynthesis
3.2.5	Folic acid	Other biosynthesis
3.2.6	Heme, porphyrin, cytochrome biogenesis	Electron transfer
3.2.7	Lipoate	Aerobic respiration/TCA cycle
3.2.8	Menaquinone, ubiquinone	Electron transfer
3.2.9	Molybdopterin	Protein biosynthesis/modification
3.3.0	Central intermediary metabolism	Intermediary metabolism
3.3.1	2'-Deoxyribonucleotide metabolism	Nucleic acid replication/repair/synthesis
3.3.10	Nucleotide hydrolysis	Nucleic acid replication/repair/synthesis
3.3.11	Nucleotide interconversions	Nucleic acid replication/repair/synthesis
3.3.12	Oligosaccharides	Other biosynthesis
3.3.13	Phosphorus compounds	Other biosynthesis
3.3.14	Polyamine biosynthesis, Putrescine	Amino acid biosynthesis
3.3.15	Pool, multipurpose conversions of intermediate metabolism	Intermediary metabolism
3.3.16	S-adenosyl methionine	Amino acid biosynthesis
3.3.17	Salvage of nucleosides and nucleotides	Nucleic acid replication/repair/synthesis
3.3.18	Sugar-nucleotide biosynthesis, conversions	Nucleic acid replication/repair/synthesis
3.3.19	Sulfur metabolism	Intermediary metabolism
3.3.2	Amino sugars	Amino acid biosynthesis
3.3.20	Amino acids	Amino acid biosynthesis
3.3.21	Other	Other
3.3.22	Nitrogen metabolism (urease)	Nitrogen metabolism
3.3.3	Entner-Douderoff	Aerobic respiration/TCA cycle
3.3.4	Gluconeogenesis	Aerobic respiration/TCA cycle

3.3.5	Glyoxylate bypass	Aerobic respiration/TCA cycle
3.3.6	Incorporation metal ions	Protein biosynthesis/modification
3.3.7	Misc. glucose metabolism	Aerobic respiration/TCA cycle
3.3.8	Misc. glycerol metabolism	Aerobic respiration/TCA cycle
3.3.9	Non-oxidative branch, pentose pwy	Aerobic respiration/TCA cycle
3.4.0	Degradation of small molecules	Degredation protein/amino acids
3.4.1	Amines	Amino acid biosynthesis
3.4.2	Amino acids	Amino acid biosynthesis
3.4.3	Carbon compounds, Cachetuate catabolism	Intermediary metabolism
3.4.4	Fatty acids	Other biosynthesis
3.4.5	Other	Other
3.5.0	Energy metabolism, carbon	Aerobic respiration/TCA cycle
3.5.1	Aerobic respiration	Aerobic respiration/TCA cycle
3.5.2	Anaerobic respiration	Other
3.5.3	Electron transport, cytochromes + redox	Electron transfer
3.5.4	Fermentation	Other
3.5.5	Glycolysis	Aerobic respiration/TCA cycle
3.5.6	Oxidative branch, pentose pathway	Aerobic respiration/TCA cycle
3.5.7	Pyruvate dehydrogenase	Aerobic respiration/TCA cycle
3.5.8	TCA cycle	Aerobic respiration/TCA cycle
3.5.9	ATP-proton motive force	Transporters
3.6.0	Fatty acid biosynthesis	Other biosynthesis
3.6.1	Fatty acid and phosphatidic acid biosynthesis	Other biosynthesis
3.7.0	Nucleotide biosynthesis	Nucleic acid replication/repair/synthesis
3.7.1	Purine ribonucleotide biosynthesis	Nucleic acid replication/repair/synthesis
3.7.2	Pyrimidine ribonucleotide biosynthesis	Nucleic acid replication/repair/synthesis
3.8.0	Secondary metabolism	Intermediary metabolism
3.8.1	Polyketide synthases	Intermediary metabolism

3.8.2	Non-ribosomal peptide synthases	Intermediary metabolism
4.0.0	Cell envelope	Cell membrane/envelope
4.1.0	Membrane/exported/lipoproteins	Cell membrane/envelope
4.1.1	Inner membrane	Cell membrane/envelope
4.1.2	Murein sacculus, petidoglycan	Cell membrane/envelope
4.1.3	Outer membrane	Cell membrane/envelope
4.1.4	Surface polysaccharide & antigens, EPS	Cell membrane/envelope
4.1.5	Surface structures	Cell membrane/envelope
4.1.6	G+ membrane	Cell membrane/envelope
4.1.7	G+ exported/lipoprotein	Cell membrane/envelope
4.1.8	G+ surface anchored	Cell membrane/envelope
4.1.9	G+ peptidoglycan, teichoic acid	Cell membrane/envelope
4.2.0	Ribosome constituents	Ribosomal protein
4.2.1	Ribosomal and stable RNAs	Ribosomal protein
4.2.2	Ribosomal proteins - synthesis, modification	Ribosomal protein
4.2.3	Ribosomes - maturation and modification	Ribosomal protein
5.0.0	Extrachromosomal	Nucleic acid replication/repair/synthesis
5.1.0	Laterally acquirred elements	Other
5.1.1	Colicin-related functions	Other
5.1.2	Phage-related functions and prophages	Other
5.1.3	Plasmid-related functions	Other
5.1.4	Transposon-related functions	Other
5.1.5	Pathogenicity island-related functions	Other
6.0.0	Regulation	Regulation
6.1.0	Two component system	Regulation
6.1.1	Sensor kinase	Regulation
6.1.2	Response regulator	Regulation
6.1.3	Sensor kinase/response regulator	Regulation

	fusion	
6.2.0	RNA polymerase core enzyme binding	Nucleic acid replication/repair/synthesis
6.2.1	Sigma factor	Regulation
6.2.2	Anti sigma factor	Regulation
6.2.3	Anti sigma factor antagonist	Regulation
6.3.0	Defined families	Other
6.3.1	AsnC?	Other
6.3.10	DeoR?	Other
6.3.11	LuxR? (GerR)	Other
6.3.12	MerR?	Other
6.3.13	ArsR?	Other
6.3.14	PadR?	Other
6.3.15	LytR	Other
6.3.2	AraC?	Other
6.3.3	GntR?	Other
6.3.4	IclR?	Other
6.3.5	LacI?	Other
6.3.6	LysR?	Other
6.3.7	MarR?	Other
6.3.8	TetR?	Other
6.3.9	ROK	Other
6.4.0	Protein kinases	Amino acid biosynthesis
6.4.1	Serine/threonine	Amino acid biosynthesis
6.4.2	Tyrosine	Amino acid biosynthesis
6.5.0	Others	Other
7.0.0	Not classified (included putative assignments)	Other
7.1.1	DNA sites, no gene product	Other
7.2.1	Cryptic genes	Other
8.0.0	Unique to P. fluorescens	

Supplementary Tables 2,3,4: Raw proteomics data

Due to the size of the datasets from the MS/MS analysis, the raw data has been made available online at https://rhizosphere.org/lab-page/open-access-data/

Supplementary Table 2: Free-living FixA

Supplementary Table 3: Free-living FixB

Supplementary Table 4: Bacteroid FixB