Investigating the links between N_2 fixation and metabolism in rhizobia

Nick Crang

St Hugh's College

University of Oxford

A thesis submitted for the degree of Doctor of Philosophy

Trinity Term 2019

I declare that this is my own account of my research and that this work has not been submitted for a degree at any other university. I would like to acknowledge that in Chapter 4 the bioinformatic analysis used to identify the suppressor mutant SNP was carried out by Dr Beatriz Jorrin of the Oxford Department of Plant Sciences, and the transmission electron microscopy of my prepared bacteroid samples was done by Dr Euan James of the James Hutton Institute.

Abstract

In this thesis the metabolic flow of carbon and electrons enabling N_2 fixation in the agriculturally relevant model organisms *Rhizobium leguminosarum* and *Azorhizobium caulinodans* was investigated. Building upon functional genomic data-sets including transcriptomics, metabolomics and mutagenic screens, the core metabolic reactions needed for dicarboxylate ORS571 and RL3841. Isocitrate dehydrogenase was found to be essential for N_2-fixation and operation of the TCA-cycle in RL3841, contradicting previously published INSeq data. This was due to TA insertions in gene termini misleading the Hidden Markov Model used to asses INSeq data. As a result of this, new bioinformatic pipelines have been developed to avoid such errors. The redox sink poly-3-hydroxy butyrate (PHB) was shown to be non-essential for N_2 fixation in ORS571. Previously published results showing PHB synthesis to be essential are actually due to missregulation of *phaR* translation in *phaC* mutants used to block PHB synthesis, rather than the loss of PHB as a redox store. This bring ORS571 in line with other species of rhizobia, where PHB defective mutants retained their ability to fix N_2. Two alternative metabolic pathways, the malonate shunt in ORS571 and the isocitrate shunt in RL3841 have both been investigated and shown to be inoperative at biologically relevant levels in their respective organisms. In ORS571, the entirety of the carbon flux between succinate and acetyl-CoA has been demonstrated to proceed via NAD^+ dependent malic enzyme (DME) and phosphoenolpyruvate carboxykinase, allowing existing models of carbon flux to be revised to remove the erroneous malonate shunt. In addition, DME has been demonstrated to be essential for N_2 fixation, but dispensable for growth on succinate, indicating a disconnect between the requirements of N_2 fixation and dicarboxylate metabolism.

Acknowledgements

Well where do I begin? This work would not have been possible without the contributions of a great many people after all. However, I have to start somewhere and without a doubt that should be with my supervisor, Philip Poole.

Phil, we've certainly had our ups and downs over the years, but your in-depth knowledge of metabolism and undimmed passion for your work and life in general have been a constant throughout my time here. This project would not have been possible without your advice and guidance, and Oxford would not have been even close to the same without you. Thank you, Phil.

In a similar vein I have to extend my thanks to the entirety of the Poole group, though special mention must go to Alison, Bea, Carmen, Marta and Vinoy for the hours spent teaching, advising and listening to me complain. In addition, particular thanks to Rachel for convincing me to come to the Poole lab in the first place, it's a recommendation I've never regretted following. Well, not seriously anyway!

Thank you also to the mysterious cookie leaver, whoever you may be, for a pick me up at just the right moment. Many thanks to Chris, for much needed Sunday lunch, sympathy and chemdraw expertise.

My thanks also to Helen and Lida, god knows how we got lucky enough to have you in our lab, but I'm certainly grateful that we did!
Thanks to Khushboo and Rachel, upon who's preliminary work this thesis has been built.
Thank you Daina, for helping me learn so much about myself during my time here.

In honour of *Azorhizobium caulinodans* who I've worked with so closely over the years, and in honour of Oxford university "Azorhizobium parvus spurius est"

Finally, thank you to my parents, Rob and Sam, always there with unwavering support in difficult times, even when they didn't fully understand what the latest lab disaster was. Thanks Mum, thanks Dad, I wouldn't be the man I am today without you.

Abbreviations

Ac: Acetate

AcCoA: Acetyl-CoA

ADP: Adenosine diphosphate

AKG: α-Ketoglutarate

AMP: Adenosine monophosphate

Amp: Ampicillin

ANOVA: Analysis of variance

ATP: Adenosine triphosphate

Ala: Alanine

Asn: Asparagine

Arg: Arginine

Asp: Asparate

BNF: Biological Nitrogen fixation

Cit: Citrate

CoA: Coenzyme A

Cys: Cysteine

CTP: Cytidine triphosphate

DOT dissolved oxygen tension

DCW: Dry cell weight

DOF: Degrees of freedom

DNA: Deoxyribonucleic acid

DHAP: Dihydroxyacetone phosphate

DME: NAD^+ dependant Malic enzyme

ED: Entner–Doudoroff

EMP: Embden-Meyerhof-Parnas

EPS: Exo-polysaccharide

E4P: Erythrose-4-phosphate

FAD: Flavine adenine dinucleotide

$FADH_2$: Dihydroflavine-adenine dinucleotide

Fum: Fumarate

F6P: Fructose-6-phosphate

FTHF: Formyltetrahydrofolate

G6P: Glucose-6-phosphate

Gent: Gentamycin

Glu: Glutamate

Gly: Glycine

Gln: Glutamine

GC: Gas chromatography

GC-MS: Gas chromatography mass spectrometry

GDP: Guanosine diphosphate

GFP: Green fluorescent protein

GTP: Guanosine triphosphate

GAP: Glyceraldehyde-3-phosphate

GS/GOGAT: Glutamine synthetase/ Glutamate synthase

His: Histidine

ICit: Isocitrate

Icd: Isocitrate dehydrogenase

IGV: Integrated genome viewer

Ile: Isoleucine

IMP: Inosine monophosphate

INCA: Isotopomer network compartmental analysis

INSeq: Insertion sequencing

Kan: Kanamycin

Kb: kilo bases

KDPG: 2-Keto-3-deoxy-6-phosphogluconate

LB: Luria-Bertoni

Leu: Leucine

Lys: Lysine

LL_DAP: L,L-diaminopimelate aminotransferase

LPS: Lipopolysaccharide

Mal: Malate

MSDH: Malonate semialdehyde dehydrogenase

Met: Methionine MI:

MSTFA: N-methyl-N-trimethylsilyl trifluroacetamide

MS: Mass spectrometry

N: Nitrogen

N_2: Atmospheric dinitrogen gas

NAD: Nicotinamide adenine dinucleotide

NADH: Dihydronicotinamide adenine dinucleotide

NADP: Nicotinamide adenine dinucleotide phosphate

NADPH: Dihydronicotinamide adenine dinucleotide phosphate

Nit: Nitrofurantoin

OAA: Oxaloacetate

ORS571: *Azorhizobium caulinodans*

PPP: Pentose phosphate pathway

PCR: Polymerase chain reaction

Pfam: Protein family

PEP: Phosphoenolpyruvate

PEPck: Phophoenolpyruvate carboxykinase

Phe: Phenylalanine

PG3: 3-Phosphoglycerate

PG6: 6-Phosphogluconate

PHB: Polyhydroxybutyrate

Pro: Proline

Pyr: Pyruvate

qRT-PCR: quantitative reverse transcription PCR

RL3841: *Rhizobium leguminosarum bv viciae 3841*

RNA: Ribonucleic acid

RNAseq: RNA sequencing

ROS: Reactive oxygen species

Ru5P: Ribulose-5-phosphate

R5P: Ribose-5-phosphate

SEM: Standard error of the mean

SD: SucCoA: Succinyl CoA

SDS-PAGE: sodium dodecyl sulphate poly-acrylamide gel electrophoresis

Spec: Spectinomycin

Strep: Streptinomycin

TBDMS: N-tert-butyldimethylsilyl-N-methyl trifluoroacetamide

TCA: Tricarboxylic acid cycle

Tet: Tetracyclin

TEM: Transmission electron microscopy

TIC: Total ion counts

Thr: Threonine

TME: $NADP^+$ dependent Malic enzyme

TNseq: Transposon sequencing

Trp: Tryptophan

TY: Tryptone-yeast

Tyr: Tyrosine

UMP: Uridine monophosphate

Val: Valine

X5P: Xylulose-5-phosphate

Xgal: 5-bromo-4-chloro-indolyl- β-D-galactoside

Table of Contents

Chapter 1

Introduction

1.1 Biological N_2 fixation

Biological nitrogen fixation (BNF) is the reduction of atmospheric dinitrogen (N_2) to ammonia (NH_3) by the nitrogenase enzyme complex. This process is both highly energy demanding and involves components that are irreversibly deactivated by O_2 (Oldroyd and Dixon 2014). A group of bacteria termed diazotrophs contain the FeMo nitrogenase complex, though the conditions under which they "fix" atmospheric N_2 to ammonia varies between species. Some are capable of free living fixation while others require a symbiotic partner plant to provide suitable conditions (Webb 2016). The most important form of BNF for agriculture is the association between the diazotrophs known as Rhizobia and the plants of the *Fabaceae* family. Rhizobial-plant chemical cross talk induces the formation of specialised root structures for the rhizobia to colonise, and carbon is supplied to the bacteria in the form of organic acids such as malate. These structures, termed root nodules, insulate the rhizobia from atmospheric O_2 levels through the use of leghaemoglobin. In return the rhizobia fix atmospheric nitrogen and export it to the plant in the form of ammonia (Udvardi and Poole 2013; Oldroyd et al. 2011). Given that yields of crops such as pearl millet can be doubled with the application of as little as 40kg/hectare of nitrogen fertiliser (Bationo and Ntare 2000) and BNF in some species has been estimated to be equivalent to 120 kg/hectare, the potential yield increase made possible by even a partial transfer of N_2 fixing symbiosis is of great significance for global food security (Bationo and Mokwunye 1991; Oldroyd 2012).

1.2 The need for biological N_2 fixation

As the global trend of population increase continues, there is a very real prospect of a Malthusian population crisis. To this end we need to dramatically increase food production over the next 50 years whilst simultaneously reducing our energy inputs for crop production (Tester and Langridge 2010). The last major leap in food production was made 50 years ago during the "Green Revolution" through the combination of smut resistance, dwarfing and large-scale chemical fertiliser application. Since then heavy use of chemical fertiliser has become a standard fixture of intensive farming, achieving yields undreamed of a century ago and enabling our current unprecedented population surge (Miflin 2000). This comes at a price however, as the reactive nitrogen species at the core of these fertilisers are produced

via the Haber-Bosch process. This process requires both high temperatures and pressures making it energy intensive and is providing a considerable contribution to global greenhouse gas emissions both from the CO_2 released due to the energy required to produce it, and because the N_2O released during denitrification (Shcherbak et al. 2014). The warming potential of N_2O is considerably greater than CO_2 as it has an atmospheric lifespan of more than a century giving it a global warming potential 264 times that of an equivalent -mass of CO_2 in that time-frame (IPCC 2013). Although reactive nitrogen species are required by the plants to produce nucleic acids and protein, excessive application can fail to be taken up by the plant leading to further environmental costs (Xu et al. 2012). The leaching of reactive nitrogen species into local groundwater, and run off into nearby water courses along with subsequent damage to the aquatic ecosystems are a far too common example of this (Schlesinger 2009). In addition to the environmental cost, it is a bulky commodity, making it difficult to transport to areas with poorly developed infrastructure. This is problematic given that some of the most nutrient poor soils in the world are present in sub-Saharan Africa and represent even more of a limitation on crop production than lack of water (Mueller et al. 2012). Thus, the areas that would benefit most from such fertilisers have very limited access to them (Oldroyd 2012; Tilman et al. 2011). Despite the limitation on the global distribution of chemical fertiliser, their use is still prevalent enough to perturb the global nitrogen cycle. This effect is severe enough that their reduction is one of the suggestions made by (Rockstrom et al. 2009) to achieve a "safe operating space for humanity".

One dramatic saving that can be made in both energy input and use efficiency is via the utilisation of BNF instead of synthetic nitrogen fertiliser. Even areas that are currently able to effectively make use of chemical fertiliser would benefit from the increased nitrogen uptake efficiency of BNF compared to chemical fertiliser. Rates of N uptake from chemical fertiliser applied to soil have been estimated from 5-50% uptake efficiency (Oenema et al. 2009; Raun and Johnson 1999; Erisman et al. 2007) whereas by contrast in BNF virtually all the fixed N_2 is provided directly to the plant (Dixon and Kahn 2004). BNF simultaneously saves the energy cost of synthesising the reactive nitrogen species ammonia on an industrial scale whilst minimising the loss of nitrogen compounds into the environment. This is made possible by the direct exchange of fixed nitrogen between the crop plant and their root nodule dwelling symbiotic partners. This process is currently limited by the nature of the legume-rhizobia symbiosis, to crop plants of the *Fabaceae* family, which does not include the cereal crops rice, wheat and maize. These cereal crops account for the largest proportion

of human caloric consumption so changes that increase their production have major implications for the health of the food supply rendering it both more sustainable and productive in areas where crop growth is limited by soil N content (Haberer et al. 2016; Pimentel and Paoletti 2009).

Taken together, these factors in favour of BNF could go a long way towards meeting the required levels of food production needed in the next 50 years, provided we can expand the range of crops capable of utilising it. Several efforts are underway currently to transfer this capability to non-legume crops, making use of a variety of different approaches. These include transferring the symbiosis pathways to the cereal plants, transferring N_2 fixation to existing endophytes or the plants themselves and creating N_2 fixing synthetic biology based designer bacteria (Rogers and Oldroyd 2014; Geddes et al. 2015). These goals can be made achievable only through a thorough understanding of both the N_2 fixation and the legume-rhizobia symbiosis system.

1.3 The legume-rhizobia symbioses

1.3.1 Initiation of symbiosis

The legume-rhizobia symbioses requires attachment and colonisation of the host plant roots and begins with a complex chemical cross-talk between the participants (Downie 2010). This cross-talk is essential for ensuring the plant is not tricked into admitting a pathogenic bacterium rather than a symbiosis partner. The plant secretes root exudates, such as flavonoids, into the surrounding soil often including chemo -attractants to facilitate rhizobial chemo-taxis back towards the roots. Once there, the rhizobia attach to the surface of the root, initiate root hair curling and enter the root. The rhizobia proceed towards the root cortex along a structure known as an infection thread. Currently it is unclear whether the infection thread is traversed by movement of the rhizobia or by replication of the rhizobia proceeding down the length of the infection thread. As the infection thread proceeds towards the root inner cortex cells differentiate first into the nodule primordium and then onto become the

nodule meristem. The meristem continues to divide resulting in the formation of a protective root nodule structure by the plant (Timmers et al. 1999).

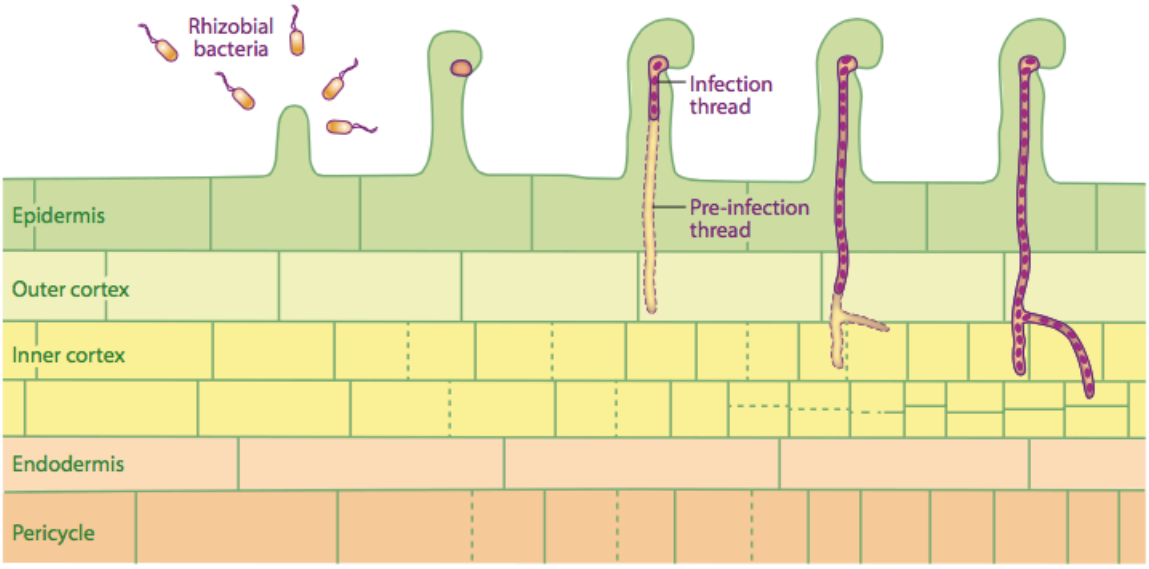


Figure 1.1 Nodulation and infection thread formation through the plant epidermis and cortex

The route of the infection thread is predicted by dense cytoplasmic subdomains with aligned cytoskeleton, known as pre-infection threads. Reproduced from (Oldroyd et al. 2011)

The rhizobia derived from the infection thread in turn differentiate, whereby they undergo a dramatic change in gene expression and phenotype becoming bacteroids. These have been labelled 'ammonioplasts' , responsible for N_2 fixation but no longer capable of free living survival due to the shutdown of genes associated with chemotaxis, growth, ribosomal protein synthesis and nucleic acid synthesis (Terpolilli et al. 2016; Oldroyd et al. 2011; Ampe et al. 2003). Whilst these survival related genes are downregulated, the *nif* and *fix* genes required for N_2 fixation are dramatically upregulated. Bacteroids are isolated from the plant by the symbiosome membrane, which enables the exchange of nutrients between the plant and the bacteroids to be controlled.

An alternative mode of bacterial infection called lateral root invasion or crack entry is also observed in some rhizobia that induce stem nodulation on the host plant. This type of intercellular colonization does not require *nod* factors. Instead cell death at the lateral root base is induced via ethylene and reactive O_2 species and infection pockets are then formed as rhizobia proliferate into the gaps produced by this targeted cell death. These are in turn

internalized in the nodule primordial cells through infection threads which require the common sym pathway (Capoen et al. 2010).

1.3.2 Determinate and indeterminate nodules

There are two primary forms of nodules that have been characterised: determinate and indeterminate (Gage 2004). Indeterminate nodules derive from inner cortical cells, have persistent meristematic activity, resulting in distinct zones of development and an elongated, cylindrical structure (Oldroyd et al. 2011; Oke and Long 1999). These types of nodules are most commonly found on plants native to temperate climates, and their N_2 fixing activity is concentrated in a single zone within the nodule. Determinate nodules by contrast are derived from the root central cortex, lose their meristematic activity shortly after nodule formation, resulting in a spherical nodule structure with no distinct developmental zones. N_2 fixation is instead distributed throughout the nodule. Determinate nodules are primarily found in tropical legumes (Oke and Long 1999; Oldroyd et al. 2011). After nodule senescence, bacteria from both types of nodule are able to recolonise the soil, ready to begin the legume infection cycle anew (Paau et al. 1980). In the case of determinate nodules this is done by de-differentiation of bacteroids within the nodules back to their free-living state, whereas indeterminate nodules rely on the release of rhizobia that remained within the infection thread (Müller et al. 2001; Paau et al. 1980; Williams et al. 2008).

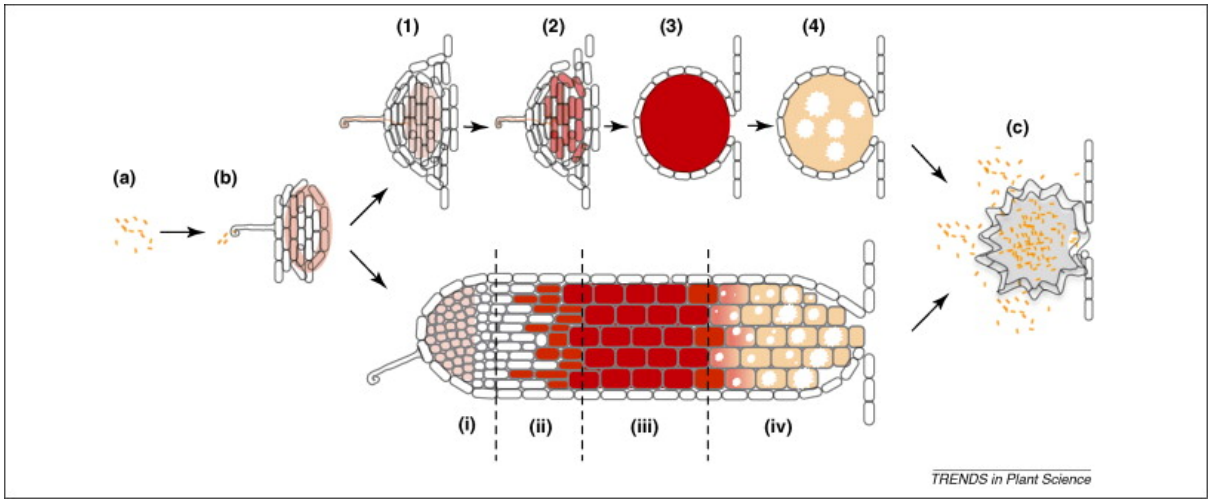


Figure 1.2 Development of determinate and indeterminate nodules

(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 (central cortex or inner cortical central cortex respectively) 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 recolonize the rhizosphere. Reproduced from (Schumpp and Deakin 2010)

Regardless of which method of entry, or type of nodule is formed during symbiosis, nodule formation is controlled by a combination of host, rhizobia and environmental factors including environmental stresses such as soil salinity and drought, as well as host responses such as levels of ROS production or rhizobial production of antioxidants (Chang et al. 2009; D'Haeze et al. 2003).

1.3.3 The nitrogenase reaction

Once N_2 fixation is underway within the nodules, using the nitrogenase reaction, the bacteroids convert stable atmospheric N_2 into biologically active NH_3, which is exported from the bacteroids and into the plant. The nitrogenase reaction is catalysed by the nitrogenase complex, which is controlled by the *nifHDK* genes. The complex itself consists of a MoFe containing nitrogenase, encoded by *nifD* and *nifK,* and a nitrogenase reductase encoded by *nifH*. Nitrogenase is a heterotetramer consisting of two pairs of metallocluster

containing proteins; the P–cluster and the iron molybdenum FeMo-co cluster. Nitrogenase reductase is homodimer containing a single [4Fe-4S] cluster (Igarashi and Seefeldt 2003). The dinitrogenase reductase transfers electrons to the dinitrogenase component at the expense of MgATP to fuel the reduction of N_2 to NH_3.

The overall stoichiometry of the nitrogenase reaction is shown in Equation 1.1 This reaction is highly reductive and energy intensive in nature, requiring eight electrons and sixteen ATP molecules per molecule of N_2 fixed. H_2 is evolved as part of the nitrogenase reaction and represents a considerable loss of efficiency for the reaction, as energy generated from plant supplied photosynthate is wasted on this H_2 evolution (Allen et al. 1991). To compensate for this, many rhizobia express hydrogenases enabling them to generate ATP through the oxidation of the evolved H_2, recapturing some of this lost energy and maximising the flow of electrons towards reduced N_2 (Schubert and Evans 1976). This hydrogenase system has also been suggested as an O_2 protection mechanism for the nitrogenase complex, in addition to preventing the accumulation of H_2 at the active site of the nitrogenase complex (Allen et al. 1991).

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

Equation 1.1 The nitrogenase reaction

N_2 fixation reaction catalysed by nitrogenase enzyme (Dixon and Kahn 2004)

When rhizobia establish a symbiosis within legume root nodules they face a major challenge in balancing the high demand for ATP required by the N_2 fixation process with the minimal O_2 levels necessitated by the nitrogenase complexes' O_2 sensitivity, whose iron-sulphur cluster becomes denatured when allowed to react with O_2 (Ludwig 2004; Udvardi and Poole 2013).

Indeed, such is the O_2 sensitivity of the nitrogenase complex that ORS571 remains the only rhizobial species yet discovered that is capable of growth under atmospheric O_2 levels whilst meeting all of its N_2 needs via BNF (Dreyfus et al. 1983; de Vries et al. 1984).

1.3.4 Nodule adaptation to enable effective N_2 fixation

The nodule therefore is required to provide an environment that enables bacteroids to fix N_2 effectively by buffering the concentration of free O_2 using an O_2 carrying haemoprotein 'leghaemoglobin' (Ott et al. 2005). In addition to this buffering, diffusion of O_2 is disrupted by the physical barrier that is the nodule outer layers, as well as the physical consumption of O_2 by the mitochondria of the plant (Udvardi and Poole 2013). Combined together, these result in a concentration of free O_2 measured in the nanomolar range (Hunt and Layzell 1993; Batut and Boistard 1994; White et al. 2007) and stimulates the onset of both and the transcription of the 'nif' and 'fix' genes required for N_2 fixation as these are regulated by the O_2 status of the cell and N_2 fixation (Mus et al. 2016; Fischer 1994). In order to cope with the microaerobic environment a bacteroid specific haem-copper terminal cbb3-oxidase FixNOPQ is expressed, which unlike other rhizobial oxidases, has a Km low enough to allow it to operate effectively despite the minimal O_2 available. In this way, the proton motive force and ATP production are optimised under these low O_2 conditions (Mandon et al. 1994).

Transport of the ammonia out of the bacteroid is believed to occur via passive diffusion in its neutral lipophilic form (Mus et al. 2016). As specific ammonia transport channels are down-regulated in bacteroids, it is believed that this diffusion may occur directly through the cell membranes, or through non-specific protein channels (Udvardi and Poole 2013). Once the ammonia enters the symbiosome space between the bacteroids and the plant, it is believed to become protonated to NH_4^+ by the acidic environment present there. It is understood that there are two possible routes to cross the symbiosome membrane and enter the plant, either via a discrete NH_3 channel or through a cation channel capable of moving K^+, Na^{2+} or NH_4^+ (Mus et al. 2016). After crossing the symbiosome membrane the ammonia can be transported freely throughout the plant via the xylem or phloem.

In exchange for this secreted ammonia, photosynthates, primarily the C_4 dicarboxylate malate are exported from the plant to the bacteroids (Udvardi and Poole 2013). The bacteria can utilise these carbon sources to provide the energy needed for both N_2 fixation as well as the cell survival (White et al. 2007) These dicarboxylates are exported across the symbiosome membrane where they are taken up by the bacteria via a high affinity dicarboxylic acid transport system (Finan et al. 1981, 1983; Ronson et al. 1984). Three

primary routes are available in rhizobia to metabolise carbon sources: the tricarboxylic acid (TCA) cycle, the Entner-Doudoroff (ED) pathway (a rhizobial variant of the glycolytic pathway) and the pentose phosphate pathway (Poole and Allaway 2000). The C_4 dicarboxylates provided by the plant feed into the TCA-cycle, generating ATP and reductant molecules as it cycles between carbon compounds.

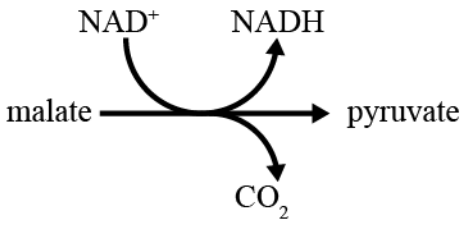
In addition to the exchange of carbon and nitrogen compounds across the symbiosome membrane, most rhizobia are also reliant on plants for the provision of key components of the nitrogenase complex. Both iron and sulphur as well as the cofactor homocitrate must be transported from the plant to the rhizobia, making the majority of rhizobia reliant on the plant for N_2 fixation to occur, though ORS571 is capable of free-living N_2 fixation in part because it contains a functional *nifV* (Clarke et al. 2014; Lee et al. 2008). Other key cellular components such as amino acids, O_2 and phosphorous also have to cross the symbiosome membrane (White et al. 2007; Terpolilli et al. 2012; Brear et al. 2013).

1.4 Free living and symbiotic rhizobial metabolism

Major differences are readily apparent between the metabolism of rhizobia growing in the free-living state compared to those in symbiosis. When grown in free living conditions rhizobia utilise a wide variety of carbon sources, and can be classified by their preferred carbon source. Fast growing rhizobia utilise a wide range of carbon sources including hexoses, pentoses, di-saccharides, tri-saccharides and organic acids, while slow growers exhibit optimum growth on sugar alcohols and aromatic compounds (Stowers 1985). By contrast, in symbiotic conditions the primary carbon sources available to all types of rhizobia are C_4 dicarboxylates. Whilst hexoses, pentoses and saccharides are primarily utilised by the ED and pentose phosphate pathway, usage of dicarboxylates is dependent on the combined activities of the TCA-cycle and malic enzyme (Mckay et al. 1988). As a consequence of this, the TCA-cycle is considered essential for N_2 fixation in most rhizobia including *R. leguminosarum* (Terpolilli et al. 2016). Mutagenesis studies indicate that a variation on the cycle utilising alternatives to the known enzymes may function in *Bradyrhizobium japonicum* as isocitrate dehydrogenase and 2-ketoglutarate dehydrogenase mutants retain symbiotic N_2 fixation (Shah and Emerich 2006; Green and Emerich 1997).

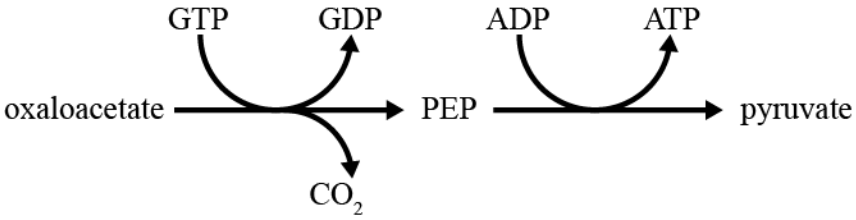
As the dissolved O_2 tension (DOT) faced by rhizobia in bacteroids is less than 20 nanomolar, it becomes impossible to couple glycolysis and oxidative phosphorylation through the conventional TCA cycle (Ludwig 2004). It is suggested that although the TCA-cycle coupled to oxidative phosphorylation is the primary source of ATP generation under aerobic conditions, when under microaerobic conditions it operates as a whole at greatly reduced levels. This minimal flow through the TCA-cycle is required to produce necessary metabolic intermediates. In a process known as 'oxidative metabolic gearing', under microaerobic conditions the primary flow is from malate towards the polyhydroxyalkanoates, generating partially oxidised intermediates that accumulate within the cells rather than being excreted like fermentation products (Ludwig 2004). This gearing system allows the rapid adaptation of rhizobia to growth in aerobic and micro-aerobic conditions with the minimal change in gene expression and co-factor synthesis. This is possible because the partially oxidised intermediates generated under microaerobic conditions can be fully oxidised and utilised as energy sources under aerobic conditions. Although this metabolic gearing allows the rate of substrate consumption to remain constant across a wide range of DOTs, the rate of oxidative phosphorylation will be reduced at low DOTs resulting in an increase in generation time under low DOT conditions (Ludwig 2004).

In order for the TCA-cycle to operate it requires the provision of acetyl-CoA and the reoxidation of reduced co-factors e.g. NADH to NAD^+. In free-living conditions hexoses and saccharides can be catabolised to generate pyruvate via the ED pathway, which can in turn allow the production of acetyl-CoA through the activity of the pyruvate dehydrogenase complex. In symbiotic conditions however, C_4 dicarboxylates are the primary carbon source provided to the bacteroids. To generate acetyl-CoA from these substrates they must first be converted to pyruvate. This can occur either through the actions of malic enzyme (Equation 1.2), or the combined activity of phosphoenolpyruvate (PEP) carboxykinase and pyruvate kinase (Equation 1.3), which result in the production of pyruvate from malate and oxaloacetate respectively. As the TCA cycle requires a supply of acetyl-CoA to function and since malate is a TCA-cycle intermediate there is a split in processing, whereby some is converted to pyruvate to provide acetyl-CoA for the cycle, and some is utilised in the TCA-cycle directly.



Equation 1.2 Malic enzyme

Malic enzyme activity for pyruvate generation



Equation 1.3 PEPck & Pyruvate kinase

Combined PEP carboxykinase and pyruvate kinase activity utilised for pyruvate generation

Different forms of malic enzyme have been identified, one NAD^+ dependent (DME) and the other $NADP^+$ dependent (TME). In *S. meliloti* the TME, encoded by *tme,* was found to be none essential for N_2 fixation, whereas DME, encoded by *dme* was essential (Driscoll and Finan 1997). While in *S. meliloti* the activity of PEP carboxykinase (PEPck) alone is unable to provide enough pyruvate to enable fixation in the absence of DME, in RL3841 limited N_2 fixation is observed when either PEPck or DME is mutated (Finan 1991; Mulley et al. 2010). It was also observed that rates of fixation from PEPck mutants were greater than those from DME mutants, suggesting that DME is the primary pathway utilised for pyruvate production (Mulley et al. 2010).

In addition to the TCA-cycle, gluconeogenesis has been suggested to have a key role in symbiotic N_2 fixation, so some of the reduction in rates of N_2 fixation observed in the PEPck mutants may be due to the lack of gluconeogenic activity (Udvardi and Poole 2013).

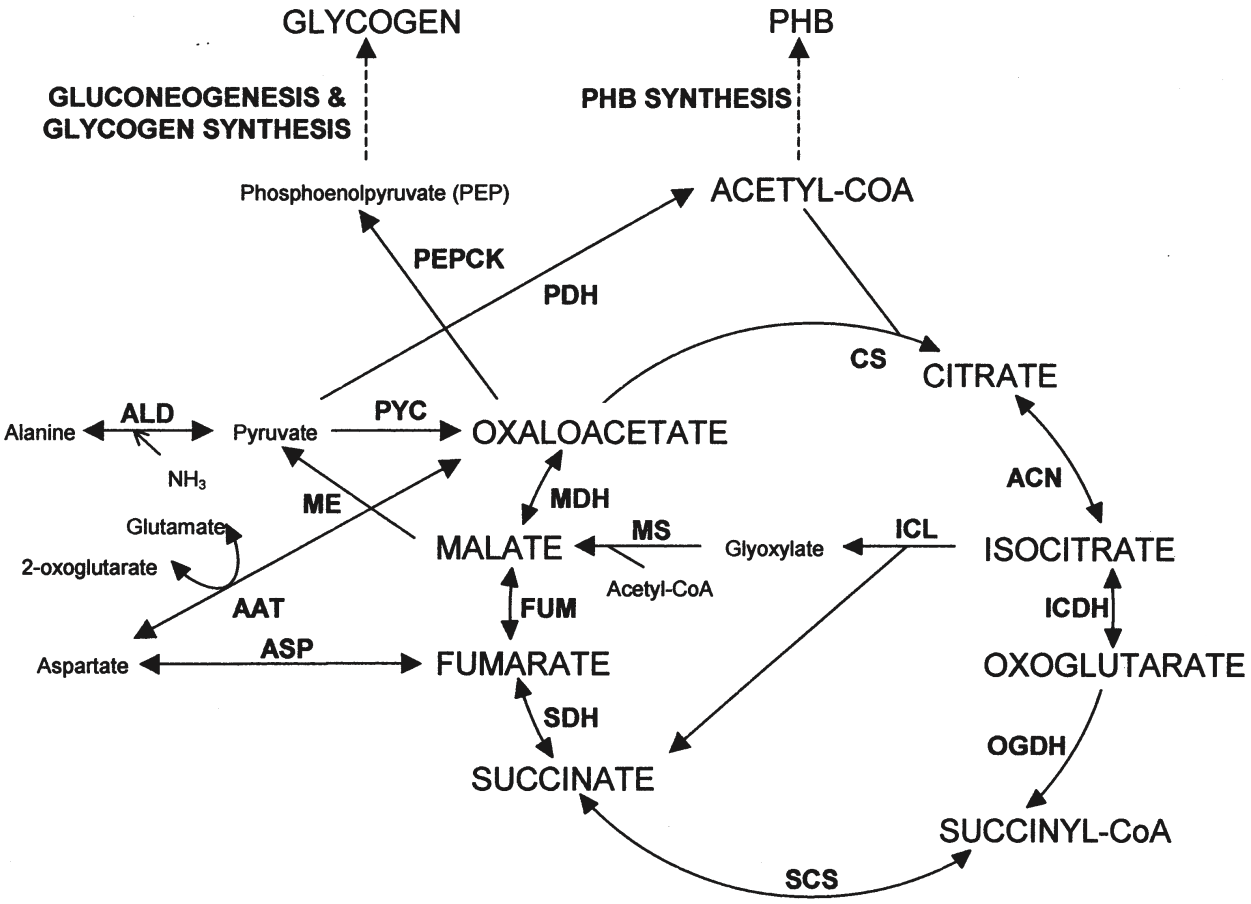


Figure 1.3 The TCA-cycle and possible pathways of regulation in the bacteroid

Arrows indicate enzymatic reactions, metabolites are labelled with full names, enzymes are given acronyms. **AAT,** Aspartate aminotransferase; **ACN,** Aconitase; **ALD,** Alanine dehydrogenase; **ω-AM,** ω-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. Reproduced from (Lodwig and Poole 2003)

1.5 Nitrogen metabolism

Free-living rhizobia take up N sources from the environment and assimilate them for use through either glutamate dehydrogenase or the combined activity of glutamine synthetase and glutamate synthase (GS/GOGAT). The GS/GOGOAT pathway has been found to be favoured under N limiting conditions in *R. leguminosarum, R. trifolli* and *R. japonicum* (Brown and Dilworth 1975) however during N_2 fixation this pathway is down-regulated suggesting bacteroids are unable to assimilate N. As a result of this, the bacteroids become auxotrophic for amino acids and require them to be supplied by the host plant during symbiosis (Udvardi and Poole 2013; Mulley et al. 2011). Because of this amino acid auxotrophy, mutation of amino acid transporters Aap and Bra leads to defective growth and bacteroid development in *R. leguminosarum* (Hosie et al. 2002).

1.6 Storage polymers

1.6.1 Synthesis of storage polymers

Although most carbon provided to bacteroids is utilised for energy production, a great deal is used for the production of carbon storage molecules such as lipids, glycogen or poly-3-hydroxybutyrate (PHB) (Lodwig et al. 2005; Trainer and Charles 2006). As well as providing a carbon store for times of stress, synthesis of these storage polymers serves as a useful sink for excess reductant, such as that resulting from aerobic metabolism under O_2-limiting micro-aerobic conditions (Mandon et al. 1998; Trainer and Charles 2006; Terpolilli et al. 2016). Under micro-aerobic conditions NAD(P)H generated by central carbon metabolism increases as there is limited O_2 to act as an electron acceptor. This build-up of NAD(P)H leads to inhibition of NADP Isocitrate dehydrogenase and citrate synthase, halting the activity of the TCA-cycle (Senior and Dawes 1971; Jackson and Dawes 1976). This is avoided, by instead channelling the electrons from the NAD(P)H towards synthesis of carbon storage polymers like PHB. The preferred storage polymer varies between rhizobial species, with glycogen synthesis mutants producing no reduction in symbiotic efficiency in *R. leguminosarum*, whilst the same mutants in *S. meliloti* resulted in reduced

N_2 fixation (Wang et al. 2007; Lodwig and Poole 2003; Lodwig et al. 2005). In a similar manner, PHB synthesis has been shown to be essential for N_2 fixation in ORS571 (Mandon et al. 1998) but non-essential in *R. leguminosarum* (Lodwig et al. 2005). As carbon storage polymer synthesis requires both carbon and energy sources there is the potential for their synthesis to compete with N_2 fixation for resources.

1.6.2 PHB metabolism

PHB is a carbon storage molecule used by rhizobia such as *Sinorhizobium meliloti* as a contingency against times of carbon limitation and accumulates in many species when under O_2 limitation (Ratcliff and Denison 2010; Trainer and Charles 2006). PHB is one of the three terminal products that acetyl CoA can be converted into alongside lipid and citrate, indicating that a lack of PHB would have a major effect on the metabolic flux of the cell. It is however possible that these alternative reductant pathways will able to compensate for the lack of PHB as seen in *R. etli* (Cevallos et al. 1996).

PHB is synthesised by the condensation of two acetyl-CoA molecules to form acetoacetyl-CoA which is subsequently reduced to form β-hydroxybutyryl-CoA. These residues are the monomers which are polymerized to form the polymer PHB (Mandon et al. 1998). The accumulation of PHB takes place in the form of storage granules and can be large enough for pure granules of PHB within the cells to be visualised by electron micrographs (Terpolilli et al. 2016).

While PHB production is stimulated by growth under nitrogen limiting, carbon excess conditions, PHB catabolism can be stimulated by growth under carbon limiting conditions or by the addition of ammonia or other alternative sources of N (Stam et al. 1986). It has been observed that PHB is not degraded in ORS571 under O_2 limiting growth conditions even in the absence of alternative carbon sources, indicating that the catabolism of PHB is aerobic. It has been suggested that the PHB store may provide a competitive benefit both in the rhizosphere and during the initial colonisation of the host legume (Stam et al. 1986).

In addition to the obvious competitive benefit of a carbon store in times of starvation, there is evidence that it provides protection from other stresses such as osmotic shock and UV irradiation (Kadouri et al. 2003). It also serves a key role in balancing the redox state during

N_2 fixation by serving as a sink for the oxidation of excess reductant (Terpolilli et al. 2016). Although not all rhizobial species accumulate PHB during symbiosis, previous work in ORS571 has suggested it is required for maintaining the reducing power necessary for N_2 fixation as a *phbC* mutant was found to be unable to fix under any conditions or to produce *nifA* (Trainer and Charles 2006; Mandon et al. 1998). Although many nodules were still produced, they were all white clearly lacking the protective leghaemoglobin of active nodules. However, ORS571 is an exceptional case, as N_2 fixation in most species of rhizobia is unaffected by the loss of PHB synthesis (Wang et al. 2007; Quelas et al. 2013; Lodwig et al. 2005). It has even been observed in *R. etli* that PHB mutants, although displaying decreased competitiveness, had an increased rate of symbiotic N_2 fixation rate, which has been suggested to represent the increased usage of N_2 fixation as an electron sink in the absence of the major sink represented by PHB (Cevallos et al. 1996) .

1.7 R. leguminosarum biovar. viciae 3841

1.7.1 Taxonomy

Rhizobium leguminosarum is an α-proteobacteria from the family *Rhizobiaceeae* and the genus *Rhizobium* closely related to *Agrobacterium tumefaciens. Rhizobium leguminosarum bv. viciae 3841*(RL3841) is capable of symbiotic N_2 fixation when in association with its host legume *Pisum sativum, (*garden pea) though multiple strains of *R. leguminosarum* have been isolated, with differing symbiosis partners. The genetic differences between these strains are largely confined to their accessory plasmids, which have been found to vary in number from 2-6 (Laguerre et al. 1993, 1992). The symbiosis plasmid also varies in both size and replication group. An excellent example of Sym plasmid variation effect are *R. leguminosarum bv. viciae* strain *A34* which associates with *Pisum sativum* and *R. leguminosarum bv. Phaseoli* which associates with *Phaseolus vulgaris*. Both strains are chromosomally identical, differing only in the genes on their respective symbiosis plasmid (Sym plasmids) (Downie et al. 1983). Although RL3841is often used as a representative of *R. leguminosarum* due to its sequenced genome, it lacks some genes found in other strains of *R. leguminosarum* such as the *hup* genes encoding the hydrogen uptake system. Optimal growth of this strain is seen at 28 ◦C on TY and UMS media.

1.7.2 Genetic organisation

R. leguminosarum bv. viciae 3841 has a 7.75 Mb genome that comprises a circular chromosome and six circular plasmids, pRL12 (0.87 Mb), pRL11 (0.68 Mb), pRL10 (0.49 Mb), pRL9 (0.35 Mb), pRL8 (0.15 Mb) and pRL7 (0.15 Mb), containing a total of 7,263 protein encoding genes, ~65% of which are chromosomal. All six plasmids contain the same type of replication system based on *repABC*. The essential protein-encoding genes are largely chromosomal, but most functional classes occur on plasmids as well. Nearly all the symbiosis and N_2 fixation genes (*nod, nif, fix*) are found on the 0.48 Mb plasmid pRL10, though a second copy of the *cbb3* high affinity terminal oxidase encoded by *fixNOPQ* is also found on pRL9 along with three of five homologs of the regulator *fixK***.** Plasmid pRL7 contains a very high proportion of pseudogenes, mobile elements and uncharacterised protein coding genes. The genes for the TCA-cycle, ED and PP pathway are primarily present on the chromosome, along with DME, PEP carboxykinase and pyruvate kinase (Young et al. 2006)

1.8 Azorhizobium caulinodans ORS571

1.8.1 Taxonomy

A. caulinodans is an α-proteobacterium from the family *Xanthobacteraceae* the genus *Azorhizobium* most closely related to the genus *Bradyrhizobium* amongst the rhizobia (Tsukada et al. 2009; Elmerich et al. 1982). Despite this relation, *A. caulinodans* grow much more rapidly than *Bradyrhizobium.* Unlike other rhizobia, ORS571 do not metabolise sugars or alcohols instead utilising organic acids as their optimum carbon source (Dreyfus et al. 1988; Pauling et al. 2001). In liquid media the cells are small bacilli, 1-2 microns long, and produce a flagellum. ORS571 also displays a remarkable tolerance for a range of temperatures from 12- 44 ◦C, though optimum growth in the lab is usually seen at 37 ◦C (Scott and Ludwig 2004).

1.8.2 Genetic organisation

ORS571 has a single large circular chromosome (5.37 Mb) containing 4714 protein coding genes (Lee et al. 2008). These genes encode for proteins including the functions of biofilm formation, chemotaxis, N_2 fixation, metabolism, symbiotic interactions, synthesis of surface poly-saccharides, regulation and transport. The genes for the TCA-cycle, ED and PP pathway are present, along with DME, PEP carboxykinase and pyruvate kinase. ORS571 cannot perform glycolysis due to a lack of phospho-fructokinase. In addition to DME, a second malic enzyme (Azc_0119) is annotated in ORS571, though this has strictly limited observed activity as a purified enzyme (Zhang et al. 2012).

Unusually amongst rhizobia, ORS571 contains a functional copy of the homocitrate synthase gene *nifV* which means that ORS571 is not reliant on a symbiotic partner to supply homocitrate, allowing ORS571 to synthesise a functional nitrogenase complex without establishing a symbiosis. As well as genes for N_2 fixation, ORS571 contains the *hup, hyp* and *hoxA* genes enabling it to oxidise the H_2 produced as a by-product of N_2 fixation in order to produce additional ATP (Lee et al. 2008).

Within the chromosome there is a discrete 87.6 Kb 'symbiosis island' containing the *nod* genes, along with genes involved in chemotaxis, amino acid uptake and a type IV secretion system. No additional plasmids are present within ORS571 (Elmerich et al. 1982; Lee et al. 2008).

1.8.3 N_2 fixation

As previously stated, ORS571 is capable of synthesising a functional nitrogenase complex without a symbiotic partner thanks to a functional homocitrate synthase gene *nifV*. Homocitrate is a critical co-factor for the formation of MoFe cluster. This allows ORS571 to fix N_2 in free living low O_2 conditions when no alternative source of N is present. A duplication event has resulted in two *nifH* genes within ORS571, with *nifH1* primarily utilised during free-living N_2 fixation and *nifH2* primarily for symbiotic N_2 fixation (Norel and Elmerich 1987; Iki et al. 2007). Expression of the N_2 fixation genes is controlled in a manner analogous with that of *S. meliloti* via the activity of the transcriptional activator NifA, which is in turn controlled by a combination of extracellular O_2 levels, via the activity

of the *fixLJK* genes, and both intra and extra-cellular N via the two component nitrogen regulatory proteins NtrBC and NtrXY (Kaminski et al. 1991; Pawlowski et al. 1991; Lee et al. 2008).

In free living conditions maximal N_2 fixation rates were observed at a DOT in the 10 µM range, with fixation rates decreasing when below 1 µM (Gebhardt et al. 1984). This raises questions as to why the O_2 level in nodules is buffered down to 0.02 µM (Bergersen et al. 1986) and why differing levels of N_2 fixation are observed at the same O_2 level in symbiotic vs free-living N_2 fixation. The reason for this difference is suggested to be the use of different micro-aerobic respiratory pathways in free-living vs symbiotic conditions. Of the four terminal oxidases present within ORS571 (*cytaa3, cytd, cyto* and *fixNOPQ)* both *fixNOPQ* and *cytd* were found to be essential for fixation in symbiotic conditions whereas in free-living conditions *fixNOPQ* is dispensable (Mandon et al. 1994). In both free-living and symbiotic N_2 fixation, ORS571 expresses hydrogenases, encoded by the *hup* and *hyp* gene cluster, increasing the efficiency of N_2 fixation by allowing the generation of ATP from the oxidation of H_2 evolved during the nitrogenase reaction (de Vries et al. 1984; Das and Lodha 1998).

1.8.4 Symbiosis

ORS571 is capable of symbiotic N_2 fixation in partnership with the fast-growing annual tropical host legume *Sesbania rostrata* (Dreyfus et al. 1988; Suzuki et al. 2007), either forming root nodules through the classic *nod* factor dependent mechanism, or producing stem nodules via a form of crack entry. Stem nodules are a characteristic adaptation of *S. rostrata* to the waterlogged soils in which it thrives. Several genes have been found to be indispensable for successful symbiosis and nodulation including the *nod* genes, *nif*, *fix*, lipopolysaccharide and exopolysaccharide biosynthesis, chemotaxis, sulphur and carbon uptake genes. All such key genes were found to be highly up-regulated in bacteroids compared to free living conditions (Suzuki et al. 2007; Tsukada et al. 2009).

Prior studies have identified no difference in the N_2 fixation efficiency of either stem or root nodules and their fixation rates compare favourably with that of soybeans leading to the use of *S. rostrata* for crop rotation as a green manure in rice paddies (Borah 2017; Lee et al. 2008; Dreyfus et al. 1985).

In addition to this effective symbiosis, ORS571 is also able to colonise other plant species including *Arabidopsis thaliana, Brassica napus,* wheat and *Oryza sativa* using the *nod*-independent crack entry method, though these do not result in N_2 fixation (O'Callaghan et al. 2000; Gough et al. 1997; Sabry et al. 1997; Jain and Gupta 2003). This indicates that ORS571 is currently capable of forming only endophytic relationships with non-legume plants.

1.9 Functional genomic approaches

In order to effectively engineer an organism, for example to extend a symbiosis or introduce the ability to fix N_2, a clear understanding is needed of how these desired phenotypes are generated. The falling cost of DNA sequencing provided by next generation sequencing technologies has resulted in an excess of DNA sequence data for many organisms, including the rhizobial species RL3841 and ORS571 (Lee et al. 2008; Young et al. 2006). These genome sequences allow us to identify potential genes, and even suggest functions for them based on homology to existing known sequences. However, these annotations are not always accurate and many 'hypothetical protein' coding genes are found where no known function can be inferred.

In biological systems, there are multiple stages of information flow starting from the DNA, progressing through RNA to proteins and metabolites finally ending at an observed phenotype (Figure 1.4). A phenotype can be regulated or otherwise affected by changes at any stage of this information flow. Functional genomics is the term given to the investigative approaches that seeks to combine together information from these stages with DNA sequencing information to allow us to identify genes that encode functionally important proteins and how these interact. There are four broad categories of investigative techniques for functional genomics; genomics investigates at the DNA level, transcriptomics at the RNA level, proteomics at the protein level and metabolomics, at the metabolite level (Lardi and Pessi 2018; van Opijnen and Camilli 2013; Gray et al. 2015).

At the genetic level, studies operate by producing genetic changes and observing the resulting change in phenotype. A classic example of this is the inference of gene-function via the generation of loss-of-function mutants. Transcriptomic approaches instead focus on the quantification of RNA transcripts, allowing the determination of the conditions under which a gene or set of genes is active or the response of gene expression to changing conditions (Lardi and Pessi 2018). Proteomic approaches allow for similar investigations of changing expression, but instead focus on actual protein content rather than RNA transcripts. Metabolomics on the other hand looks at metabolite levels within the system demonstrating the effect of enzyme activity on small molecules. A variant of metabolomics that investigates the metabolic fluxes within a system is known as fluxomics.

In order to gain an accurate understating of gene function, information from these different approaches must be integrated. This can prove difficult to interpret as the information, for example on gene expression levels, will not necessarily agree with the proteomic data observed. These ambiguities can provide more details about the system when correctly interpreted. For example, if a transcriptomic study indicated a gene was highly expressed under low O_2 conditions, but little protein were found under those same conditions, it would suggest that this gene was regulated post-transcriptionally. Other ambiguities could also arise due to genetic redundancy or protein degradation, again providing additional information about the system.

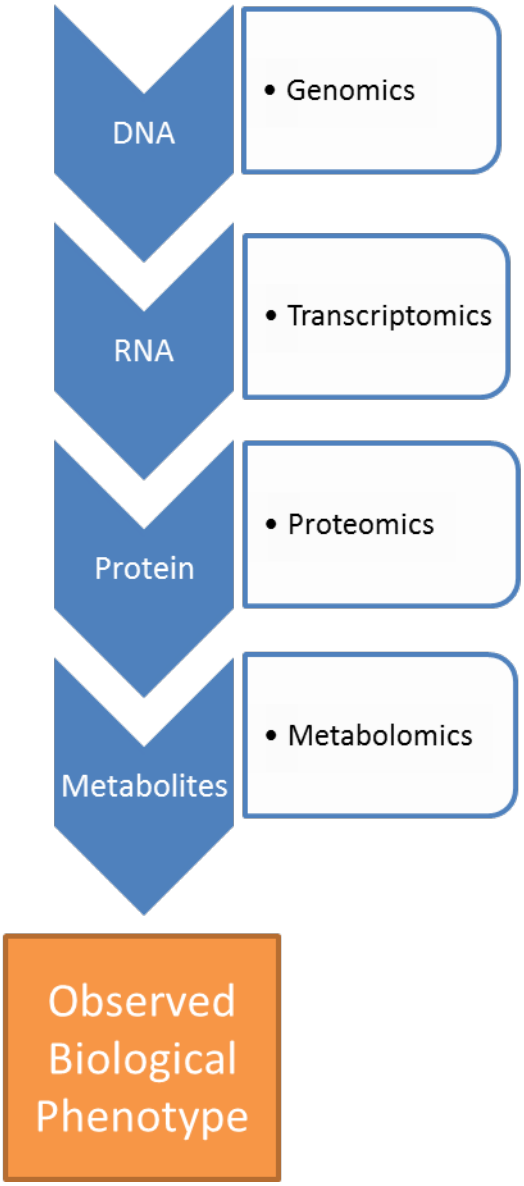


Figure 1.4 Information flow in a biological system

Overview of the accepted information flow in a biological system, starting from DNA and ending with an observable phenotype. Stages are indicated by blue chevrons and the systems biology approach used to investigate the stage is indicated by the companion box.

1.9.1 Genetic approaches

Genetic approaches aim to deduce gene function through direct manipulation at the DNA level. This is most classically seen in the generation of loss-of-function mutants via single gene mutagenesis. The roles of many key genes have been found this way including the C_4 dicarboxylic acid transport system (Dct) of rhizobia, where loss of function of single genes e.g. *dctA* resulted in a strain unable to transport C_4 dicarboxylates into the bacteroids and defective in N_2 fixation (Finan et al. 1983; Reid and Poole 1998; Finan et al. 1988; Engelke et al. 1987). Such single gene mutations can be generated by a variety of techniques but generally fall into one of two categories', single integration or double integration mutagenesis.

In a single integration mutant, a section of non-coding DNA is introduced into the coding region of a gene resulting in a disruption of the gene and usually a non-functional gene product. Such mutants can be generated rapidly, but can easily be lost through DNA recombination unless there is a selection pressure to maintain them. This pressure is often applied through the introduction of an antibiotic resistance cassette as part of the DNA used to disrupt the target gene. This selection pressure cannot always be maintained however, such as when rhizobia are applied to plants, rendering it impractical for certain studies.

Double integration mutants by contrast result in the complete deletion of a target DNA sequence preventing the possibility of spontaneous DNA recombination producing wild type reversion mutants. This is the gold standard for mutagenesis, but is slightly more labour intensive than single integration mutagenesis and cannot be adapted for whole genome mutant screening.

Although useful, single gene mutagenesis studies are time and labour intensive and they cannot easily identify functional genes when there is genetic redundancy i.e. multiple gene copies or multiple potential pathways which serves to mask the phenotypic effect of the mutation. This is a particularly common problem for metabolic mutants due to the multiplicity of alternative pathways available to compensate for the loss of single mutations.

An approach that is becoming more common is whole genome screening using libraries of mutants (Lardi and Pessi 2018). These mutant libraries can be generated in a number of

ways including chemical and UV mutagenesis, however, a particularly powerful new tool is the use of transposons to generate libraries of single integration mutants (Barquist et al. 2013). The transposons generate single integration mutations at predicable sites throughout the genome based on the transposon utilised. The DNA sequence flanking the transposon can be used to identify where in the genome the mutation has occurred. Through this transposon system, libraries of mutants can be made containing disruptions of every gene in the genome of the target organism. These libraries can then be grown in conditions of interest and DNA extracted from the starting library and the post selection library. These can be sequenced using next-generation sequencing technologies and the number of mutants of each gene present before and after selection can be compared. In this way the effect of every gene in genome for growth under the target condition can be determined (Perry et al. 2016). This is a process known as Tn-Seq, though many variants exist depending on the transposon utilised. In some cases barcoded transposon libraries have been generated (termed Bar-Seq), allowing genes to be identified via Sanger sequencing rather than next gene sequencing (Wetmore et al. 2015).

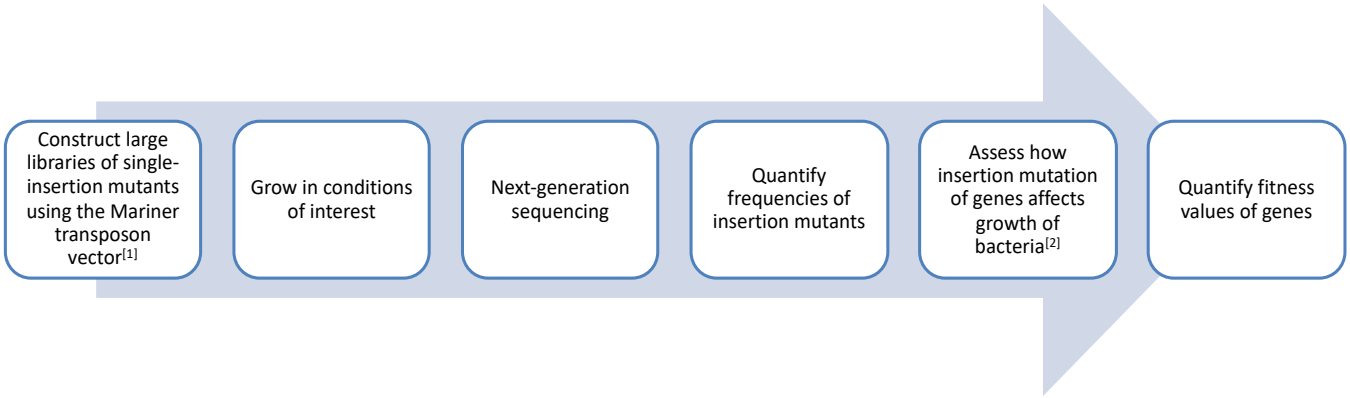


Figure 1.5 Work flow of INSeq method

Reproduced from (Wheatley 2018).

A TnSeq variant of particular relevance to this work makes use of the mme-I Tn5 *mariner* transposon to generate the mutant library in a process known as INSeq (Figure 1.5). This type II restriction endonuclease cuts downstream of the *mariner* transposon, taking a 16 bp flanking region of the insertion site. This tag can then be used to identify the site of insertion (Perry et al. 2016; Perry and Yost 2014). A suicide vector is used to introduce the transposon to ensure only a single insertion takes place. The *mariner* transposon has been found to produce unbiased, stable mutants in the *Rhizobiaeceae* family (Perry and Yost 2014).

In order to determine the fitness values of each gene, the frequency of each mutation is compared before and after selection to see if there is a difference between the observed and expected number of reads per gene. This expected number of reads is calculated based on sequencing depth, number of insertion sites per gene, and the mutant library size. A range of statistical tools are used to take into account the number of insertion sites available within each gene, to prevent genes with few available insertion sites being mislabelled as having a greater fitness effect (Wang et al. 2009). The statistical models generally seek to assign these fitness value to the genes and sort them into categories based of 'essential' and 'non-essential' though the use of Bayesian calculations (DeJesus et al. 2013). Though limited to a binary view of essentiality, such models are still extremely powerful and can even identify the specific domains within a gene that are essential (DeJesus et al. 2013). There are some more nuanced forms that identify genes as 'essential', 'growth defective', ' growth neutral' and 'growth advantage' by making use of a four state Hidden Markov Model (HMM) (Wheatley et al. 2017).

The HMM splits the sequenced genome into alternating sequences of essential and non-essentials regions, in which a sequence of observed values (the Tn insertions within insertion sites) is explained by the underlying state sequence (the 'essentiality' of the insertion region). This model makes use of read counts for individual insertion sites to infer the probability distribution over states and infer the most likely state sequence, including a conditional probability of a state dependent on the previous neighbouring site. This serves to smooth the read count data, reducing the effect of outlier data. For example insertion sites in non-essential regions which contain no insertions are tolerated because their neighbouring sites do contain insertions (DeJesus and Ioerger 2013). A stated advantage of the HMM is that it allows data from insertion sites in the N and C termini of genes to be utilised, which is discarded in other models due to the uncertainty of the phenotypic effect of insertions in these region (Gawronski et al. 2009).

The assignment of genes to the four states of essentiality (essential, growth defective, non-essential, growth advantage), relies on user assigned boundaries set for read count likelihood functions. Growth-essential genes will have no, or very few, insertions, growth-neutral genes will sit within a mean parameter of reads between the boundaries of growth advantage and growth defective. The growth advantage state is set five time over the mean read count for a statistically significant stretch of insertion sites, while the growth defective boundary

is less than 1/100th of the mean read count for a statistically significant stretch of insertion sites.

Regardless of the method chosen, all statistical models will make some assumptions, most commonly that the insertion occurs randomly and that gene count in the sequencing is proportional to the number of mutants present in the library (DeJesus et al. 2017).

Mutant library screens are powerful tools, but they also have drawbacks such as the strong selection in favour of second site mutants, and the possibility of bias in the initial library against low fitness mutants. It is also difficult to target subsets of genes for investigation using this approach as the mutagenesis methods result in mutations throughout the genome (Wheatley 2018).

1.9.2 Transcriptomic approaches

As previously stated, transcriptomic approaches can be used to identify which genes are actively expressed under specific conditions. The transcriptome refers to all the expressed genes within a cell at a given time, and this will vary with the growth conditions. Transcriptomic analysis can result in comparative data sets, allowing the relative expression of different genes to be compared, providing an indication of the relative importance of a gene under the stated condition. This transcriptional data can prove particularly useful in instances where there are multiple genes encoding proteins with similar functions, as the gene most highly expressed under the condition of interest is likely to be more important in that condition than the alternative copies. The most common approaches for investigating transcription are microarrays, qRT-PCR and RNA-Seq (Schena et al. 1995; Wang et al. 2009).

All techniques require the initial conversion of mRNA within the cell to stable cDNA. Both microarrays and qRT-PCR make use of gene specific probes allowing their expression level to be assessed. Both techniques have different uses, with qRT-PCR being more time and labour intensive, it is generally used to investigate a small number of genes at a time, whereas a microarray is a screening tool used to identify the transcription level of all known genes in an organism under the condition of interest. However, as each probe needs to be designed to target a specific gene, these techniques can only be used to investigate known gene transcripts. Additionally, the probes may also prove to be ineffective if poorly

designed and can give inaccurate measures of gene expression level if hybridisation of the cDNA sequence to the probe is poor, or off target.

Microarrays are rarely used now that the cost of sequencing has fallen, however many such data sets exist for rhizobia, providing valuable gene expression data for central carbon metabolism and colonisation amongst other things (Karunakaran et al. 2009; Ramachandran et al. 2011) .

RNA-Seq by contrast does not make use of pre-designed probes and instead sequences the entire cDNA transcriptome of the organism under the condition of interest, allowing the expression of previously unknown genes to be measured (Wang et al. 2009). Transcriptomes are available for RL3841 and ORS571 (Roworth 2017; Rachwal et al. 2015).

1.9.3 Proteomic approaches

Proteomics is the umbrella term for investigation of proteins, with the proteome representing all the proteins within a cell at a given time. This is not a fixed entity, and the composition of the proteome will vary with the conditions the cell is exposed to (Graves and Haystead 2002). In addition to providing information on protein expression level, a proteomic investigation also allows the identification of protein post-translational modifications, as well as protein sub-cellular localisation and interaction partners. Post-translational modification can profoundly affect the functionality of a protein, and cannot be identified by genomic or transcriptomic studies (Graves and Haystead 2002). Although the transcriptome and proteome can provide similar data, this is not guaranteed to be so as there are additional factors that affect the expression of proteins which do not affect the expression of gene transcripts such as post-transcriptional regulation.

In the past the proteome of a cell was determined by 2D gel electrophoresis, which allowed the proteins to be separated by both size and isoelectric charge. Relative abundance of proteins within the cell could be visualised via western blot or protein staining of the gel and the protein constituents could be examined by peptide fragment analysis using mass-spectrometry (Lardi and Pessi 2018). This approach has been used to investigate the symbiosis of *Sinorhizobium meliloti* and *Medicago* and was able to demonstrate the metabolic shift that took place in *S. meliloti* upon bacteroids formation is reflected in an

increase in N_2 fixation proteins, stress-related proteins, chaperonins and amino acid transporters (Natera et al. 2000; Djordjevic 2004; Djordjevic et al. 2003)

A more sensitive form of proteomic analysis can be achieved using liquid chromatography to purify the proteins rather than 2D gel electrophoresis combined with tandem mass spectroscopy analysis in a 'shotgun proteomics' approach (Washburn et al. 2001).

1.9.4 Metabolomic approaches

Metabolites are small molecular intermediates and products of metabolism and the metabolome is the combined total of metabolites in a cell at a particular time. As with the transcriptome and proteome this will vary with the time and growth conditions of the organism. A metabolomics study aims to investigate the presence and abundance of metabolites in a system under a set condition (Lardi and Pessi 2018). A metabolic profile of an organism will give a snapshot of the physiological state of the cell in that moment, providing a useful addition to transcriptomic and proteomic data. When the three are combined you have a record of what genes are active, at what relative levels and what effects these active genes have resulted in within the cell.

Mass spectrometry and nuclear magnetic resonance (NMR) are commonly used techniques to both quantify metabolites and identify their structures. Both techniques can be used for either targeted or untargeted metabolomics. In targeted metabolomics, the focus is on the measurement of a defined subset of chemically characterised metabolites, to investigate particular processes such as the TCA-cycle or gluconeogenesis. In untargeted metabolomics there is no such filter, and the entire metabolome is sampled in an unbiased manner (Lardi and Pessi 2018). This approach has the advantage of allowing the discovery of novel metabolites, and gaining new information about unexplored systems (Scalbert et al. 2009), though it lacks the resolution available to a targeted metabolomics investigation.

Steady state metabolic flux analysis is a form of metabolomics investigation that looks to identify the flow of a defined target molecule through a system, allowing the relative utilisation of this by different pathways to be observed. Such studies have been conducted in defined sub-section of plants, mammalian cell cultures and a wide range of prokaryotes

(Allen and Young 2013; Roscher et al. 2000; Young 2013; Hiller and Metallo 2013; Walther et al. 2012; Jiang et al. 2016; Kruger and Ratcliffe 2009). An example of this using isotopic ^{13}C labelled succinate is seen in both aerobically growing RL3841 (Terpolilli et al. 2016) and in ORS571 (Borah 2017). In ORS571 this was done in aerobic, micro-aerobic, free-living N_2 fixing and symbiotic N_2 fixing conditions. In both studies, the relative enrichment of different metabolites with ^{13}C was used to inform a model of metabolic flux in their target species under their target conditions. This flux map suggested the relative utilisation of known metabolic pathways under these different sets of conditions, and indicated dramatic changes in metabolism took place for each condition, such as a greatly reduced TCA-cycle flux during symbiosis (Borah 2017). Such maps can be extremely useful for metabolic engineering as they can suggest likely competing side reactions, as well as the primary pathways utilised to generate metabolites of interest. They have been used to great effect in this way in the crop species *Zea mays* where the primary flux supplying both carbon and NADPH to fatty acid metabolism, key for the accumulating seed storage reserves of fatty acid, was identified as NADP dependent malic enzyme (Paula Alonso et al. 2010).

Although such flux maps have been used to identify novel pathways, notably the GAS pathway in *Mycobacterium tuberculosis* (Beste et al. 2011), they are generally limited by both our knowledge of the metabolic pathways present within an organism, and the resolution of the labelling compound. When metabolites are utilised by multiple pathways, it can be very difficult if not impossible to resolve which pathways are being utilised.

1.10 Research objectives

Maintenance of redox balance is crucially important in rhizobia, both for growth and for enabling effective N_2 fixation. In this work we aim to dissect the relative importance of the TCA-Cycle and redox storage for enabling N_2 fixation in rhizobia. ORS571 is an excellent model organism for such investigation as N_2 fixation can be measured both in free living and in symbiotic conditions. The latter aim requires that the pathway of dicarboxylate metabolism is properly established in ORS571 so metabolic flux maps can be properly constructed. RL3841 was also targeted for investigation due to its agricultural importance and the intriguing hypothesis generated by the recently acquired INSeq data of (Wheatley 2018).

1.10.1 Chapter 3: Role of Isocitrate dehydrogenase activity in RL3841

The research objective for this chapter was to follow up on the INSeq results of (Wheatley et al. 2017) that indicated isocitrate dehydrogenase (ICD) was non-essential for growth in either free-living or symbiotic conditions for RL3841. These results had been partially validated with a single integration mutant and together with the results of (McDermott and Kahn 1992; Shah and Emerich 2006) suggested the presence of a TCA-cycle variant that did not require ICD activity which we had termed the isocitrate dehydrogenase shunt. Our intention here was to fully characterise the phenotype of an *icd* mutant, and subsequently identify members of the shunt using a combination of INSeq to identify candidate genes and targeted mutagenesis to validate these results.

1.10.2 Chapter 4: Re-evaluation of the role of poly-3-hydroxybutyrate in ORS571

In this chapter we investigated a novel phenotype observed from a *phaC* mutant strain (OPS0865) of ORS571 defective in PHB production and free-living N_2 fixation as well as displaying greatly reduced in growth on multiple carbon sources. This strain underwent *in planta* selection resulting in restored N_2 fixation in both free-living and symbiotic conditions, along with restoration of growth to a near wild type growth rate, whilst remaining defective in PHB production. A combination of bioinformatics, mutagenesis, qRT-PCR and translational fusions were used to elucidate the causal mutation for the improved growth and restored N_2 fixation observed in this suppressor mutant strain (OPS0921).

1.10.3 Chapter 5: Investigation of the existence and potential role of the malonate shunt in ORS571

A pyruvate independent, malonate semialdehyde dehydrogenase dependent pathway proposed by (Pauling et al. 2001) and termed the 'malonate shunt' was indicated by the metabolic flux analysis of (Borah 2017) to be highly utilised in symbiotic N_2 fixing conditions, but not free living N_2 fixing conditions in ORS571. We aimed to identify the components of this malonate shunt, and investigate the separation of C_4 dicarboxylate metabolism from the requirements of N_2 fixation in this strain.

Chapter 2

Materials & Methods

2.1 Bacterial strains, plasmids and primers

2.1.1 Strains

The bacterial strains used in this thesis are listed in Table 2.1.1. All strains were stocked in 15% glycerol and flash-frozen in liquid nitrogen for storage at -80 ºC.

Table 2.1.1 Strains used in this work.

Species	**Strain**	**Description**	**Chapter used in**	**Reference**
E. coli	DH5 alpha	Competent E. coli strain for use in transformations, carrying the following mutations; *F-deoR endA1 recA1 relA1 gyrA96 hsdR17(rk-mk+) supE44 thi-1 - phoA Δ(lacZYA-argF) U169 Φ80lacZΔM15 λ.*	Chapter 3-5	Bioline
E. coli	BL21 (DE3)	Competent *E. coli,* widely used for T7 based protein expression	Chapter 3, 5	New England Biolabs
E. coli	803	E. coli strain for Rhizobium conjugations (alternative to DH5alpha) hsdS (rk-,mk-,gal-,met-) In this work contained pRK2013 helper plasmid	Chapter 3-5	(Wood 1966)1
Rhizobium leguminosarum	3841	*Rhizobium leguminosarum* biovar *viciae* 3841 (Rlv3841); Str^r derivative of strain 300. Nodulates *Viciae* family legumes such as pea (*Pisum sativum*) and Vetch (*Vicia cracca*)	Chapter 3	(Johnston and Behringer 1975)
Azorhizobium caulinodans	ORS571	*WT Azorhizobium caulinodans*	Chapter 4-5	(Dreyfus et al. 1983)
Rhizobium leguminosarum	OPS1561	pK19 mutant of the *RL2631* Isocitrate Dehydrogenase gene of R. leguminosarum 3841. Generated from pK19mob based plasmid containing an *icd* fragment amplified with oxp1411 + 1412 Mapped with primers oxp1413/1414 and pK19A/B.	Chapter 3	(Wheatley 2018)

Rhizobium leguminosarum	OPS1620	Single integration mutant on LB of RL3841 *icd* gene (*RL2631*) inserted an omega Spec cassette. Generated using plasmid pOPS1116 + sucrose counterselection. Sequence mapped with oxp0929+pOT Forfarr Enzymatically confirmed 60% ISDH activity	Chapter 3	This work
Azorhizobium caulinodans	OPS0865	pK19mob mutant of *phbC* (pOPS0391 in ORS571). Sequence mapped with oxp1043 +pK19A/B	Chapter 4	(Borah 2017)
Azorhizobium caulinodans	OPS0921	OPS0865 (*phaC* pK19 mutant) variant purified from Sesbania root nodules, Fix + phenotype in free living and symbiosis Isolate 1. Presence of pOPS0391 confirmed by sequence mapping with oxp1043 + pK19B	Chapter 4	This work
Azorhizobium caulinodans	OPS0922	OPS0865 (*phaC* pK19 mutant) variant purified from Sesbania root nodules, Fix + phenotype in free living and symbiosis Isolate 2 Presence of pOPS0391 confirmed by sequence mapping with oxp1043 + pK19B	Chapter 4	This work
Azorhizobium caulinodans	OPS0923	OPS0865 (*phaC* pK19 mutant) variant purified from Sesbania root nodules, Fix + phenotype in free living and symbiosis Isolate 3 Presence of pOPS0391 confirmed by sequence mapping with oxp1043 + pK19B	Chapter 4	This work
Azorhizobium caulinodans	OPS1298	pK19 mutant of *phaA* (Azc_0306), an acetyl coA acetyl transferase (pOPS0894 into ORS571). Confirmed by sanger sequencing with pK19B	Chapter 4	This work
Azorhizobium caulinodans	OPS1426	plasmid pOPS1029 conjugated into ORS571. Sequence mapped with pOT For	Chapter 4	This work
Azorhizobium caulinodans	OPS1427	plasmid pOPS1030 conjugated into ORS571. Sequence mapped with pOT For and pOT Rev	Chapter 4	This work
Azorhizobium caulinodans	OPS1432	plasmid pOPS1029 conjugated into OPS0865. Sequence mapped with pOT for and pOT Rev	Chapter 4	This work

Azorhizobium caulinodans	OPS1433	plasmid pOPS1030 conjugated into OPS0865. Sequence mapped with pOT for and pOT Rev	Chapter 4	This work
Azorhizobium caulinodans	OPS2077	OPS0865 cured of its integrated pK19mob via multiple sub culturing on TY, confirmed via PCR across the *phaC* (PHB synthase) gene with oxp3394+3395 (1900bp in wild type)	Chapter 4	This work
Azorhizobium caulinodans	OPS1951	Omega Spec mutant of malic enzyme like gene *Azc_0119,* generated using pOPS1240, sequence mapped from gDNA in forward and reverse direction using oxp3095/3096 and pot Forfar	Chapter 5	This work
Azorhizobium caulinodans	OPS2025	pOPS1348 conjugated into ORS571 to complement a DME malic enzyme mutant (*Azc_3656*). Sequence mapped using primers oxp0283 and 0284	Chapter 5	This work
Azorhizobium caulinodans	OPS2026	pOPS1348 conjugated into OPS1448 to complement a DME malic enzyme mutant (*Azc_3656*). Sequence mapped using primers oxp0283 and 0284	Chapter 5	This work
Azorhizobium caulinodans	OPS2027	pOPS1348 conjugated into OPS1952 to complement a DME malic enzyme mutation in a double mutant background (*Azc-0119* and *Azc_3656*). Sequence mapped using primers oxp0283 and 0284	Chapter 5	This work
Azorhizobium caulinodans	OPS2028	pOPS1347 conjugated into ORS571 to complement a malic enzyme like mutant (*Azc_0119*). Sequence mapped using primers oxp0283 and 0284	Chapter 5	This work
Azorhizobium caulinodans	OPS2029	pOPS1347 conjugated into OPS1951 to complement a malic enzyme like mutant (*Azc_0119*). Sequence mapped using primers oxp0283 and 0284	Chapter 5	This work

Azorhizobium caulinodans	OPS2030	pOPS1347 conjugated into OPS1952 to complement a malic enzyme like mutation (*Azc_0119*) in a double mutant background (*Azc_0119* and *Azc_3656*). Sequence mapped using primers oxp0283 and 0284	Chapter 5	This work
Azorhizobium caulinodans	OPS2089	Markerless deletion of *Azc_4119* gene using sucrose selection from pOPS1421, confirmed via PCR across the gene with primers oxp2335 +2336 (800bp in mutant, 2200bp in wild type)	Chapter 5	This work
Azorhizobium caulinodans	OPS2090	Markerless deletion of *Azc_3656* DME malic enzyme using sucrose selection from pOPS1354 conjugated into ORS571, confirmed via PCR across the gene with oxp3390+3391 (300bp band in mutant, 2500 in wild type)	Chapter 5	This work
Azorhizobium caulinodans	OPS2091	Markerless deletion of PEP carboxykinase using sucrose selection from pOPS1354 conjugated into ORS571, confirmed via PCR across the gene with oxp3767 +3768(800bp band in mutant, 2700 in wild type)	Chapter 5	This work
Azorhizobium caulinodans	OPS2092	Markerless double mutant of malonate semialdehyde dehydrogenase mutant (*Azc_0575*, *Azc_4119*), generated by conjugation and sucrose selection of pOPS1352 into OPS2089, confirmed by PCR across the *Azc_0575* gene using oxp3947+3948 (200bp in mutant)	Chapter 5	This work
Azorhizobium caulinodans	OPS2093	Markerless double mutant of *Azc_3656* and PEP carboxykinase, generated by conjugation and sucrose selection of pOPS1421 into OPS2090 and confirmed by PCR across the PEP gene with oxp3767+3768 (800bp in mutant, 2700 in wild type)	Chapter 5	This work
Azorhizobium caulinodans	OPS2094	Double mutant, Markerless *Azc_3656* DME malic enzyme and omega spec *Azc_0119* DME-like enzyme generated by conjugation and sucrose	Chapter 5	This work

		selection of pOPS1240 into OPS2090confirmed by PCR into the omega cassette with pOT Forfar and oxp3091/3094		
Azorhizobium caulinodans	OPS2095	Empty pOGG093 conjugated into ORS571 (pL1V-Lv1-TetAR-RK2-par-ELT4 (plasmid for level 1 golden gate cloning) (Checked by sequencing MMS 04/03/2019)	Chapter 5	This work
Azorhizobium caulinodans	OPS2096	*Azc_3656* DME malic enzyme gene and native promoter from *A. caulinodans (*pOPS1348) conjugated into ORS571. Sequence mapped using primers oxp0283 and 0284	Chapter 5	This work
Azorhizobium caulinodans	OPS2097	*PEPck* gene +native promoter from *A. caulinodans* (pOPS1547) conjugated into ORS571. Sequence mapped using primers oxp0283 and 0284	Chapter 5	This work
Azorhizobium caulinodans	OPS2098	*Azc_3656* DME malic enzyme gene and native promoter from *A. caulinodans* (pOPS1348) conjugated into the markerless malic enzyme mutant OPS2090. Sequence mapped using primers oxp0283 and 0284	Chapter 5	This work
Azorhizobium caulinodans	OPS2099	*PEPck* gene +native promoter from *A. caulinodans* (pOPS1547) conjugated into markerless *PEPck* mutant OPS2091. Sequence mapped using primers oxp0283 and 0284	Chapter 5	This work
Azorhizobium caulinodans	OPS2100	*Azc_3656* DME malic enzyme gene and native promoter from *A. caulinodans* (pOPS1348) conjugated into the markerless double malic enzyme + *PEPck* mutant OPS2093	Chapter 5	This work
Azorhizobium caulinodans	OPS2101	*PEPck* gene +native promoter from *A. caulinodans* (pOPS1547) *PEPck* mutant OPS2093. Sequence mapped using primers oxp0283 and 0284	Chapter 5	This work

Azorhizobium caulinodans	OPS2122	Markerless triple mutant *Azc_3656* DME malic enzyme and both malonate semialdehyde dehydrogenases (*Azc_0575* and *Azc_4119*), conjugated into OPS2092, confirmed via PCR across *Azc_3656* with oxp3990+3991 (300bp in mutant,2500 in wild type)	Chapter 5	This work

2.1.2 Plasmids

The plasmids used in this thesis are listed in Table 2.1.2.

Table 2.1.2 Plasmids used in this work.

Plasmid name	Description	Antibiotic Resistance markers	Chapter	Reference
pOPINE	T7 IPTG inducible expression vector from the POPIN series. C terminal 6xHis	Amp	Chapters 3,5	(Berrow et al. 2007)
pOGG093	pL1V-Lv1-TetAR-RK2-par-ELT4 (plasmid for level 1 golden gate cloning	Tet	Chapter 5	(Geddes et al. 2019)
pK19mob	Mobilisable vector used for integration mutagenesis;	Kan/Neo	Chapters 3,4	(Schafer et al. 1994)
pHP45Ω-Spec	pRB322 derivative carrying Ω interposon Spec resistance	Spec	Chapters 3,5	(Fellay et al. 1987)
pK19mobSacB	pK19 based mobilisable plasmid for double recombination in Rhizobium. Suicide vector on sucrose.	Kan/Neo	Chapters 3,5	(Schafer et al. 1994)
pOPS1115	1000bp left and right flanking regions of *R.L 3841* Isocitrate Dehydrogenase (RL 2631) gene (amplified with oxp2719+2720 and oxp2721+2722) InFusion cloned into smaI linearised pK19mobSacB. Sequence mapped with oxp2719 and 2722	Kan/Neo	Chapter 3	This work
pOPS1116	1000bp flanking regions of *R.L 3841* Isocitrate Dehydrogenase (RL2631) gene	Kan/Neo, Spec	Chapter 3	This work

	(amplified with oxp2719+2763 and oxp2764+2722) with internal omega Spec cassette InFusion cloned into smaI linearised pK19mobSacB. Sequence mapped plasmid using pK19A/B + pOT Forfar			
pOPS1235	RL3841 *ttuc* Tartrate dehydrogenase HD cloned into pOPINE with primers oxp3090 + oxp3091, confirmed by sequencing with POPIN mapping primers	Amp	Chapter 3	This work
pOPS1236	RL3841 *RL0995* Tartrate dehydrogenase HD cloned into pOPINE with primers oxp3085 + oxp3086, confirmed by sequencing with POPIN mapping primers	Amp	Chapter 3	This work
pOPS1237	RL3841 *pRL120629* Bifuncitonal Tartrate dehydrogenase HD cloned into pOPINE with primers oxp3081 + oxp3082, confirmed by sequencing with POPIN mapping primers	Amp	Chapter 3	This work
pOPS1239	RL3841 *leuB* Isopropylmalate dehydrogenase HD cloned into pOPINE with primers oxp3073 + oxp3074, confirmed by sequencing with POPIN mapping primers	Amp	Chapter 3	This work
pOPS1029	*phaR* translational fusion of promoter region from ORS571 and 100 bases of *phaR* (oxp2419+oxp2420) InFusion cloned into pRU1097 (linearised by amplification with oxp2421+2422) Sequence mapped with pOT for and pOT rev	Gent	Chapter 4	This work
pOPS1030	*phaR* translational fusion of promoter region from OPS0921 and 100 bases of *phaR* (oxp2419+oxp2420) InFusion cloned into pRU1097 (linearised by amplification with oxp2421+2422) Sequence mapped with pOT for and pOT rev	Gent	Chapter 4	This work

pOPS0894	pK19 mob containing internal fragment of the *A. caulinodans* acteyl coA acetyltransferase gene *Azc_0306 (phaA*). Generated using primers oxp2172 +2173. Sequence mapped with pK19A/B	Kan/Neo	Chapter 4	This work
pOPS1117	*A. caulinodans Azc_0119* potential Oxaloacetate decarboxylase cloned into the pOPINE protein expression vector via InFusion cloning, sequenced using common primers popin f and popin R	Amp	Chapter 5	This work
pOPS1118	*A. caulinodans Azc_4119* Malonate semialdehyde dehydrogenase cloned into the pOPINE protein expression vector via InFusion cloning, sequenced using common primers popin f and popin R	Amp	Chapter 5	This work
pOPS1145	*A. caulinodans Azc_0575* Malonate semialdehyde dehydrogenase cloned into the pOPINE protein expression vector via InFusion cloning, sequenced using common primers popin f and popin R	Amp	Chapter 5	This work
pOPS1240	1500bp Left and Right border of *Azc_0119* Potential Oxaloacetate Decarboxylase from ORS571 InFusion cloned into pK19mobSacB with omega spec cassette added by HindIII digestion and ligation. Confirmed by sequencing with pK19A/B + pOT Forfar	Strep, Kan/Neo	Chapter 5	This work
pOPS1347	*Azc_0119* gene and native promoter from *A. caulinodans* derived using primers oxp3392 and 3393 cloned int *BsaI* digested pOGG093 using InFusion HD cloning. Sequence mapped using primers oxp0283 and 0284. To be used for complementation of *Azc_0119* mutant	Tet	Chapter 5	This work

pOPS1348	Azc_3656 gene and native promoter from *A. caulinodans* derived using primers oxp3394 and 3395 cloned int *BsaI* digested pOGG093 using InFusion HD cloning. Sequence mapped using primers oxp0283 and 0284. To be used for complementation of Azc_3656 mutant	Tet	Chapter 5	This work
pOPS1352	~1000bp LB and RB of Azc_0575 malonate semi-aldehyde dehydrogenase gene InFusion cloned into *smaI* digested pK19mobSacB, confirmed via sequencing with pK19A + pK19B	Kan/Neo	Chapter 5	This work
pOPS1353	~1000bp LB and RB of *Azc_4119* malonate semialdehyde dehydrogenase gene InFusion cloned into *smaI* digested pK19msB, confirmed via sequencing with pK19A + pK19B	Kan/Neo	Chapter 5	This work
pOPS1354	~1000bp LB and RB of *Azc_3656* DME malic enzyme gene InFusion cloned into *smaI* digested pK19msB, confirmed via sequencing with pK19A + pK19B	Kan/Neo	Chapter 5	This work
pOPS1355	A. *caulinodans Azc_3656* DMA malic enzyme cloned into the pOPINE protein expression vector via inFusion cloning, sequenced using common primers popin f and popin R. Enzyme purified and tested	Amp	Chapter 5	This work
pOPS1421	~1000bp LB and RB of *A. caulinodans PEPck* gene InFusion cloned into *smaI* digested pK19msB, confirmed via sequencing with pK19A + pK19B	Kan/Neo	Chapter 5	This work
pOPS1547	*PEPck* + native promoter InFusion cloned into *BsaI* digested pOGG93, amplified with oxp3767+3768 and sequenced with oxp0283+0284	Tet	Chapter 5	This work

2.1.2 Primers

The primers used in this thesis are listed in Table 2.3. All primers were synthesised by Eurofins MWG Operon

Table 2.1.3 Primers used in this work.

Name	Sequence 5' →3'	Description	Chapter
oxp0283	AGCGTTCTGAACAAATCC	Antisense screening primer for pOGG093	Chapter 3,5
oxp0284	GGCGGATTTGTCCTACT	Sense screening primer for pOGG093	Chapter 3,5
pK19A	ATCAGATCTTGATCCCCTGC	Antisense screening primer for pK19mob/pK19mobSacB	Chapter 3-5
pK19B	GCACGAGGGAGCTTCCAGGG	Sense screening primer for pK19mob/pK19mobSacB	Chapter 3-5
popin F	GACCGAAATTAATACGACTCACTATAGGG	Sense screening primer for pOPIN vectors	Chapter 3,5
popin R	CCAGCCACCACCTTCTGATA	Antisense screening primer for pOPIN vectors	Chapter 3,5
pOT for	CGGTTTACAAGCATAAAGC	Sense screening primer for pOT vectors	Chapter 3,5
pOT rev	CATTTTTTCTTCCTCCACTAGTG	Antisense screening primer for pOT vectors	Chapter 3,5
pot Forfar	GACCTTTTGAATGACCTTTA	Alternative sense screening primer for pOT vectors, used to screen for presence of Ω Spec cassette	Chapter 3,5
oxp1410	TTTGGATCCGCCGAAACCGCCGCA	Antisense mapping primer for *RL2631* (*icd*)	Chapter 3
oxp1411	TTTTTGAATTCCAGAACCTGCAGAAGGCCAT	Sense primer for *RL2631* (*icd)* internal fragment, adding homology to *EcoRI* digested pK19mob for InFusion cloning	Chapter 3
oxp1412	TTTTTGAATTCCTTGTTCATGTATCTGCCTC	Antisense primer for *RL2631* (*icd*) internal fragment, adding homology to *EcoRI* digested pK19mob for InFusion cloning	Chapter 3

oxp1413	CCGGCTCGCATCGAAAA	Sense mapping primer for *RL2631* (*icd*)	Chapter 3
oxp2719	CTCTAGAGGATCCCCAGATC GACAGATACCGCCG	Sense primer for Rl2631 (*icd*) pK19mobSacB left 1000bp flanking region, adding homology to *SmaI* digested pK19mobSacB for InFusion cloning	Chapter 3
oxp2763	CCAAGCTTAAGCTTATCTGC CTCGTTCGAGCTTC	Antisense primer for *Rl2631 (icd*) pK19mobSacB left 1000bp flanking region, adding *HindIII* site to allow T4 ligation to *HindIII* digested Ω Spec cassette	Chapter 3
oxp2720	GAAAGGCCATCTGCCTCGTT CGAGCTTC	Antisense primer for Rl2631 (*icd*) pK19mobSacB left 1000 bp flanking region, adding homology to right 1000 bp flanking region for InFusion cloning.	Chapter 3
oxp2764	ATAAGCTTAAGCTTGGCCTT TCCGAGATAACATC	Sense primer for *Rl2631 (icd)* pK19mobSacB right 1000bp flanking region, adding *HindIII* site to allow T4 ligation to *HindIII* digested Ω Spec cassette	Chapter 3
oxp2721	AGGCAGATGGCCTTTCCGAG ATAACATC	Sense primer for Rl2631 (*icd*) pK19mobSacB right 1000 bp flanking region, adding homology to left 1000 bp flanking region for InFusion cloning.	Chapter 3
oxp2722	TCGAGCTCGGTACCCTTAGC ACTTCTCTCAGGCGG	Antisense primer for *Rl2631 (icd)* pK19mobSacB right 1000bp flanking region, adding homology to *SmaI* digested pK19mobSacB for InFusion cloning	Chapter 3
oxp2880	CCTTCGAACCCTCACCTGAC	Antisense mapping primer of *RL2631 (icd*)	Chapter 3
oxp3073	TCACCACCATCACAGCAGCG ATGACAGCACGCAATCTTTT CC	Sense primer for *RL4707 (leuB)* adding homology for InFusion cloning into *KpnI+HindIII* digested pOPINE	Chapter 3

oxp3074	TTAAACTGGTCTAGAAAGCTGGCGGAAAGCGTCTTGAA	Antisense primer for *RL4707 (leuB)* adding homology for InFusion cloning into *KpnI+HindIII* digested pOPINE	Chapter 3
oxp3081	TCACCACCATCACAGCAGCGATGAAAACCTACAAAATCGCCC	Sense primer for *pRL120629* adding homology for InFusion cloning into *KpnI+HindIII* digested pOPINE	Chapter 3
oxp3082	TTAAACTGGTCTAGAAAGCTGCCGAGTGCCGAAAGC	Antisense primer for *pRL120629* adding homology for InFusion cloning into *KpnI+HindIII* digested pOPINE	Chapter 3
oxp3085	TCACCACCATCACAGCAGCGATGCGTGAATATGCAATCGC	Sense primer for *RL0995* adding homology for InFusion cloning into *KpnI+HindIII* digested pOPINE	Chapter 3
oxp3086	TTAAACTGGTCTAGAAAGCTGACGTTCGACGCGTAGATC	Antisense primer for *RL0995* adding homology for InFusion cloning into *KpnI+HindIII* digested pOPINE	Chapter 3
oxp3089	TCACCACCATCACAGCAGCGATGACATCCCCATCCGGAGA	Sense primer for *pRL110153 (ttuC)* adding homology for InFusion cloning into *KpnI+HindIII* digested pOPINE	Chapter 3
oxp3090	TTAAACTGGTCTAGAAAGCTGGCGTTGGATCCAAAGATTGC	Antisense primer for *pRL110153 (ttuC)* adding homology for InFusion cloning into *KpnI+HindIII* digested pOPINE	Chapter 3
oxp3243	TCACCACCATCACAGCAGCGATGAACAAGATCAAGGTCGCCA	Sense primer for *RL2631 (icd)* adding homology for InFusion cloning into *KpnI+HindIII* digested pOPINE	Chapter 3
oxp3244	TTAAACTGGTCTAGAAAGCTCGCCATGGCCTTCTGC	Antisense primer for *RL2631 (icd)* adding homology for InFusion cloning into *KpnI+HindIII* digested pOPINE	Chapter 3
oxp2132	TTCTGGTTCCACCGTGAGC	Sense mapping primer for *phaC* suppressor mutant point mutation	Chapter 4
oxp2133	GGCCGTGAAGGGGATCCTC	Antisense mapping primer for *phaC* suppressor mutant point mutation	Chapter 4

oxp2172	TGATTACGCCAAGCTGCCGAGGCCCTTGTTGAC	Sense primer for *Azc_0306 (phaA)* internal fragment, adding homology to *HindIII* digested pK19mob for InFusion cloning	Chapter 4
oxp2173	GCAGGCATGCAAGCTATGGGCTCGCTCGACCTT	Antisense primer for *Azc_0306 (phaA*) internal fragment, adding homology to *HindIII* digested pK19mob for InFusion cloning	Chapter 4
oxp2419	GTGGAGGAAGAAAAATGGGCTTCTCTCCTGGGG	Sense primer for *Azc_0307 (phaR*) adding homology for InFusion cloning into linearised pRU1097	Chapter 4
oxp2420	TTCTTCTCCTTTACGTTCGAGGGTCACATAGGTACTG	Antisense primer for *Azc_0307 (phaR*) adding homology for InFusion cloning into linearised pRU1097	Chapter 4
oxp2421	CGTAAAGGAGAAGAACTTTTCACTGG	Antisense inverse primer for linearisation of pRU1097	Chapter 4
oxp2422	TTTTTCTTCCTCCACTAGTCTCTCT	Sense inverse primer for linearisation of pRU1097	Chapter 4
Oxp2475	GAACGGCTACGAGTGGAAC	Sense primer for qRT-PCR *Azc_2328 (gyrA)*	Chapter 4
Oxp2476	GGAGAAGGTCTGCACCATG	Antisense primer qRT-PCR *Azc_2328 (gyrA)*	Chapter 4
oxp2973	GTAGAAACGGATGATCTGC	Sense primer for qRT-PCR *Azc_0307 (phaR)*	Chapter 4
oxp2974	GTCTATGACGCGAAGAAC	Antisense primer for qRT-PCR *Azc_0307 (phaR)*	Chapter 4
oxp2975	ATGGGCTCGCTCGACC	Antisense primer for qRT-PCR *Azc_0306 (phaA)*	Chapter 4
oxp3105	ATGAGCCAGTCCACCCATG	Sense primer for qRT-PCR *Azc_0306 (phaA)*	Chapter 4
oxp2848	GTCGACTCTAGAGGATCCCCGTGTTCTGTTGGATGGGCG	Sense primer for *Azc_0575* (MSDH) pK19mobSacB left 1000bp flanking region,	Chapter 5

		adding homology to *SmaI* digested pK19mobSacB for InFusion cloning	
oxp2849	GGATCGCATCCCCGGGCGCTCGGCCTCCCG	Antisense primer for *Azc_0575* (MSDH) pK19mobSacB left 1000bp flanking region, adding homology to right flanking region N terminal end for InFusion cloning	Chapter 5
oxp2850	AGCGCCCGGGGATGCGATCCTCCCGATGAT	Sense primer for *Azc_0575* (MSDH) pK19mobSacB right 1000bp flanking region, adding homology to left border C terminal end for InFusion cloning	Chapter 5
oxp2851	TGAATTCGAGCTCGGTACCCCCTGATCCCGCCGTTC	Antisense primer for *Azc_0575* (MSDH) pK19mobSacB right 1000bp flanking region, adding homology to *SmaI* digested pK19mobSacB for InFusion cloning	Chapter 5
oxp2883	TCACCACCATCACAGCAGCGATGCGGGAGATGGGC	Sense primer for *Azc_0575* (MSDH) adding homology for InFusion cloning into *KpnI+HindIII* digested pOPINE	Chapter 5
oxp2884	TTAAACTGGTCTAGAAAGCTATTCATGGTCGGGATGACG	Antisense primer for *Azc_0575* (MSDH) adding homology for InFusion cloning into *KpnI+HindIII* digested pOPINE	Chapter 5
oxp2885	TCACCACCATCACAGCAGCGATGGACACCATCG	Sense primer for *Azc_4119* (MSDH) adding homology for InFusion cloning into *KpnI+HindIII* digested pOPINE	Chapter 5
oxp2886	TTAAACTGGTCTAGAAAGCTGGACATCACGGG	Antisense primer for *Azc_4119* (MSDH) adding homology for InFusion cloning into *KpnI+HindIII* digested pOPINE	Chapter 5
oxp2887	TCACCACCATCACAGCAGCGATGGCGGACAC	Sense primer for *Azc_0119* (DME-like enzyme) adding homology for InFusion cloning into *KpnI+HindIII* digested pOPINE	Chapter 5

oxp2888	TTAAACTGGTCTAGAAAGCTATAGTGGGGCTTGTA	Antisense primer for *Azc_0119* (DME-like enzyme) adding homology for InFusion cloning into *KpnI+HindIII* digested pOPINE	Chapter 5
oxp3091	GTCGACTCTAGAGGATCCCCGACTGGCTCAGAAGACGTTGA	Sense primer for *Azc_0119* (DME-like enzyme) internal fragment, adding homology to *HindIII* digested pK19mob for InFusion cloning	Chapter 5
oxp3092	CGGTTTACAAGCATAAAGCTTGAGCACTTTCCCATCCTGA	Antisense primer for *Azc_0119* (DME-like enzyme) internal fragment, adding homology to *HindIII* digested pK19mob for InFusion cloning	Chapter 5
oxp3093	CGGTTTACAAGCATAAAGCTCATACTCTATTCTGTCTCGAGG	Sense primer for *Azc_0119* (DME-like enzyme) adding homology for InFusion cloning into linearised pRU1097	Chapter 5
oxp3094	TGAATTCGAGCTCGGTACCCCCATGACGGACAGGTGACAA	Antisense primer for *Azc_0119* (DME-like enzyme) adding homology for InFusion cloning into linearised pRU1097	Chapter 5
oxp3095	CATACGTTCCCGGCCGAATA	Sense mapping primer for *Azc_0119* (DME-like enzyme)	Chapter 5
oxp3096	CATTGCCCGCGATCTTGATG	Antisense mapping primer for *Azc_0119* (DME-like enzyme)	Chapter 5
oxp3390	TGCCGAATTCGGATCCGGAGCCTCCACGTGGAACCGAATC	Sense primer for *Azc_3656* (DME) + native promoter adding homology to *BsaI* digested pOGG093 for InFusion cloning	Chapter 5
oxp3391	TCAACAGGAGTCCAAGAGCGCGGCTCGGCACCTTCC	Antisense primer for *Azc_3656* (DME) + native promoter adding homology to *BsaI* digested pOGG093 for InFusion cloning	Chapter 5
oxp3392	TGCCGAATTCGGATCCGGAGTGAACCGGGCGCCTCC	Sense primer for *Azc_0119* (DME-like enzyme) + native promoter adding homology to *BsaI* digested pOGG093 for InFusion cloning	Chapter 5

oxp3393	TCAACAGGAGTCCAAGAGCGGAGGACATGACACGTCTCCA	Antisense primer for *Azc_0119* (DME-like enzyme) + native promoter adding homology to *BsaI* digested pOGG093 for InFusion cloning	Chapter 5
oxp3671	CTCTAGAGGATCCCCTCAGCGGGCGAGAAAGGTCC	Sense primer for *Azc_4119* (MSDH) pK19mobSacB left 1000bp flanking region, adding homology to *SmaI* digested pK19mobSacB for InFusion cloning	Chapter 5
oxp3672	GATGACCGCGCCGTTCCTTCA	Antisense primer for *Azc_4119* (MSDH) pK19mobSacB left 1000bp flanking region, adding homology to right flanking region N terminal end for InFusion cloning	Chapter 5
oxp3673	GAACGGCGCGGTCATCGGGTCA	Sense primer for *Azc_4119* (MSDH) pK19mobSacB right 1000bp flanking region, adding homology to left border C terminal end for InFusion cloning	Chapter 5
oxp3674	TCGAGCTCGGTACCCCGATTTGCCCACCGTCGCC	Antisense primer for *Azc_4119* (MSDH) pK19mobSacB right 1000bp flanking region, adding homology to *SmaI* digested pK19mobSacb for InFusion cloning	Chapter 5
oxp3675	GGTCGCGGGCATGGCAATG	Sense mapping primer - *Azc_4119* (MSDH)	Chapter 5
oxp3676	ACCGCACCATCGTGACCGG	Antisense mapping primer - *Azc_4119* (MSDH)	Chapter 5
oxp3677	CTCTAGAGGATCCCCTCGTAGATGGGCAGCGGGC	Sense primer for *Azc_3656* (DME) pK19mobSacB left 1000bp flanking region, adding homology to *SmaI* digested pK19mobSacb for InFusion cloning	Chapter 5
oxp3678	GCCCTTGCGTTTGACGCCATGTCCTCC	Antisense primer for *Azc_0575* (MSDH) pK19mobSacB left 1000bp flanking region, adding homology to right flanking region N terminal end for InFusion cloning	Chapter 5

oxp3679	CGTCAAACGCAAGGGCTGCCTTC	Sense primer for *Azc_0575* (MSDH) pK19mobSacB right 1000bp flanking region, adding homology to left border C terminal end for InFusion cloning	Chapter 5
oxp3680	TCGAGCTCGGTACCCGATGCCGCGAGTGCCGTTC	Antisense primer for *Azc_0575* (MSDH) pK19mobSacB right 1000bp flanking region, adding homology to *SmaI* digested pK19mobSacB for InFusion cloning	Chapter 5
oxp3681	TGCTGGAGCCGCTCGATGG	Sense mapping primer - *Azc_3656* (DME)	Chapter 5
oxp3682	TCATCGCGGCCCCCACATC	Antisense mapping primer - *Azc_3656* (DME)	Chapter 5
oxp3685	ACCATCACAGCAGCGATGTCGAACATTTCCGAGGA	Sense primer for *Azc_3656* (DME) adding homology for InFusion cloning into *KpnI+HindIII* digested pOPINE	Chapter 5
oxp3686	CTGGTCTAGAAAGCTGGCTTCGCTCTGGGCTTCCA	Antisense primer for *Azc_3656* (DME) adding homology for InFusion cloning into *KpnI+HindIII* digested pOPINE	Chapter 5
oxp3701	ACTCTAGAGGATCCCCTTTCCAACGGCGGTGCG	Sense primer for *Azc_4063* (*PEPck*) pK19mobSacB left 1000bp flanking region, adding homology to *SmaI* digested pK19mobSacB for InFusion cloning	Chapter 5
oxp3702	GTCAAACGAGGAACATCTCCTGTTGCGCG	Antisense primer for *Azc_4063 (PEPck*) pK19mobSacB left 1000bp flanking region, adding homology to right flanking region N terminal end for InFusion cloning	Chapter 5
oxp3704	TCGAGCTCGGTACCCGCGGGTGGTCAAAGAAGCC	Antisense primer for *Azc_4063 (PEPck*) pK19mobSacB right 1000bp flanking region, adding homology to *SmaI* digested pK19mobSacB for InFusion cloning	Chapter 5
oxp3706	TCGCCAAATCTCGGTTGCC	Sense mapping primer - *Azc_4063 (PEPck*)	Chapter 5

oxp3707	TGCACCGTCATCACCGC	Antisense mapping primer - *Azc_4063(PEPck*)	Chapter 5
oxp3759	AGATGTTCCTCGTTTGACCC AGCGG	Sense primer for *Azc_4063 (PEPck*) pK19mobSacB right 1000bp flanking region, adding homology to left border C terminal end for InFusion cloning	Chapter 5
oxp3767	TGCCGAATTCGGATCCGGAG GACACCTCTTTCAAGCGTCG	Sense primer for *Azc_4063 (PEPck*) + native promoter adding homology to *BsaI* digested pOGG093 for InFusion cloning	Chapter 5
oxp3768	TCAACAGGAGTCCAAGAGC GCTACCGCTGGGTCAAACGA	Antisense primer for *Azc_4063 (PEPck*) + native promoter adding homology to *BsaI* digested pOGG093 for InFusion cloning	Chapter 5
oxp3947	CGGAATGGGGAAATGGTGC G	Sense mapping primer for deletion of *Azc_0575* (MSDH) from the genome	Chapter 5
oxp3948	GCGCCCAACCAGAACACAA G	Antisense mapping primer for deletion of *Azc_0575* (MSDH) from the genome	Chapter 5

2.2 Chemicals & kits

Antibiotics ampicillin (Amp), gentamycin (Gent), kanamycin (Kan), neomycin (Neo), nitrofurantoin (Nit), streptomycin (Strep), spectinomycin (Spec) and tetracycline (Tet) were purchased from Sigma Aldrich. Plasmid isolation, gel extraction and PCR purification kits were from the GeneJET series and were purchased from Thermo Scientific. The DNeasy Blood & Tissue Kit for genomic DNA isolation were purchased from Qiagen. The enzymes and buffers for molecular biology work were purchased from New England Biolabs (NEB) and the primers were obtained from Eurofins Genomics. Go taq green master mix and Phusion PCR kits for PCR amplification were purchased from Promega and NEB respectively.

2.3 General molecular biology techniques

2.3.1 DNA isolation

Genomic DNA was isolated from RL3841 and ORS571 cultures using DNeasy Blood and Tissue Kit (Qiagen), following the manufacturer's protocol. Plasmid DNA was extracted from *E. coli* using either the GeneJET Plasmid Miniprep Kit (Thermo Scientific) or QIAprep Spin Miniprep Kit (Qiagen). Both kits were used following the manufacturer's 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 manufacturer's protocols. DNA isolation from electrophoresis gels was carried out using GENEJet Gel Extraction Kit (Thermo Scientific) or QIAquick Gel Extraction Kit (Qiagen) according to manufacturer's protocols. Isopropanol was used to improve efficiency of purification of small products.

2.3.2 Restriction digest

Restriction enzyme digests were carried out on purified DNA using restriction endonucleases and respective buffers (Roche, Thermo Scientific or NEB) following manufacturer's instructions. DNA digests waere purified either directly from the reaction mix or through gel extraction agarose gel electrophoresis.

2.3.3 PCR amplification

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 50 µL reactions using Phusion® High-fidelity PCR Master Mix (Thermo Scientific) with DMSO. Thermocycler conditions were set according to the manufacturer's instructions for polymerases, with Tm determined by Geneious R9. 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 pipette tip. A typical colony PCR screening using GoTaq used 35 cycles of 95 °C denaturation for 30 second followed by 30 seconds of annealing between 56-59 °C, and extension at 70 °C for one minute per kb. Amplification with Phusion had a typical

annealing temperature of 61-63 °C, and both denaturation and annealing time was reduced to 15 seconds. Extension time was also reduced to 30 seconds per kb.

2.3.4 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 pH 8.4 (400 mM Tris acetate, 1 mM EDTA,). Gels were run at 120 mV for 30 minutes. Agarose percentage was raised for better resolution when isolating small DNA fragments, and running time was altered accordingly. DNA was stained using Sybr® Safe (Invitrogen). Stained agarose gels were visualized using the GelDoc EZ System (BioRad)

2.3.5 General DNA sequencing

Sanger sequencing was of plasmids and PCR fragments was performed by Eurofins genomics. Samples of plasmid or PCR product were pre-mixed with primers according to the manufacturer's protocol.

2.4 Media

Tryptone-yeast (TY) media was used as a rich-media growth source for Rhizobium strains. This consisted of: 3 g L^{-1} yeast extract, 5 g L^{-1} tryptone, 1.3 g L^{-1} $CaCl_2$. For solid medium, 17.5 g L^{-1} agar was added before autoclaving.

Universal minimal salts (UMS) media was used a minimal-media growth source for Rhizobium strains. Preparation for one litre of media is as follows: 0.5 g L^{-1} $MgSO_4 \cdot 7H_2O$, 0.2 g L^{-1} NaCl, 4.19 g L^{-1} MOPS, 0.5mM K_2HPO_4, and 1 mL of trace elements (0.375 g L^{-1} EDTA-Na_2, 016g L^{-1} $ZnSO_4 \cdot 7H_2O$, 0.2 g L^{-1} $NaMoO_4$, 0.25 g L^{-1} H_3BO_3, 0.2g L^{-1} $MnSO_4 \cdot 4H_2O$, 0.02 g L^{-1} $CuSO_4 \cdot 5H_2O$, 1 g L^{-1} $CoCl_2 \cdot 6H_2O$). pH was adjusted to 7 before autoclaving. Media was supplemented after autoclaving with 1 mL of calcium stock solution (75 g L^{-1} $CaCl_2 \cdot 2H_2O$), 1 mL of iron stock solution (12 g L^{-1} $FeSO_4 \cdot 7H_2O$ dissolved in 1M HCl), and 1 ml of vitamin stock solution (1 g L^{-1} thiamine hydrochloride, 2 g L^{-1} D-pantothenic acid calcium salt, 100 mg L^{-1} biotin). UMS was then supplemented with

variable carbon and nitrogen source additions, depending on the experimental requirements. Concentrations of added carbon and N sources used in UMS are given in table 2.4.1.

Table 2.4.1 Carbon and N sources for media

Concentrations of carbon and N sources added to UMS

Carbon/nitrogen supplementation	**Concentration in UMS**	**Stock**
Acetate	50 mM	Sodium acetate ($C_2H_3NaO_2$)
Ammonium	10 mM	Ammonium chloride (NH_4Cl)
Arabinose	20 mM	L-(+)-Arabinose ($C_5H_{10}O_5$)
Glucose	10 mM	D-(+)-Glucose ($C_6H_{12}O_6$)
Glutamate	10 mM	L-Glutamic acid ($C_5H_9NO_4$)
Pyruvate	30 mM	Sodium pyruvate ($C_3H_3NaO_3$)
Succinate	20 mM	Sodium succinate dibasic hexahydrate ($C_4H_4Na_2O_4 \bullet 6H_2O$)

Luria-Bertani broth (LB) was used for the growth of *E. coli* strains. This consisted of: 10 g L^{-1} tryptone, 5 g L^{-1} yeast extract, 5 g L^{-1} NaCl. For solid medium, 14 g L^{-1} agar was added before autoclaving.

Antibiotics were added to media in concentrations shown in Table 2.4.2 to provide desired strain selection. Because Rlv3841 is a spontaneous streptomycin resistant mutant streptomycin (500 μg mL^{-1}) was used for standard selection (Johnston and Behringer 1975). For standard selection of ORS571 ampicillin (100 μg mL^{-1}) was used.

For *E. coli* transformation where screening was carried out via blue-white colony screening, X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) was added to LB agar at 40 μg mL^{-1}.

Table 2.4.2 Antibiotic working concentrations

Working concentrations of antibiotics used in this work. a: higher selection is used for interposon mutants in pK19mobSacB, b: higher selection is used for interposon mutants in pK19mob

Antibiotic	**Antibiotic suspended in**	***R. leguminosarum bv viciae. 3841***	***A. caulinodans* ORS571**	***E. coli***
Ampicillin (Amp)	Water	100 µg mL^{-1}	100 µg mL^{-1}	100 µg mL^{-1}
Gentamycin (Gent)	HEPES 10mM, pH 7.0	25 µg mL^{-1}	25 µg mL^{-1}	10 µg mL^{-1}
Kanamycin (Kan)	Water	50 µg mL^{-1}	50 µg mL^{-1}	20 µg mL^{-1}
Neomycin (Neo)	HEPES 10mM, pH 7.0	80/160[a]/240[b] µg mL^{-1}	80/160[a]/240[b] µg mL^{-1}	20 µg mL^{-1}
Nitrofurantoin (Nit)	DMF		10 µg mL^{-1}	
Spectinomycin (Spec)	HEPES 10mM, pH 7.0	100 µg mL^{-1}		50 µg mL^{-1}
Streptomycin (Str)	HEPES 10mM, pH 7.0	500 µg mL^{-1}		25 µg mL^{-1}
Tetracycline (Tet)	Methanol	5 µg mL^{-1}		10 µg mL^{-1}

2.5 Bacterial growth conditions and assays

2.5.1 Bacterial growth conditions

All *Rhizobium* species used in this thesis were grown at 28 °C in TY or UMS media. Liquid cultures were aerated in a 28 °C shaker at 200-225rpm. *E. coli* strains were grown at 37 °C in LB media, and liquid cultures were aerated in a 37 °C shaker at 150rpm. When N limited growth was required, cultures were grown in UMS without the addition of an N source.

2.5.2 Generation time of rhizobial strains

To assess growth, an inoculum culture was grown and one mL of OD_{600} 0.01 inoculum was used to inoculate 50 mL of the appropriate media (in 250 mL conical flasks) at 28 ºC in an orbital shaker set at 200 rpm. Unless stated otherwise the inoculum was grown in TY. The media used to asses growth was either TY or UMS supplemented with 1000 X stock vitamins and the desired carbon/nitrogen sources (Table 2.4.1). The cultures were grown at either atmospheric conditions (21% O_2) or at 3% O_2 or at 1% O_2 in an O_2 cabinet (Belle Technology). Optical density measurements (OD_{600}) were taken at appropriate time intervals, four-hour for RLv3841 and ORS571, six-eight hours for slow growth mutants of either stain, until growth reaches stationary phase, measurements were taken of biological triplicates at a minimum. The mean generation time (MGT) was calculated as the number of hours it takes the population to double while in exponential growth phase**.**

2.5.3 Droplet plates of rhizobial growth

Square 120 x 120 mm petri dishes of UMS amp were prepared, containing 300 µM nicotinic acid and 10 mM ammonia as a source of nitrogen. The plates contained either acetate, arabinose, pyruvate or succinate as sole carbon source. Single colonies of each strain to be tested were used to inoculate TY slopes, which were incubated at 37°C for three days. Three biological replicates of each strain were prepared in this way. These slopes were washed with three ml TY and diluted down to OD_{600} 0.1. A five µl droplet of each diluted culture was spotted onto a 120 by 120 mm petri dish at 1 cm intervals. These droplet plates were incubated for 60 hours at 28 °C at 21% O_2 or at 28 °C at 1% O_2 in the O_2 cabinet. After

incubation whole plate images were taken by John Baker (Department of Plant Sciences, Oxford).

2.6 Seed sterilisation and germination

2.6.1 Germination of *Pisum sativum*

P. sativum cv. Aveola seeds were surface sterilised in 95% ethanol for 30-60 seconds and immediately washed in sterile H_2O. Seeds were then surface sterilised again 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

2.6.2 Germination of *Sesbania rostrata*

Seeds were sterilised for 10 minutes in concentrated sulphuric acid to break seed dormancy then washed in 400 mL sterile MilliQ water. Seeds were left imbibing overnight in a sealed falcon tube of sterile water. Seeds were then germinated during two days on two percent water agar containing a double layer of autoclaved filter paper soaked in MilliQ water.

2.7 Plant growth conditions

2.7.1 *Pisum sativum* growth conditions

One litre pots were filled with medium vermiculite, and 300 mL of a nitrogen-free rooting solution was added to each pot. 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 prevent 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 resulting in one pea per pot. All plants were harvested at 21-28 days post-inoculation.

2.7.2 *Sesbania rostrata* growth conditions

Germinated seeds were planted individually in one litre pots containing autoclaved fine vermiculite and 300 ml of nitrogen free rooting solution. Inoculum was prepared from slope cultures. Bacterial cultures were incubated for three days at 37 °C on TY slopes containing relevant antibiotics, resuspended in 3 ml of TY and the resuspension diluted 1/100. One mL of this diluted inoculum was applied to each germinated seed. Four plants were inoculated for each strain. Plants were watered with sterile water every two days and grown at 22°C in a controlled environment room (hours of light/no light) for 28 days.

2.7.3 Recipe for N free rooting solution

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 $CuSO_4$.

2.8 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, the total nodules for each plant weighed and collected in Falcon tubes for use in future assays if required. When stored, nodules were kept at -80°C. To confirm bacterial identity and plasmid retention bacteria were re-isolated from nodules. 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% ethanol. The macerated nodule was streaked onto TY plates using a sterile loop. Nodule streaks were incubated at 28°C for 3-4 days in the case of RL3841 or 2-3 days at 37 °C for ORS571 until single colonies grew. These single colonies were patched onto relevant antibiotics to determine if the correct markers were retained.

2.9 Acetylene reduction

2.9.1 Plant assay – symbiotic N_2 fixation

Pea plants were placed in 250 mL Duran bottles (total volume 320 mL), *S. rostrata* were placed in 500 mL Duran bottles (total volumes 610 mL) due to the length and fragility of their stems, with a small piece of wet tissue paper 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. 8 mL air was removed from each bottle using a syringe and needle and replaced by 6.4 mL acetylene (15.25 mL air, 12.2 mL acetylene for 500 mL Duran) to result in concentration of 2% acetylene within the gas phase of each bottle. When testing peas, bottles were incubated at room temperature for one hour. When testing *S. rostrata* plants were incubated for two hours. One mL samples of the gas phase inside the bottle were taken. Needles were left in corks to prevent air escaping from the syringes during sampling. Each 1 mL sample was run through a Clarus® 480 gas chromatogram (Perkin-Elmer) in order to determine acetylene and ethylene levels in the bottles, and this was converted to a rate of nitrogenase activity $plant^{-1}$ $hour^{-1}$.

2.9.2 Slope culture – free-living ORS571 N_2 fixation

Due to the residual N content of agar, slopes were prepared for this assay using 0.6 % agarose. Otherwise the slopes were UMS, containing 20 mM succinate, 300 µM nicotinic acid and 10 mM glutamate. Ten mL of this was used for each slope, which were prepared in glass universal tubes and sealed with lids containing a rubber seal allowing for the addition and removal of gases using a syringe. Single colonies of strains to be tested were isolated on TY slopes, grown for two days at 37 °C prior to use, resuspended in UMS and diluted down to OD_{600} 0.1. Subsequently, 25 µl of each culture was streaked onto an individual slope and the bottles were sealed. The slopes were transferred to the oxygen cabinet, set to 3% O_2, opened and left to incubate for 40 hours at 37 °C. Before being removed from the cabinet the bottles were resealed. For acetylene reduction we removed 670µl of air and added 560 µl of acetylene to give a final acetylene concentration of 2 %. The sealed tubes were then incubated at 37 °C for 3-17 hours. Low levels of fixation are visible after one hour, but the raw data increased greatly over time. Three x one mL samples of the gas phase from each tube were taken and measured on the GC as for the plants. After measurement on

the GC, the bacterial slopes were resuspended in 3-5 mL UMS and the OD_{600} was measured. Acetylene reduction was adjusted to per unit OD_{600} and then to per mg of protein allowing acetylene reduction to be expressed as µmol ethylene produced $hour^{-1}$ $(mg\ protein)^{-1}$. To convert from per unit OD_{600} to per unit protein the correction factor of 0.112 mg protein unit OD_{600}^{-1} previously calculated for ORS571 by B. Geddes was used. Due to the variable results obtained with this assay, due to the highly variable nature of ORS571, strains were measured with a minimum of 4 biological replicates.

2.9.3 Liquid culture - free-living ORS571 N_2 fixation

ORS571 (10ml) cultures were grown overnight in TY to mid-log phase, pelleted and washed twice to remove ammonia. These cultures were then resuspended in 500 µl ammonia free UMS and used to inoculate 10ml of UMS in 50 mL baffled flasks, containing 20 mM succinate/30 mM pyruvate and 300 µM nicotinic acid. These flasks were incubated in the O_2 cabinet at 37 °C, 150 rpm and 3% O_2 for 4 hours. At the end of this incubation air was removed from the flask an acetylene added to a final concentration 2% acetylene and the flasks were incubated for another 2 hours at 37 °C and 150 rpm. Three x one mL samples of the gas phase from each tube was taken and measured on the GC as for the plants. After measurement on the GC, the OD_{600} was measured. Acetylene reduction was adjusted to per unit OD_{600} and then to per mg of protein allowing acetylene reduction to be expressed as µmol ethylene produced $hour^{-1}$ $(mg\ protein)^{-1}$. To convert from per unit OD_{600} to per unit protein the correction factor of 0.112 mg protein unit OD_{600}^{-1} previously calculated for ORS571 by B. Geddes was used.

2.10 Whole genome sequencing

2.10.1 Sample preparation

A single colony each of ORS571, OPS0865, OPS0921, OPS0922 and OPS0923 were individually resuspended in 100 µl sterile TY. These were streaked out on TY agar plate from the 100 µl culture. Around 1/3 of the plate was made a lawn of bacteria and then the rest streaked out to determine that the culture is pure. The plates were incubated for two days at 37 °C. A large sterile loop was then used to take all bacterial culture off the plate and

mix into the barcoded bead tube supplied by MicrobesNG. The tubes were mixed by inverting 10 times, sealed and sent at room temperature to MicrobesNG for sequencing.

2.10.2 SNP identification (performed by B. Jorrin)

Reads from *phaC* mutant (OPS865) and suppressor mutants (*OPS0921*, *OPS0922* and *OPS0923*) were aligned against the chromosome of ORS571 (NCBI:txid438753) using Bowtie2 v2.1.0 in "very-sensitive" parameters in end-to-end mode (Langmead and Salzberg 2012). The output of the alignment was transformed using Samtools v1.4 (Li et al. 2009) and Single Nucleotide Polymorphisms (SNPs) were detected with VarScan v2.4.3 (Koboldt et al. 2013).

2.11 Cloning

2.11.1 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 manufacturer's protocol. Ligations were carried out at room temperature for 30 minutes**.**

2.11.2 InFusion HD cloning

HD cloning (previously known as InFusion cloning) was carried out using InFusion® HD Cloning Kit (ClonTech). Primers were designed with a 15-20 bp overhang corresponding to the ends of the destination vector, designed and tested in silico using Geneious. HD cloning was carried out in a thermocycler according to the manufacturer's protocols.

2.11.3 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 OD600 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.

For transformation into *E. coli* 50 µL of cells were thawed on ice and 5 µL of plasmid or ligation reaction added. 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 X-gal where necessary and plates were incubated overnight at 37 °C

2.11.4 Conjugation from *E. coli* into RLv3841

Triparental mating was used to conjugate chromosomal integration plasmids from *E. coli* into RL3841. Helper plasmid pRK2013 (Ditta et al. 1980) was used in conjugations. 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 strain containing pRK2013 were inoculated in 10 mL LB containing appropriate antibiotics. Overnight cultures were sub-cultured by inoculating 500 µL into 5 mL fresh LB containing appropriate antibiotics and grown for 3-4 hours at 37°C shaking at 150 rpm until OD_{600} 0.4-0.6. *E. coli* strains were pelleted at a relative centrifugal force of 3894 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 sterilised pair of tweezers was then used to place the filter in one ml TY and dilutions of this from 1/10 – 1/10,000 were plated on TY plates 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*.

2.11.5 Conjugation from *E. coli* into ORS571

The procedure used for ORS571 was modified from that of RL3841 as follows; the resuspended pellet of donor, rhizobia and pRK2013 helper plasmid was plated directly onto TY without a nitrocellulose filter and incubated at 37 °C overnight. The bacterial colony was then scraped up with a sterile loop and resuspended in 1ml of TY and dilutions of this from 10^{-1} to 10^{-3} were plated on TY plates containing appropriate antibiotics, and incubated for 2-3 days at 37 °C. Nitrofurantoin at 10 μg mL^{-1} was used to select against *E. coli*.

2.12 Mutagenesis techniques

2.12.1 Mutagenesis by pK19mob-integration (single cross-over mutant)

Cloning into pK19mob is carried out using InFusion® HD cloning (2.3.6). 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 was used to clone the internal fragment into the digested pK19mob vector. The product of the InFusion® reaction was transformed into chemically competent DH5α and plated on LB containing kanamycin 50 μg mL^{-1} and X-gal to allow blue-white selection to choose correct colonies. Correct colonies were confirmed using restriction digest and sequencing (Eurofins MWG Operon). pK19A/B primers were used for the sequencing reaction.

Once correct colonies had been obtained from *E. coli,* the correct pK19mob plasmid was conjugated into RL3841 or ORS571 as required. Conjugants were selected on TY containing 500 μg mL^{-1} streptomycin and 240 μg mL^{-1} neomycin for RL3841 or 100 μg mL^{-1} ampicillin and 240 μg mL^{-1} neomycin for ORS571. Successful conjugation was confirmed by mapping the insertion with pK19A/B and a gene specific mapping primer. This PCR product could then be confirmed via Sanger sequencing (Eurofins).

2.12.2 Mutagenesis by pK19mobSacB (double cross-over mutant)

This strategy was used to isolate stable mutants and was more suitable when the mutant being generated has a fitness penalty. The strategy was the same in RL3841 and ORS571, though the incubation time was increased from one day to two for *R. leguminosarum* liquid cultures, and from three to five days for sucrose selection. The incubation temperatures used were 37 °C for ORS571 and 28 °C for RL3841. The plasmid pK19mobSacB was digested with *SmaI*. Primers were designed to amplify regions of approximately 1000 bp on the left and right borders of the gene to be deleted. These primers were designed with a 20 bp overhang to the N terminus of the left border and the C terminus of the right border, with homology to *SmaI* digested pK19mobSacB. The primers were also designed to add 20 bp overhangs with self-homology to the remaining termini of the left and right border. These regions of homology allowed InFusion HD cloning reaction to insert the amplified left and right border regions into the *SmaI* digested pK19mobSacB. The product of this reaction was transformed into chemically competent DH5α, and plated on LB containing kanamycin 50 μg mL^{-1} and X-gal to allow blue-white selection to choose correct colonies. Correct colonies were confirmed using restriction digest and sequencing (Eurofins MWG Operon). pK19A and pK19B primers were used for the sequencing reactions.

Once correct colonies had been obtained from *E. coli,* the correct pK19mobSacB plasmid was conjugated into RL3841 or ORS571 as required. Conjugants were selected on TY containing 500 μg mL^{-1} streptomycin and 160 μg mL^{-1} neomycin for RL3841 or 100 μg mL^{-1} ampicillin and 160 μg mL^{-1} neomycin for ORS571. Two correct colonies were identified via mapping with pK19A/B and an appropriate mapping primer and subsequently grown in ten ml of TY liquid media without selection. Five μl of this culture were sub-cultured into a fresh ten ml of TY and allowed to grow to stationary phase. Dilutions of this culture were made from 10^{-1} to 10^{-3} and 100 μL were plated on TY containing 500 μg mL^{-1} streptomycin plus 15% sucrose for RL3841 or 100 μg mL^{-1} ampicillin plus 15% sucrose for ORS571. After species appropriate incubation, 100-200 colonies were patched on TY 500 μg mL^{-1} streptomycin and TY 160 μg mL^{-1} neomycin for RL3841 and onto TY 100 μg mL^{-1} ampicillin and TY 160 μg mL^{-1} neomycin plates for ORS571. Neomycin sensitive colonies were then screened with mapping primers flanking the gene to be deleted. A successful deletion produced a truncated PCR product relative to the wild type. Successfully mapped

deletions were streaked on fresh plates containing strain appropriate antibiotic, and re-mapped.

2.12.3 Mutagenesis by pK19mobSacB with omega interposon insertion (double cross –over mutant)

This strategy was used to isolate stable mutants when the phenotype was suspected to be highly deleterious. The strategy resulted in the insertion of an ΩSpec cassette in place of the gene of interest. The omega cassette was excised from vector pHP45Ω using *HindIII*. The plasmid pK19mobSacB was digested with *SmaI*. Primers were designed to amplify regions of approximately 1000 bp on the left and right borders of the gene to be deleted. These primers were designed to add a 20 bp overhang to the N terminus of the left border and the C terminus of the right border, with homology to *SmaI* digested pK19mobSacB. The primers were also designed to add 20 bp overhangs with homology to the *HindIII* digested ends of the ΩSpec cassette. These regions of homology allowed InFusion HD cloning reaction to insert the ΩSpec cassette between the amplified left and right border regions and insert the whole fragment into the *SmaI* digested pK19mobSacB. The product of this reaction was transformed into chemically competent DH5α and plated on LB containing spectinomycin 100 μg mL^{-1} and X-gal to allow blue-white selection to choose correct colonies. Correct colonies were confirmed using restriction digest and sequencing (Eurofins MWG Operon). pK19A/B and pOT Forfar primers were used for the sequencing reactions.
Once correct colonies had been obtained, from the *E. coli* the correct pK19mobSacB-ΩSpec plasmid was conjugated into RL3841 or ORS571 as required. Conjugants were selected on TY containing 500 μg mL^{-1}streptomycin and 100 μg mL^{-1} spectinomycin for RL3841 or 100 μg mL^{-1} ampicillin and 100 μg mL^{-1} spectinomycin for ORS571. Two correct colonies were identified via mapping with pK19A/B and an appropriate mapping primer and subsequently grown in ten ml of TY liquid media without selection. Five μl of this culture were sub-cultured into a fresh ten ml of TY and allowed to grow to stationary phase. Dilutions of this culture were made from 10^{-1} to 10^{-3} and 100 μL were plated on TY containing 500 μg mL^{-1} streptomycin, 100 μg mL^{-1} spectinomycin plus 15% sucrose for RL3841 or
100 μg mL^{-1} ampicillin, 100 μg mL^{-1} spectinomycin plus 15% sucrose for ORS571. After species appropriate incubation, 100-200 colonies were patched onto TY 500 μg mL^{-1} streptomycin, 100 μg mL^{-1} spectinomycin and TY 160 μg mL^{-1} neomycin, 100 μg mL^{-1} spectinomycin for RL3841 and onto TY 100 μg mL^{-1} ampicillin, 100 μg mL^{-1} spectinomycin

and TY 160 μg mL^{-1} neomycin, 100 μg mL^{-1} spectinomycin plates for ORS571. Neomycin sensitive colonies were then screened with mapping primers flanking the gene to be deleted plus pOT Forfar. A successful deletion could be mapped on either side of the deleted gene from genomic DNA derived from a single colony.

2.12.4 Complementation of mutants

Complementation of mutants was achieved using the stable expression plasmid pOGG093 to express the gene of interest under its native promoter. This RK2 plasmid was originally created via golden gate cloning and contains a tetracycline resistance cassette. The plasmid was digested with *BsaI*. Primers were designed to amplify the gene of interest along with the upstream intergenic region presumed to contain its native promoter. These primers added 20 bp regions to the ends of the fragment with homology to the *BsaI* digested pOGG093. These regions of homology allowed InFusion HD cloning reaction to insert the gene under the control of its native promoter into the *BsaI* digested pOGG093. The product of this reaction was transformed into chemically competent DH5α and plated on LB containing tetracycline 2 μg mL^{-1}. The colonies were screened for the presence of the gene of interest using primers oxp0282 and oxp0284. These colonies were re-streaked on LB containing tetracycline 2 μg mL^{-1} and screened again. The plasmid was confirmed by sequence mapping using the same primers. Once confirmed the complementation plasmids were conjugated into both wild type and the mutant of interest. Tetracycline 5 μg mL^{-1} was used to select for successful conjugation and they were screened with primers oxp0282 and oxp0284 to confirm.

2.12.5 Curing of pK19mob-integration mutant

A single colony of OPS0865 was grown overnight in 5 ml TY liquid culture without selection. This was sub-cultured into fresh TY and grown overnight again. The following day 20 μl of culture was plated on TY ampicillin and incubated for two days at 37 °C. Subsequently colonies were patched on TY Neo 160 and TY Amp 100. Neomycin sensitive cultures were screened for the presence of a full *phaC* gene using oxp3394 + 3395 resulting in strain OPS2077.

2.16 qRT-PCR

RNA was extracted from 10 ml ORS571 cultures using the Direct-zol miniprep kit (Zymo). These cultures were grown with 10 mM glutamate as a source of N_2 to stimulate the production of PHB. After the RNA was DNase treated again using Turbo DNA-free kit (Invitrogen). The RNA concentration was assessed via QUBIT RNA HS kit and converted to cDNA using Superscript II (Invitrogen) following the manufacturer's instructions. *gyrA* was used as housekeeping gene reference for *phaR*. Primers used for qRT-PCR were designed using Geneious, and the reactions were run on a StepOnePlus qRT-PCR machine (Applied Biosystems).

2.17 Translational fusions

2.17.1 Generating translational fusions

GFP translation fusions to phaR of the wild type and the suppressor mutant strains were made by amplifying the *phaA: phaR* intergenic region from ORS571 and OPS0921 along with the first 90 bases of the *phaR* gene using oxp2419 and oxp2420 and cloning into linearised pRU1097 by InFusion HD cloning. Linearisation was achieved by inverse PCR using primers oxp2421 and oxp2422. These constructs were conjugated into ORS571 and *phaC* mutant *OPS0865.*

2.17.2 Testing translational fusions

Single colonies were used to make TY slopes, with three biological replicates of each strain being tested. These were resuspended in 5 ml UMS and used to inoculate overnight cultures at OD_{600} OD 0.01. Cultures were grown overnight in 10 ml UMS containing 10mM glutamate, 20 mM succinate and 300 µM nicotinic acid. Cultures were grown to OD 0.4 and 200 µl aliquots were used for GFP florescence and OD_{600} using a plate reader (Promega).

2.18 TEM microscopy

2.18.1 Preparation of sample

Fresh fixative was prepared consisting of 2.5% glutaraldehyde in 0.1 M filter sterilized (0.22 µm filter) sodium cacodylate (pH 7).
After being picked from the plants, nodules were sterilized for two minutes in 2% bleach, then washed six times in sterile water. Nodules used for TEM microscopy were taken from plants 28 days post inoculation. Three nodules from three plants per strain were selected. Once sterilized the nodules were sliced longitudinally in half using commercial razor blades (Derby Extra) and placed immediately in Eppendorf tubes containing the fixative solution. These Eppendorfs were placed inside a larger 50 mL falcon tube to prevent leakage, stored at 4 °C and shipped on ice to Dr Euan James of the James Hutton Institute who embedded, sectioned and imaged the nodules.

2.18.2 Sample imaging (performed by E. James)

Dr Euan James imaged them using a procedure modified from that of (Beck et al. 2008) in the following manner. A JEOL JEM1400 transmission electron microscope was used, samples were not hexane treated and pyroxylin-coated grids were used.

2.19 Protein biochemistry

2.19.1 Determination of protein concentration

Two protocols were used for determination of protein concentration. A Qubit™ Protein Quantification Kit (Thermo Scientific) was used according to manufacturer's protocol. Qubit quantification was used to determine protein concentration of purified proteins for enzyme assays

For quantification to express values per mg protein, a Bradford assay was used. This assay was done using a 96 well plate and measured on a FLUOstar Omega spectrophotometric plate reader (BMG Labtech). Protein standards were prepared by dilution of 0.5 mg ml $^{-1}$

BSA stock, in the range 0 to 0.5 µg ml^{-1} . These stocks were made up to a final volume of 150 µl with milliQ water. Typically, 5µl of cell extract was measured, made up to 150 µl with water. Bradford reagent (150 µl) was added to each standard and sample, and the plate incubated at room temperature for ten minutes before the absorbance was measured at 595 nm. The standards were used to prepare a standard curve, and sample concentration was calculated via reference to the curve.

2.19.2 SDS-PAGE electrophoresis

Pre-cast SDS-PAGE gels (BioRad) were used to separate proteins by molecular weight for visualisation using staining.
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. 20 µL (0.75 mm gel) or 40 µL (1.5 mm gel) of sample was loaded. 2.5 µL BLUE Wide Range Prestained Protein Ladder (GeneFlow) 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 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 Gel Doc (BioRad) using the Coomassie Blue protocol.

2.20 Expression of tagged proteins in *E. coli*

The pOPIN system of high-throughput expression vectors was used (OPPF-UK). 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. Vector pOPINE was used in this study to apply a C-terminal His tag, requiring the removal of the STOP codon from the gene being cloned into the pOPINE vector.

Expression of pOPINE constructs was carried out using a T7-based expression protocol. Five µL of plasmid was transformed via heat shock into 50 µL BL21 (DE3) competent cells (NEB). One mL of SOC medium was added, and this was used to inoculate four ml of LB-ampicillin liquid media and grown overnight at 37 °C. The next day mid-log phase cultures were used to inoculate 50 mL LB which was incubated at 28°C and 150 rpm for 2 hours. IPTG was then added to a final concentration of 1 mM to induce the culture, which was incubated at 17 °C and 250 rpm for one day. The next day cultures were placed on ice for five minutes to arrest the growth, split into two twenty-five mL cultures and pelleted at 4500 rpm for 15 minutes at 4 °C. After decanting the supernatant, the pellets could then be flash frozen in liquid N_2 and stored at -80 °C until needed or used immediately.

When needed the pellets were thawed and washed twice with appropriate buffer e.g. ten mM HEPES pH 7.0, pelleting between washes at 6000 rpm for 5 minutes. After the final wash the pellets were resuspended in one mL of buffer and homogenised in a ribolyser (FastPrep FP120, Thermo) for 30 seconds at speed 6.5, incubated on ice for two minutes, with the whole riobolysation process being repeated three times. The beads and cell debris were removed from the sample by centrifugation for ten minutes at 15,000 g at 4 °C. The supernatant was saved and used for further protein purification.

2.21 His-spin column protein purification

His-tagged proteins were purified from clarified cell lysate, prepared as above, using His-spin protein miniprep (Zymo) columns according to the manufacturer's instructions. One ml of lysate was split across five columns to maximise yield. Due to the phosphate content of the elution buffer, the final column eluates were pooled and buffer exchanged into appropriate enzyme assay buffer using Amicon Ultra 15 mL 30K centrifugal filters according to the manufacturer's instructions. The concentration of the buffer-exchanged enzyme was quantified using a Qubit™ Protein Quantification Kit (Thermo Scientific) according to manufacturer's protocol. To aid in trouble shooting samples of the initial cells, the ribolysed cells, the supernatant, the purified protein and the buffer exchanged protein were retained and run on an SDS-PAGE gel to confirm the protein was expressed and not located in inclusion bodies.

2.22 Enzyme assays

2.22.1 General procedure for preparing cell extract for enzyme assays

Single colonies of each strain were used to make TY slopes, and the slopes used as inoculum for liquid cultures of UMS succinate. These cultures were grown to OD 0.4-0.6 and pelleted by centrifugation at 4500 rpm. Cell pellets were washed by resuspension in assay buffer. These were pelleted again at 4500 rpm and the wash step repeated twice. Washed pellets were either used immediately or snap frozen in liquid nitrogen and stored at -80 °C until needed.

When needed the pellets were resuspended in one mL of appropriate assay buffer and homogenised in a ribolyser (FastPrep FP120, Thermo) for 30 seconds at speed 6.5, then incubated on ice for two minutes, with the whole ribosylation process being repeated three times. The beads and cell debris were removed from the sample by centrifugation for ten minutes at 15,000 g at 4 °C. The protein content of the supernatant was quantified via either a qubit or Bradford assay.

2.22.2 Enzyme assays – general considerations

When using the plate reader, enzyme assays were run using a clear 96 well plate (Falcon) with a total reaction volume was 200µl. A heat killed enzyme control was made by incubating a portion of cell lysate/purified protein for 10 minutes at 90 °C and was included in every reaction. Ten µg of protein was used per reaction when using cell lysate. When using purified proteins this was lowered to between 1 µg and 200 ng.

2.22.3 Aspartate aminotransferase

The assay was performed in 100 mM HEPES (pH 7.2), 0.2 mM NADH, 0.05 mM pyridoxal phosphate, 10 mM oxoglutarate, 10 µg cell lysate and started with the addition of 50 mM aspartate. This was incubated at 37 °C for two hours, then heat killed at 90 °C for 10 minutes. The enzyme activity was measured by glutamate production versus a heat killed control. The glutamate formed was quantified using a glutamic dehydrogenase assay and compared to a standard curve of 0-120 nM of glutamate. The dehydrogenase assays were carried out

in 60 mM hydrazine, 75 mM glycine buffer (pH 9.0) with 3 units of glutamic dehydrogenase and was started by the addition of 10 mM NAD^+. The reaction was performed in a 250 µl volume and incubated for 90 minutes at 37 °C and the change in absorbance at OD_{340} was measured on a FLUOstar Omega spectrophotometric plate reader (BMG Labtech).

2.22.4 Beta-alanine aminotransferase

The initial enzyme assay was performed in a one ml volume consisting of 10 mM Tris buffer (pH8), 20 mM NADH, 0.05 mM pyridoxal phosphate, 150 µl of clarified cell lysate, and 50 mM of either pyruvate or oxoglutarate. This reaction was incubated for four hours at 37 °C, then heat killed at 90 °C for 15 minutes. Protein concentrations of these heat killed assays were determined using Qubit quantification. Activity was measured via either alanine production, when donating to pyruvate, or glutamate production, when donating to oxoglutarate via alanine/glutamate dehydrogenase assays. One µg of total cell extract was used for each dehydrogenase reaction. The dehydrogenase assays were carried out in 60 mM hydrazine, 75 mM glycine buffer (pH 9.0) with 3 units of alanine or glutamic dehydrogenase and was started by the addition of 10 mM NAD^+. The reaction was performed in a 250 µl volume and incubated for 90 minutes at 37 °C and the change in absorbance at OD_{340} was measured on a FLUOstar Omega spectrophotometric plate reader (BMG Labtech). Standard curves of alanine (0-30 nM) and glutamate (0-120 nM) were prepared.

2.22.5 Malolactic enzyme assay

The malolactic enzyme assay was adapted for the pH ranges of rhizobia from the method used by (Schümann et al. 2013) to monitor malic enzyme activity. The reaction was started by addition of 1.5 mM final concentration NAD^+ to 100 mM HEPES pH 8.0, 1mM Mn^{2+}, 10 mM malate, 1 µg purified Azc_0119 DME-like enzyme. After 30 minutes of incubation at 37 ◦C, this assay was heat killed at 70 ◦C for one minute and stored on ice. The lactate content of the samples was then measured using a lactate dehydrogenase assay and compared to a standard curve prepared from L-lactate 0-1 5 mM. A 0.5 M glycine, 0.4 M hydrazine buffer (pH 9.0) was used, along with 20 µl of sample, 2 mM NAD^+ and two units of lactate dehydrogenase (Sigma). As the sample contained some NAD^+ from the initial assay, this reaction was begun by addition of the lactate dehydrogenase. The reactions were incubated at 37 ◦C for 30 minutes and its progress was measured by the change in absorbance

at OD 340 nm. No evidence was seen for malolactic enzyme activity, although an interaction between the Mn^{2+} ions from the assay mix and the phosphate ions in the His spin column elution buffer was observed. This interaction was lost after the purified DME-like enzyme (Azc_0119) was buffer exchanged into 100 mM HEPES, pH 8.0 prior to its use in the reaction. In addition to the monitoring of the OD 340 nm absorbance, samples of the incubated assay were derivatised with MSTFA and GCMS was used to ascertain whether pyruvate or lactate were being produced from this reaction (Figure 5.7 A).

2.22.6 Malonate semi-aldehyde dehydrogenase

Malonate semi-aldehyde was prepared fresh each day from diethoxypropionate following a procedure modified from (Hayaishi et al. 1961). A 1% solution of diethoxypropionate was brought to pH 2 with 2 M H_2S0_4 and was kept at 45 °C for 30 minutes. The solution was then brought to pH 6.0 with 2 M KOH and was diluted with 0.5 volumes of cold water. This was necessary due to the rapid degradation of malonate semialdehyde, even when stored at 4 °C. This procedure does not result in a pure sample but was reported to give yields of consistently 70 % with respect to malonate semi-aldehyde. The final concentration produced was approximately 50 mM with respect to malonic semialdehyde as the sample was diluted with 0.5 volumes of water, not 2.5 volumes as stated in the original protocol.

This assay was conducted either using 10 µg of purified MSDH (Azc_0575/ Azc_4119) or cell lysate. MSDH was inferred as malonate semialdehyde-dependent reduction of NAD^+. Reactions contained: 0.1M Tris/HCl pH 8.5, 10mM $MgCl_2$ 0.20mM coenzyme A, 1 mM dithiothreitol, 10 µg protein and ~ 10 mM malonate semialdehyde. Reactions were initiated with addition of 1.5 mM NAD^+ and monitored by spectrophotometric absorbance at 340 nm which was measured on a FLUOstar Omega spectrophotometric plate reader (BMG Labtech).

2.22.7 DME assay

The malic enzyme assay was established using 0.5µg of purified DME (Azc_3656) as a positive control and was performed in 100 mM Tris (pH 8), 1.5 mM Mn^{2+}, 10 mM malate and 1.5 mM NAD^+ in a 200µl volume. The reaction was started by addition of the NAD^+ and the change in absorbance at OD_{340} was measured on a FLUOstar Omega spectrophotometric plate reader (BMG Labtech). Both Vmax and Km were determined with respect to malate by measuring the rate of reaction at 1.5, 5, 10 and 20 mM malate, and using the obtained rates to produce a Lineweaver-Burke plot.

2.22.8 Oxaloacetate decarboxylase assay

This method was modified from the assay of (Sender et al. 2004) to be optimised for rhizobia pH, 100 mM Tris (pH 8), 1.5 mM Mn^{2+}, 10 mM oxaloacetate and 1.5 mM NAD^+ in a 200µl volume. Oxaloacetate levels were monitored by absorption at 280 nm for thirty minutes, however the natural rate of decarboxylation rendered this form of monitoring inconclusive. This assay was repeated and samples were run on the GCMS.

The assay was run for half an hour, and then heat killed, derivatised with MSTFA and run on the GCMS. As the nature of the reaction was uncertain, different combinations of cofactors were utilised to make sure no reaction was missed. All combination of NAD (NAD^+, NADH, $NADP^+$, NADPH) and ionic co-factors (Mn^{2+} and Mg^{2+}) were tested for oxaloacetate decarboxylase activity (Figure 5.7).

2.22.9 GCMS

One mL of enzyme assay/sample were dried using a speed vacuum centrifuge. This dried material was then mixed with the solvent pyridine and derivatised with the chemical MSTFA to improve the suitability, efficiency and detection of metabolites in GCMS.

The dried samples were mixed with 25 µl of pyridine and mixed in a thermomixer at 900 rpm, 37°C for 30 min. The samples were briefly spun and 35 µl of MSTFA, derivatization reagent was added. The samples were mixed for another 30 min at 900 rpm and 37°C. After this incubation, samples were briefly spun, transferred to GC-MS vials and sealed. 1 µl of this sample was automatically injected into the GC (Agilent 5977B MSD) at an injection port temperature of 50 °C. For the malolactic enzyme assay samples, the oven temperature

was constant at 50 °C for 2 mins, followed by an increase to 150 °C at 20 °C minute $^{-1}$. The temperature was then ramped to 300 °C for 5 minutes at 8 °C minute^{-1}. For all subsequent assay samples, the run conditions were altered to an initial oven temperature at 50 °C remaining constant for 2 mins, followed by an increase to 150 °C at 10 °C minute $^{-1}$. The temperature was then ramped to 300 °C at 15 °C minute^{-1}. Flow rate was between 1 and 1.2 mL min^{-1}. Peak data was visualised using Agilent MassHunter Qualitative Analysis Navigator.

2.23 Computational analysis

2.23.1 *In-silico* cloning

All cloning reactions were tested in silico using Geneious R8. Sanger sequencing of primers and genomes was aligned against references with the same program when performing sequence confirmation.

2.23.2 Bioinformatic analysis

Global nucleotide and protein alignments were carried out using BLASTn and BLASTp (NCBI) respectively. Local alignments were carried out using Geneious R9 and 10, using the Geneious alignment algorithm. Sequences were obtained from the NCBI database

2.23.3 Statistical analysis and data-handling

Data handling was predominantly carried out in MS Excel. Graphs were generated using GraphPad Prism 7/8. Appropriate statistical tests were carried out using GraphPad Prism 7/8.

Chapter 3

Role of Isocitrate Dehydrogenase in RL3841

3.1 Introduction

Isocitrate dehydrogenase (*icd*) is a key gene of the decarboxylating arm of the TCA cycle, catalysing the oxidative decarboxylation of isocitrate to alpha-ketoglutarate and CO_2 (Figure 3.1). This step of the cycle is often thought of as irreversible in a biological context due to the large negative standard Gibbs free energy change of the reaction. It is also a branch point for the TCA cycle, with the alpha-ketoglutarate it produces able to progress the TCA cycle or be removed from the cycle to allow the production of key metabolites such as glutamine (Lodwig and Poole 2003).

Multiple forms of this enzyme exist, some dependent on NAD^+ and other dependent on $NADP^+$. In either case Mn^{2+} or Mg^{2+} are required as enzyme co-factors. These co-factors are reduced to NADH/NADPH during the reaction, allowing for a simple method to assay this reaction by spectrophotometric absorbance at 340nm. This gene is of particular interest for the *Rhizobiales* due to the contradictory phenotypes thus far observed from *icd* mutants in different rhizobia species. Despite the obligate aerobic nature of all rhizobia, knockout mutants in *icd* have been successfully generated in several species. These mutants also result in a variable phenotype, with an *icd* mutant in *Sinorhizobium melilotii* (McDermott and Kahn 1992) producing a non-N_2 fixing, glutamate auxotroph. By contrast in *Bradyrhizobium japonicum* (Shah and Emerich 2006) an *icd* mutant was capable of wild type levels of N_2 fixation and wild type growth rates. It too however, was a glutamate auxotroph. The fact that these mutants were able to survive suggested that an alternative pathway to at least a portion of the TCA cycle is present within the two species. This alternative pathway is here termed the isocitrate shunt. A transposon mutagenesis study was conducted by Wheatley *et al.*, 2017, utilising a TnSeq variant termed INSeq. In this study, an initial mutant library was generated using the mariner transposon and grown on TY. This mutant library was than subject to selection via growth in minimal media with either glucose or succinate as sole carbon source at both 21 and 1% O_2. In all four growth conditions *icd* (RL2631) was classified as non-essential for growth (Wheatley et al. 2017). This suggested that an isocitrate shunt is also present in this species.

As a first step towards elucidating this isocitrate shunt, it was decided that a mutant should be made to validate the results of the INSeq experiment. Once validated, this mutant could be used as the starting point for a fresh INSeq experiment to identify candidate genes for the isocitrate shunt, with the rational that isocitrate shunt genes would be essential in the absence of *icd*. A single integration mutant of the *icd* gene was made using pK19mob by a former

PhD student. This mutation was confirmed to carry the kanamycin resistance cassette and was found by them to have no measurable isocitrate dehydrogenase (ICD) activity. Additionally, no increase in isocitrate lyase activity was seen relative to the wild type but this mutant has otherwise not been characterised (Wheatly, 2019).

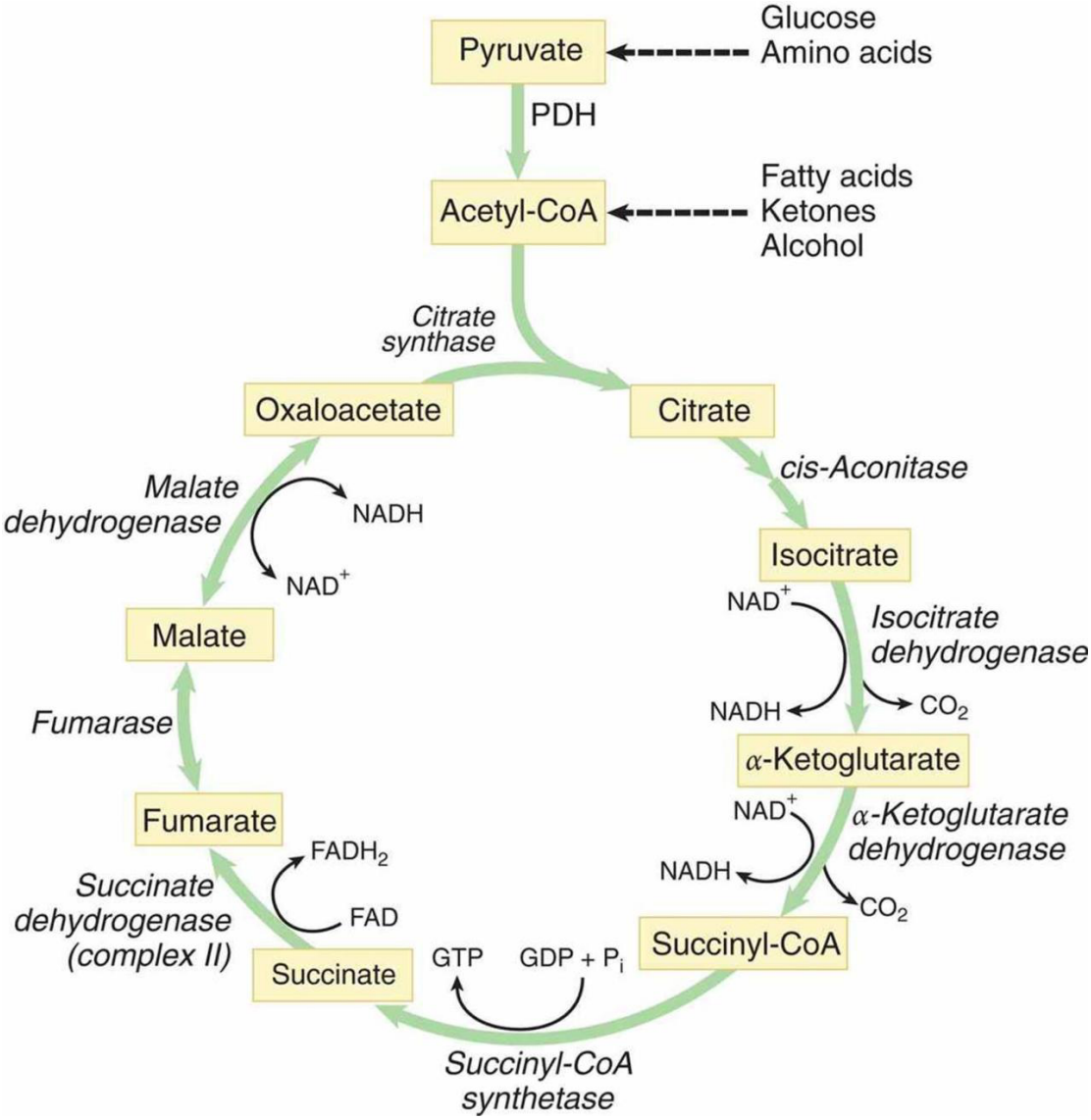


Figure 3.1 TCA cycle summary diagram

Yellow boxes indicate metabolites, green arrows are enzymes and are labelled in black italics next to the reaction that they catalyse. Co-factors are indicated in black by black arrows. PDH stands for pyruvate dehydrogenase. Black text followed by dashed arrows indicate feedstocks supplied via multiple different multi-step pathways. Reproduced from https://ramneetkaur.com/krebs-cycle-tca-cycle-mnemonic/ accessed 6.2.18

3.2 Generation and characterisation of an *icd* mutant

The single integration mutant of the *icd* gene had been generated using pK19 mutant mob by a previous PhD student and named OPS1561. This *icd* mutant had previously been tested for ICD activity using crude cell lysate and measuring the change in $NADP^+$ to NADPH ratios via absorbance at 340 nm. No activity was observed in this strain at the time. However, upon repetition in work for this thesis it became apparent that the *icd* mutant (OPS1561) retained approximately 40% of the wild type ICD activity (0.48 µmol NADPH $minute^{-1}$ mg $protein^{-1}$ vs 0.84 µmol $NADPH^{-1}$ mg $protein^{-1}$) (Figure 3.2). This was a surprising result, but the retention of the kanamycin cassette suggested the mutation was still present, and implied that the observed activity was the result of alternative isocitrate dehydrogenase to the *icd* gene present within RL3841.

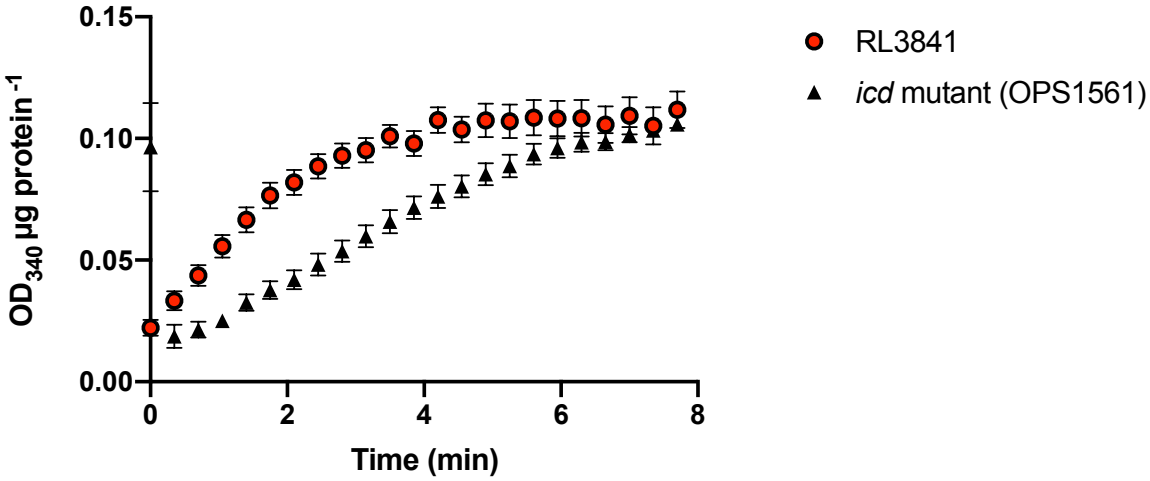


Figure 3.2 ICD activity of wild type and initial *icd* mutant

ICD activity of 1 µg cell lysate indicated via change in absorbance at 340 nm. Cultures were grown on UMS succinate and harvested between OD_{600} 0.4 and 0.6. **R.L3841** is wild type *R. leguminosarum*, **icd mutant** is OPS1561, an integration mutant of *icd*. Points are mean values with bars of SEM, N = 3.

Given the clear reduction in ICD activity observed within the mutant, along with its retention of kanamycin resistance, it was decided that both the growth and symbiotic N_2 fixation rates of the *icd* mutant (OPS1561) should be measured. As RL3841 is capable only of symbiotic N_2 fixation, this was done by inoculating sterilised *Pisum sativum* seeds with either wild type RL3841 or the putative *icd* mutant (OPS1561). Wild type levels of N_2 fixation were apparent from both strains indicating that the observed reduction in ICD activity from the

icd mutant strain (OPS1561) had not affected its ability to fix N_2 (Figure 3.3). The growth rates of the *icd* mutant was measured through growth curves against wild type on pyruvate, succinate, arabinose and glucose as sole carbon sources with 10 mM ammonia as sole N_2 source. No significant difference in generation time was seen on any carbon source (Table 3.1). Additionally, the mutant did not require supplementation of the media with glutamate, indicating it was not a glutamate auxotroph. This suggested that in addition to not affecting N_2 fixation, this reduction in ICD activity had minimal effect on the metabolic health of the *icd* mutants. We decided the next step should be the identification of the alternative sources of ICD activity responsible for the 40% of residual ICD activity in OPS1561.

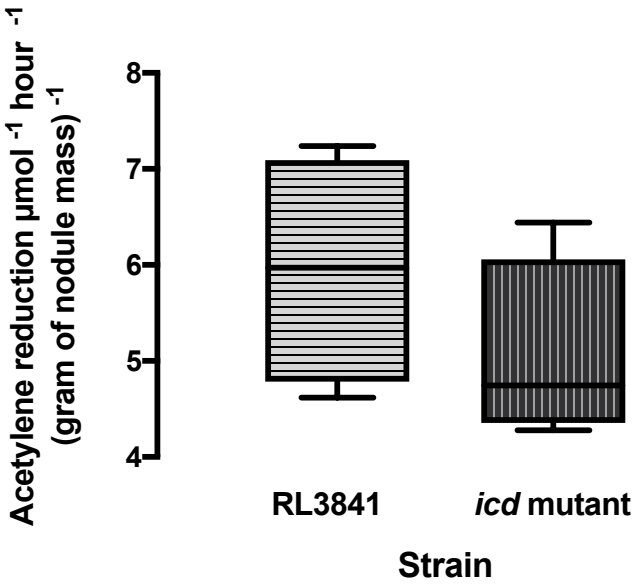


Figure 3.3 Symbiotic N_2 fixation rates of RL3841 and initial *icd* mutant

Symbiotic fixation rates observed 28 days post-inoculation on *P. sativum*. Acetylene reduction rates are corrected per gram of nodule mass. **RL3841** is the wild type strain, ***icd mutant*** is the *icd* mutant strain OPS1561. Boxes are of the interquartile range, with the median indicated within and bars displaying the minimum and maximum values, N=4.

Table 3.1 Generation time of RL3841 and initial *icd* mutant on a variety of carbon sources

Generation time in hours +/- SEM observed from 50 ml UMS cultures with 10 mM ammonia as sole N_2 source and either 30 mM pyruvate, 20 mM succinate, 10 mM glucose or 20 mM arabinose as sole carbon source. **R.L 3841** is wild type R. leguminosarum, ***icd mutant*** is the *icd* mutant OPS1561. N=3.

	Generation time (hour $^{-1}$)	
Carbon Source	R.L 3841	*Icd* mutant
Arabinose	3.85 +/- 0.1	3.6 +/- 0.3
Glucose	4.59 +/- 0.3	4.33 +/- 0.1
Pyruvate	4.38 +/- 0.1	4.37 +/- 0.0
Succinate	4.32 +/- 0.1	4.05 +/- 0.1

3.3 Bioinformatic identification of alternative icd genes

To identify candidate genes responsible for the alternative ICD activity a search was performed for genes containing the ICD domain identified by the Pfam database within their structure using integrated microbial genome viewer. The Pfam database makes use of a hidden markov model (HMM) to identify sequences associated with particular protein families. The isocitrate dehydrogenase (ICD) Pfam consists of isocitrate dehydrogenase, isopropylmalate dehydrogenase and tartrate dehydrogenase. The *Icd* protein from *R. leguminosarum* matched the ICD Pfam from amino acid 5 to 394 from 403 total amino acids. This strategy proved effective and provided us with four candidate genes. Our four target genes, were labelled as being tartrate dehydrogenases, isopropylmalate dehydrogenases or dual function enzymes for both reactions (Table 3.2).

To independently verify these results, BLAST searches were done within RL3841 and *B. japonicum* to assess the similarity of these genes to the known *Icd* from RL3841. The percentage identity to the *icd* genes were all in the region of 30-38% (Table 3.2). Given the number of apparent duplicate enzymes present in the genome it seemed probable that some had been functionally mis-annotated since all functional labels were derived from sequence homology and isocitrate dehydrogenase, isopropylmalate dehydrogenase and tartrate dehydrogenase are members of the same Pfam. The decision was made to begin two approaches, firstly to generate single integration knockout mutants of all 4 alternative ICD

candidate genes using pK19mob and subsequently testing the crude cell lysate of these mutants for reduced ICD activity. Our second approach was to purify all five proteins incluing the known ICD enzyme (encoded by RL2631) to test them for ICD activity directly.

Table 3.2 Potential alternative ICD encoding genes

Gene and protein data obtained of the ICD Pfam proteins in *R. leguminosarum*. Percentage identity of each was compared to the RL3841 Icd using a BLASTp analysis. Protein size was calculated from protein sequence using the https://www.bioinformatics.org/sms/prot_mw.html protein calculator (accessed 23.8.19)

Gene Name	Annotated Function	Percentage identity to R. leguminosarum *Icd*	Protein size kDa^{-1}
RL2631 (icd)	Putative isocitrate dehydrogenase [NADP]	100	45.10
pRL120629	Bifunctional tartrate dehydrogenase/isopropylmalate dehydrogenase	38	37.01
RL0995	Tartrate dehydrogenase,	35	38.75
RL4707 (leuB)	Isopropylmalate dehydrogenase	34	39.67
pRL110153 (ttuc)	Tartrate dehydrogenase	35	39.13

3.4 Purification and testing of alternative icd genes

The pOPINE vector was our first choice for protein expression, given the low chance of interference with enzyme activity from the His tag, coupled with the rapid purification made possible by benchtop His-Trap spin columns (Zymo). All five genes of interest, four potential ICD genes and the known ICD (*icd*), were introduced into the pOPINE vector by In-Fusion cloning and confirmed by Sanger sequencing.

These five vectors were conjugated into BL21 cells and expression was induced by addition of IPTG. Crude cell lysate was purified using His-Trap spin columns and the resulting protein eluents were quantified with a Bradford assay. Successful purification was tested by running the His Spin column elute on an SDS gel and visualising with Coomassie stain. The

quantity of protein to be used in the assay was determined by initial tests using purified Icd alone leading to 200 ng of each protein being used to perform the final ICD assays.

The ICD assay made abundantly clear that no ICD activity was present in any of the purified proteins, with the exception of Icd, the known ICD enzyme (Figure 3.4). This was unexpected, but incontrovertible. With a measured activity rate for Icd of 13 µmol of NADH min^{-1} mg $protein^{-1}$ and no measurable reaction for any of the other purified proteins.

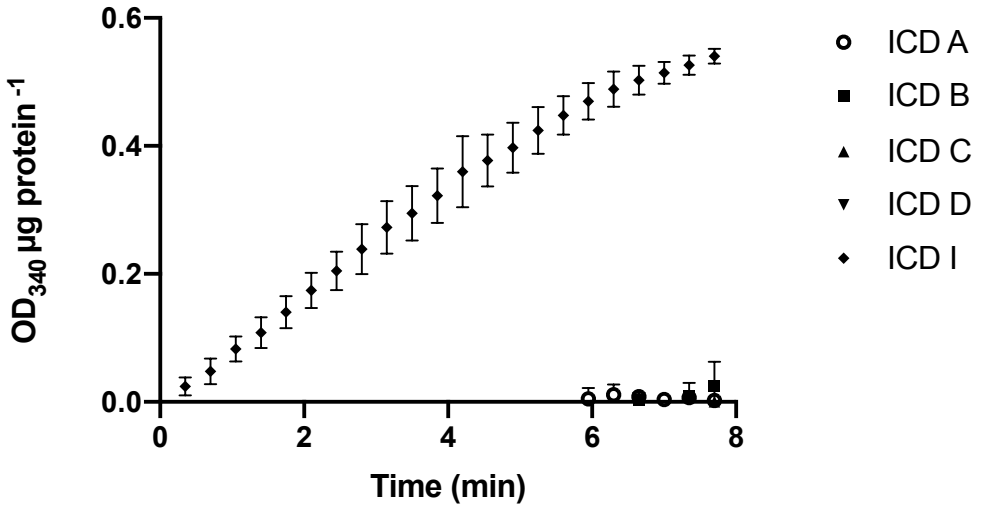


Figure 3.4 ICD activity of purified potential ICD enzymes

Isocitrate dehydrogenase activity purified potential ICD proteins calculated from absorbance at 340 nm, corrected to 1 µg of enzyme. **ICD A** is pRL120629, **ICD B** is pRl110597, **ICD C** is RL0995, **ICD D** is RL4707, **ICD E** is pRL110154, **ICD I** is purified Icd (the known ICD gene of RL3841. Points are mean values with bars of SEM, N = 3

Although isolation of single integration mutants for all four genes using pK19mob was begun, the ICD assay results using purified proteins were obtained before they could all be mapped. As no discernible ICD activity was observed for any of the potential ICD proteins it was decided that it would not be a productive use of time to measure the ICD activity of mutants strains. The lack of obvious ICD activity in the putative ICD proteins led us to question the original *icd* mutant. Given the occasional unreliability of single integration mutations, sometimes resulting in knock-downs or reversion mutations the decision was

made to create a double integration mutant to remove the *icd* gene entirely, removing this potentially confounding variable.

3.5 Generation of an *icd* double integration mutant

To generate the double integration mutant, it was decided to make use of pK19mobSacB to replace the *icd* gene (absolute position 2,771,242-2,773,690) with an in frame spectinomycin resistance cassette flanked by omega terminators. Successful single integration of the plasmid was selected for using neomycin selection from the pK19mobSacB plasmid backbone. Two colonies with neomycin resistance were selected and grown in liquid TY culture without antibiotic selection, before being sub-cultured and grown again in fresh TY. A sample of these liquid cultures was plated on sucrose counter-selection plates containing spectinomycin. The sucrose selection makes retention of the pK19mobSacB plasmid backbone lethal, whilst the spectinomycin selection makes reversion to wild type lethal. In this way, selection for a double recombination event, replacing the *icd* gene with the spectinomycin cassette and losing the pk19mobSacB plasmid is selected for. Single colonies from the sucrose + spectinomycin selection plates were patched onto spectinomycin plates and neomycin plates, to screen for loss of the pK19mobSacB plasmid. Over 600 colonies were screened in this way, from a total of four colonies initially transformed with the pK19mobSacB plasmid, and no double cross-over mutations were isolated. One single integration mutant, with spectinomycin resistance and neomycin resistance, was isolated and termed OPS1620. When mapped it was found that the spectinomycin cassette had integrated on the left hand border of the *icd* gene (Figure 3.5). Interestingly, this single integration mutant also resulted in a reduction of ICD activity, though just under 80% of wild type activity was retained (Figure 3.6). The intransigence of generating a total icd gene deletion, combined with the lack of ICD activity in the alternative purified proteins led us to undertake a reanalysis of the original INSeq data.

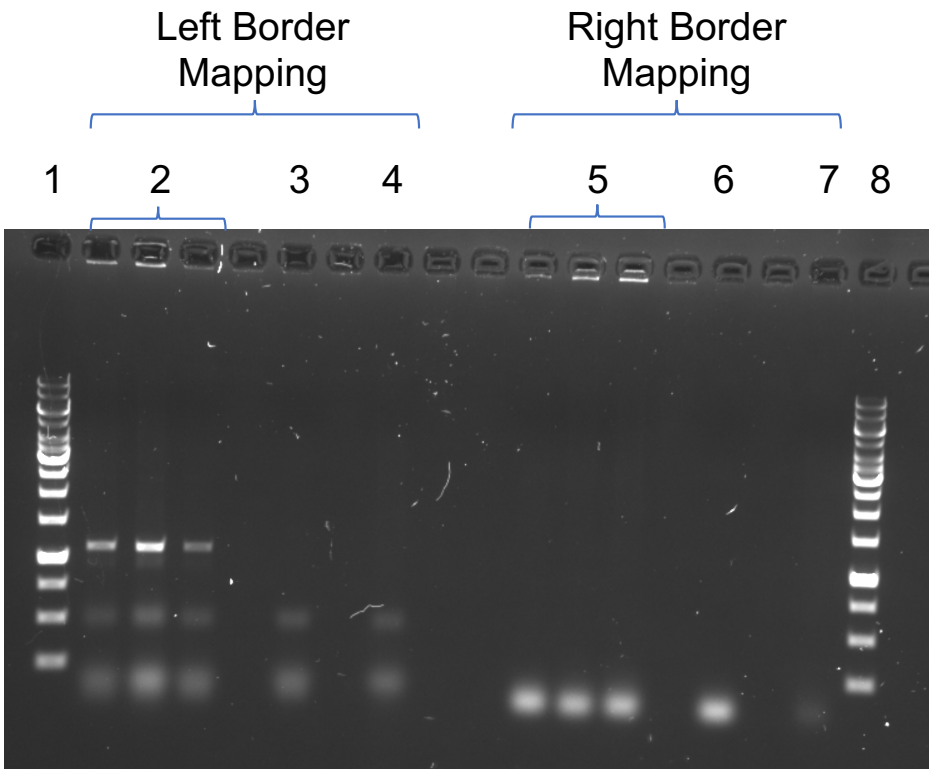


Figure 3.5 PCR mapping of attempted *icd* double integration mutant

Mapping PCR of attempted double integration mutant of *icd* in RL3841. Mapping of the left border was done with oxp0929 and pOT For-Far, mapping of the right border was done using oxp0930 and pOT For-Far. Section 2 are amplified from colonies of OPS1620, section 3 is from a wild type control colony of R.L 3841, section 4 is a no template control, section 5 are from colonies of OPS1620, section 6 is from a wild type control colony of R.L 3841, section 7 is a no template control. Section 1 and 8 are 1 Kb plus DNA ladders (Thermo).

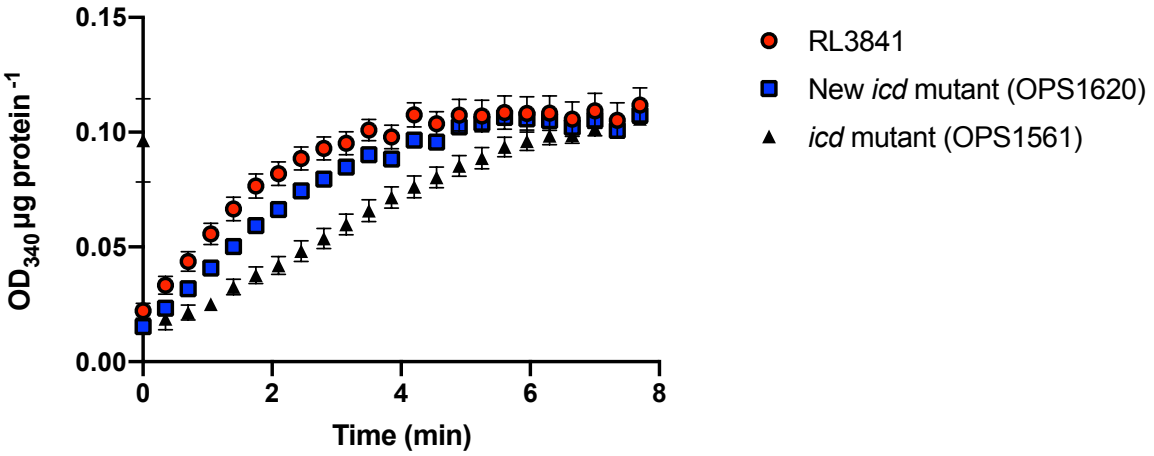


Figure 3.6 ICD activity of wild type and *icd* mutants

ICD activity of 1 µg cell lysate calculated from absorbance at 340 nm. Cultures were grown on UMS succinate and harvested between OD_{600} 0.4 and 0.6. **RL3841** is wild type RL3841, ***icd* mutant** is OPS1561, an integration mutant of *icd*. N**ew *icd* mutant** is *OPS1620,* a single integration mutant on the left flank of the *icd* gene rather than internal to the gene. Points are mean values with bars of SEM, N = 3.

3.6 Re-examination of the INSeq data

Upon re-examination of the INSeq data using integrated genome viewer (IGV) it became apparent that an error had occurred when processing the original sequencing data. The *icd* gene is in fact essential in RL3841 and the previous classification by the hidden markov model (HMM) of the gene as non-essential was a statistical error. The reason for this is that a small region of the C- terminus, a region highly unlikely to affect function, was capable of accepting transposon insertions at two TA insertion sites. This resulted in a high enough level of insertions within the gene to allow the mis-classification of the gene as non-essential (Figure 3.7). The inability to generate a knockout, as opposed to a knock down mutation in this gene is explained by the lack of toleration for insertions anywhere else within the gene.

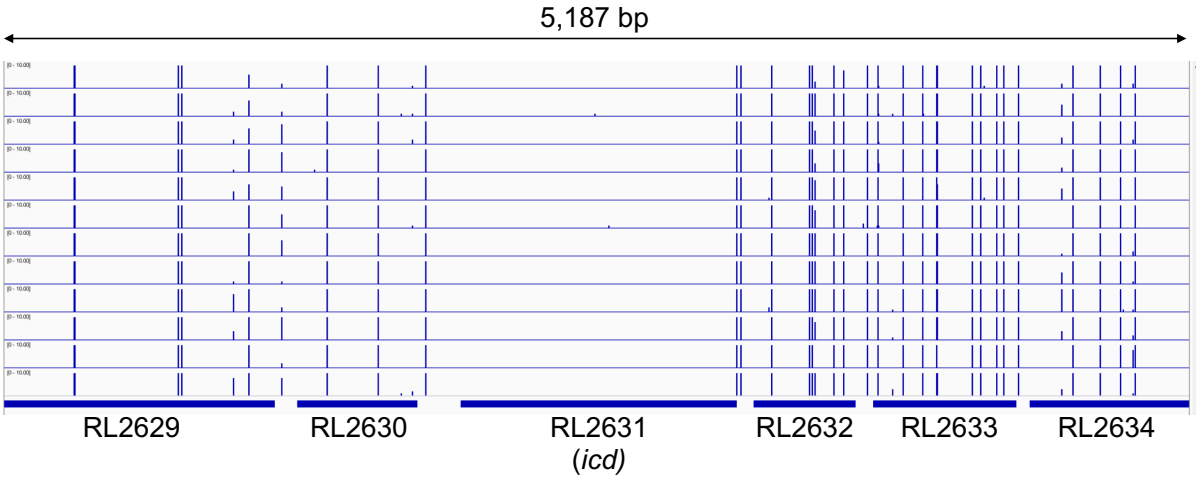


Figure 3.7 IGV overview of the INSeq data in the *icd* gene region

Each horizontal line represents the results of a sequencing run, and each vertical line indicates a TA site containing transposon insertions. The height of the lines indicates the number of insertions at that site. A gene map is indicated along the bottom, with gene positions shown by dark blue boxes and labelled with gene numbers.

3.7 Discussion

The original *icd* insertion mutant OPS1561 was found upon re-testing to have retained ~40% of wild type ICD activity (Figure 3.2). This reduced ICD activity did not result in a measurable reduction in symbiotic N_2 fixation rate or generation time on arabinose, glucose, pyruvate or succinate (Table 3.1, Figure 3.3). As the kanamycin resistance cassette was retained by this strain, even after passage through the plant, it was determined that there were likely alternative ICD encoding genes still present in OPS1561, accounting for the remaining observed ICD activity in this strain. All 5 members of the ICD Pfam in RL3841 (including *icd*) were identified (Table 3.2) and expressed in *E. coli* using the pOPINE vector series. These four potential alternative ICD genes were annotated by sequence homology as either tartrate dehydrogenases, isopropylmalate dehydrogenases or bi-functional enzymes with both activities (Table 3.2). Of the purified enzymes only the known ICD enzyme Icd had any measurable ICD activity in solution (Figure 3.4). This lack of activity from the alternative ICD enzymes suggested that either the residual ICD activity was due to an entirely novel form of ICD enzyme that did not share similarity in its annotation to the existing ICD Pfam and thus was not identified during our bioinformatic analysis, or the

original *icd* single integration mutant OPS1561 was a knock-down rather than a knockout mutant. To test this, attempts were made to generate a double integration mutant to delete the *icd* gene to ensure a true knockout strain of *icd* based ICD activity. However, this too resulted in the generation of single integration mutants (Figure 3.5), which retained an even greater amount of ICD activity (Figure 3.6). When the original INSeq data was reanalysed using IGV it was observed that no significant insertions took place within the main body of the *icd* gene, though a high insertion density was observed at two sites on the C-terminus of the gene (Figure 3.7). The high insertion density of this region would have been enough to lead the HMM to miss-classify *icd* as non-essential, despite the lack of insertions within the functional main body of the gene. When the data was reanalysed with the HMM set to ignore insertions within the first or last 100 bp of a gene *icd* was classed as essential. This would explain why the isolation of a double integration mutant of *icd* proved impossible. A reanalysis of the entire INSeq data-set was run by Dr Vinoy Ramachandran, using a modified program that trimmed the reads, discounting transposon insertions in TA sites at the termini. This analysis identified ~ 35 genes that had been incorrectly classified in the original analysis including *icd* and *glnb.*

The necessity of the *icd* gene in RL3841 could be demonstrated by inserting a second copy of the gene on a plasmid. If a merodiploid strain of *icd* is made then a double integration mutant of the chromosomal *icd* gene should be possible using pK19mobSacB. However, given our repeated inability to generate a true knock-out mutant of Icd by sac mutagenesis and the careful reanalysis of the INSeq data this seemed unnecessary.

The original single insertion mutant OPS1561 was never precisely mapped, though the kanamycin resistance cassette has been confirmed to be present. When combined with the obvious lethality of insertions within the main body of the gene indicated by the INSeq data (Figure 3.7) this suggests that some form of rearrangement took place in order to prevent a total loss of function mutant in the *icd* gene being produced. While such a rearrangement would be rare, the double selection pressure of the kanamycin and the lethal nature of the mutant would have resulted in very strong selection pressure in favour of such a re-arrangement. This selection pressure could be seen again when attempting to generate the double integration mutant, when single integration could take place either in the left or the right hand flanking regions (absolute positions_2,770,471-2,771,242 and 2,772,479-2,773,690). Single integration mutants were observed integrated on both the left and the

right flanking regions of the *icd* gene. It is believed that the reduction in ICD activity observed in the single integration mutants derived from pK19mobSacB are due to the introduction of the terminators flanking the spectinomycin resistance cassette prior to the *icd* gene. This likely reduces the amount of successful transcription of the *icd* gene, though qRT-PCR would be required to confirm this.

Whilst the purification of the putative ICD proteins did not reveal any novel ICD activity, the purification of the Icd protein allows the confirmation of its ICD activity (Figure 3.4), which had previously been inferred purely from homology. The remaining putative ICD proteins despite being members of the ICD Pfam did not demonstrate any ICD activity, making it likely that they are instead ispopropylmalate or tartrate dehydrogenases as they were annotated. It remains unclear why multiple copies of both these enzymes are found within *R. leguminosarum* however.

We can conclude that while *icd* is not necessary in certain rhizobia species such as *B. japonicum* and *S. meliloti* (McDermott and Kahn 1992; Shah and Emerich 2006), it is essential for growth in RL3841. Any alternative isocitrate shunt present in these two species is clearly not active at any significant level in RL3841 though it may be present in other rhizobia species. This work also highlights the caution that must be taken when interpreting INSeq data, particularly when insertions sites are present at the C and N termini of the genes in question. As a result of our findings, the pipeline for analysis of INSeq data in our lab has been altered to exclude insertions within the first and last 100 bp of each gene. In this way, insertions that are less likely to affect the functionality of a gene can be excluded from the analysis. When the existing INSeq data was reanalysed using this new pipeline, *icd* was reclassified

Chapter 4

Re-evaluation of the Role of Poly-3-Hydroxybutyrate in ORS571

4.1 Introduction

Biological N_2 fixation by diazotrophic bacteria, gives them the ability to convert atmospheric dinitrogen into ammonia using the enzyme nitrogenase (Oldroyd et al. 2011). ORS571 is a species of the order Rhizobiales able to fix N_2 both under free-living conditions as well as during symbiosis with its host plant *Sesbania rostrata* (Elmerich et al. 1982; Dreyfus et al. 1983, 1988). This is unusual for N_2-fixing rhizobia as most can only do so symbiotically (Terpolilli 2016).

Given that the nitrogenase enzyme is permanently inactivated by O_2, a low O_2 environment is essential for successful fixation to occur (Lodwig and Poole 2003). This is a problem as all rhizobia, including ORS571, are obligate aerobic organism using the tricarboxylic acid (TCA) cycle as their primary means of energy generation. In low O_2 conditions there may be insufficient O_2 to act as a terminal acceptor for electron transport to balance reductant formation. One of the main ways to remove an excess of reductant is channelling reductant into the generation of carbon storage molecules such as lipids or poly-3-hydroxybutyrate (PHB) (Terpolilli et al. 2016).

PHB is a carbon storage molecule synthesised by most rhizobia (Trainer and Charles 2006) using molecules of acetyl-CoA, in a multi-stage reaction, the polymerisation step of which is catalysed by PHB synthase (PhaC) (Schubert et al. 1988; Sharma and Dhingra 2016) (Figure 4.1).

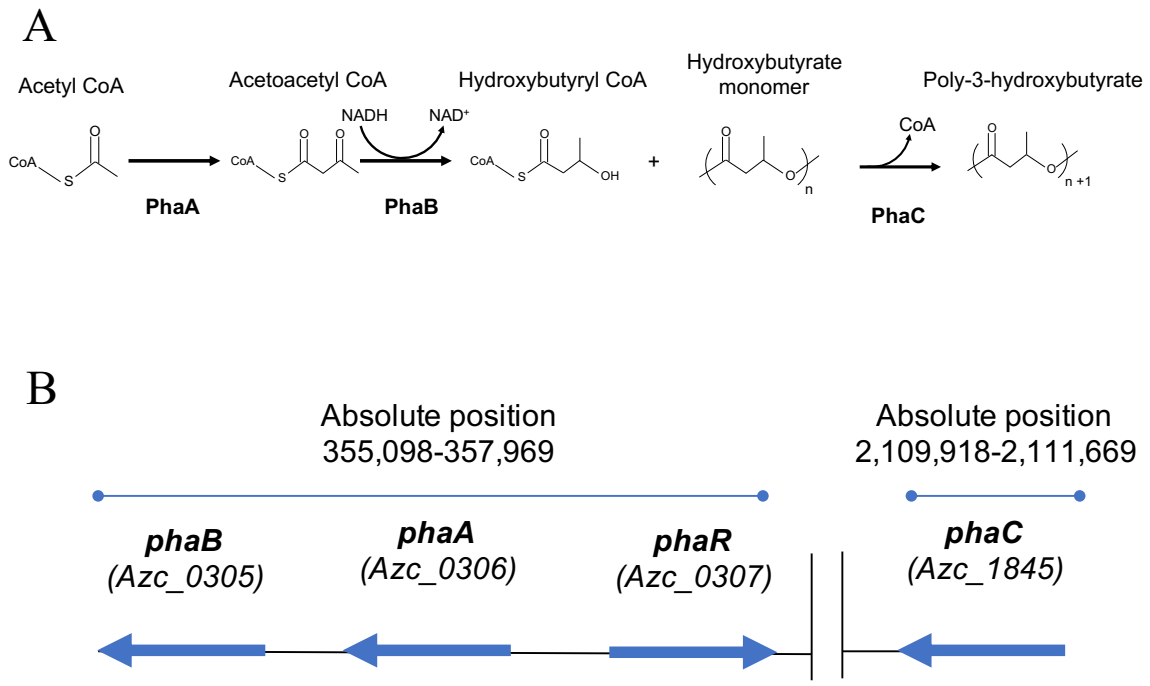


Figure 4.1 Overview of genes and reactions involved in PHB synthesis

A; Stages of PHB synthesis. Condensation of two Acetyl-CoA molecules by PhaA to acetoacetyl-CoA, followed by reduction to hydroxybutyryl CoA by PhaB. The hydroxybutyrate acts as the monomer for PHB. The hydroxybutyrate is bonded to a seconded hydroxybutyrate by PhaC to form the hydroxybutyrate polymer, releasing the CoA in the process. This polymer can then be extended by PhaC catalysed interactions with additional hydroxybutyryl CoA.

B; Genome architecture of the PHB operon in ORS571. Absolute genome positions for the two regions involved in this operon along with gene numbers are indicated. The *phaR* gene encodes a regulatory protein that controls expression of the PHB operon. PhaA is an acetyl-CoA acetyl-transferase, PhaB is an acetoacetyl-CoA reductase and PhaC is a PHB synthase.

In addition to its use as a reductant sink, the carbon storage provided by PHB has been shown to support fixation during periods of carbon starvation and it has been suggested that is has a role in initial nodule formation (Gerson et al. 1978; Wang et al. 2007). Many species of rhizobia, including ORS571, *Bradyrhizobium japonicum, Rhizobium etli, and Rhizobium tropici* , accumulate PHB both under free living conditions (Stam et al. 1986; Tombolini and Nuti 1989) and within the N_2 fixing nodule form of rhizobia known as bacteroids (Ndoye et al. 1994; Karr et al. 1984). Generally speaking rhizobia that form determinate nodules accumulate the most PHB as bacteroids (Wong and Evans 1971; Bergersen et al. 1991) even though it can still be produced by some that form indeterminate nodules such as Rlv3841(Terpollili 2016).

Mutagenesis studies have been performed in a number of species to determine the importance of PHB for N_2 fixation. The results of these studies have been varied; with *phaC* mutants of the species *R. etli* fixing at higher levels than wild type (Cevallos et al. 1996), while reduced symbiotic efficiency was seen in *Sinorhizobium meliloti* in symbiosis with *Medicago truncatula*, but not with *Medicago sativa* (Wang et al. 2007). No change at all was seen in the fixation rates of *phaC* mutants of *R. leguminosarum* on *Pisum sativum* or *Phaseolus vulgaris* (Lodwig et al. 2005).

The formation of PHB is regulated by the interaction of PhaR, an auto-regulating negative regulator (Maehara et al. 2002) and phasin proteins (PhaP). Phasins are present in both bacteria and archaea (Kim et al. 2017; Jang et al. 2017; Cai et al. 2015), where they are sequestered to the surface of polyhydroxyalkanoate (PHA) granules, such as PHB, reducing their concentration within the cytoplasm despite their continuous synthesis. When the granules reach a critical size, PhaP displaces PhaR from the surface due to the greater binding affinity of PhaP to PHB relative to PhaR. This results in a build-up of cytoplasmic PhaR, which then binds to the promoters of the *phaR, phaP and phaC* genes, preventing their transcription and halting the production of further PHB (Potter et al. 2002; Kuchta et al. 2007). The binding mechanism of phasins to the PHA granules is currently unknown, but is believed to be via the interaction of the hydrophobic and hydrophilic faces of the phasin α helix insulating the growing PHA granule from contact with the cytosol (Maestro and Sanz 2017).

Previous knockout mutations of the *phaC* gene indicated it to be essential for N_2 fixation in ORS571 (Mandon et al. 1998) indicating that PHB synthesis was essential for fixation in this species. In this work we confirm that a *phaC* mutant cannot fix N_2 in the laboratory but show that PHB synthesis *per se* is unnecessary for effective fixation by ORS571 in free-living or symbiotic conditions. Instead, we show that *phaR* misregulation explains the reduced growth and lack of fixation in a *phaC* mutant.

4.2 Initial investigation of a *phaC* mutant

4.2.1 Testing the free-living fixation of a *phaC* mutant

A *phaC* mutant strain had previously been isolated via single integration using pK19mob which had a non-fixing phenotype in free-living liquid culture (Borah 2017). This experiment was repeated using free-living UMS slope cultures grown under nitrogen-free, carbon excess conditions (20 mM succinate) at 3% O_2 and again a clear non-fixing phenotype was evident in the *phaC* mutant (Figure 4.2). This *phaC* mutant also exhibited dramatically reduced growth on pyruvate and acetate at 21% O_2, though growth on succinate was unaffected (Figure 4.3, 4.4, Table 4.1). A severe growth penalty was observed on all four carbon sources at 1% O_2, though again succinate allowed the greatest amount of growth. Although growth on succinate had the shortest generation time it still increased dramatically from 3.2+/- 0.2 hours at 21% O_2 to 19.3+/-1 .5 hours at 1%, compared to the 3.2+/-0.1 and 7.5+/-1 .2 seen in the wild type (Table 4.1). This indicates the *phaC* mutation results in a major growth penalty to ORS571 which is exacerbated under O_2 limiting conditions.

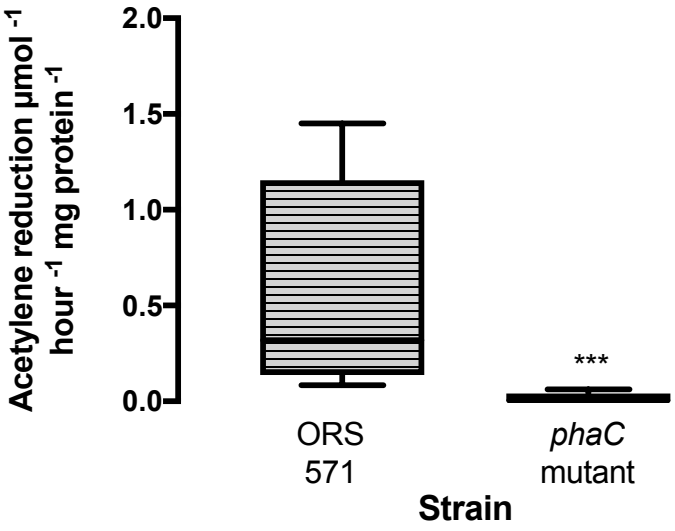


Figure 4.2 Free living fixation rates of ORS571 vs the *phaC* mutant (OPS0865)

Fixation rates calculated from acetylene reduction assay per mg of bacterial protein in the sample. Medians are significantly different P <0.0001, Mann Whitney Test, Boxes are of the interquartile range, with the median indicated within and bars displaying the minimum and maximum values, N=12.

4.2.2 Characterising the growth of a *phaC* mutant

To confirm the integration of pK19 into *phaC* is responsible for this altered phenotype, pK19mob in *phaC* was cured from OPS0865 by growth of the strain on TY liquid media in the absence of neomycin selection for 24 hours. This liquid culture was sub-cultured into fresh TY media and subsequently plated on TY agarose. Colonies were patched from this plate onto TY Amp and TY Neo plates, and colonies that had lost the neomycin resistance were screened via amplification of the *phaC* gene using primers oxp1043 and pK19B. Loss of neomycin resistance, and the return of a 1.5 Kb band indicated the successful curing of the plasmid integron. This cured strain was stocked and named OPS2077. After pK19mob was cured wild type growth rates were restored on all carbon sources for OPS2077 (Figure 4.3, 4.4). This observation is in line with the results of (Mandon et al. 1998) suggesting that PHB is required for N_2 fixation in ORS571.

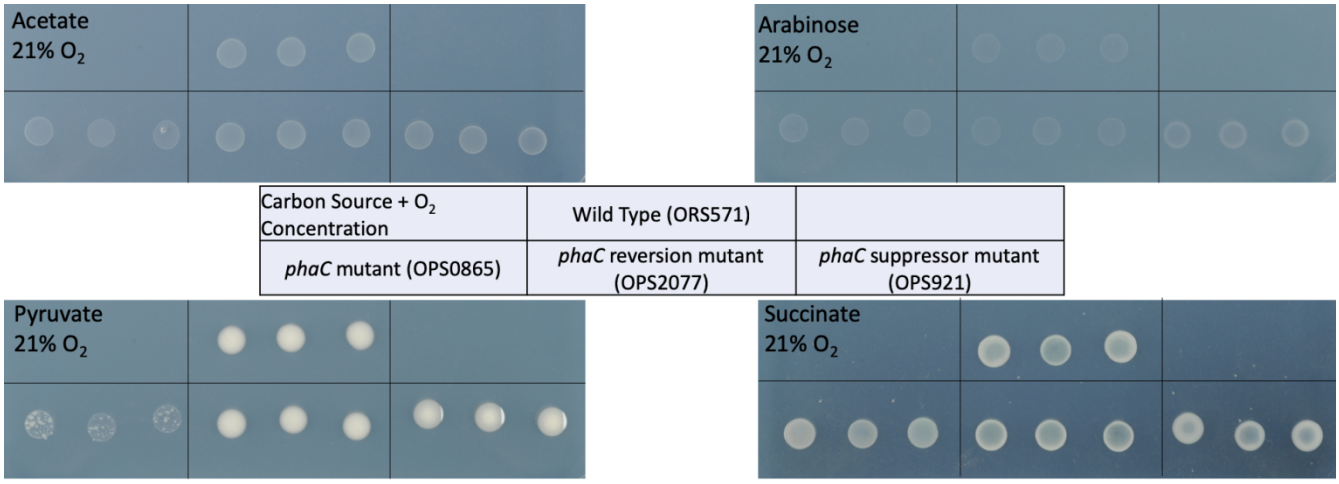


Figure 4.3 Atmospheric growth of ORS571 and PHB mutants

Summary growth table of wild type (ORS571), the *phaC* mutant (OPS0865) the *phaC* revertant (OPS2077 and the putative suppressor mutant OPS0921) under atmospheric O_2 conditions at 28 °C on four different sole carbon sources. Bacteria (5 ul at OD_{600} 0.1 dilution) were drop plated onto UMS ammonia (10 mM) with acetate, arabinose, pyruvate or succinate as sole carbon source. N = 3.

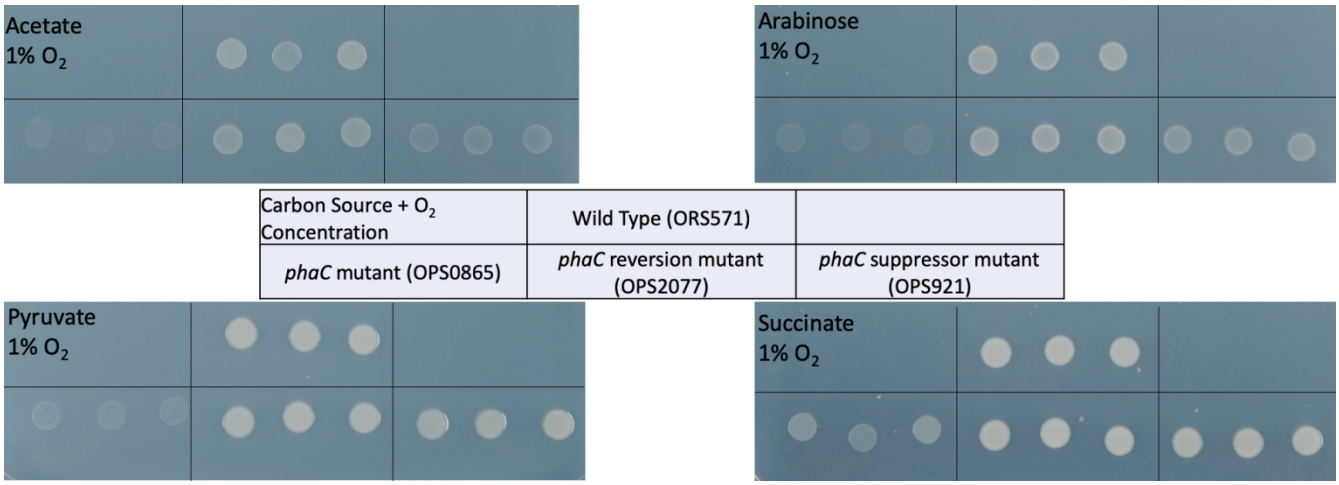


Figure 4.4 Microaerobic growth of ORS571 and PHB mutants

Summary growth table of wild type (ORS571), the *phaC* mutant (OPS0865) the *phaC* revertant (OPS2077) and the putative suppressor mutant OPS0921 at 1% O_2 at 28 °C on four different carbon sources. Bacteria (5 ul at OD_{600} 0.1 dilution) were drop plated onto UMS ammonia (10 mM) with acetate, arabinose, pyruvate or succinate as sole carbon source. N = 3.

4.3 Investigation of a *phaA* mutant

4.3.1 Generation of a *phaA* mutant

As the *phaC* mutation was demonstrated to be deleterious to growth we decided to target an earlier stage of the PHB synthesis pathway to ensure the previously observed phenotype was due to lack of PHB and not a site specific effect of the *phaC* mutation. Our chosen target was the *phaA* gene, encoding an acetyl-CoA acetyl transferase, the first step in the PHB synthesis pathway.

To generate this mutant, a single crossover was made using pK19mob to insert the pK19mob backbone containing a neomycin resistance cassette into the *phaA* gene (absolute position 356,160-356,753) This mutation was confirmed to have gained resistance to neomycin, and was successfully mapped using primers oxp2139+pK19B and confirmed via sequence mapping of this amplified fragment. This strain was named OPS1298.

When tested for free living fixation using the slope culture method previously described, the *phaA* deletion mutant unexpectedly proved capable of wild type rates of fixation despite the expected loss of PHB synthesis (Figure 4.5). It was decided therefore to test the symbiotic fixation rates of both the OPS0865 and OPS1298 strains.

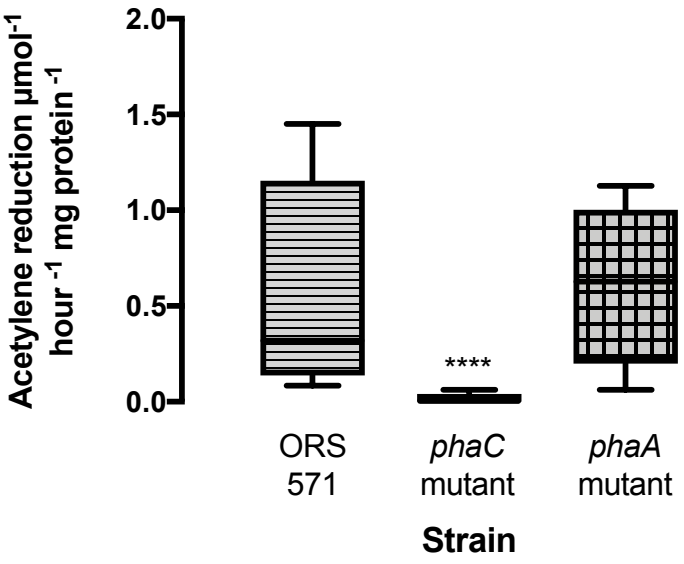


Figure 4.5 Free-living N_2 fixation rates for ORS571, *phaC* and *phaA* mutants

Free-living N_2 fixation rates by ORS571 vs the *phaC* mutant (OPS0865) and the *phaA* mutant (OPS1298). Fixation rates calculated from acetylene reduction assay per mg of bacterial protein in the sample. Medians for OPS1298 and ORS571 are significantly different from OPS0865, P <0.0001, P <0.0001, Mann Whitney Test. Boxes are of the interquartile range, with the median indicated within and bars displaying the minimum and maximum values, N=12.

4.3.2 Symbiotic characterisation of *phaC* and *phaA* mutants

Although the loss of N_2 fixation by free-living ORS571 had been demonstrated to be due to the *phaC* mutation, its symbiotic N_2 fixation phenotype remained unclear. Previously, a *phaC* mutant strain was found to primarily produce small, non-fixing nodules (0.02 +/- 0.02 nmol of acetylene/min/mg of nodule) containing no bacteria (Mandon et al. 1998). However, it was noted at the time that some plants also contained "a few large N_2 fixing nodules" which were theorised to be revertants, the bacteria from these nodules were not isolated and tested individually for fixation rate or nodule size and no tests were reported to confirm their reversion (Borah 2017; Mandon et al. 1998). We therefore inoculated *S. rostrata* plants with both the wild type ORS571 strain, the *phaC* mutant strain OPS0865 and *phaA* mutant strain OPS1298.

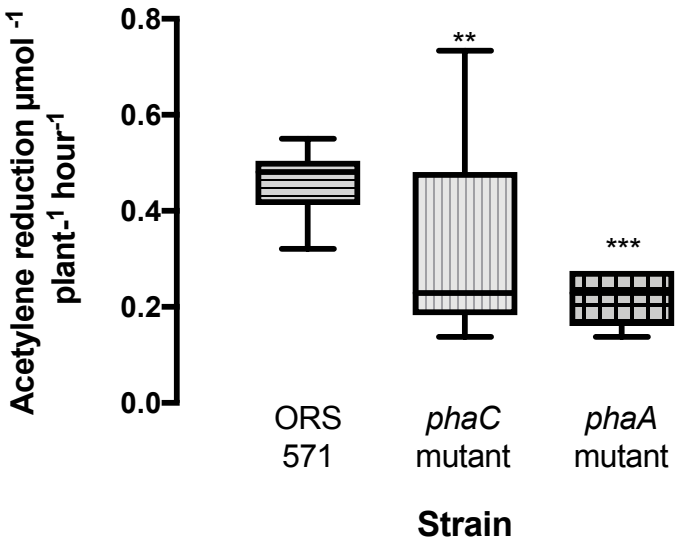


Figure 4.6 Symbiotic N_2 fixation rates of ORS571, *phaC* mutant and *phaA* mutant

Symbiotic N_2 fixation rates observed 28 days post-inoculation on *S. rostrata*. ORS571 is the wild type strain, the *phaC* mutant is strain OPS0865 and the *phaA* mutant is strain OPS1298. Per plant fixation rate is significantly greater when inoculated with ORS571 than the *phaC* mutant (P=0.0016) and the *phaA* mutant (P= <0.0001). There was also a significant difference between the *phaC* and *phaA* mutants (P=0.0044) two tailed Mann-Whitney. Boxes are of the interquartile range, with the median indicated within and bars displaying the minimum and maximum values, N=4.

We were surprised to observe that all strains reduced acetylene, although the *phaA* mutant did so at a significantly reduced rate. Given the variability of fixation rates seen between *S. rostrata* plants additional repeats would be necessary to confirm the reduced fixation observed in the *phaA* and *phaC* mutant strains are truly significant (Figure 4.6). Given the clear N_2 fixing phenotype already observed in the free-living cultures of the *phaA* strain, we first focused on the most unexpected result; the near wild type fixation rate of the *phaC* mutant. In addition, we show below that the *phaA* mutant is still capable of reduced PHB synthesis (section 4.4.2).

The N_2 fixation of the *phaC* mutant strain was so unexpected, that fresh plants were put up for both wild type and the *phaC* mutant strain. The acetylene reduction rates observed in these plants were corrected per gram of nodule mass and no significant difference was seen between the fixation rates of the two strains. This confirmed the original observation of fixation in the *phaC* strain, and suggested wild type levels of fixation.

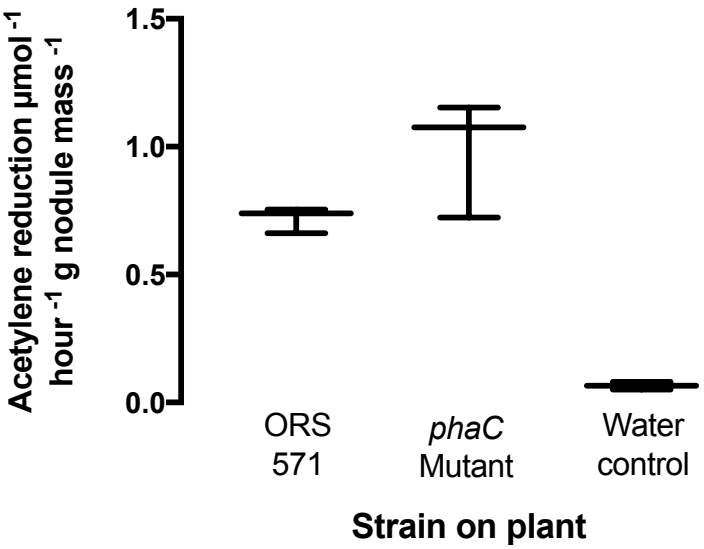


Figure 4.7 Replicate symbiotic N_2 fixation rates of ORS571 and *phaC* mutant

Symbiotic fixation rates observed 28 days post-inoculation on *S. rostrata*. ORS571 is the wild type strain, the phaC mutant is strain OPS0865 and the water control is a set of plants inoculated with water alone. No significant difference observed ORS571 and OPS0865 (P=0.4) two tailed Mann-Whitney. Boxes are of the interquartile range, with the median indicated within and bars displaying the minimum and maximum values, N=3.

Our initial suspicion was that the wild type symbiotic fixation rates observed for strain OPS0865 was due to a spontaneous reversion mutant being favoured by the plant, resulting in the loss of the *phaC* mutation.

4.4 Identification and characterisation of a *phaC* suppressor mutation

4.4.1 Identification of a *phaC* suppressor mutation

To test whether a spontaneous reversion mutant had indeed occurred, bacteria were isolated from six separate nodules of the plants inoculated with the *phaC* mutant strain. This was done by first surface sterilising the nodules in sodium hypochlorite, then crushing the nodules and plating them on non-selective TY agarose. Single colonies from the crushed nodules were subsequently patched on both TY amp and TY neo. Of 30 colonies patched, all retained neomycin resistance. These colonies were then screened for the presence of the pK19 mutation with the chromosomal primer oxp1043 and pK19B, amplifying out from the

pK19mob plasmid backbone. This gave no band in the wild type, and a 1.5 Kb fragment as expected from the original *phaC* mutant strain OPS0865 (Borah 2017). Glycerol stocks were made of three of the mapped strains isolated from nodules, which were named OPS0921, OPS0922 and OPS0923. The free-living N_2 fixation of strain OPS0921 was measured using the slope culture assay, along with the original *phaC* mutant (OPS0865), the *phaA* mutant OPS1077 and the wild type ORS571 (Figure 4.8). The nodule isolated *phaC* mutant strain OPS0921 fixed N_2 at wild type levels, along with the *phaA* mutant, whilst the original *phaC* mutant OPS0865 remained incapable of free living N2 fixation. Thus, passage through the plant was sufficient to restore N_2 fixation to a *phaC* mutant strain. Since we had confirmed the retention of the *phaC* mutation in this strain, we considered it to be a putative suppressor mutant. Our next focus was on determining the method of action of this mutation.

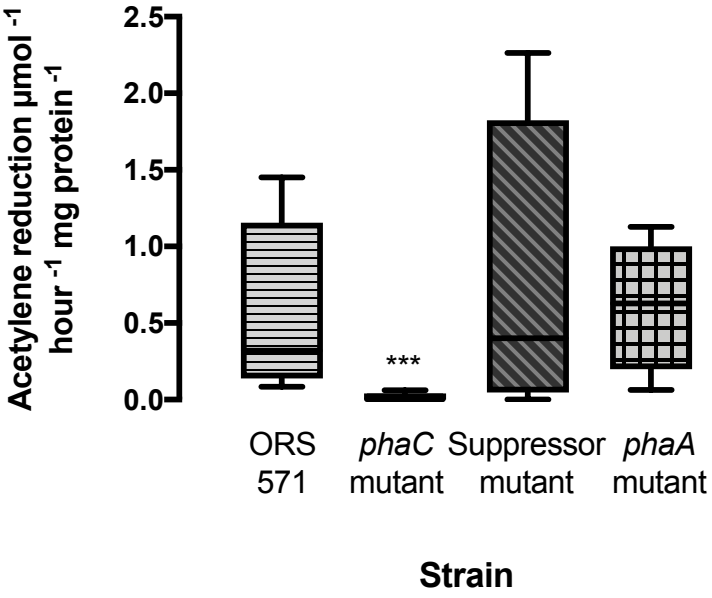


Figure 4.8 Free-living N_2 fixation rates for ORS571, *phaC*, *phaC* suppressor and *phaA* mutants

Free living fixation rates of ORS571 vs the *phaC* mutant (OPS0865), putative *phaC* suppressor mutant (OPS0921) and the *phaA* mutant (OPS1298). Fixation rates calculated from acetylene reduction assay per mg of bacterial protein in the sample. Medians for OPS0921, OPS1298 and ORS571 are significantly different from OPS0865, P<0.0064, P<0.0001, P <0.0001, Mann Whitney Test. Boxes are of the interquartile range, with the median indicated within and bars displaying the minimum and maximum values, N=12.

4.4.2 Quantification of PHB via electron microscopy

To quantify the PHB in the strains we made use of transmission electron microscopy on sections of *S. rostrata* nodules. This allows the visualisation of PHB granules as white electron transparent granules inside bacteroids. Such granules were readily apparent in the wild type strain, and in far more limited quantity in the *phaA* mutant strain. However, no granules were observed in the *phaC* or putative *phaC* suppressor mutant (Figure 4.9). This indicated to us that whatever had allowed the restoration of N_2 fixation in strain OPS0921, it had not restored PHB production. From this, and the acetylene reduction data, we were able to conclude that PHB production *per se* is not necessary for N_2 fixation in either symbiotic or free-living conditions for ORS571. It also shows that there must be another PhaA-like activity in ORS571 as some PHB is made (Fig 4.9D).

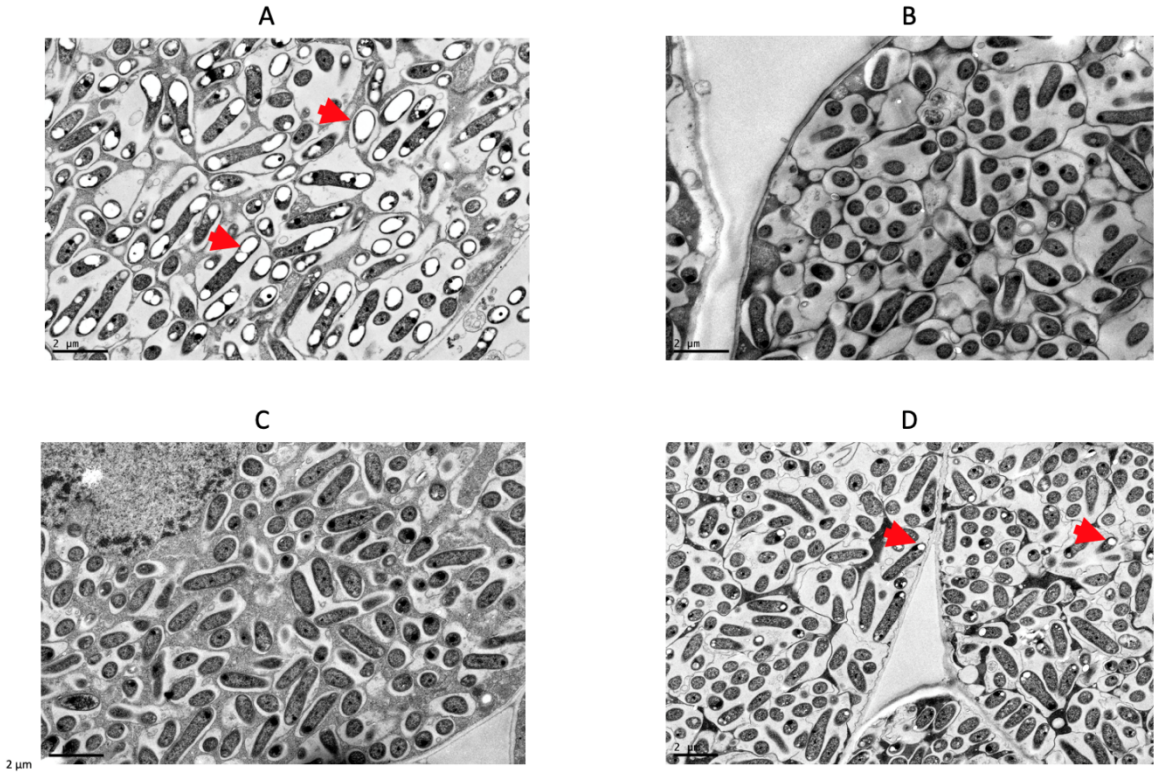


Figure 4.9 TEM images of ORS571, *phaC* mutant, *phaC* suppressor mutant and *phaA* mutant bacteroids

PHB quantification of strains via transmission electron microscopy of bacteroids within nodule sections taken from 28 day old plants. PHB appears as white masses within the cells, example granules are indicated by red arrows.

A: Wild Type (ORS571), PHB granules clearly visible. B: *phaC* mutant (OPS0865), C: *phaC* Suppressor mutant OPS0921, D: *phaA* mutant (OPS1077).

4.4.3 Growth characterisation of the putative suppressor strain OPS0921

Since strains OPS0865 and OPS0921 had such a clear phenotypic difference in terms of free-living N_2 fixation, their growth was assessed to identify any further differences (Figure 4.3, Figure 4.4, Table 4.1).

Table 4.1 Effect of O_2 level on growth of ORS571, *phaC* and *phaC* suppressor mutant on succinate

Generation times of the wild type ORS571, phaC mutant strain OPS0865 and the suppressor mutant strain OPS0921. The values presented are mean +/- SEM. Cultures were grown on UMS succinate and ammonia at 28°C, 150 RPM, N=3.

	O_2 Concentration	
	21%	**1%**
Strain	**Generation Time (hour)**	
ORS571 (Wild Type)	**3.15 +/- 0.12**	**7.51 +/- 1.19**
OPS0865 (*phaC* mutant)	**3.16 +/- .20**	**19.30 +/- 1.49**
OPS0921 (Suppressor Mutant)	**3.32 +/-0.23**	**17.19 +/- 1.23**

When grown on succinate at 21% O_2 neither mutant displayed significantly reduced growth compared to wild type, however the generation time for both increased dramatically when subjected to growth at 1% O_2 (Table 4.1). This is consistent with the synthesis of PHB under low O_2 contributes to redox balance in ORS571.

To assess how growth varied between carbon sources, droplet growth plates were made for both strains utilising acetate, arabinose, pyruvate and succinate as their sole carbon sources (Figure 4.3, 4.4).

Different growth phenotypes were seen for each of ORS571, OPS0865 and OPS0921 at both 21 and 1% O_2, with OPS0865 in particular showing reduced growth relative to the wild type,

and the putative suppressor OPS0921 producing an intermediate growth phenotype. Of particular note was the difficulty OPS0865 had growing on pyruvate as a sole carbon source at either O_2 tension compared to the other strains and represents the biggest comparative reduction in growth. This limited growth is likely to represent a flux balance problem given that pyruvate can be used to generate acetyl-CoA directly, which then feeds directly into the TCA cycle. As a result of this it would force the TCA cycle to perform a complete revolution and generate four moles of reductant for every one mole of pyruvate utilised.

Succinate by contrast, can generate small amounts of reductant via malic enzyme, or malate dehydrogenase activity, which can then be converted to sugars via gluconeogenesis. These sugars can then either be converted to storage molecules or catabolised and fed into the TCA cycle at a more controlled rate. As a result, it displayed the smallest growth penalty of any carbon source for the *phaC* mutant (OPS0865). The reduced growth on acetate in the suppressor stain (OPS0921), and the near absence of growth seen in the *phaC* mutant (OPS0865) are likely the result of their shared lack of PHB synthesis. Since acetate is directly converted into acetyl-CoA it would, like pyruvate, also force the rapid generation of reductant through TCA cycling. The wild type strain (ORS571) and revertant strain (OPS2077) do not display any growth impairment on acetate or pyruvate as they are capable of utilising PHB as a reductant sink. Given the patchy nature of OPS0865 growth on pyruvate at 21% O_2, compared to the near wild type levels of growth seen in OPS0921, it seems likely that these represent the emergence of similar suppressor mutants.

The clear growth differences between stains OPS0865 and OPS0921 confirmed that some form of secondary mutation had occurred. Given the retention of both neomycin resistance and the pK19mob insertions, as well as the lack of PHB synthesis in OPS0921, it was apparent that there must be a suppressor mutation present in this strain in addition to the original *phaC* mutation. To identify the nature of this mutation whole genome sequencing (MicrobesNG) was utilised.

4.5 Identification and analysis of the causal suppressor mutation

4.5.1 Identification of potential SNPS

To aid in the mutant identification, strains OPS0921, OPS0922 and OPS0923 were sequenced and compared to both OPS0865 and ORS571, with the assumption that any mutation shared by the OPS092N strains but absent from OPS0865 and ORS571 were likely candidates to investigate as the causal mutations of the observed phenotypes.

Additionally, as part of our analysis of the sequencing data we also checked for the presence of the pK19 integration remaining within OPS092N strains. In doing so the plasmid backbone was clearly identifiable within the assembled contigs, providing further evidence that the *phaC* mutation was retained within the three OPS092N strains.

Our analysis indicated that a single SNP was present in all three OPS092N strains and absent from both ORS571 and OPS0865. This was a G to A transition (absolute position 357,354) thirteen bases upstream of the *phaR* ATG start codon (Figure 4.1, 4.10), a general regulator of low O_2 growth as well as PHB metabolism. Given the location of the SNP, it is likely to be within the ribosome binding site (RBS) of the *phaR* gene and thus most probably has a translational effect on *phaR*. Hereafter, this SNP will be referred to as the 0921 SNP and further analysis was performed on strain OPS0921 only.

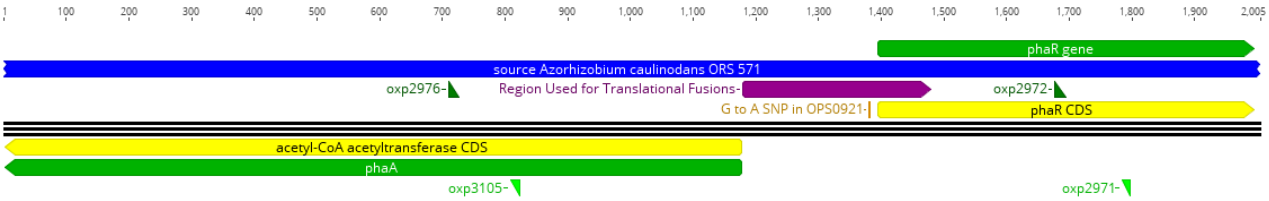


Figure 4.10 Gene map of *phaA*:*phaR* region indicating translational fusions and qRT-PCR primers

Gene map displaying the *phaA* and *phaR* genes along with their intergenic regions. Genes are indicated in dark green, CDS in yellow. The intergenic region, along with the first 90 bases of the *phaR* gene (357,152-357,459) used to create translation fusions are indicated in purple. The location of the G-A SNP is shown in orange at absolute position 357,354. Primer pairs for qRT-PCR are indicated by green triangle along with their orientation.

4.5.2 Determining the translational effect of the mutation on *phaR*

To confirm the mutation produced a translational effect on *phaR* expression, GFP reporter fusions were produced containing the *phaA-phaR* inter-genic region and the first 90 bases of the *phaR* gene fused in-frame to GFP. Two versions of this were produced, one containing the wild type intergenic region, and the other containing the G to A SNP identified in OPS0921 (Figure 4.10). These were termed the 571 and 921 reporter plasmids, and both were transformed separately into ORS571 and OPS0865. All strains were then grown under carbon excess, nitrogen starvation conditions on UMS succinate with glutamate as sole N source to stimulate PHB production.

The wild type 571 *phaR* RBS reporter produced a six-fold greater amount of GFP in strain OPS0865 compared to that in strain ORS571 (Figure 4.11 A). This increase in PhaR/GFP translation was likely to be in response to the lack of PHB in strain OPS0865. Thus, the translation of *phaR* is reduced by the presence of PHB, so in the absence of PHB the regulator PhaR accumulates. This may have serious side effects on the large PhaR regulon in rhizobia (Quelas et al. 2016). By contrast, the *0921 SNP phaR RBS* reporter gave a consistently low signal in both strains (Figure 4.11 A), indicating it was insensitive to the presence or absence of PHB in the cell. Thus, the ribosome binding site mutation in OPS0921 has locked the translation of *phaR* to a low level, independent of the PHB level. This suggests the restoration of fixation and improved growth seen in strain OPS0921 results from preventing excess expression of *phaR* in a *phaC* background.

4.5.3 Determining the transcriptional effect of the mutation in OPS0921 by qRT-PCR

The *gyrA* gene was selected as a housekeeping gene, and the transcription of both *phaA* and *phaR* was measured in both ORS571 and OPS0865 by qRT-PCR. Both strains were grown under carbon excess, nitrogen starvation conditions on UMS succinate with glutamate as sole N source to stimulate PHB production. As expected, there was no significant difference in the transcription of *phaR* in either strain and although a small drop was seen in *phaA* expression but this was statistically insignificant ($p>0.05$) (Figure 4.11 B). This is consistent with the misregulation of PhaR resulting from a translational effect. This make sense as

PhaR regulates the expression of *phaA* and is itself regulated in part by the PHB content of the cell.

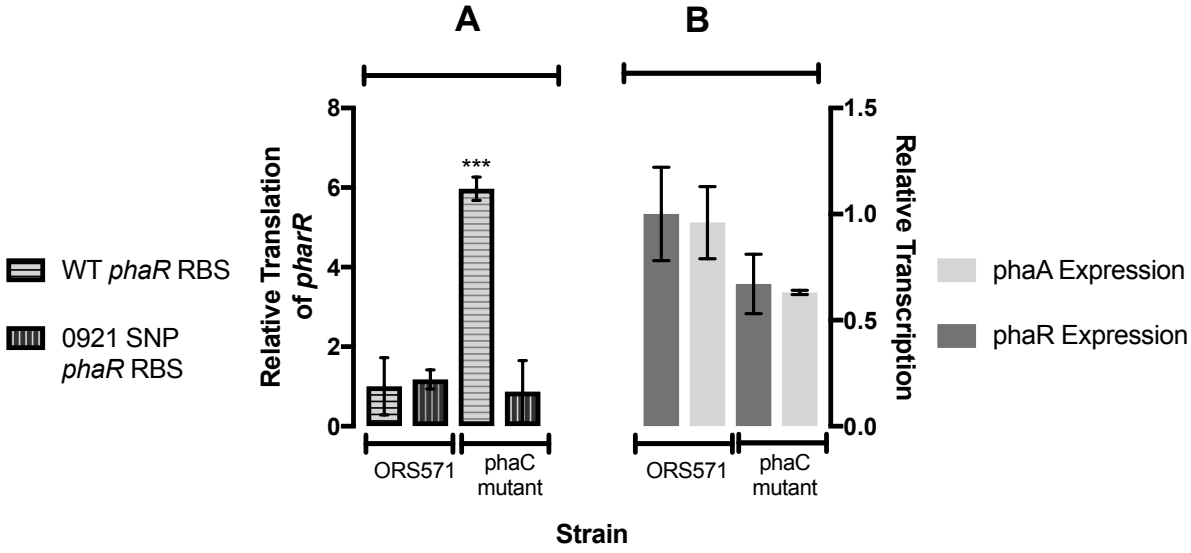


Figure 4.11 Translational reporters of PhaR levels and qRT_PCR of *phaA* and *phaR* transcription in ORS571 and the *phaC* suppressor mutant

A- Translation rates of GFP reporters bearing the wild type (571) or mutant SNP (921) of the phaR RBS site. Data is relative to the translation rate of the wild type *phaR* RBS reporter in wild type strain ORS571. Turkey's multiple comparison test indicates a significant difference between wild type *phaR* RBS reporter in ORS571 vs OPS0921 $P<0.0001$, d.f 89. No significant difference is seen between any combination of the other three samples. Bars indicate the mean with bars of SEM, N=8

B- qRT-PCR data indicating the expression of the *phaA* and *phaR* genes from strains ORS571 and OPS0865 relative to the expression of housekeeping gene *gyrA*. No significant difference is seen between expression of either gene in either strain. No significant difference is seen between any combination of the other three samples. Bars indicate the mean with bars of SEM, N=3

4.6 Discussion

We have demonstrated that a *phaA* mutant strain of ORS571 is still capable of N_2 fixation in both free living and symbiotic conditions despite a dramatic reduction in PHB production observable within fixing bacteroids (Figure 4.6, 4.8, 4.9). Given that wild type levels of fixation were observed in free-living conditions (Figure 4.8), it is likely that the lower fixation rate seen per plant (Figure 4.6) is the result of reduced colonisation ability leading to reduced nodule number rather than reduced fixation. This would be in line with previous research in *Sinorhizobium meliloti,* where reduced colonisation ability has been observed in PHB mutants (Cevallos et al. 1996; Aneja et al. 2005). The fact that PHB is still produced in the *phaA* mutant strain is unsurprising, as in addition to the Azc_0306 gene deleted from the strain, a duplicate acetyl-CoA acetyltransferase (Azc_4535) is present in ORS571, though our results suggest Azc_0306 is the primary one utilised for PHB production, given the dramatic decrease in PHB accumulation in the *phaA* mutant strain OPS1298 (Figure 4.9).

The intermediate growth phenotype observed of the suppressor mutant, lower than wild type and the revertant strain on all carbon sources, but better on all except succinate than the original *phaC* mutant, indicates that there are factors affecting the growth other than metabolic flux (Figures 4.3, 4.4). The *phaC* mutant (OPS0865) had a reduced growth rate on acetate, arabinose, and pyruvate at atmospheric O_2 levels, but not succinate (Figure 4.3). This was expected given the high levels of acetyl-CoA generated by growth on these three carbon sources and is in line with previous PHB mutants in *Rhizobium etli* and RL3841 (Cevallos et al. 1996; Lodwig et al. 2005). The observed growth penalty was far greater under conditions of low O_2, as expected for a mutation affecting such an important redox sink, with reduced growth apparent even on succinate (Table 4.1). OPS0865 was also unable to fix nitrogen in free-living conditions (Figure 4.2). However, the isolation of a *phaC* suppressor mutant (OPS0921) from actively N_2 fixing root nodules of a host *S. rostrata* plant (Figure 4.7, 4.8), combined with the TEM microscopy images confirmed it lacks any trace of PHB within its bacteroids (Figure 4.9). This demonstrates that the lack of PHB *per se* does not prevent N_2 fixation in ORS571.

The location of the SNP within the RBS of the major regulatory gene *phaR* in the suppressor mutant strain suggested an alternative hypothesis; that the lack of PHB resulted in the misregulation of PhaR levels. The negative regulator PhaR is known to affect the expression levels of over 200 genes in other rhizobial species (Quelas et al. 2016). It is an auto-regulatory protein, whose expression is controlled by the sequestration of free PhaR to the surface of polyhydroxyalkanoates such as PHB (Maehara et al. 2002). In the *phaC* mutant strain (*OPS0865*), PHB is absent. While the *phaR* auto-regulates its own expression, PhaR binds to the *phaR* operator with no greater affinity than to a non-specific DNA sequence (Yamada et al. 2013), suggesting a high level of expression is needed for significant binding and auto-regulation. Given the large number of genes regulated by PhaR, and its complex self-regulatory control mechanism, any mutation affecting PhaR levels could result in major phenotypic changes (Nishihata et al. 2018).

Such changes in the translational activity of *phaR* were demonstrated by the use of GFP reporter fusions, with a six-fold increase in translation from the wild type *phaR* RBS when the reporter was in the *phaC* mutant (OPS0865) than when the same reporter was in the wild type strain (ORS571). In contrast, the 0921 SNP RBS reporter gave uniformly low translation rates in both ORS571 and OPS0865 (Figure 4.11A). This means a much higher basal translational level is seen in the *phaC* mutant (*OPS0865*) than in the wild type (Figure 4.11 A), which would result in the down-regulation of over two hundred genes including the key regulator $fixK_2$ (Quelas et al. 2016). This down-regulation, combined with the lack of the major reductant sink PHB would explain the extremely poor growth rate achieved by the *phaC* mutant strain (*OPS0865*). The partial rescue of growth rate in the suppressor strain (*OPS0921*) was achieved by reducing translation of *phaR* to basal levels disconnected from PHB levels (Figure 4.9, 4.10 and 4.11 A). This would have prevented it from causing this widespread downregulation of genes we believe occurred in the *phaC* mutant strain (*OPS0865*), but growth rates could not be restored to wild type levels due to the suppressor stain still lacking PHB as a reductant sink.

To confirm that the difference seen in translation rates was not due to a change in transcription levels qRT-PCR was performed on both ORS571 and OPS0865 (Figure 4.11B). This demonstrated there was no significant difference in transcription of either the *phaR* or *phaA* genes between the two strains. The lack of a significant reduction in *phaA* transcription in strain OPS0865 is surprising given the six-fold increase in translation rate

of PhaR seen in this strain. We would have expected to observe a much greater reduction in *phaA* transcription and this suggests that any reduction in expression levels directly caused by PhaR would be modest. The numerous genes encoding regulatory proteins whose expression is controlled by PhaR levels e.g. *fixk2* could serve to amplify the effect of this modest reduction, contributing to the reduced growth seen in OPS0865 vs OPS0921 (Figure 4.2, 4.3) (Quelas et al. 2016).

Plants were inoculated with strain OPS0865 on twelve separate occasions and each time resulted in N_2 fixation *in planta.* Colonies isolated from these nodules (OPS092N) had an enhanced growth rate compared to the inoculant strain OPS0865. Together with the original growth penalty clearly visible in OPS0865 on all carbon sources under low O_2 conditions (Figure 4.4), this demonstrates that despite PHB not being required for fixation, there is still very strong selection acting against a PHB mutant strain in ORS571. The importance of the PHB reductant sink for growth, as opposed to the PhaR overexpression can be dissected by comparison of growth rates between the wild type, OPS0865 and OPS0921 strains.

The reason for the greatly reduced growth seen by strain OPS0865 on pyruvate and acetate is likely the result of several factors. The first part of the growth penalty can be explained by the increased severity of growth reduction observed in both the OPS0865 and OPS0921 at 1% vs 21% O_2 when compared to strains ORS571 and OPS2077. At 1% O_2 there is an increased need for a reductant sink, making PHBs role in growth far more apparent. Indeed, this role as reductant sink was originally proposed to justify the necessity of PHB synthesis for N_2 fixation (Mandon et al. 1998). The fact that their growth on succinate was unaffected at 21% O_2 and reduced to lesser extent than on pyruvate or acetate at 1% also suggests that PHB is important due to its role as a reductant sink. Unlike acetate and pyruvate, which require a full turn of the TCA cycle for each molecule utilised, succinate feeds directly into the reducing arm of the cycle, and therefore does not have to utilise a full revolution of the cycle to generate energy. This means a greatly decreased amount of reductant is required to generate energy from succinate as a carbon source. Therefore, growth on succinate would not require as great a redox sink as pyruvate or acetate.

The role of PHB as a redox sink explains a great deal of the difference in growth seen in strains OPS0865 and OPS0921 and the wild type on these different carbon sources, as increased NADH/NAD^+ ratios cause suppression of key metabolic enzymes including pyruvate dehydrogenase (Encarnación et al. 2002; Cevallos et al. 1996). However, it does not explain why a differential growth impairment remains between OPS0865 and OPS0921 as both lack PHB as a reductant sink. Our hypothesis is that the remaining growth difference is due to the fact that many of the genes regulated by PhaR encode proteins that are of particular importance for low O_2 metabolism, including *pdhB*, a key part of the pyruvate dehydrogenase complex (Quelas et al. 2016). This would explain why the growth penalty is more severe in OPS0865 than OPS0921, as OPS0865 has a much more severe misregulation of PhaR levels

The fact that reduction in *PhaR* levels alone was sufficient to restore wild type N_2 fixation rates but only partially rescue growth in the suppressor strain (*OPS0921*) indicates that the loss of fixation in the *phaC* mutant strain (*OPS0865*) (Figure 4.2) was most likely due to growth impairment rather than the lack of the reductant sink provided by PHB. This brings ORS571 in line with most other rhizobial species are able to fix nitrogen without PHB synthesis (Lodwig et al. 2005; Cevallos et al. 1996; Quelas et al. 2013) , making use of other reductant sinks such as lipid and possibly glycogen (Lodwig et al. 2005; Wang et al. 2007). The reason why other rhizobial species did not produce a similar altered *PhaR* phenotype upon loss of PHB synthesis is unclear. Given the frequency of suppressor mutation formation in ORS571 when applied to plants (11 of 12 plants across three experiments) and even occasionally on plates, it is possible that since other rhizobia can only fix when in symbiosis, undetected suppressor mutants have occurred in the other rhizobial species. It is easy to imagine a scenario in which such mutants could have been rationalised as revertant mutations if fixation had been observed *in planta* and the mutant was not re-isolated and mapped from the root nodules. Alternatively, ORS571 may have a higher basal expression level of *phaR* compared to other rhizobial species due its propensity to accumulate high levels of PHB (Stam et al. 1986). This would in turn make it more susceptible to the down-regulatory effects of excessive *phaR* expression.

It has previously been suggested that under low O_2 conditions N_2 fixation itself could serve as sink for excess electron production (Resendis-Antonio et al. 2007). Indeed, N_2 fixation

from a PHB deficient mutant was seen to extend N_2 fixation in *Rhizobium etli* (Cevallos et al. 1996) and increased H_2 evolution from nitrogenase activity has been observed in PHB mutants grown on acetate in *Rhodospirillum rubrumi* (Hustede et al. 1993). Previous work has been done to create a metabolic flux map of the *phaC* mutant (Borah 2017), but this will have been affected by the confounding variable of the surge in PhaR levels and the widespread metabolic gene misregulation that resulted in a non-fixing phenotype for the strain. An obvious example would be the expression of the pyruvate dehydrogenase gene family, which have been shown to be affected by both PhaR levels, as well as the NAD^+/NADH ratios within the cell (Nishihata et al. 2018; Encarnación et al. 2002; Cevallos et al. 1996).

A flux map made using the *phaC* suppressor mutant strain OPS0921 would allow an investigation of redox conditions in both free-living and symbiotic nitrogen-fixing symbiosis in the absence of PHB without the confounding variable of increased PhaR levels. This would give us a clearer understanding of how reductant sinks are prioritised in ORS571 and may suggest if it is possible for such a mutation to increase N_2 fixation. Although the peak N_2 fixation rates observed for the suppressor strain OPS0921 has been considerably greater than that of the wild type ORS571 (Figure 4.8), the suppressor also has far greater variation in values. For this reason, there does not appear to be a significant difference between the fixation rates in the absence of PHB, however it does suggest it may be possible to observe this under the right conditions. A more limited alternative to a flux map would be to measure the levels of NAD^+ to NADH in ORS571, OPS0865 and OPS0921. This would determine if the increased PhaR levels had further affected the redox balance of the cell in strain OPS0865 compared to the absence of the PHB sink alone in OPS0921.

In future it would be interesting to conduct an RNAseq experiment to compare the transcription of the ORS571, OPS0865 and OPS0921 strains, in order to gain a clearer understanding of which genes are regulated by *phaR* in ORS571. The levels of any such genes should be significantly altered from the wild type in the OPS0865 strain, but at wild type levels in strain OPS0921. It could additionally indicate whether it is acting as a positive or negative regulator of gene expression as both effects have been reported in *B. japonicum* (Nishihata et al. 2018).

It is also interesting to note that strain OPS0921 was restored to wild type levels of fixation, despite still having reduced levels of growth when compared to strain ORS571. This implies a disconnection between the health of the cell and its efficacy as a nitrogen fixer. A competition assay between the OPS0921 and ORS571 could allow us to see if plant sanctioning will select against the metabolically impaired strain, or if it will respond to levels of N_2 fixation alone. This could be observed as reduced size or number of nodules when compared to that of the wild type strain or when inoculated alone. It is likely that a PHB mutant will suffer a severe penalty to its competitive ability, given that PHB defective mutants have been reported widely in the literature to have delayed nodule formation along with reduced occupancy when in competition with wild type in *S. meliloti* (Wang et al. 2007; Aneja et al. 2005; Cevallos et al. 1996). In addition to a competitive penalty, the strains themselves may prove less robust as PHB has been theorised to support bacteroids during periods of reduced carbon flux such as seed development or overnight (Bergersen et al. 1991).

Overall, it is clear that while PHB is not strictly necessary for N_2 fixation in ORS571, it still has an important role as both a redox sink and for maintaining proper levels of PhaR. Interruption of either of these roles can result in a severe growth penalty, particularly when grown on carbon sources that produce high levels of acetyl-CoA like pyruvate or acetate. The complex interplay between regulation of PHB synthesis and redox levels still remains open to investigation.

Chapter 5

Role of the Malonate Shunt and dicarboxylate metabolism in ORS571

5.1 Introduction

The ability of ORS571 to fix N_2 in both symbiotic and free-living conditions allows us the possibility to separate dicarboxylate metabolism from N_2 fixation in a way not possible for other obligate symbiotic fixers. In free-living conditions ORS571 can be grown on a wide variety of carbon sources and still fix N_2, whereas during symbiosis rhizobia are provided with C_4 dicarboxylates as their primary carbon source (Lodwig and Poole 2003; Poole and Allaway 2000). This means that during symbiosis any perturbation to dicarboxylate metabolism is highly likely to affect N_2 fixation. If, however, the same mutation affects N_2 fixation in free-living conditions when ORS571 is grown on alternative carbon sources then it is likely that the mutation will have produced a more generally deleterious effect on ORS571 rather than a perturbation of the dicarboxylate metabolism. An example of such a generally deleterious mutation is the *phaC* mutant in Chapter 4.

In order for the obligate aerobic rhizobia to survive in the micro-aerobic conditions of root nodules, they need to carefully balance flux through the TCA cycle. Enough flux must pass through in order to generate sufficient energy for N_2 fixation and basic survival functions, but too much flux will result in a build-up of excess reductant due to the extremely limited O_2 available to serve as electron acceptors for oxidative phosphorylation. There are mechanisms to deal with this excess reductant, such as the formation of carbon storage molecules as reductant sinks e.g. PHB, lipids or glycogen (Terpolilli et al. 2016; Lodwig et al. 2005). There is also evidence that modifications to the TCA cycle as whole occur, with the de-carboxylating arm of the cycle being down-regulated relative to the reductive arm (Lodwig and Poole 2003). In this way, the majority of the carbon flux from C_4 dicarboxylates does not result in a full turn of the TCA cycle. Instead a single molecule of reductant can be generated via malate dehydrogenase activity, whilst the majority of the malate can be converted to pyruvate via NAD^+ dependent malic enzyme (DME) activity or the combined activity of PEP carboxykinase (PEPck) and pyruvate kinase (Mulley et al. 2010), which in turn can be used to create acetyl-CoA, which can be stored as PHB or lipid. These energy stores can then be broken down to feed acetyl-CoA into the TCA cycle in a controlled manner to generate energy without generating an excess of reductant. However, this method is reliant on the conversion of C_4 dicarboxylates to acetyl-CoA, a process which is generally accepted to require the formation of pyruvate as an intermediate. This made the results of (Pauling et al. 2001) all the more surprising when they isolated a pyruvate

dehydrogenase (PDH) complex mutant of ORS571 capable of growth on C_4 dicarboxylates. They went on to isolate temperature sensitive mutants through growth on a valine rich selection media, and found growth on this media at the permissive temperature was reliant on malonate semialdehyde dehydrogenase (MSDH) activity. From this they went on to propose what they termed the Malonate Shunt, a PDH independent pathway for the generation of acetyl-CoA from C_4 dicarboxylates. A summary of the potential acetyl-CoA producing pathways can be seen in figure 5.1.

The shunt as it was originally proposed terminated with the CoA dependent production of acetyl-CoA from malonate semialdehyde through the activity of malonate semi-aldehyde dehydrogenase. To reach malonate semialdehyde from the starting point of malate, it was proposed to proceed via malate to oxaloacetate, to aspartate, onto beta-alanine and finally to malonate semialdehyde and subsequently acetyl CoA (Figure 5.2).

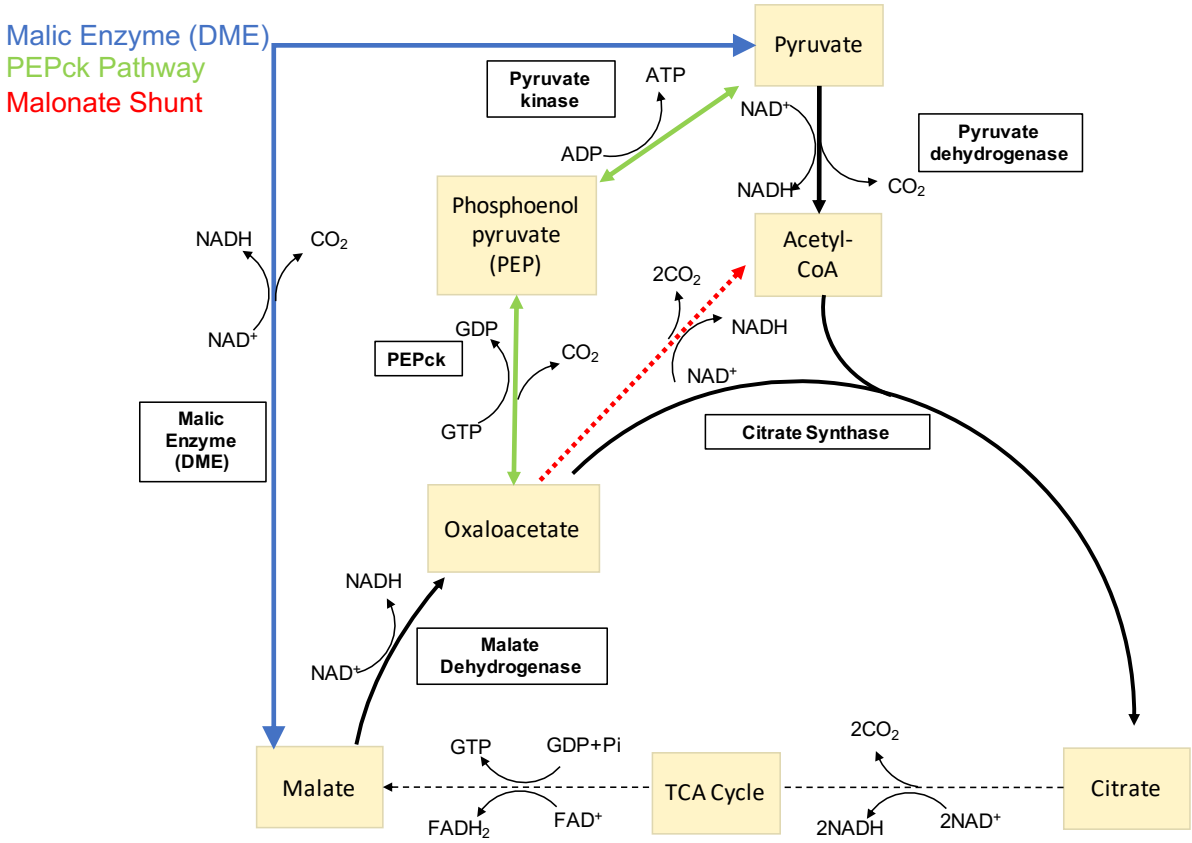


Figure 5.1 Summary of Acetyl-CoA generating reactions feeding into the TCA cycle.

Solid lines indicate single enzyme activities, dashed lines indicate a summary pathways involving multiple enzyme activities. Co-factors and side products are indicaeed in black alongside each reaction. Yellow boxes indicate chemicals, black boxes indicate enzymes. The dashed red arrow indicates the overall pathway of the malonate shunt which is illustrated in more detail in Figure 5.2. The green arrows indicate the combined PEPck pathway and the blue arrow indicates DME malic enzyme activity.

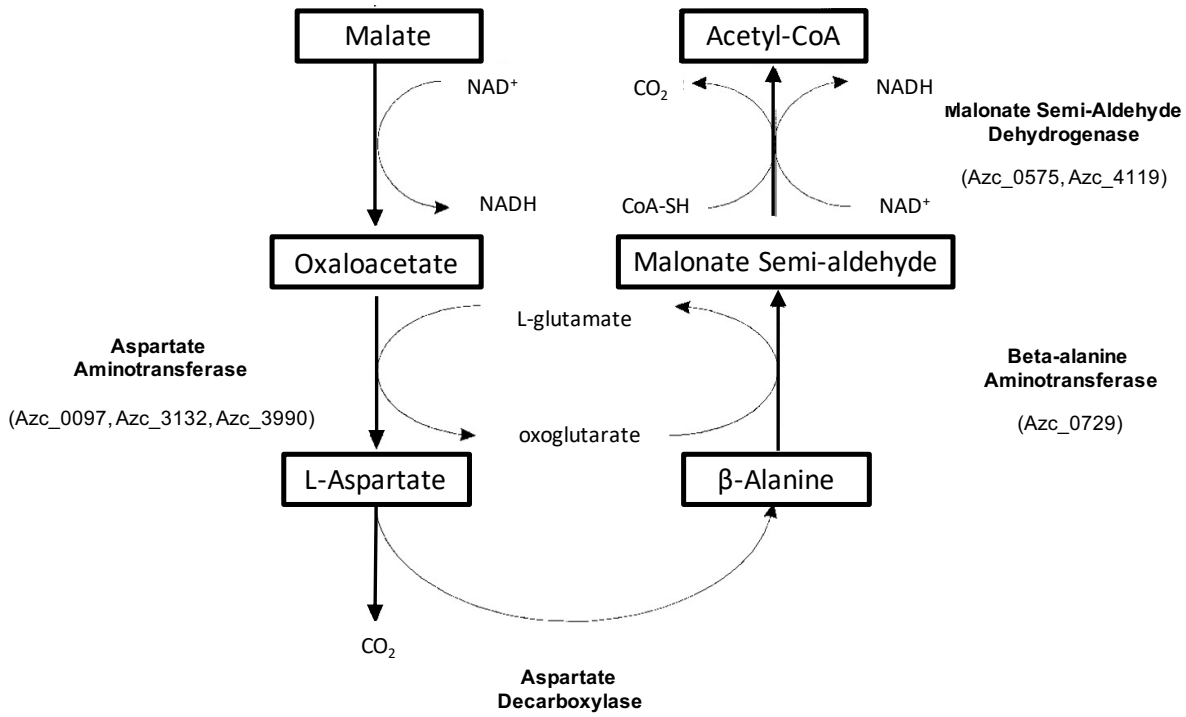


Figure 5.2 Proposed pathway of the malonate shunt

Malonate shunt pathway as proposed by (Pauling et al. 2001). Metabolites involved in the shunt are indicated by black boxes. Black arrows indicate chemical reactions. The enzyme functions required to produce one metabolite from another are indicated in bold above potential ORS571 genes suggested to have that function identified during our bioinformatic analysis. Figure modified from (Pauling et al. 2001).

Our interest in the shunt began with the metabolic flux maps generated by (Borah 2017). In that work, metabolic flux maps were generated for ORS571 under free-living and symbiotic N_2 fixing conditions using ^{13}C isotopically labelled succinate. Critically, the maps indicated a major difference in flux under symbiotic and free-living conditions, with a large flux flowing through the malonate shunt during symbiotic N_2 fixation (Figure 5.3). This implied a distinct difference between the two states existed and made dissecting the nature and purpose of the malonate shunt of major importance in ORS571. This map was somewhat unexpected, as PDH mutants which would rely on the malonate shunt for acetyl-CoA generation, were incapable of symbiotic N_2 fixation *in planta* but could fix nitrogen at low levels in free-living conditions (Pauling et al. 2001). To address this contradiction, we investigated how malate and oxaloacetate are metabolised by ORS571 via all known pathways including the proposed malonate shunt.

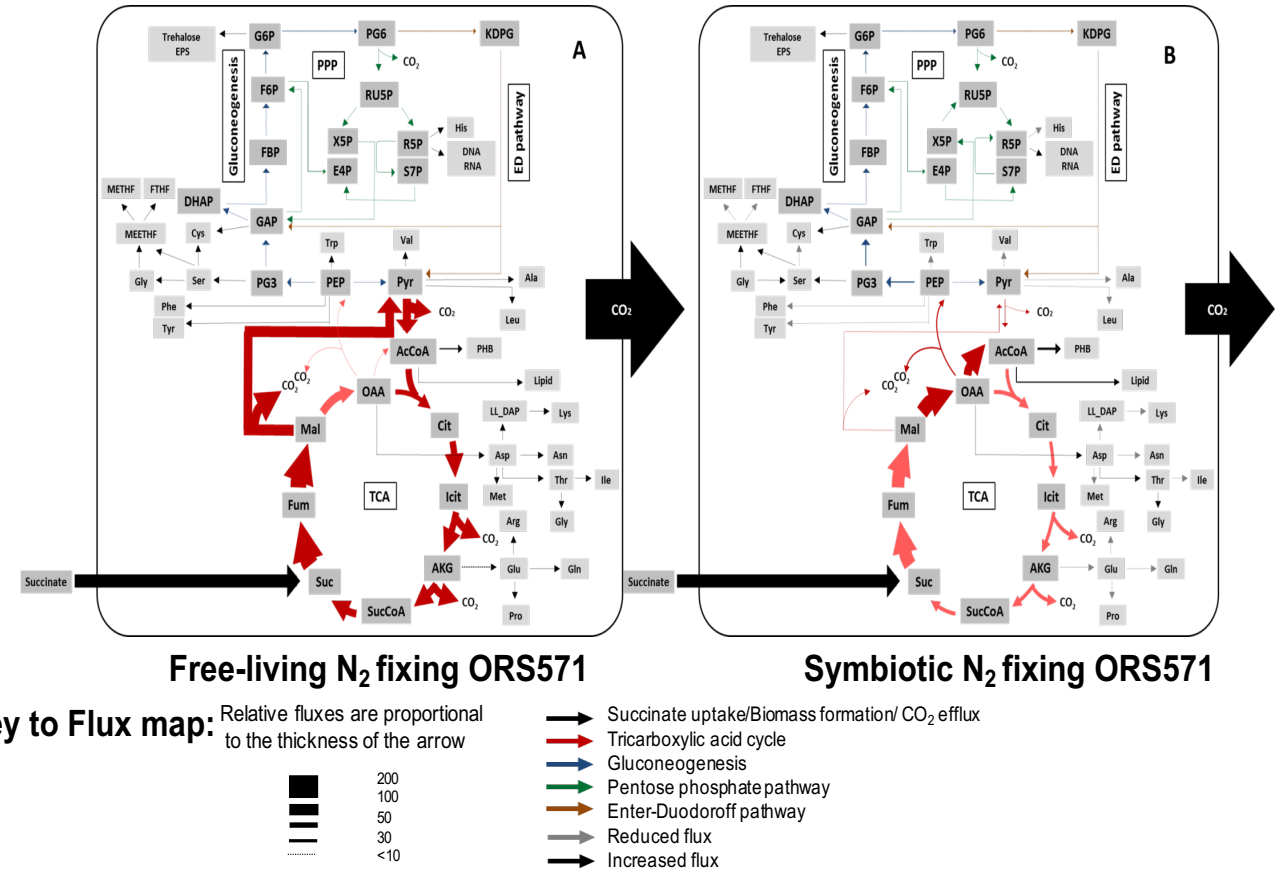


Figure 5.3 Flux maps of symbiotic and free-living N_2 fixing ORS571

Comparison of flux maps of free living N_2 fxing and symbiotic N_2 fixing ORS571. The flux map represents statistically valid best-fit models based on the redistribution of label following metabolism of ^{13}C labelled succinate by N_2 fixing ORS571. The thickness of each arrow in the flux maps is proportional to the relative molar flux. Metabolic intermediates and biomass outputs are shown in rectangular boxes with distinct colour keys. Reproduced from (Borah 2017)

5.2 Identification of potential shunt genes

We began our investigation by identifying potential genes involved in the malonate shunt. A theoretical pathway for the malonate shunt was provided by (Pauling et al. 2001) indicating several functions that would be required, including aspartate aminotransferase, aspartate decarboxylase, beta alanine aminotransferase and malonate semialdehyde dehydrogenase. This was used a starting point for identifying genes potentially related to the shunt. The microarray data of (Tsukada et al. 2009) was then used to look for genes with these suggested, or related activities, that were upregulated in bacteroids compared to free living nitrogen fixing conditions (Table 5.1, Figure 5.2). Additionally, the flux map data produced by (Borah 2017) suggested a clear phenotype for the shunt, given the apparent preponderance of the shunts utilisation in symbiotic conditions vs free-living. This suggested phenotype could then be combined with the transposon mutagenesis study of (Suzuki et al. 2007) to identify genes that caused a loss of fixation in symbiotic, but not free-living conditions. In this way, candidate genes were identified for almost every step of the proposed pathway including aspartate aminotransferase (Azc_3132), beta alanine aminotransferase (Azc_0729) and MSDH (Azc_0575, Azc_4119). However, candidate genes for the aspartate decarboxylase step could not be identified.

Table 5.1 Putative malonate shunt genes

Putative malonate shunt genes were identified through a combination of annotated gene function, expression change in bacteroids vs free-living growth in TY media and the presence or absence of a non-fixing symbiotic phenotype. The non-fixing symbiotic phenotype was predicted from the flux maps of (Borah 2017), which indicated high importance for the shunt in symbiosis but not free-living conditions.

Gene	Annotated Gene Function	Fold-expression change in bacteroids vs TY	None-fixing phenotype in bacteroids, but not free – living conditions
Azc_0097	Putative aspartate aminotransferase	Up 0.4x	
Azc_0119	NAD dependent malic enzyme	Up 6.6x	
Azc_0575	Methylmalonate semialdehyde dehydrogenase	Down 2.3x	
Azc_0729	Beta-alanine aminotransferase	Down 0.19x	
Azc_1184	Betaine aldehyde dehydrogenase	Down 0.4x	**Yes**
Azc_1847	Aminotransferase class I	Down 0.05%	**Yes**
Azc_3132	Aspartate aminotransferase	No Change	
Azc_3423	DegT aminotransferase	Up 5.5x	
Azc_3656	NAD dependent malate dehydrogenase (DME)	Down 0.1x	**Yes**
Azc_3990	Aspartate aminotransferase	Up 3.7x	
Azc_4119	Malonate semi-aldehyde dehydrogenase	Down 0.5x	

5.2.1 Bioinformatics demonstrating lack of aspartate decarboxylase

Candidate genes were found for almost every step of the suggested pathway however, one step could not be identified. Aspartate decarboxylase, produced from the highly conserved *panD* gene, is not present in ORS571, and BLAST analysis did not reveal any proteins with significant identity to PanD. While this does not preclude the presence of a hitherto unidentified family of aspartate decarboxylase proteins within ORS571, it suggests the putative malonate shunt may not proceed by the pathway as proposed Pauling *et al.*, 2001. It is noteworthy that the only experimental evidence for the putative malonate shunt pathway of Pauling *et al.*, 2001 was malonate semialdehyde dehydrogenase dependent production of

acetyl CoA present in the *pdh* mutant grown at the permissive temperature and absent in the cultures transferred to the restrictive temperature .

5.2.3 Testing putative shunt enzyme activities

Assays were therefore developed for all the steps of the putative malonate shunt, starting with the first enzyme aspartate aminotransferase (Figure 5.2). It was successfully assayed from the wild type strain with an average measured activity of 0.311µmol glutamate/mg of protein/minute. In the aspartate aminotransferase reaction, the amine group from aspartate is donated to oxoglutarate forming oxaloacetate and glutamate, allowing aspartate aminotransferase activity to be determined through glutamate production versus a heat killed control. The assay was run using 10 mM oxoglutarate, 10 µg cell lysate and started with the addition of 50 mM aspartate. The glutamate formed was quantified using a glutamic dehydrogenase assay and compared to a standard curve of 0-120 nM of glutamate. The second step of the pathway is aspartate decarboxylase but given the absence of any obvious candidate protein and the difficulty of the assay this was omitted.

The third enzyme of the pathway, beta-alanine aminotransferase did not donate from beta-alanine to oxoglutarate as required by the shunt. However, it did donate an amino group from beta-alanine to pyruvate to form alanine with an activity of 0.93 nmol of alanine/mg of protein/minute. The initial enzyme assays were performed using 50 mM beta-alanine with either 50 mM pyruvate or oxoglutarate and the products (alanine or glutamate) were measured with alanine or glutamate dehydrogenase. Standard curves of alanine (0-30 nM) and glutamate (0-120 nM) were prepared.

5.3 Investigating the importance of the DME-like Azc_0119

5.3.1 Bioinformatic analysis of Azc_0119

Given the uncertainty in the shunt engendered by the lack of a clear aspartate decarboxylase gene or beta alanine aminotransferase donating to oxoglutarate donation, we began to consider alternatives to the pathway originally proposed.

During our bioinformatic screening, our attention was drawn to gene Azc_0119 , which was indicated to be six-fold upregulated in bacteroids (Tsukada et al. 2009). This gene was labelled as an NAD^+ dependent malic enzyme, but shared only a 27.8 % sequence identity to the known and characterised DME of ORS571 *(Azc_3656*). This suggested that while it was related to this gene, it was highly unlikely to have precisely the same activity. A BLASTp search was conducted and resulted in an 82% match to the oxaloacetate decarboxylase gene from the Betaproteobacterium *Aquitalea magnusonii.* This bioinformatic data, along with the attempts by (Zhang et al. 2012) to characterize the DME activity of both Azc_3656 and Azc_0119 make it highly probable that this gene has an alternative primary decarboxylating function and the remaining 2% DME activity observed by (Zhang et al. 2012) is a legacy of its origin.

Several possible alternative functions for this enzyme were considered. Given the similarity in structure between malic acid and oxaloacetic acid, it was impossible to be sure which is the more likely substrate. The most likely co-factors for any reaction would be Mn^{2+}/Mg^{2+} ions, along with NAD+ or NADP+. The structural similarity of these substrates also raised the intriguing possibility of Azc_0119 being a form of malolactic enzyme. These enzymes were originally derived from DME, and use both Mn^{2+} ions and NAD^+ as co-factors. Though NAD^+ is required it has no net reduction during this reaction (Caspritz and Radler 1983). In theory this could provide ORS571 with a novel method in rhizobia of dealing with excess reductant by converting malic or oxaloacetic acid to lactic acid and excreting the excess as occurs in malolactic bacteria.

While intriguing, this was far from the only hypothesis. Acting as an oxaloacetate decarboxylase it was equally possible for this gene to produce pyruvate or malonic semialdehyde. If the latter proved to be the case, then the malonate shunt could consist of

only two genes, Azc_0119 and one or both or the malonate semialdehyde dehydrogenases (Azc_0575 and Azc_4119). Alternatively, the same decarboxylating activity could result in the conversion of malonic acid to either pyruvate or 2-hydroxy-propionic acid, which itself could be easily converted to malonate semialdehyde (Figure 5.4).

While all these possibilities were intriguing, they had to be tested. Using the most likely set of ionic cofactors utilised by both malic and malolactic enzyme (Mn^{2+}, Mg^{2+}) combined with the most likely electron acceptors (NAD^{+}, $NADP^{+}$) along with the more unlikely (NADH, NADPH) both substrates were tested and the products produced checked using GCMS, in addition to the monitoring of the enzymatic reactions (Caspritz and Radler 1983; Lonvaud-Funel and Strasser de Saad 1982). Before the enzyme activity could be tested, first pure enzyme samples needed to be obtained.

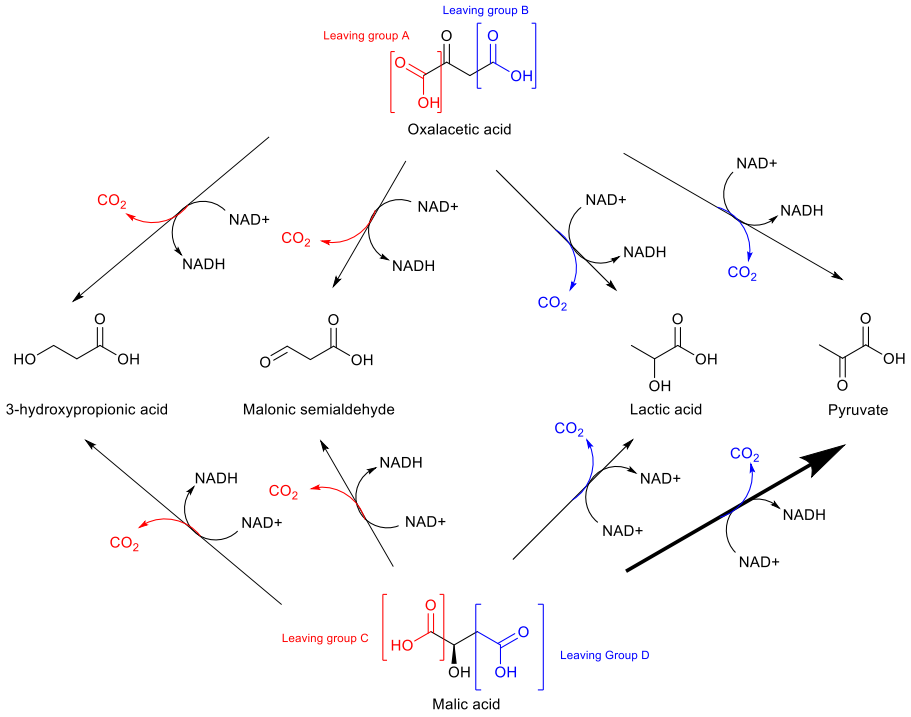


Figure 5.4 Skeletal structures of potential DME-like enzyme (Azc_0119) substrate and products

Names of the chemicals appear below the structures. The carboxyl leaving group required produce each product is colour coded in either blue or red. All reactions are believed to require Mn/Mg^{2+} as co-factors along with NAD^{+}. Black arrows indicate potential reactions, the bold arrow indicates the reaction previously confirmed by (Zhang et al. 2012).

5.3.2 pOPINE expression of DME- like (Azc_0119) and DME (Azc_3656)

The pOPINE vector was used for expression of both Azc_0119 and Azc_3656. Primers oxp2887 +2888 were used to amplify Azc_0119 (absolute position 142,680-144,353), whilst oxp3685+3686 were used to amplify Azc_3656 (absolute position 4,199,096-4,201,372). In both cases the STOP codon was removed to allow for the attachment of the C-terminal His tag. Each gene was InFusion (ClonTech) cloned into pOPINE vectors linearised with HindIII and KpnI. Purified Azc_3656 enzyme was subsequently used as a positive control for the DME assay. The Azc_3656 expressing plasmid was named pOPS1355 and the Azc_0119 expressing plasmid was named pOPS1117.

5.3.3 Purification of DME -like (Azc_0119) and DME (Azc_3656)

The pOPINE vectors were expressed in BL21 chemically competent cells and induced by IPTG induction. After harvesting, the proteins were ribolysed and the cell lysate was purified by both His spin column and AKTA column purification. Although both enzymes were purified the induction was more successful for Azc_3656 (Figure 5.5 lanes 4-5) than Azc_0119 (Figure 5.5 lanes 9-10)

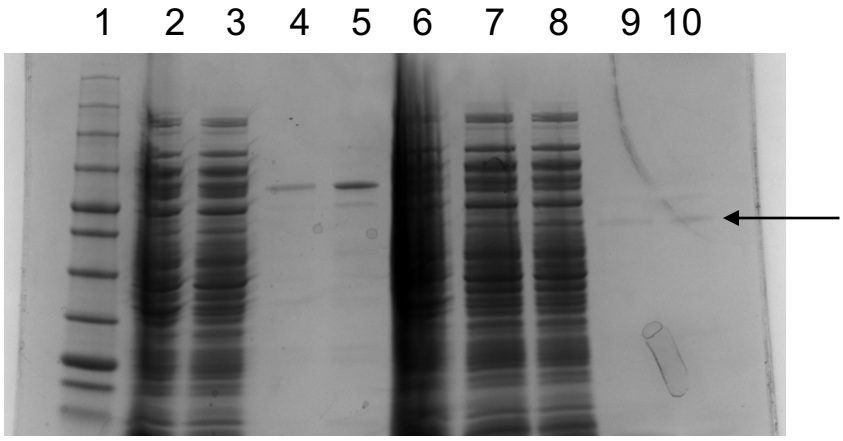


Figure 5.5 Purification of DME & DME-Like enzymes

Coomassie gel of induced BL21 cells containing DME (Azc_3656) and Malic enzyme -like (Azc_0119) expressing pOPINE vectors. Lane 2-6 contain the Azc_3656 expressing pOPINE, lanes 7-10 contain the Azc_0119 expressing pOPINE. 1 – Geneflow BLUeye prestained protein ladder, 2- Non-ribolysed induced cells, 3 – ribolysed induced cells, 4- His Spin purified, 5 Amicon concentrated purified protein, 6- Flow through, 7- non-riboylsed induced cells, 8- ribolysed induced cells, 9- His Spin purified, 10- Amicon concentrated purified protein. Purified Azc3656 protein clearly visible in lanes 4 and 5. Black arrow indicates purified Azc_0119 in lanes 9 and 10.

5.3.4 Malolactic enzyme assay

The malolactic enzyme assay was adapted for the pH ranges of rhizobia from the method used by (Schümann et al. 2013) to monitor DME activity. The reaction was started by addition of 1.5 mM final concentration NAD^+ to 100 mM HEPES pH 8.0, 1mM Mn^{2+}, 10 mM malate, 1 µg purified Azc_0119 DME-like. The lactate content of the samples was then measured using a lactate dehydrogenase assay and compared to a standard curve prepared from L-lactate 0-15 mM. As the sample contained some NAD^+ from the initial assay, this

reaction was begun by addition of the lactate dehydrogenase. The reactions were incubated at 37 ˚C for 30 minutes and its progress was measured by the change in absorbance at OD 340 nm. No evidence was seen for malolactic enzyme activity, although an interaction between the Mn^{2+} ions from the assay mix and the phosphate ions in the His spin column elution buffer was observed. This interaction was lost after the purified malic enzyme -like (Azc_0119) was buffer exchanged into 100 mM HEPES, pH 8.0 prior to its use in the reaction. In addition to the monitoring of the OD 340 nm absorbance, samples of the incubated assay were derivatised with MSTFA and GCMS was used to ascertain whether pyruvate or lactate were being produced from this reaction (Figure 5.7 A). No lactate was apparent on the GCMS.

5.3.5 DME assay

The assay was established using 0.5µg of purified DME (Azc_3656) as a positive control and was performed in 100 mM Tris (pH 8), 1.5 mM Mn^{2+} 10 mM malate and 1.5 mM NAD^{+} in a 200µl volume. The reaction was started by addition of the NAD^{+} and the change in absorbance at OD_{340} was measured on a FLUOstar Omega spectrophotometric plate reader (BMG Labtech). Clear activity 0.95 mM $minute^{-1}$ (mg protein)$^{-1}$ was seen from Azc_3656 with a calculated Km of 7.8 mM with respect to malate (Fig 5.6), similar to the 2.7 mM value previously observed (Zhang et al. 2012), however, no reaction was apparent with purified Azc_0119 at this enzyme concentration.

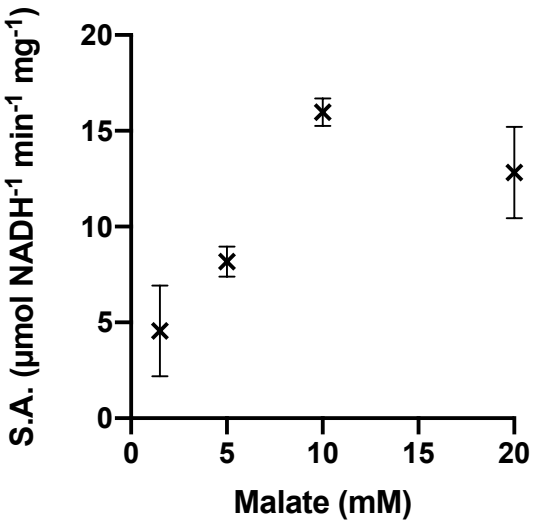


Figure 5.6 A – Specific activity (S.A.) of DME with respect to malate

DME specific activity (S.A.) measured using 0.5 µg of purified DME enzyme with varying concentration of malate from 1.5 to 20 mM. Mean values are plotted with bars of SEM, N=3.

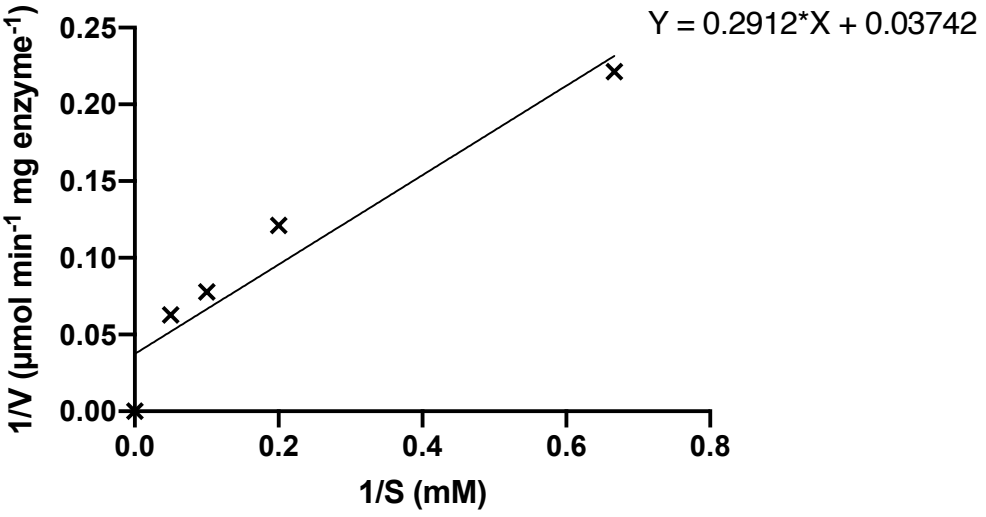


Figure 5.6 B – Lineweaver-Burke Plot of DME with respect to malate

Generated from the S.A. data pictured in figure 5.6 A and used to calculate both the Vmax Y intercept = 1/Vmax) and Km (Gradient * Vmax) of DME with respect to malate.

5.3.6 Oxaloacetate decarboxylase assay

This method was modified from the assay of (Sender et al. 2004) to be optimised for rhizobia pH, 100 mM Tris (pH 8), 1.5 mM Mn^{2+}, 10 mM oxaloacetate and 1.5 mM NAD^{+} in a 200µl volume. Oxaloacetate levels were monitored by absorption at 280 nm for thirty minutes, however the natural rate of decarboxylation rendered this form of monitoring inconclusive. We decided to repeat the assay and use GCMS to attempt ascertain what product if any was being produced by the reaction. If something other than pyruvate was being produced by an enzyme catalysed reaction, it should produce an alternate peak pattern to that observed in the absence of the enzyme. As the nature of the reaction was uncertain, different combinations of cofactors were utilised to make sure no reaction was missed. All combination of NAD (NAD^{+}, NADH, $NADP^{+,}$ NADPH) and ionic co-factors (Mn^{2+} and Mg^{2+}) were tested for oxaloacetate decarboxylase activity (Figure 5.7). The background rate of spontaneous decarboxylation made this assay extremely difficult to determine. However, comparison of the background rate of decarboxylation to pyruvate production with the enzyme suggests that the enzyme has little to no effect. It does not appear that anything was produced in any reaction other than pyruvate. It should be noted however, that since we did not have a pure sample of malonate semialdehyde it was not possible to clearly distinguish between this and pyruvate on the GCMS.

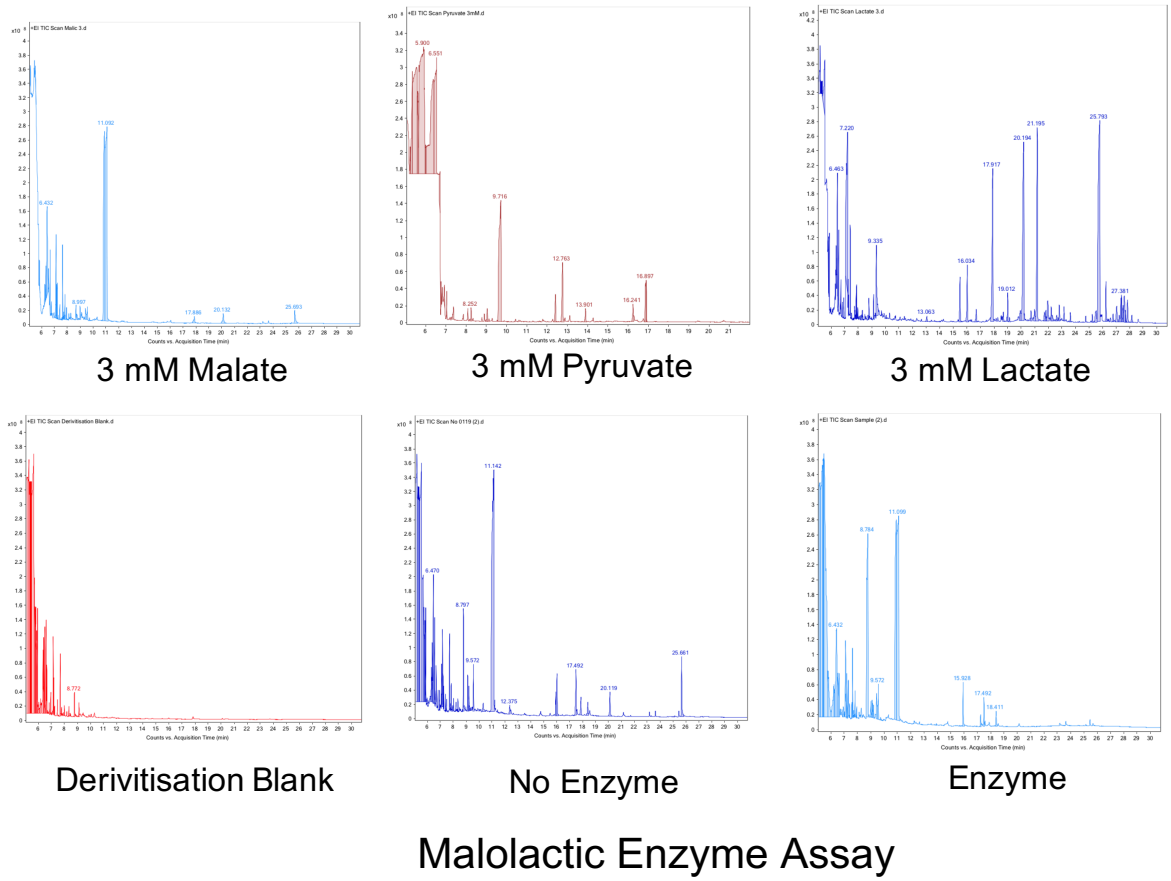


Figure 5.7 A - Malolactic enzyme assay

Malate, **Pyruvate** and **Lactate** are reference standards each at 3 mM concentration. **Derivatisation Blank** was MSTFA. **No Enzyme** was a sample of the malolactic enzyme assay performed with no added Azc_0119 , **Enzyme** was the same reaction with 10 µg Azc_0119 added.

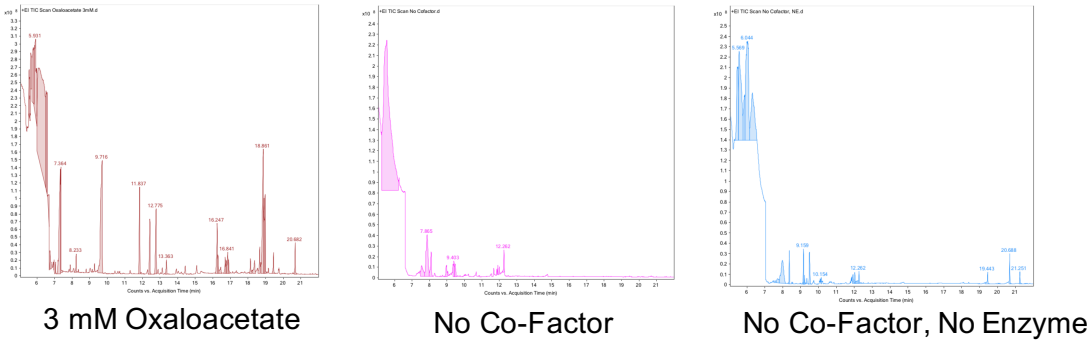


Figure 5.7 B Oxaloacetate decarboxylase assay – No Co-Factor Variant

Oxaloacetate was a reference standard at 3mM concentration. **No Co-Factor** was an oxaloacetate decarboxylase assay of reaction buffer, oxaloacetate and 10µg of Azc_0119 with no ionic or NAD based cofactors. **No Co-Factor, No Enzyme** was reaction buffer and oxaloacetate.

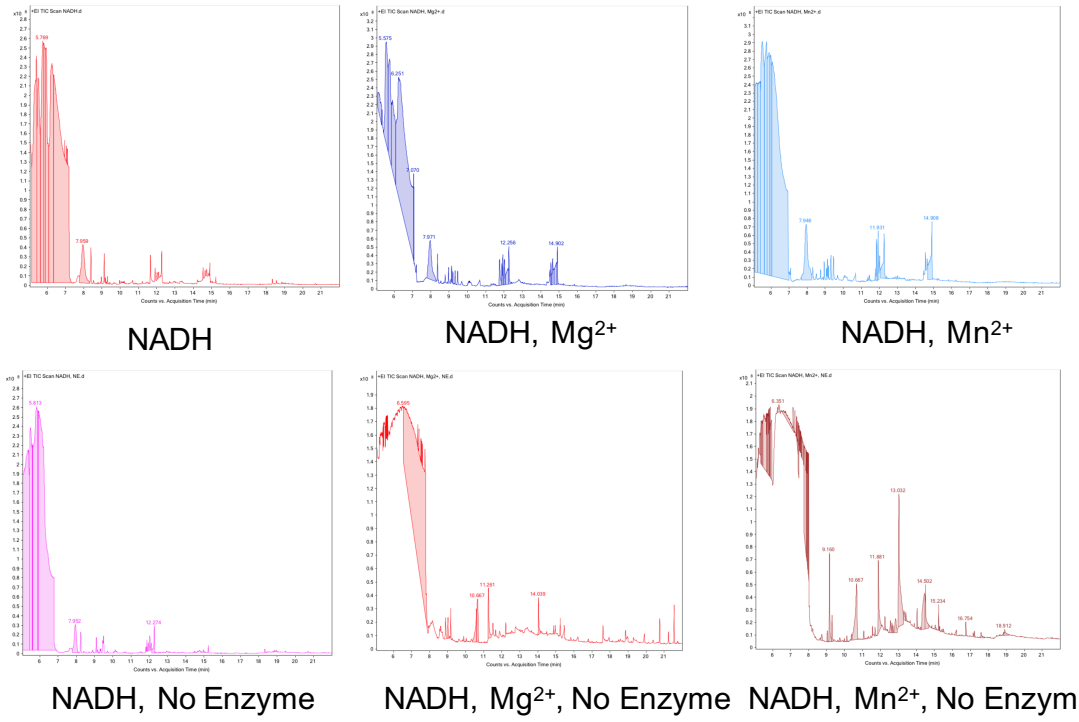


Figure 5.7 C Oxaloacetate decarboxylase assay - NADH Variants

All samples contain reaction buffer, and oxaloacetate; **NADH**, (**NADH, Mg^{2+}**), **(NADH, Mn^{2+})** contained 10µg Azc_0119 and the stated co-factors. **(NADH, No Enzyme)**, **(NADH, Mg^{2+}, No Enzyme), (NADH, Mn^{2+}, No Enzyme)** contained only the stated co-factors.

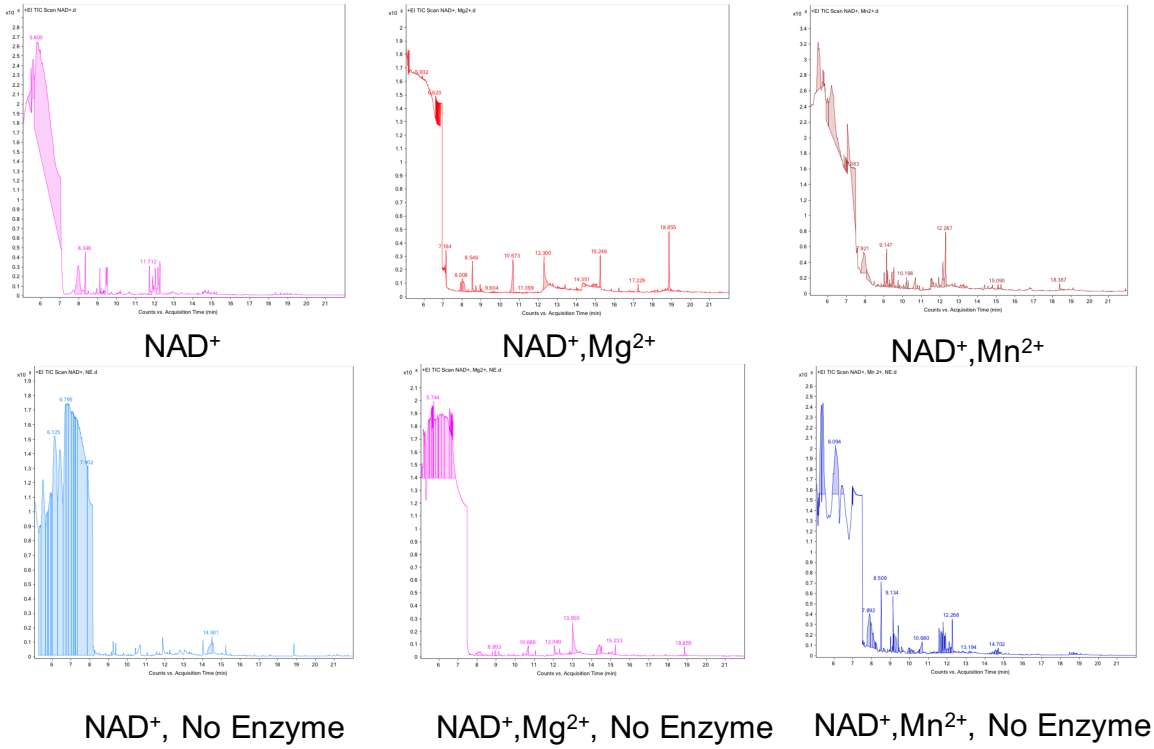


Figure 5.7 D Oxaloacetate decarboxylase assay - NAD+ Variants

All samples contain reaction buffer, and oxaloacetate; **NAD^+**, **(NAD^+, Mg^{2+}), (NAD^+, Mn^{2+})** contained 10µg Azc_0119 and the stated co-factors. **(NAD^+, No Enzyme)**, **(NAD^+, Mg^{2+}, No Enzyme), (NAD^+, Mn^{2+}, No Enzyme)** contained only the stated co-factors.

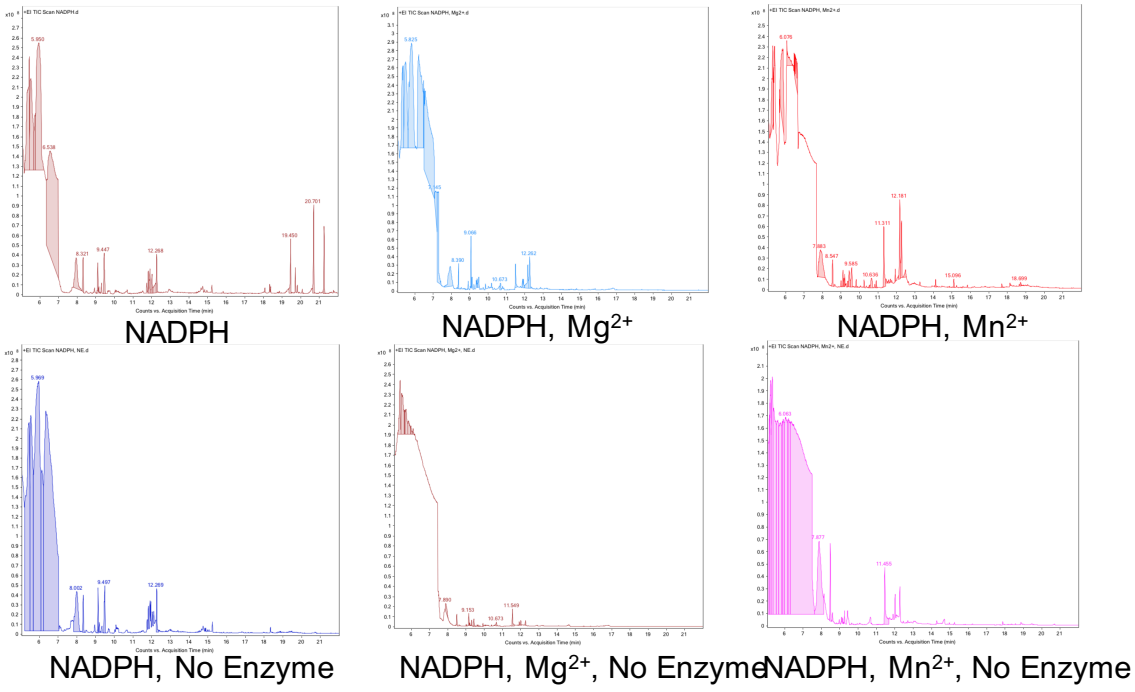


Figure 5.7 E Oxaloacetate decarboxylase assay - NADPH Variants

All samples contain reaction buffer, and oxaloacetate; **NADPH**, (**NADPH, Mg^{2+}**), **(NADPH, Mn^{2+})** contained 10µg Azc_0119 and the stated co-factors. **(NADPH, No Enzyme)**, (**NADPH, Mg^{2+}, No Enzyme), (NADPH, Mn^{2+}, No Enzyme)** contained only the stated co-factors.

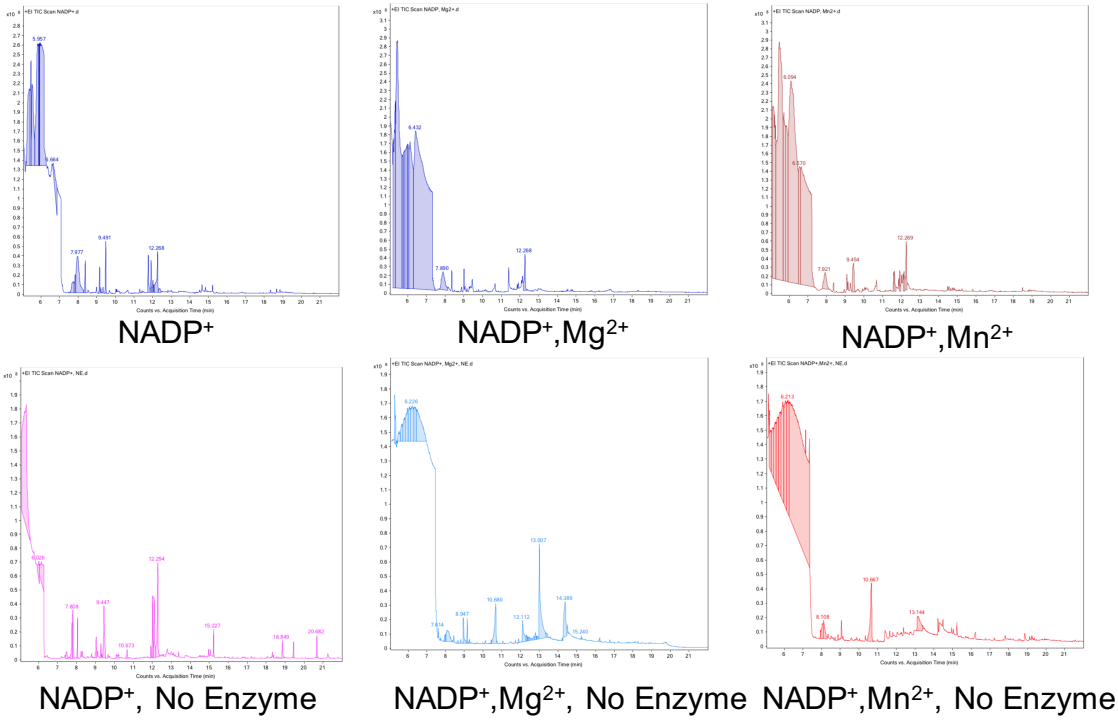


Figure 5.7 F Oxaloacetate decarboxylase assay - NADP+ Variants

All samples contain reaction buffer, and oxaloacetate; **NADP$^+$**, (**NADP$^+$, Mg^{2+}**), **(NADP$^+$, Mn^{2+})** contained 10µg Azc_0119 and the stated co-factors. **(NADP$^+$, No Enzyme)**, **(NADP$^+$, Mg^{2+}, No Enzyme), (NADP$^+$, Mn^{2+}, No Enzyme)** contained only the stated co-factors.

5.3.7 DME-like (Azc_0119) mutant

As the enzymatic assays did not reveal any novel activity, we decided to take a genetic approach to determine the importance of this enzyme within any malonate shunt. An in frame double integration deletion mutant was isolated using pK19msB to replace Azc_0119 (absolute position 142,680-144,356) with a spectinomycin resistance gene flanked by two omega terminators. This mutant (OPS1951) was confirmed via sequence mapping with primers oxp3095/3096 + pot Forfar from the spectinomycin cassette. OPS1951 exhibited no clear growth phenotype and proved capable of symbiotic N_2 fixation (Figure 5.8). This suggested one of three possibilities: Azc_0119 was not a part of the malonate shunt; Azc_0119 is a part of the malonate shunt however its lack of activity can be compensated for by DME or other pathways unknown; or the malonate shunt does not exist. To determine which of these possibilities was true, it was decided that we needed to knock out the most likely path to be compensating for the lack of malonate shunt activity, the DME Azc_3656.

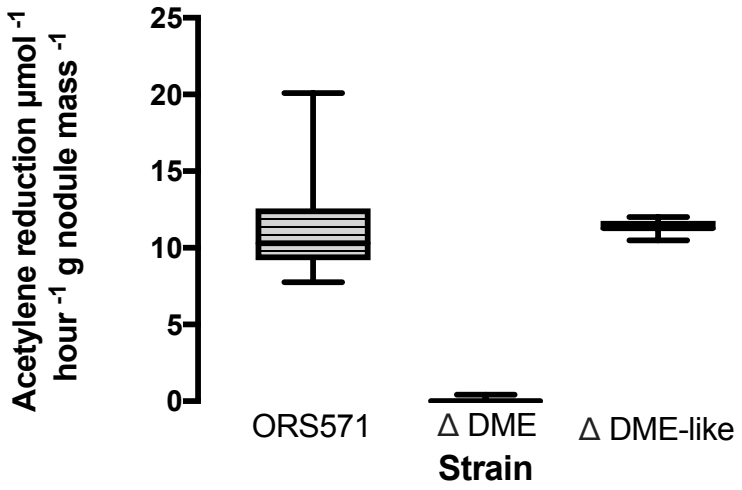


Figure 5.8 Symbiotic N_2 fixation of ORS571, DME and DME-like mutant

Symbiotic fixation rates observed 28 days post-inoculation on *S. rostrata*. Acetylene reduction rates are corrected per gram of nodule mass. **ORS571** is the wild type strain, the **DME mutant** is strain OPS2090, the **DME-like mutant** is strain OPS2094. Boxes are of the interquartile range, with the median indicated within and bars displaying the minimum and maximum values, N=4.

5.3.8 Generation of a double malic enzyme (Azc_3656) and DME-like (Azc_0119) mutant

An in-frame marker-less mutant of Azc_3656 termed OPS2090 was isolated using plasmid pOPS1354 and selection on pyruvate supplemented sucrose media to remove any growth penalty suffered by a DME mutant. Gene deletion was confirmed via PCR across the gene region with oxp3390+3391 resulting in a 300 bp band in the mutant, as opposed to 2500 bp observed in the wild type. This single DME mutant was demonstrated to be defective in symbiotic N_2 fixation (Figure 5.8). An Azc_0119 mutant was then produced in strain OPS2090 using plasmid pOPS1240 to produce a double mutant named OPS2094. To confirm there were not additional sources of DME activity in the strains, a DME assay was conducted on ORS571, OPS1951, OPS2090 and OPS2094 using 10 µg of protein from a total cell lysate. In addition, potassium cyanide at 1 mM working concentration was added to the reaction to knock down electron transport chain oxidation of NADH. The reaction mix and method were otherwise identical to that used for the pure protein described earlier in this chapter. The deletion of Azc_3656 in OPS2090 resulted in a near total loss of DME activity, with no difference in the NADH levels of OPS2090, OPS2094 and the heat killed cell lysate controls until nearly three hours into the assay. By contrast the single Azc_0119 mutant exhibited greater than wild type levels of activity, though this is most likely due to differences in the timing of culture being harvest rather than a true increase (Figure 5.9).

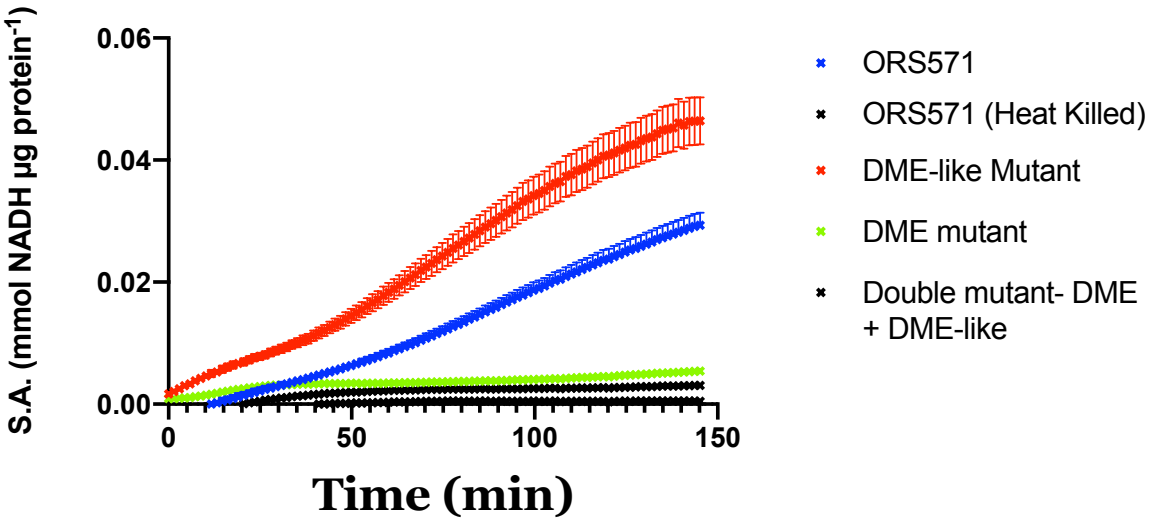
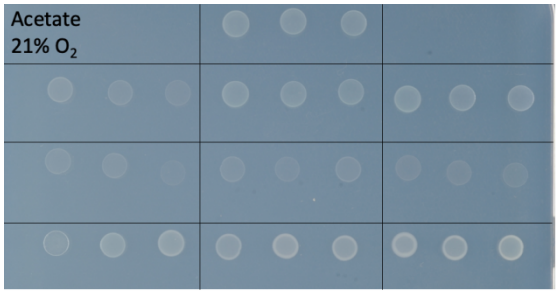
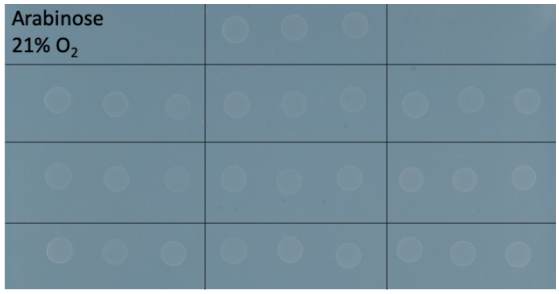


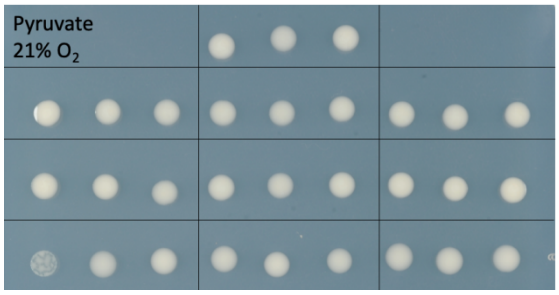
Figure 5.9 DME activity of ORS571 and DME mutant variants

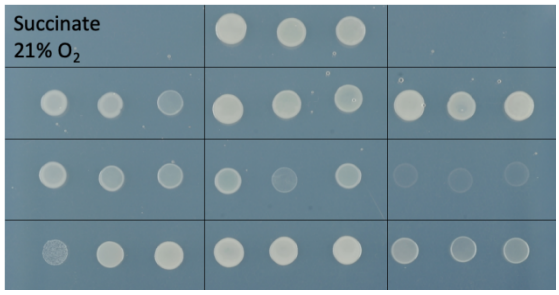
DME activity of 10 µg cell lysate observed from absorbance at 340 nm corrected to per µg. Cultures were grown on UMS succinate and harvested between OD_{600} 0.4 and 0.6. **ORS571** is wild type *A. caulinodans*, **ORS571 (Heat Killed)** is wild type *A. caulinodans* heat killed for 10 minutes at 90 °C, **DME-like Mutant** is OPS1951, a deletion of Azc_0119 , **DME mutant** is OPS2090, a deletion of Azc_3656, **Double Mutant – DME + DME-like** is OPS2094, a deletion of both Azc_0119 and Azc_3656. Points plotted are means with bars of SEM, N = 3.

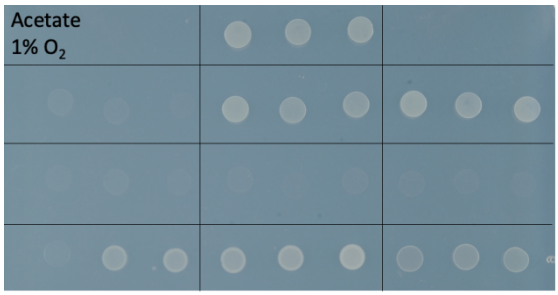


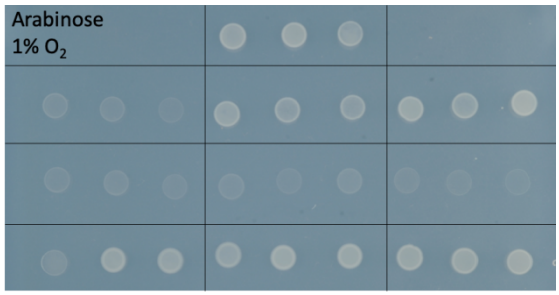


A

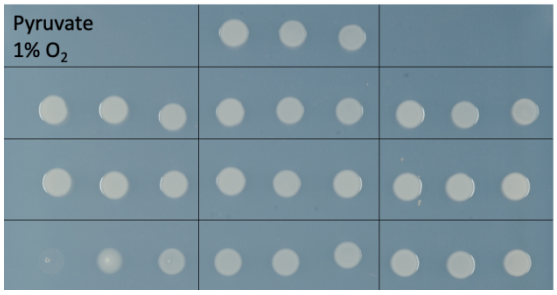


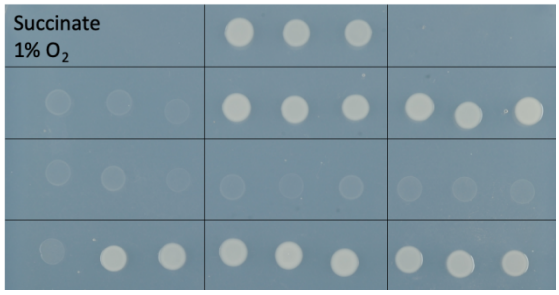






B





Sole carbon source + O_2 Concentration	C	
	ORS571 (Wild Type)	
OPS2090 (**Δ DME Enzyme**)	**OPS2091** (Δ PEPck)	**OPS2089** (Δ MSDH)
OPS2094 (Δ DME + Δ DME-like)	**OPS2122** (Δ DME, Δ MSDH)	**OPS2093** (Δ DME, Δ PEPck)
OPS2098 (Δ DME, DME complementation -	**OPS2100** (Δ DME, Δ PEPck, DME complementation)	**OPS2101** (Δ DME, Δ PEPck, PEPck complementation)

Figure 5.10 Growth of DME, DME-like, PEPck and MSDH mutants in aerobic and microaerobic conditions

A: Growth of cultures at 21% O_2 on UMS, 10 mM ammonia and the indicated sole carbon source
B: Growth of cultures at 1% O_2 on UMS, 10 mM ammonia and the indicated sole carbon source
C: Plate Layout. **ORS571**- wild type *A. caulinodans*, **OPS2090**- DME mutant, (ΔAzc_3656), **OPS2091**- PEPck mutant (ΔPEPck), **OPS2089** –Double MSDH mutant (ΔAzc_0575, ΔAzc_4119), **OPS2094** – Double mutant, DME and DME-like enzyme (ΔAzc_3656, ΔAzc_0119), **OPS2122** – Triple mutant, DME and MSDH (ΔAzc_3656, ΔAzc_0575, ΔAzc_4119), **OPS2093** – Double mutant, DME and PEPck (ΔAzc_3656, ΔPEPck), **OPS2098** – complementation of DME mutant OPS2090 with DME under native promoter (pOPS1348), **OPS2100-** complementation of double mutant OPS2094 with DME under native promoter (pOPS1348) **OPS2101**- complementation of double mutant OPS2094 with PEPck under native promoter (pOPS1421). N =3.

The free-living growth and symbiotic nitrogen fixing activity of the single DME mutant (OPS2090) and the double mutant of the DME and DME-like (OPS2094) were compared and no clear difference could be seen between them (Figure 5.10). Both mutants had reduced growth on acetate and succinate compared to wild type (ORS571) however, no growth penalty was apparent on pyruvate. Growth on arabinose was minimal for all strains at 21% O_2. When the O_2 tension was lowered to 1%, the growth on acetate and succinate was reduced further relative to wild type (ORS571) though again growth on pyruvate proceeded at wild type levels. For reasons unknown, growth on arabinose at 1% O_2 resulted in improved growth for the wild type strain, whereas there was a slight reduction in growth for both the DME and double DME-DME-like mutants. The loss of N_2 fixing ability in the single DME (Azc_3656) mutant (Figure 5.8) strongly indicates that the malonate shunt does not play a significant role in N_2 fixation, however it does not preclude the malonate shunts existence, particularly since the role of the DME-like (Azc_0119) in this shunt is theoretical. The continued growth on succinate of the double DME DME-like (Azc_36656 + Azc_0119) mutant indicated one of three possibilities: the malic enzyme–like was not involved in the malonate shunt; there was a third pathway capable of compensating for the lack of both DME and the malonate shunt and thus enabling growth; or finally, the malonate shunt does not exist at biologically relevant levels.

5.4 Mutagenesis based approaches to investigate the Malonate Shunt

In order to definitively determine whether a malonate shunt is important to N_2 fixation, we needed a definitive way to knock it out. Given that (Pauling et al. 2001) demonstrated the MSDH dependent production of CoA was linked in their temperature sensitive *pdhB* mutants, this was determined to be the key step to target. BLASTp searches were conducted using both Azc_0575 and Azc_4119 and besides these two known MSDH proteins, no proteins were found with identity greater than 35.4% were discovered. This gave us confidence in focusing our efforts on these two genes. An assay was developed based on the methods of (Hayaishi et al. 1961) and (Pauling et al. 2001). This was tested on wild type cell lysate and determined to be effective.

5.4.1 Testing the MSDH activity of the putative MSDH genes Azc_0575 and Azc_4119)

The two MSDH enzymes listed in the ORS571 genome (Azc_0575 and Azc_4119) were purified using the pOPINE expression system. Azc_0575 was amplified from absolute position 637,555-639,048 using primers oxp2883+2884 and Azc_4119 was amplified from absolute position 4,698,570-4,700,064 using primers oxp2885+2886. Both were InFusion cloned into pOPINE vectors to create vectors pOPS1145 (containing Azc_0575) and POPS1118 (containing Azc_4119), which were sequence mapped using POPIN F and POPIN R primers. These plasmids were transformed into chemically competent Bl21 (DE3) *E. coli* cells (New England Biolabs), grown up to 50 ml OD_{600} 0.4-0.6 and induced with 1mM IPTG. After overnight induction the cultures were pelleted, and crude cell lysate obtained by ribolysation. The tagged proteins were purified using His Hi-Trap Spin columns (Zymo) and quantified using a Qubit assay (Invitrogen). Once the purified protein was quantified, 10µg of enzyme was used to perform the MSDH assay. Samples of denatured enzyme (10 minutes at 90 °C) were used as negative controls for each enzyme (Figure 5.11). Each of the purified enzymes was determined to have MSDH activity.

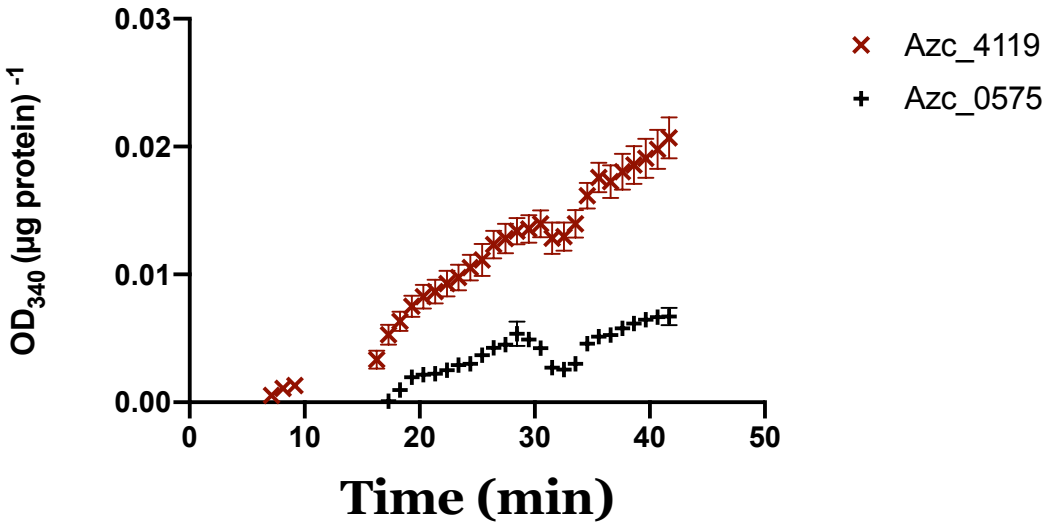


Figure 5.11 MSDH activity of purified potential MSDH enzymes

MSDH activity of 10 µg of purified MSDH enzymes Azc_0575 and Azc_4119 calculated from absorbance at 340 nm. Points plotted are mean with bars of SEM, N = 2

5.4.2 MSDH assay of MSDH single and double mutants

A marker-less single mutant of Azc_4119 (strain OPS2089) was isolated using the double integration of the plasmid pOPS1421 to delete from absolute position 4,698,570- 4,700,066. This strain was used to create a double mutant strain (OPS2092) of Azc_0575 and Azc_4119 using plasmid pOPS1352 to delete the entirety of Azc_0575 from absolute position 637,555-639,051. Both gene deletions were confirmed by PCR spanning the original site of each gene. For Azc_0575 this was oxp3947+3948 and for Azc_4119 this was oxp2335 +2336. Both mutagenesis plasmids pOPS1421 and pOPS1352 were derived from pK19msB and produced in-frame marker-less deletion mutants. The single and double mutants were then assayed for MSDH activity vs wild type, and the double mutant was demonstrated to have lost all MSDH activity (Table 5.2). From this, we were able to conclude we now had a method to effectively knock out the malonate shunt.

Table 5.2 MSDH activity from crude cell lysate of MSDH mutants

MSDH activity of 10 µg of cell lysate calculated from absorbance at 340 nm after 60 minutes incubation at 37 °C. **ORS571** is wild type *A. caulinodans*, **ORS571 (Heat Killed)** is wild type *A. caulinodans* heat killed for 10 minutes at 90 C, **Δ4119** is OPS2089 (AZC_4119 deleted), **MSDH Mutant** is OPS2092 (Azc_0575, Azc_4119 deletion). Values are mean +/- SEM, N=3.

Strain	Malonate semialdehyde dehydrogenated (nmol) min^{-1} $(\mu g\ protein)^{-1}$
ORS571	0.20 +/- 0.02
ORS571 (Heat Killed)	0.07 +/-0.00
Δ4119	0.1+/-0.02
MSDH Mutant (Δ4119, Δ0575)	0 +/-0.02

5.4.3 Characterisation of a MSDH double mutant

This double mutant was tested for symbiotic N_2 fixation and growth and found to suffer no penalty to either growth or rate of fixation (Figure 5.10, Figure 5.12). This allowed us to conclude that the malonate shunt is not an important pathway for dicarboxylate metabolism relative to the crucial DME reaction.

5.4.4 Generation and characterisation of a DME, MSDH triple mutant

By transforming the plasmid OPS2122 into the double MSDH mutant strain OPS2092 we were able to create a mutant defective in both DME activity and the malonate shunt. As expected for a DME mutant, this was found to be defective in N_2 fixation (Figure 5.12). Once again, this mutant was found to exhibit the same growth impairment seen in the single DME mutant strain (Figure 5.10). Critically however, this triple mutant was still capable of growth on dicarboxylates (Figure 5.10). This left one of two possibilities: a third pathway was compensating for the lack of the malonate shunt and DME; or the malonate shunt does not exist at biologically relevant levels and this unknown pathway is in fact the only other pathway.

The most likely candidate for this alternative pathway was hypothesised to be via combined activity of PEPck and pyruvate kinase as identified in RL3841 by (Mulley et al. 2010).

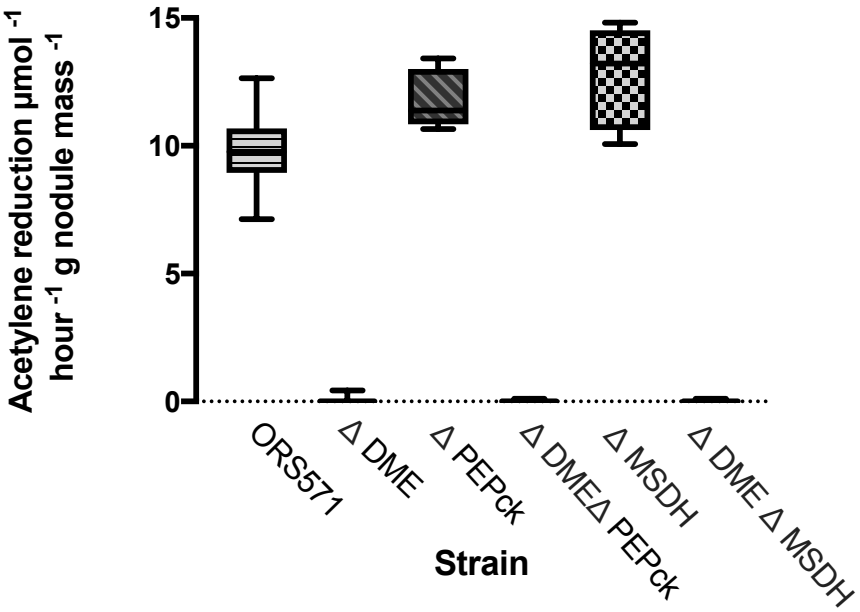


Figure 5.12 Symbiotic N_2 fixation rates of DME, PEPck, MSDH and combined mutants

Symbiotic fixation rates observed 28 days post-inoculation on *S. rostrata*. Acetylene reduction rates are corrected per gram of nodule mass. **ORS571** is the wild type strain of *A. caulinodans*, the **DME Mutant** is strain OPS2090 (Azc_3656 deletion), the **PEPck Mutant** is OPS2091 (PEPck deletion), the **Double Mutant DME + PEPck** is OPS2093 (Azc_3656, PEPck deletion), the **MSDH Mutant** is OPS2092 (Azc_0575, Azc_4119 deletion), the **Triple Mutant –MSDH + DME** is OPS2122 (Azc_0575, Azc_4119 , Azc_ 3656). Boxes are of the interquartile range, with the median indicated within and bars displaying the minimum and maximum values, N=4.

5.4.5 Generation and characterisation of PEP carboxykinase (PEPck) mutant

To mutate PEPck (absolute position 4,633,135-4,634,748) an in-frame marker-less single mutant was produced using plasmid pOPS1354 to produce strain OPS2091. The growth and N_2 fixing ability of this strain were assessed and no change was apparent for either between a PEPck mutant and wild type (Figure 5.10, Figure 5.12). The wild type growth of the PEPck mutant on succinate with no added sugars was unexpected given the role of PEPck in gluconeogenesis. This ability to grow without added sugars suggests another pathway is present within ORS571 to allow gluconeogenesis to occur, such as a pyruvate phosphate dikinase or a PEP synthase.

5.4.6 Generation and characterisation of a PEPck, DME double mutant

An in-frame marker-less double mutant of PEPck and DME was produced by introducing plasmid pOPS1421 into strain OPS2090. This double mutant (OPS2093) was found to be unaffected in its growth on pyruvate and capable of N_2 fixation when grown on pyruvate as a sole carbon source (Figure 5.10, 5.13 A). However, it proved incapable of growth on succinate as a sole carbon source and gave a non-fixing symbiotic phenotype (Figure 5.10, 5.12 From this we were able to conclude three things: both the PEPck and DME are actively used to enable dicarboxylate growth in ORS571; only DME is strictly necessary for N_2 fixation and under normal conditions the malonate shunt does not exist in ORS571 at biologically relevant levels. In addition to its lack of growth on dicarboxylates, the double DME, PEPck mutant was defective in nodulation, with only a single nodule being formed across four plants.

5.4.7 Acetylene reduction of mutants grown on pyruvate vs succinate

A free-living acetylene reduction experiment was performed on all single and double mutants grown on pyruvate to demonstrate that the previously observed loss of fixation in DME mutants on succinate was a direct result of their reduced capacity to produce pyruvate. Wild type fixation rates were observed in all strains grown on pyruvate as a sole carbon source. This indicates that, despite retaining growth on C_4 dicarboxylates via PEPck, the DME mutants were unable to generate sufficient pyruvate for fixation when grown on C_4 dicarboxylates.

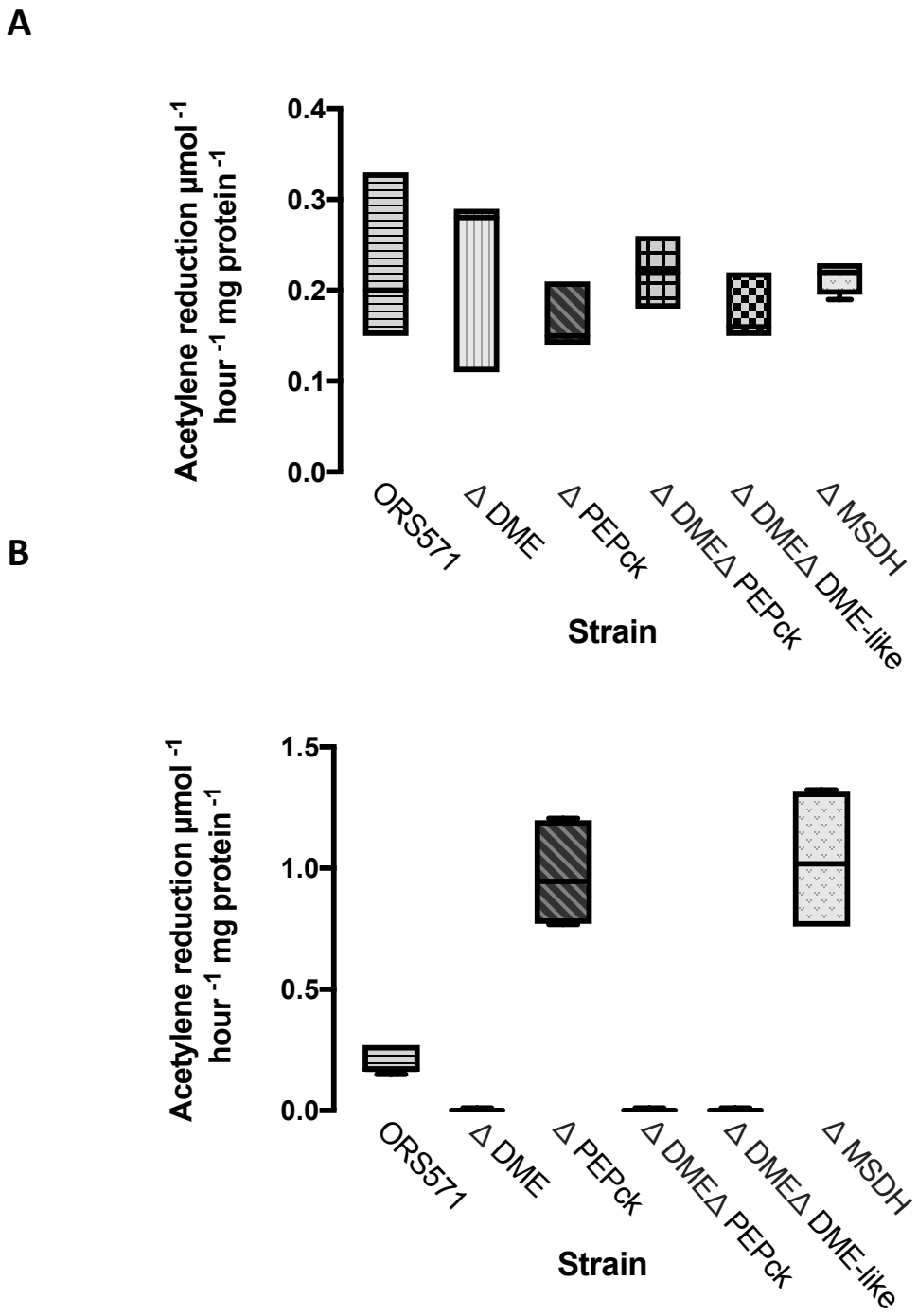


Figure 5.13 Free living N_2 fixation rates of DME, PEPck and combination mutants when grown on pyruvate or succinate as sole carbon source

A: Free living rates of acetylene reduction of cultures grown on UMS pyruvate. **ORS571** is the wild type strain of *A. caulinodans*, the **DME Mutant** is strain OPS2090 (Azc_3656 deletion), the **PEPck Mutant** is OPS2091 (PEPck deletion), the **Double Mutant DME + PEPck** is OPS2093 (Azc_3656 + PEPck deletion), **Double Mutant – DME + Malic enzyme– like** is OPS2094 (Azc_3656, Azc_0119 deletion) the **MSDH Mutant** is OPS2092 (Azc_0575, Azc_4119 deletion). Boxes are of the interquartile range, with the median indicated within and bars displaying the minimum and maximum values, N=4.

B: Free living rates of acetylene reduction of cultures grown on UMS succinate. **ORS571** is the wild type strain of *A. caulinodans*, the **DME Mutant** is strain OPS2090 (Azc_3656 deletion), the **PEPck Mutant** is OPS2091 (PEPck deletion), the **Double Mutant DME + PEPck** is OPS2093 (Azc_3656 + PEPck deletion), **Double Mutant – DME + Malic enzyme– like** is OPS2094 (Azc_3656, Azc_0119 deletion) the **MSDH Mutant** is OPS2092 (Azc_0575, Azc_4119 deletion. Fixation rates calculated from acetylene reduction assay per mg of bacterial protein in the sample. Boxes are of the interquartile range, with the median indicated within and bars displaying the minimum and maximum values, N=4.

5.4.8 Complementation of DME and PEPck mutants

To complement the DME mutants, the DME gene (Azc_3656) was amplified along with its native promoter (absolute position 4,199,002-4,201,526) using primers oxp3390 +3391 and cloned into the stable plasmid backbone pOGG093 producing the complementation plasmid pOPS1348. The PEPck complementation plasmid (pOPS1421) was made the same way, using primers oxp3767+3768 to amplify the *PEPck* gene and its native promoter (absolute position 4,635,096-4,633,113). The complementation of the single DME (OPS2090) with pOPS1348 (containing Azc_3656) was called OPS2098. The complementation of the double mutant in DME and PEPck (OPS2093) with pOPS1348 was called OPS2100, and when complemented with pOPS1421 (PEPck) it was called OPS2101. In the double mutant (OPS2093), growth on dicarboxylates was restored by complementation with either pOPS1348 or pOPS1421 (Figure 5.10).

5.5 Discussion

Although the aspartate aminotransferase assay indicated this activity is present in ORS571 the beta alanine aminotransferase assays demonstrated that there was no significant donation from beta alanine to oxoglutarate to form glutamate. Coupled with our initial bioinformatic analysis, this led us to question the existence of the shunt in the form originally proposed by (Pauling et al. 2001) (Figure 5.2). The lack of an obvious *panD* gene made the presence of aspartate decarboxylase activity in ORS571 an unlikely prospect, and this gene was key for the formation of beta-alanine from aspartate. However, the presence of multiple MSDH genes suggested that the MSDH activity might be active in ORS571, though neither were markedly upregulated in bacteroids (Tsukada et al. 2009). We therefore changed the focus of our investigation into the malonate shunt towards how malonate semialdehyde is

generated. Our attention was drawn to the DME-like enzyme Azc_0119 since it is both upregulated in bacteroids and structurally similar to a known oxaloacetate decarboxylase, giving the potential for the direct formation of malonate semialdehyde from oxaloacetate (Figure 5.4). Since attempts had been made to characterise the enzyme in the past (Zhang et al. 2012), it was known that the enzyme had at least a limited amount of affinity for malate as a substrate. The possibility was considered that this gene may in fact be a form of malolactic enzyme, due to the close relationship between oxaloacetate decarboxylase, malic enzyme and malolactic enzyme (Groisillier and Lonvaud-Funel 1999). When this enzyme was purified and tested for oxaloacetate decarboxylase, DME and malolactic enzyme activity no clear activity was seen based on the GCMS data (Figure 5.7). In addition to this, a single mutant of the DME-like enzyme Azc_ (0119) had no effect on symbiotic N_2 fixation (Figure 5.8). This suggests one of three possibilities; the DME-like enzyme (Azc_0119) is not involved in the malonate shunt, the malonate shunt is not essential for N_2 fixation, or the malonate shunt does not exist. To begin testing this a double mutant of DME (Azc_3656) and DME-like enzyme (Azc_0119) (OPS2094) was isolated and compared with single mutants of DME (OPS2090) and the DME-like enzyme (OPS1951). To confirm the mutants, DME assays were performed on the three mutants and ORS571 (Figure 5.9). Minimal, but measurable, DME activity was apparent in the DME mutant (OPS2090), but was absent in the double DME + DME-like enzyme mutant (OPS2094) when cell lysate was tested (Figure 5.9). From this we concluded that both DME (Azc_3656) and the DME-like (Azc_0119) produce DME activity, though the overwhelming majority of the observed DME activity came from Azc_3656.

No clear difference in growth was seen between the double mutant (OPS2094) and the DME mutant (OPS2090) (Figure 5.10), however, the mutation of DME alone was demonstrated to be sufficient to knock out N_2 fixation in both free-living (cells grown on succinate) and symbiotic conditions (Figure 5.8, 5.13 B). This suggested two things; either the malonate shunt has little effect on the growth of ORS571 on C_4 dicarboxylates or the DME-like enzyme (Azc_0119) is not a part of the malonate shunt. To conclusively demonstrate which the case was, a definitive method of knocking out the malonate shunt was required. This was achieved by deleting the two MSDH genes (Azc_0575 and Azc_4119) resulting in strain OPS2092, a strain with no measurable MSDH activity (Table 5.3). This mutant had no effect on N_2 fixation or growth in any condition (Figure 5.10, Figure 5.12, Figure 5.13) showing clearly that the malonate shunt it not essential for either growth on C_4 dicarboxylates or N_2

fixation. To make sure the growth was not being compensated for by another pathway, a triple mutant of DME (Azc_3656) and both MSDH enzymes (Azc_0575, Azc_4119) were isolated (OPS2122). This mutant was defective in N_2 fixation, as all DME mutants are in ORS571 (Figure 5.12, Figure 5.13 B). However, it maintained its ability to grow on C_4 dicarboxylates (Figure 5.10). The source of this residual growth was determined to be PEPck. A single mutant of PEPck was able to grow and fix N_2 at wild type levels, however, a double mutant of PEPck and DME was unable to grow on C_4 dicarboxylates and was ineffective at nodulation. Together, this led us to conclude that the malonate shunt is not biologically relevant for growth or N_2 fixation and growth on C_4 dicarboxylates is enabled solely by the combined DME and PEPck activity. The successful complementation of mutants demonstrated that the loss of fixation and reduced growth rates observed in the DME mutants is due solely to their loss of DME activity (Figure 5.10)

The wild type fixation rates achieved by DME mutants grown on pyruvate demonstrate that dicarboxylate metabolism is not necessary for fixation per-se (Figure 5.13 A), though this certainly is the case in symbiotic N_2 fixation. The fact that DME mutants are incapable of fixation despite their ability to grow on dicarboxylates suggests there may be additional reasons for the necessity of DME activity during symbiosis. This may represent an upper limit to the carbon flux which can be handled by the PEP shunt compared to DME, as would be suggested by the lack of obvious growth penalty in a PEPck single mutant relative to wild type at 21% or 1% O_2. If this is indeed the case, it is possible that the lack of fixation represents an insufficient quantity of energy production by the pathway to meet the need of N_2 fixation. Given that fixation has been demonstrated to be possible for both DME mutants (OPS2090) and double mutants of DME and PEPck (OPS2094) when grown on pyruvate, it is unlikely that reductant balance issues are to blame for the lack of fixation in the DME mutant strains. The fact that no growth issues are apparent in DME mutants when grown solely on pyruvate, even at 1% O_2, is further evidence that the lack of effective growth is due to insufficient carbon flux to pyruvate from succinate rather than excess reductant production. This flux theory is supported by the reduced growth on succinate of DME mutants relative to PEPck mutants. PEPck mutants exhibit near wild type growth under all tested O_2 concentration. Additionally, when the double DME, PEPck mutant was complemented with a multicopy plasmid form of PEPck it had increased growth on succinate relative to the DME mutant (Figure 5.10), indicating that the flux capacity of PEPck to pyruvate increased. Since during complementation PEPck gene is expressed on a

plasmid rather than chromosomally, the greater copy number enables a greater than wild type expression level. This would suggest that PEPck is the rate limiting step of pyruvate production in these DME mutant strains, and would support the hypothesis that the reduced growth and lack of N_2 fixation supported by the PEP shunt is due to a low flux capacity of the shunt at wild type levels. An intriguing experiment presents itself of attempting to increase the flux capacity of this pathway by introducing additional copies of PEPck, pyruvate kinase or both. In addition to exploring the growth effect of this PEP shunt overexpression, we could also test if this would enable N_2 fixation in the absence of DME, as has been observed in RL3841 (Mulley et al. 2010). As a companion experiment, additional copies of DME could be introduced, which we would expect to increase the maximum possible rate of N_2 fixation in ORS571. By combining the results of these experiments, we could ensure pyruvate production from C_4 dicarboxylates is not limited by enzyme availability, which would hopefully result in a measurable change in maximum N_2 fixation. We would expect there to be a maximum operating capacity for both DME and PEPck and it may be that this is more dependent on the availability of co-factors such as ATP and NAD^+ than on the total amount of PEPck/DME enzyme present. A comparison of the relative expression levels of PEPck in both ORS571 and RL3841 would be a good starting point to determine what level of PEPck expression is likely needed to achieve N_2 fixation.

Given the marked difference in growth evident for the DME mutants strain at 21% vs 1% O_2 (Figure 5.10) and the apparently binary choice of pyruvate production pathways from C_4 dicarboxylates, indicated by the lack of growth in the double PEPck-DME mutant, it would be intriguing to investigate the relative utilisation of the two pathways. From our growth experiments it appears that DME is the most critical pathway for growth under low O_2 conditions, but it appears that either can be used in atmospheric O_2 conditions. A metabolic flux map (such as that performed by Borah, 2017 , Figure 5.3) experiment comparing the DME mutant and the PEPck mutant to wild type under these different O_2 tensions could provide some intriguing results. Additionally, the knowledge that only a single pathway was being utilised would assist greatly in improving the resolution of the flux maps in this area, as the intersection of multiple metabolic paths renders a degree of uncertainty to the flow of metabolites connecting succinate and acetyl-CoA.

The questions posed by the DME-like enzyme Azc_0119 remain tantalising ones. While on the one hand we obtained evidence a small amount of DME activity (Figure 5.9), at a higher level than that previously reported (Zhang et al. 2012), approximately 10% of the Azc_3656 DME), it remains extremely unlikely that this is the sole or even primary function of this gene given its upregulation in bacteroids (Tsukada et al. 2009). On the other hand, mutation of this gene resulted in no measurable effect on either growth or symbiotic N_2 fixation. Therefore, although this gene function may be useful in bacteroids it is clearly not essential. Purification of greater quantities of the enzyme could allow for isothermal calorimetry or NMR spectroscopy to determine whether oxaloacetate or malate are the primary substrate of the enzyme and demonstrate conclusively what products can be formed.

A question raised by this work is what caused the MSDH dependent temperature sensitive mutants in the original work by (Pauling et al. 2001), capable of growth on valine as a sole C and N source, that led them to propose the malonate shunt? While we cannot explain the growth on dicarboxylates of the original PDH transposon mutant screened by (Pauling et al. 2001) as both DME and PEPck result in pyruvate formation, the importance of MSDH activity indicated by the temperature sensitive mutants is likely due to the secondary mutations required to allow growth on valine.

As part of the experimental set-up used to generate PDH temperature sensitive mutants derived from the original PDH mutant strain, valine was utilised as a sole carbon and nitrogen source for growth of ORS571. Since the valine degradation pathway can result in the production of methyl-malonate semialdehyde we hypothesised that this may result in increased MSDH activity. To test this, we attempted to grow ORS571 in UMS ammonia with valine as a sole carbon source, as well as on UMS with valine as sole carbon and nitrogen source. In both cases, no growth was apparent after incubation at 37 °C, 150 RPM for 48 hours. Given the severity of the growth penalty caused by a PDH mutant, coupled with this lack of growth on valine, such conditions are highly likely to generate secondary mutations. We propose that the importance of MSDH activity, and consequently the malonate shunt as a whole, may be an artefact of the extreme conditions used by (Pauling et al. 2001) to generate their temperature sensitive *pdhB* mutants. A background mutation allowing for either increased uptake of valine, more rapid valine breakdown, increased MSDH activity or a combination of the three could explain the need for MSDH activity to effectively grow on valine. Given the current understanding of the valine utilisation in ORS571 this must remain speculation at this stage.

As for the DME-like Azc_0119, it is possible that it operates in conjunction with the PEP shunt in a manner independent of the putative malonate shunt. It is possible that a double mutant of PEPck and Azc_0119 could result in a small but measurable reduction in fixation rates which is not apparent when mutated singly. If the effect size is small this could prove to be a difficult effect to observe in fixation using our current free-living fixation experimental set up, however a growth curve on succinate vs wild type may be more sensitive.

Overall, we can confidently state that under standard conditions the malonate shunt has no biological relevance for N_2 fixation or growth on C_4 dicarboxylates. Both the DME and PEPck pathways are capable of providing growth on C_4 dicarboxylates however, only DME is critical for N_2 fixation. The reason why PEPck alone cannot support effective N_2 fixation in either remains unknown, but it is likely due to a limited flux capacity of the shunt.

Chapter 6

Discussion

6.1 General overview

This work builds upon functional genomic datasets including transcriptomics, metabolomics and mutagenesis screens to resolve several seemingly anomalous results previously observed from the metabolism of ORS571. In doing so, it has allowed the effective revision of existing metabolic flux models, whilst laying the groundwork required for further studies of this agriculturally useful organism. The work on isocitrate dehydrogenase is self-explanatory and will not be expanded upon further in this section.

6.2 PHB as a redox store and regulator of *phaR*

In Chapter 4, PHB synthesis, which in ORS571 uniquely amongst rhizobia, was believed to be essential for N_2 fixation (Mandon et al. 1998) was demonstrated to be unnecessary to achieve wild type rates of N_2 fixation in either free-living or symbiotic conditions. Furthermore, the growth of a PHB defective suppressor mutant (OPS0921) recovered to near wild type levels on a wide range of carbon sources. The causal mutation that led to the restoration of both growth and N_2 fixation has been identified as a G to A SNP in what we understand to be the RBS of the regulatory *phaR* gene (absolute position 357,354). Translational fusions have demonstrated that this SNP has rendered the *phaR* translation levels of this suppressor mutant strain insensitive to the absence of PHB. By contrast, the wild type reporter indicated a six-fold increase in *phaR* translation in the absence of PHB. Together with the partial restoration of growth observed in OPS0921, this suggests that it is not PHBs role as a redox sink that makes it important in N_2 fixation, but its role in the regulation of *phaR*. The importance of PhaR would be a fruitful area for further investigation. Given the complex regulation of *phaR*, PhaR has been shown to act as both a repressor and an activator of gene expression as well as regulating the expression of the *phaR* gene itself in response to PHB levels, it would be useful to determine the exact nature of the global shift in gene expression produced by the misregulation of *phaR* translation in OPS0865 via an RNA-Seq study. Determination of which downstream regulators it acts upon to produce the severely deleterious growth phenotype observed in strain OPS0865 could begin with the RNA-Seq study, then expanded via targeted mutagenesis. It might be possible to produce similar phenotypes to the *phaR* misregulation through direct targeting of some of the downstream regulatory genes it controls, or through mutation of other global regulators of carbon metabolism such as *ptsN*. Instead of producing PHB defective mutants

through targeting synthesis genes, these global regulators could be targeted and the PHB production of such mutants measured.

The dramatic effect of *phaR* misregulation does not mean that the role of PHB as a redox sink doesn't matter however; the residual growth penalty in OPS0921 compared to wild type demonstrates that the PHB redox sink is highly beneficial for growth under low oxygen conditions. However, PHB as a redox sink is not essential for N_2 fixation. This should not be surprising, since there are several alterative redox sinks to PHB as both lipid synthesis, glycogen synthesis and even N_2 fixation itself can act as electron sinks. Measurement of the lipid and glycogen levels in both PHB defective strains should be undertaken, to see if there is an increase in the relative production of either, which would indicate an increase in the utilisation of these redox sink in the absence of PHB. If the levels remain the same as wild type, then it suggests that there may be an unidentified redox sink active in ORS571.

Previously, the PHB mutant strain OPS0865 was reported to have reduced rates of N_2 fixation, however these results were obtained by measurements of plants two weeks post inoculation (Borah 2017). By contrast the wild type fixation rates observed in Chapter 4 were measured four weeks post inoculation. Since suppressor mutations were obtained almost every time strain OPS0865 was applied to plants over the course of this thesis, it seems probable that also occurred in the original plant experiments of (Borah 2017). The reduced fixation rates she observed therefore likely represent a delayed nodulation effect as less actively fixing nodules were established after two weeks compared to the wild type. This apparent reduction in fixation rates vanished after four weeks as nodules containing bacteroids already containing suppressor mutations had a chance to effectively establish themselves. It is also possible that the nodules were a mix of OPS0865 and suppressor mutants, giving the appearance of an intermediate fixation rate, but that seems unlikely given the preponderance of suppressor mutants observed during every experiment within this thesis.

6.3 Separation of dicarboxylate metabolism and N_2 fixation, the malonate shunt

Our investigation into the malonate shunt in Chapter 5 allowed us to conclude that the metabolic requirements for nitrogen fixation and growth on dicarboxylates in ORS571 are not the same. Both mutants are completely defective in growth on dicarboxylates, and mutants reduced in their growth on dicarboxylates were nonetheless capable of wild type levels of N_2 fixation when grown on pyruvate as their sole carbon source. DME mutants (OPS2090, OPS2093, OPS2094, OPS2122) were demonstrated consistently to be essential for N_2 fixation in both symbiotic and free-living conditions when grown on succinate. In addition to this, growth was reduced relative to ORS571 on all tested carbon sources except pyruvate. This residual growth was eliminated altogether in a double mutant of DME and PEPck (OPS2093). However, wild type levels of growth and N_2 fixation could be achieved from both DME and DME-PEPck double mutants grown on pyruvate. The lack of N_2 fixation observed in DME mutants despite their residual growth on dicarboxylates is indicative of a lower limit to the flux capacity of the PEPck pathway used to generate pyruvate in these strains. It also suggests that the lack of N_2 fixation from the DME mutant strains is a result of an insufficient level of pyruvate generation in this strain. While this would prevent fixation in symbiotic conditions, where dicarboxylates are the primary form of carbon provided to the bacteria, it does indicate that this is not essential. If a synthetic symbiosis was established, alternative carbon sources could potentially be utilised to support effective N_2 fixation.

Although the PEPck pathway clearly operates to generate pyruvate from succinate in ORS571, it is equally clear that it is not the primary pathway utilized by this strain, as demonstrated by the lack of observable phenotype for a single PEPck mutant (OPS2091). This lack of phenotype, including the ability to grow on gluconeogenic substrates such as acetate and succinate, indicates that there are alternative sources of PEP for gluconeogenesis in addition to PEPck. A likely candidate for this would be pyruvate phosphate dikinase which can produce PEP from pyruvate. A BLASTp search against the known Ppdk from RL3841 identified Azc_0410 as the most likely primary *ppdk* gene with a 67 % protein sequence identity, however a second pyruvate phosphate dikinase is also present (Azc_2763) with a 39 % sequence identity. An INSeq study in the *PEPck* mutant background could both provide an overview of gluconeogenesis genes in ORS571, as well

as highlighting any potentially novel genes involved in the continued gluconeogenic growth of this mutant and indicating which of the labelled gluconeogenic genes are truly functional e.g. Azc_0410 vs Azc_2763. The generation and characterization of *ppdk* single mutants should be a matter of high priority, followed by that of a double *PEPck-ppdk* mutant. Given what is known about gluconeogenic metabolism in other rhizobia we would predict such a double mutant to be unable to grow on gluconeogenic substrates. However, if we discover this is not the case, the results of the INSeq would inform our future targets of mutagenesis, though a triple mutant of *PEPck* and both *ppdk* would be a likely starting candidate.

The lack of nodulation observed in the double DME-PEPck mutant is a surprisingly strong phenotype, as DME mutants in other rhizobial species are still capable of nodulation but are incapable of N_2 fixation. This extreme phenotype is likely due to the nature of ORS571 metabolism. As it is incapable of utilizing sugars it is entirely dependent on organic acids for growth resulting in a strain that is unable to survive on the root exudates provided by *S. rostrata.*

The DME-like enzyme Azc_0119 remains a mystery. While it is clearly up-regulated in bacteroids, and related closely enough to DME (Azc_3656) to retain limited activity, mutation of this gene (OPS1951) has no measurable effect on growth or N_2 fixation rates in free-living or symbiotic conditions. The enzymatic function also remains unaccounted for, as no measurable interaction could be detected with oxaloacetate despite the high sequence identity of the protein to an oxaloacetate decarboxylase. Furthermore, the measurable DME activity of this enzyme was extremely limited compared to the primary DME enzyme (Azc_3656) leaving no clear reason why Azc_0119 should be upregulated so highly in bacteroids. In order to clarify the purpose of this enzyme, there are a number of techniques that could be utilised, though a priority should be given to determining ligand binding through either isothermal calorimetry or NMR spectroscopy. A ligand binding study would have the benefit of identifying the primary substrate (malate vs oxaloacetate) and key co-factors required by the enzyme such $NAD(P)^+$/NAD(P)H, or ATP. An alternative approach would be to use the ligand binding study to determine only the key substrate (malate vs oxaloacetate) and to determine the additional co-factors experimentally. Determining the carbon substrate would halve the number of variant enzyme reactions to study, as only malate or oxaloacetate could be used as appropriate. If the required enzyme co-factors are particularly difficult to predict, then enzyme function could still be determined by adding

small quantities of cell extract to the purified enzyme reaction mix. Once the enzyme binding partners are clearly established, the optimal reaction conditions of Azc_0119 can then be systematically investigated e.g. temperature and pH.

Although the path to take for such an investigation into the enzyme is clear, given the lack of observed symbiotic phenotype for such a mutant unless further evidence is forthcoming identifying the function of this enzyme should not be a priority in future work.

The malonate shunt itself has proven to be something of a red herring. The pathway as suggested by (Pauling et al. 2001) has been demonstrated to be incorrect, both due to the apparent lack of an aspartate decarboxylase in ORS571 and the inability of beta-alanine aminotransferase to donate to oxoglutarate. Although further proof of the lack of aspartate decarboxylase could have been obtained by a direct enzyme assay, for example measuring the release of isotopically labeled ^{14}C from ^{14}C labelled aspartate, this was not attempted due to the lack of appropriate beta-alanine aminotransferase activity. Instead focus was shifted to the MSDH activity upon which the shunt relies with their activity demonstrated by (Pauling et al. 2001) to be required for acetyl-CoA generation in their temperature sensitive mutants. Both MSDH enzymes were identified and a double deletion mutant generated (OPS2092) with no measurable MSDH activity. This strain had no measurable effect on growth or N_2 fixation under any condition, and a triple mutant with DME (OPS2122) still retained growth on dicarboxylates comparable to a single DME mutant. This, together with the lack of growth apparent in the DME-PEPck double mutant leads us to conclude that the malonate shunt is not active at biologically relevant levels in ORS571and may in fact represent an experimental artefact in the original paper (Pauling et al. 2001).

6.4 Future directions

Together, this work has allowed the metabolic flux maps developed by (Borah 2017) to be refined, both through the discovery that PHB is non-essential, just as in other rhizobia, and that the primary flux during dicarboxylates metabolism is via DME with a branch through PEPck. Additionally, the biological irrelevance of the malonate shunt has been demonstrated allowing it to be removed entirely from the model, providing far greater resolution of this branch of metabolism. Now that a robust metabolic model has been established further comparative metabolic flux studies can be performed to accurately assess the preferred redox stores of ORS571 making use of the OPS0921 suppressor mutant. This strain would be preferable for such work to the original *phaC* mutant OPS0865 as it does not suffer from the same PhaR misregulation induced growth penalty. It could therefore explore the sole contribution of PHB itself to the metabolism of ORS571.

This flux model can be improved further through the use of INSeq. The growth essential genes of ORS571 can be identified in free living and symbiotically N_2 fixing wild type strain and the effects of such mutants on the metabolic flux of ORS571 can then be modelled providing insight into the potential underlying causes of observed mutant phenotypes. This would also serve to identify new key enzymes, as INSeq could highlight genes that have a major phenotypic impact on growth which were not present in the metabolic flux model. A combination of protein purification, enzymatic testing and single gene mutagenesis studies could then be used to characterise the effects of these newly identified genes. Potentially, new flux maps could be generated using the mutagenized strains in order to suggest the reactions uncharacterised enzymes, such as Azc_0119, or hypothetical proteins are involved in.

Another obvious candidate for future study using these tools is the true nature of the PDH independent pathway as reported by (Pauling et al. 2001). In this study it was reported that the pathway allows for growth retention and limited N_2 fixation in an ORS571 PDH mutant. Generation of an INSeq library in a PDH mutant background and comparison of this library to wild type should highlight genes that become essential in the absence of PDH. Within this subset of genes, unless there is further metabolic redundancy, we are highly likely to find members of our PDH independent acetyl-CoA generating pathway. It should of course also identify genes needed in the putative malonate shunt. This approach would be

strengthened by combining it with our refined metabolic flux model to predict the effect of gene knockouts in any of the highlighted candidate genes. Together these approaches could provide us with clear predictions of pathways to test in a systematic manner.

In addition to these more targeted investigations, the mapping tools refined here combined with the dual symbiotic and free-living N_2 fixing nature of ORS571 allows the possibility of investigating the effect of individual plant metabolites on the metabolism of ORS571 during N_2 fixation. By adding either purified metabolites, or whole nodule extract we can determine both if it is possible to induce the metabolism of ORS571 to enter a more 'symbiosis-like' state and may determine the combination of metabolites responsible for this. Amino acid secretion and ammonia uptake would be obvious targets to explore in such studies, given this critical difference between symbiotic and free-living N_2 fixation. RNA-Seq would also be a powerful tool to combine with such studies, to dissect the gene cascades involved in such state shifts.

To conclude, ORS571 remains a useful model organism with clear agricultural relevance, and this investigation has refined the existing tools developed by (Borah 2017) to further explore the possible insights into the metabolic shifts of N_2 fixation made possible by this N_2 fixing organism in its different life cycles.

Chapter 7

Bibliography

Allen, D. K., and Young, J. D. 2013. Carbon and nitrogen provisions alter the metabolic flux in developing soybean embryos. Plant Physiol. 161:1458–1475

Allen, G. C., Grimm, D. T., and Elkan, G. H. 1991. Oxygen uptake and hydrogen-stimulated nitrogenase activity from Azorhizobium caulinodans ORS571 grown in a succinate-limited chemostat. Appl. Environ. Microbiol. 57:3220–3225

Ampe, F., Kiss, E., Sabourdy, F., and Batut, J. 2003. Transcriptome analysis of Sinorhizobium meliloti during symbiosis. Genome Biol. 4

Aneja, P., Zachertowska, A., and Charles, T. C. 2005. Comparison of the symbiotic and competition phenotypes of Sinorhizobium meliloti PHB synthesis and degradation pathway mutants. Can. J. Microbiol. 51:599–604

Barquist, L., Boinett, C. J., and Cain, A. K. 2013. Approaches to querying bacterial genomes with transposon-insertion sequencing. RNA Biol. 10:1161–1169

Bationo, A., and Mokwunye, A. U. 1991. Alleviating soil fertility constraints to increased crop production in west africa - the experience in the Sahel. Fertil. Res. 29:95–115

Bationo, A., and Ntare, B. R. 2000. Rotation and nitrogen fertilizer effects on pearl millet, cowpea and groundnut yield and soil chemical properties in a sandy soil in the semi-arid tropics, West Africa. J. Agric. Sci. 134:277–284

Batut, J., and Boistard, P. 1994. Oxygen control in Rhizobium. Antonie Van Leeuwenhoek. 66:129–150

Beck, S., Marlow, V. L., Woodall, K., Doerrler, W. T., James, E. K., and Ferguson, G. P. 2008. The Sinorhizobium meliloti MsbA2 protein is essential for the legume symbiosis. Microbiology. 154:1258–1270

Bergersen, F. J., Peoples, M. B., and Turner, G. L. 1991. A role for poly-β-hydroxybutyrate in bacteroids of soybean root nodules. Proc. R. Soc. B Biol. Sci. 245:59–64

Bergersen, F. J., Turner, G. L., and Bogusz, D. 1986. Effects of O_2 concentrations and various haemoglobins on respiration and nitrogenase activity of bacteroids from stem and root nodules of Sesbania rostrata and of the same bacteria from continuous cultures. J. Gen. Microbiol. 132:3325–3336

Berrow, N. S., Alderton, D., Sainsbury, S., Nettleship, J., Assenberg, R., Rahman, N., Stuart, D. I., and Owens, R. J. 2007. A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. Nucleic Acids Res. 35

Beste, D. J. V, Bonde, B., Hawkins, N., Ward, J. L., Beale, M. H., Noack, S., Nöh, K., Kruger, N. J., Ratcliffe, R. G., and McFadden, J. 2011. ^{13}C metabolic flux analysis identifies an unusual route for pyruvate dissimilation in mycobacteria which requires isocitrate lyase and carbon dioxide fixation. PLoS Pathog. 7

Borah, K. 2017. Metabolic flux analysis of the nitrogen-fixing bacterium Azorhizobium caulinodans. D.Phil thesis, University of Oxford

Brear, E. M., Day, D. A., and Smith, P. M. C. 2013. Iron: an essential micronutrient for the legume-rhizobium symbiosis. Front. Plant Sci. 4

Brown, C. M., and Dilworth, M. J. 1975. Ammonia assimilation by Rhizobium cultures and bacteroids. J. Gen. Microbiol. 86:39–48

Cai, S., Cai, L., Zhao, D., Liu, G., Han, J., Zhou, J., and Xiang, H. 2015. A Novel DNA-binding protein, PhaR, plays a central role in the regulation of polyhydroxyalkanoate accumulation and granule formation in the haloarchaeon Haloferax mediterranei. Appl. Environ. Microbiol. 81:373–385

Capoen, W., Goormachtig, S., and Holsters, M. 2010. Water-tolerant legume nodulation. J. Exp. Bot. 61:1251–1255

Caspritz, G., and Radler, F. 1983. Malolactic enzyme of Lactobacillus plantarum: Purification, properties, and distribution among bacteria. J. Biol. Chem. 258:4907–4910

Cevallos, M. A., Encarnacion, S., Leija, A., Mora, Y., and Mora, J. 1996. Genetic and physiological characterization of a Rhizobium etli mutant strain unable to synthesize poly-beta-hydroxybutyrate. J. Bacteriol. 178:1646–1654

Chang, C., Damiani, I., Puppo, A., and Frendo, P. 2009. Redox changes during the legume-rhizobium symbiosis. Mol. Plant. 2:370–377

Clarke, V. C., Loughlin, P. C., Day, D. A., and Smith, P. M. C. 2014. Transport processes of the legume symbiosome membrane. Front. Plant Sci. 5

D'Haeze, W., De Rycke, R., Mathis, R., Goormachtig, S., Pagnotta, S., Verplancke, C.,

Capoen, W., and Holsters, M. 2003. Reactive oxygen species and ethylene play a positive role in lateral root base nodulation of a semiaquatic legume. Proc. Natl. Acad. Sci. U. S. A. 100:11789–11794

Das, A., and Lodha, M. L. 1998. Differential expression of uptake hydrogenase activity in free living cells and nodule bacteroids of Azorhizobium caulinodans. J. Plant Biochem. Biotechnol. 7:51–52

DeJesus, M. A., and Ioerger, T. R. 2013. A Hidden Markov Model for identifying essential and growth-defect regions in bacterial genomes from transposon insertion sequencing data. BMC Bioinformatics. 14

DeJesus, M. A., Nambi, S., Smith, C. M., Baker, R. E., Sassetti, C. M., and Ioerger, T. R. 2017. Statistical analysis of genetic interactions in Tn-Seq data. Nucleic Acids Res. 45

DeJesus, M. A., Zhang, Y. J., Sassetti, C. M., Rubin, E. J., Sacchettini, J. C., and Ioerger, T. R. 2013. Bayesian analysis of gene essentiality based on sequencing of transposon insertion libraries. Bioinformatics. 29:695–703

Ditta, G., Stanfield, S., Corbin, D., and Helinski, D. R. 1980. Broad host range DNA cloning system for gram-negative bacteria: construction of a gene bank of Rhizobium meliloti. Proc. Natl. Acad. Sci. U. S. A. 77:7347–7351

Dixon, R., and Kahn, D. 2004. Genetic regulation of biological nitrogen fixation. Nat. Rev. Microbiol. 2:621–631

Djordjevic, M. A. 2004. Sinorhizobium meliloti metabolism in the root nodule: A proteomic perspective. Proteomics. 4:1859–1872

Djordjevic, M. A., Chen, H. C., Natera, S., Van Noorden, G., Menzel, C., Taylor, S., Renard, C., Geiger, O., Galibert, F., Finan, T. M., Long, S. R., Puhler, A., Abola, P., Ampe, F., Barloy-Hubler, F., Barnett, M. J., Becker, A., Boistard, P., Bothe, G., Boutry, M., Bowser, L., Buhrmester, J., Cadieu, E., Capela, D., Chain, P., Cowie, A., Davis, R. W., Dreano, S., Federspiel, N. A., Fisher, R. F., Gloux, S., Godrie, T., Goffeau, A., Golding, B., Gouzy, G., Gurjal, M., Hernandez-Lucas, I., Hong, A., Huizar, L., Hyman, R. W., Jones, T., Kahn, D., Kahn, M. L., Kalman, S., Keating, D. H., Kiss, E., Komp, C., Lelaure, V., Masuy, D., Palm, C., Peck, M. C., Pohl, T. M., Portetelle, D., Purnelle, B., Ramsperger, U., Surzycki, R., Thébault, P., Vandenbol, M., Vorhölter, F. J., Weidner, S., Wells, D. H., Wong, K., Yeh, K. C., Batut, J., and Weiller, G. F. 2003. A global analysis of protein expression profiles in Sinorhizobium meliloti: Discovery of new genes for nodule occupancy and stress adaptation. Mol.

Plant-Microbe Interact. 16:508–524

Downie, J. A. 2010. The roles of extracellular proteins, polysaccharides and signals in the interactions of rhizobia with legume roots. FEMS Microbiol. Rev. 34:150–170

Downie, J. A., Hombrecher, G., Ma, Q. S., Knight, C. D., Wells, B., and Johnston, A. W. B. 1983. Cloned nodulation genes of Rhizobium leguminosarum determine host-range specificity. MGG Mol. Gen. Genet. 190:359–365

Dreyfus, B., Garcia, J. L., and Gillis, M. 1988. Characterization of Azorhizobium caulinodans gen a stem-nodulating nitrogen-fixing bacterium isolated from Sesbania rostrata. Int. J. Syst. Bacteriol. 38:89–98

Dreyfus, B. L., Elmerich, C., and Dommergues, Y. R. 1983. Free-living Rhizobium strain able to grow on N2 as the sole nitrogen source. Appl Env. Microbiol. 45:711–713

Dreyfus, B., Rinaudo, G., and Dommergues, Y. 1985. Observations on the use of Sesbania rostrata as green manure in paddy fields. MIRCEN J. Appl. Microbiol. Biotechnol. 1:111–122

Driscoll, B. T., and Finan, T. M. 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

Elmerich, C., Dreyfus, B. L., Reysset, G., and Aubert, J. P. 1982. Genetic analysis of nitrogen fixation in a tropical fast-growing Rhizobium. EMBO J. 1:499–503

Encarnación, S., Del Carmen Vargas, M., Dunn, M. F., Dávalos, A., Mendoza, G., Mora, Y., and Mora, J. 2002. Ania regulates reserve polymer accumulation and global protein expression in Rhizobium etli. J. Bacteriol. 184:2287–2295

Engelke, T., Jagadish, M. N., and Puhler, A. 1987. Biochemical and genetical analysis of Rhizobium meliloti mutants defective in C_4-dicarboxylate transport. Microbiology. 133:3019–3029

Erisman, J. W., Bleeker, A., Galloway, J., and Sutton, M. S. 2007. Reduced nitrogen in ecology and the environment. Environ. Pollut. 150:140–149

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

Finan, T. M. 1991. Complex symbiotic phenotypes result from gluconeogenic mutations in Rhizobium meliloti. Mol. Plant-Microbe Interact. 4:386

Finan, T. M., Oresnik, I., and Bottacin, A. 1988. Mutants of Rhizobium meliloti defective in succinate metabolism. J. Bacteriol. 170:3396–3403

Finan, T. M., Wood, J. M., and Jordan, D. C. 1981. Succinate transport in Rhizobium leguminosarum. J. Bacteriol. 148:193–202

Finan, T. M., Wood, J. M., and Jordan, D. C. 1983. Symbiotic properties of C_4-dicarboxylic acid transport mutants of Rhizobium leguminosarum. J. Bacteriol. 154:1403–1413

Fischer, H. M. 1994. Genetic regulation of nitrogen fixation in Rhizobia. Microbiol. Rev. 58:352–386

Gage, D. J. 2004. Infection and invasion of roots by symbiotic, nitrogen fixing Rhizobia during nodulation of temperate legumes. Microbiol. Mol. Biol. Rev. 68:280–300

Gawronski, J. D., Wong, S. M. S., Giannoukos, G., Ward, D. V., and Akerley, B. J. 2009. Tracking insertion mutants within libraries by deep sequencing and a genome-wide screen for Haemophilus genes required in the lung. Proc. Natl. Acad. Sci. U. S. A. 106:16422–16427

Gebhardt, C., Turner, G. L., and Gibson, A. H. 1984. Nitrogen-fixing growth in continuous culture of a strain of Rhizobium sp. isolated from stem nodules on Sesbania rostrata. J. Gen. Microbiol. 130:843–848

Geddes, B. A., Mendoza-Suárez, M. A., and Poole, P. S. 2019. A Bacterial Expression Vector Archive (BEVA) for flexible modular assembly of golden gate-compatible vectors. Front. Microbiol. 10

Geddes, B. A., Ryu, M. H., Mus, F., Costas, A. G., Peters, J. W., Voigt, C. A., and Poole, P. 2015. Use of plant colonizing bacteria as chassis for transfer of N_2 fixation to cereals. Curr. Opin. Biotechnol. 32:216–222

Gerson, T., Patel, J. J., and Wong, M. N. 1978. The effects of age, darkness and nitrate on poly-β-hydroxybutyrate levels and nitrogen-fixing ability of Rhizobium in Lupinus angustifolius. Physiol. Plant. 42:420–424

Gough, C., Galera, C., Vasse, J., Webster, G., Cocking, E. C., and Dénarié, J. 1997. Specific flavonoids promote intercellular root colonization of Arabidopsis thaliana by Azorhizobium caulinodans ORS571. Mol. Plant-Microbe Interact. 10:560–570

Graves, P. R., and Haystead, T. A. J. 2002. Molecular biologist's guide to proteomics. Microbiol. Mol. Biol. Rev. 66:39–63

Gray, A. N., Koo, B. M., Shiver, A. L., Peters, J. M., Osadnik, H., and Gross, C. A. 2015. High-throughput bacterial functional genomics in the sequencing era. Curr. Opin. Microbiol. 27:86–95

Green, L. S., and Emerich, D. W. 1997. Bradyrhizobium japonicum does not require α-

ketoglutarate dehydrogenase for growth on succinate or malate. J. Bacteriol. 179:194–201

Groisillier, A., and Lonvaud-Funel, A. 1999. Comparison of partial malolactic enzyme gene sequences for phylogenetic analysis of some lactic acid bacteria species and relationships with the malic enzyme. Int. J. Syst. Bacteriol. 49:1417–1428

Haberer, G., Mayer, K. F., and Spannagl, M. 2016. The big five of the monocot genomes. Curr Opin Plant Biol. 30:33–40

Hayaishi, O., Nishizuka, Y., Tatibana, M., Takeshita, M., and Kuno, S. 1961. Enzymatic studies on the metabolism of β-alanine. J. Biol. Chem. 236:781–790

Hiller, K., and Metallo, C. M. 2013. Profiling metabolic networks to study cancer metabolism. Curr. Opin. Biotechnol. 24:60–68

Hosie, A. H. F., Allaway, D., Galloway, C. S., Dunsby, H. A., and Poole, P. S. 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. J. Bacteriol. 184:4071–4080

Hunt, S., and Layzell, D. B. 1993. Gas exchange of legume nodules and the regulation of nitrogenase activity. Annu. Rev. Plant Physiol. Plant Mol. Biol. 44:483–511

Hustede, E., Steinbüchel, A., and Schlegel, H. G. 1993. Relationship between the photoproduction of hydrogen and the accumulation of PHB in non-sulphur purple bacteria. Appl. Microbiol. Biotechnol. 39:87–93

Igarashi, R. Y., and Seefeldt, L. C. 2003. Nitrogen fixation: the mechanism of the Mo-dependent nitrogenase. Crit. Rev. Biochem. Mol. Biol. 38:351–384

Iki, T., Aono, T., and Oyaizu, H. 2007. Evidence for functional differentiation of duplicated nifH genes in Azorhizobium caulinodans. FEMS Microbiol. Lett. 274:173–179

IPCC. 2013. Anthropogenic and natural radiative forcing. Pages 659–740 in: Climate Change 2013 the Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change,

Jackson, F. A., and Dawes, E. A. 1976. Regulation of the tricarboxylic acid cycle and poly β hydroxybutyrate metabolism in Azotobacter beijerinckii grown under nitrogen or oxygen limitation. J. Gen. Microbiol. 97:303–312

Jain, V., and Gupta, K. 2003. The flavonoid naringenin enhances intercellular colonization of rice roots by Azorhizobium caulinodans. Biol. Fertil. Soils. 38:119–123

Jang, S. H., Jang, H. A., Lee, J., Kim, J. U., Lee, S. A., Park, K. E., Kim, B. H., Jo, Y. H.,

and Lee, B. L. 2017. PhaR, a Negative Regulator of PhaP, modulates the colonization of a Burkholderia gut symbiont in the midgut of the host insect, Riptortus pedestris. Appl Env. Microbiol. 83

Jiang, L., Shestov, A. A., Swain, P., Yang, C., Parker, S. J., Wang, Q. A., Terada, L. S., Adams, N. D., McCabe, M. T., Pietrak, B., Schmidt, S., Metallo, C. M., Dranka, B. P., Schwartz, B., and Deberardinis, R. J. 2016. Reductive carboxylation supports redox homeostasis during anchorage-independent growth. Nature. 532:255–258

Johnston, A. W. B., and Behringer, J. E. 1975. Identification of the rhizobium strains in pea root nodules using genetic markers. J. Gen. Microbiol. 87:343–350

Kadouri, D., Jurkevitch, E., and Okon, Y. 2003. Poly beta-hydroxybutyrate depolymerase (PhaZ) in Azospirillum brasilense and characterization of a phaZ mutant. Arch Microbiol. 180:309–318

Kaminski, P. A., Mandon, K., Arigoni, F., Desnoues, N., and Elmerich, C. 1991. Regulation of nitrogen fixation in Azorhizobium caulinodans: identification of a fixK-like gene, a positive regulator of nifA. Mol. Microbiol. 5:1983–1991

Karr, D. B., Waters, J. K., Suzuki, F., and Emerich, D. W. 1984. Enzymes of the poly-hydroxybutyrate and citric acid cycles of Rhizobium japonicum bacteroids. Plant Physiol. 75:1158–1162

Karunakaran, R., Ramachandran, V. K., Seaman, J. C., East, A. K., Mouhsine, B., Mauchline, T. H., Prell, J., Skeffington, A., and Poole, P. S. 2009. Transcriptomic analysis of Rhizobium leguminosarum biovar viciae in symbiosis with host plants Pisum sativum and Vicia cracca. J Bacteriol. 191:4002–4014

Kim, Y., Lee, K., Kim, H., Cho, S., and Lee, J. 2017. Inhibition of poly 3-hydroxybutyrate (PHB) synthesis by phaR deletion in Methylobacterium extorquens AM1. Korean Chem. Eng. Res. 55:363–368

Koboldt, D. C., Larson, D. E., and Wilson, R. K. 2013. Using varscan 2 for germline variant calling and somatic mutation detection. Curr. Protoc. Bioinforma.

Kruger, N. J., and Ratcliffe, R. G. 2009. Insights into plant metabolic networks from steady-state metabolic flux analysis. Biochimie. 91:697–702

Kuchta, K., Chi, L., Fuchs, H., Pötter, M., and Steinbüchel, A. 2007. Studies on the influence of phasins on accumulation and degradation of PHB and nanostructure of PHB granules in Raistonia eutropha H16. Biomacromolecules. 8:657–662

Laguerre, G., Geniaux, E., Mazurier, S. I., Casartelli, R. R., and Amarger, N. 1993. Conformity and diversity among field isolated of Rhizobium leguminosarum bv.

viciae, bv. trifolii, and bv. phaseoli revealed by DNA hybridization using chromosome and plasmid probes. Can. J. Microbiol. 39:412–419

Laguerre, G., Mazurier, S. I., and Amarger, N. 1992. Plasmid profiles and restriction fragment length polymorphism of Rhizobium leguminosarum bv. viciae in field populations. FEMS Microbiol. Lett. 101:17–26

Langmead, B., and Salzberg, S. L. 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods. 9:357–359

Lardi, M., and Pessi, G. 2018. Lardi, M., and Pessi, G. 2018. Functional genomics approaches to studying symbioses between legumes and nitrogen-fixing rhizobia. High-Throughput.Functional genomics approaches to studying symbioses between legumes and nitrogen-fixing rhizobia. High-Throughput. 7

Lee, K. B., De Backer, P., Aono, T., Liu, C. T., Suzuki, S., Suzuki, T., Kaneko, T., Yamada, M., Tabata, S., Kupfer, D. M., Najar, F. Z., Wiley, G. B., Roe, B., Binnewies, T. T., Ussery, D. W., D'Haeze, W., Herder, J. D., Gevers, D., Vereecke, D., Holsters, M., and Oyaizu, H. 2008. The genome of the versatile nitrogen fixer Azorhizobium caulinodans ORS571. BMC Genomics. 9:271

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 25:2078–2079

Lodwig, E. M., Leonard, M., Marroqui, S., Wheeler, T. R., Findlay, K., Downie, J. A., and Poole, P. S. 2005. Role of polyhydroxybutyrate and glycogen as carbon storage compounds in pea and bean bacteroids. Mol. Plant-Microbe Interact. 18:67–74

Lodwig, E., and Poole, P. 2003. Metabolism of rhizobium bacteroids. CRC. Crit. Rev. Plant Sci. 22:37–78

Lonvaud-Funel, A., and Strasser de Saad, A. M. 1982. Purification and properties of a malolactic enzyme from a strain of Leuconostoc mesenteroides isolated from grapes. Appl. Environ. Microbiol. 43:357–361

Ludwig, R. A. 2004. Microaerophilic bacteria transduce energy via oxidative metabolic gearing. Res Microbiol. 155:61–70

Maehara, A., Taguchi, S., Nishiyama, T., Yamane, T., and Doi, Y. 2002. A repressor protein, PhaR, regulates polyhydroxyalkanoate (PHA) synthesis via its direct interaction with PHA. J Bacteriol. 184:3992–4002

Maestro, B., and Sanz, J. M. 2017. Polyhydroxyalkanoate-associated phasins as phylogenetically heterogeneous, multipurpose proteins. Microb. Biotechnol.

10:1323–1337

Mandon, K., Kaminski Pa Fau - Elmerich, C., Elmerich, C., Bacteriol, J., Kaminski, P. A., and Elmerich, C. 1994. Functional analysis of the fixNOQP region of Azorhizobium caulinodans. J. Bacteriol. 176:2560–2568

Mandon, K., Michel-Reydellet, N., Encarnacion, S., Kaminski, P. A., Leija, A., Cevallos, M. A., Elmerich, C., and Mora, J. 1998. Poly-beta-hydroxybutyrate turnover in Azorhizobium caulinodans is required for growth and affects nifA expression. J. Bacteriol. 180:5070–5076

McDermott, T. R., and Kahn, M. L. 1992. Cloning and mutagenesis of the Rhizobium meliloti isocitrate dehydrogenase gene. J. Bacteriol. 174:4790–4797

Mckay, I. A., Dilworth, M. J., and Glenn, A. R. 1988. C_4-dicarboxylate metabolism in free-living and bacteroid forms of Rhizobium leguminosarum MNF3841. Microbiology. 134:1433–1440

Miflin, B. 2000. Crop improvement in the 21st century. J Exp Bot. 51:1–8

Mueller, N. D., Gerber, J. S., Johnston, M., Ray, D. K., Ramankutty, N., and Foley, J. A. 2012. Closing yield gaps through nutrient and water management. Nature. 490:254–257

Müller, J., Boller, T., and Wiemken, A. 2001. Trehalose becomes the most abundant non-structural carbohydrate during senescence of soybean nodules. J. Exp. Bot. 52:943–947

Mulley, G., Lopez-Gomez, M., Zhang, Y., Terpolilli, J., Prell, J., Finan, T., and Poole, P. 2010. Pyruvate is synthesized by two pathways in pea bacteroids with different efficiencies for nitrogen fixation. J. Bacteriol. 192:4944–4953

Mulley, G., White, J. P., Karunakaran, R., Prell, J., Bourdes, A., Bunnewell, S., Hill, L., and Poole, P. S. 2011. Mutation of GOGAT prevents pea bacteroid formation and N2 fixation by globally downregulating transport of organic nitrogen sources. Mol. Microbiol. 80:149–167

Mus, F., Crook, M. B., Garcia, K., Costas, A. G., Geddes, B. A., Kouri, E. D., Paramasivan, P., Ryu, M. H., Oldroyd, G. E. D., Poole, P. S., Udvardi, M. K., Voigt, C. A., Ane, J. M., and Peters, J. W. 2016. Symbiotic nitrogen fixation and the challenges to its extension to non-legumes. Appl. Environ. Microbiol. 82:3698–3710

Natera, S. H. A., Guerreiro, N., and Djordjevic, M. A. 2000. Proteome analysis of differentially displayed proteins as a tool for the investigation of symbiosis. Mol. Plant-Microbe Interact. 13:995–1009

Ndoye, I., De Billy, F., Vasse, J., Dreyfus, B., and Truchet, G. 1994. Root nodulation of Sesbania rostrata. J. Bacteriol. 176:1060–1068

Nishihata, S., Kondo, T., Tanaka, K., Ishikawa, S., Takenaka, S., Kang, C.-M. M., and Yoshida, K.-I. I. 2018. Bradyrhizobium diazoefficiens USDA110 PhaR functions for pleiotropic regulation of cellular processes besides PHB accumulation. BMC Microbiol. 18:156

Norel, F., and Elmerich, C. 1987. Nucleotide sequence and functional analysis of the two nifH copies of Rhizobium ORS571. J. Gen. Microbiol. 133:1563–1576

O'Callaghan, K. J., Stone, P. J., Hu, X., Griffiths, D. W., Davey, M. R., and Cocking, E. C. 2000. Effects of glucosinolates and flavonoids on colonization of the roots of Brassica napus by Azorhizobium caulinodans ORS571. Appl. Environ. Microbiol. 66:2185–2191

Oenema, O., Witzke, H. P., Klimont, Z., Lesschen, J. P., and Velthof, G. L. 2009. Integrated assessment of promising measures to decrease nitrogen losses from agriculture in EU-27. Agric. Ecosyst. Environ. 133:280–288

Oke, V., and Long, S. R. 1999. Bacteroid formation in the Rhizobium-legume symbiosis. Curr. Opin. Microbiol. 2:641–646

Oldroyd, G. E. 2012. Solving the Nitrogen problem in food production. Guest lecture, University of York

Oldroyd, G. E. D., Murray, J. D., Poole, P. S., and Downie, J. A. 2011. The rules of engagement in the legume-rhizobial symbiosis. Annu. Rev. Genet. Vol 45. 45:119–144

Oldroyd, G. E., and Dixon, R. 2014. Biotechnological solutions to the nitrogen problem. Curr Opin Biotechnol. 26:19–24

van Opijnen, T., and Camilli, A. 2013. Transposon insertion sequencing: A new tool for systems-level analysis of microorganisms. Nat. Rev. Microbiol. 11

Ott, T., Van Dongen, J. T., Günther, C., Krusell, L., Desbrosses, G., Vigeolas, H., Bock, V., Czechowski, T., Geigenberger, P., and Udvardi, M. K. 2005. Symbiotic leghemoglobins are crucial for nitrogen fixation in legume root nodules but not for general plant growth and development. Curr. Biol. 15:531–535

Paau, A. S., Bloch, C. B., and Brill, W. J. 1980. Developmental fate of Rhizobium meliloti bacteroids in alfalfa nodules. J. Bacteriol. 143:1480–1490

Paula Alonso, A., Dale, V. L., and Shachar-Hill, Y. 2010. Understanding fatty acid synthesis in developing maize embryos using metabolic flux analysis. Metab. Eng.

12:488–497

Pauling, D. C., Lapointe, J. P., Paris, C. M., and Ludwig, R. A. 2001. Azorhizobium caulinodans pyruvate dehydrogenase activity is dispensable for aerobic but required for microaerobic growth. Microbiology. 147:2233–2245

Pawlowski, K., Klosse, U., and de Bruijn, F. J. 1991. Characterization of a novel Azorhizobium caulinodans ORS571 two-component regulatory system, NtrY/NtrX, involved in nitrogen fixation and metabolism. MGG Mol. Gen. Genet. 231:124–138

Perry, B. J., Akter, M. S., and Yost, C. K. 2016. The use of transposon insertion sequencing to interrogate the core functional genome of the legume symbiont rhizobium leguminosarum. Front. Microbiol. 7

Perry, B. J., and Yost, C. K. 2014. Construction of a mariner-based transposon vector for use in insertion sequence mutagenesis in selected members of the Rhizobiaceae. BMC Microbiol. 14

Pimentel, D., and Paoletti, M. G. 2009. Developing a 21st Century View of Agriculture and the Environment.

Poole, P., and Allaway, D. 2000. Carbon and nitrogen metabolism in rhizobium. Adv. Microb. Physiol. 43:117–163

Potter, M., Madkour, M. H., Mayer, F., and Steinbüchel, A. 2002. Regulation of phasin expression and polyhydroxyalkanoate (PHA) granule formation in Ralstonia eutropha H16. Microbiology. 148:2413–2426

Quelas, J. I., Mesa, S., Mongiardini, E. J., Jendrossek, D., and Lodeiro, A. R. 2016. Regulation of polyhydroxybutyrate synthesis in the soil bacterium Bradyrhizobium diazoefficiens. Appl Env. Microbiol. 82:4299–4308

Quelas, J. I., Mongiardini, E. J., Pérez-Giménez, J., Parisi, G., and Lodeiro, A. R. 2013. Analysis of two polyhydroxyalkanoate synthases in bradyrhizobium japonicum USDA 110. J. Bacteriol. 195:3145–3155

Rachwal, K., Matczyńska, E., and Janczarek, M. 2015. Transcriptome profiling of a Rhizobium leguminosarum bv. trifolii rosR mutant reveals the role of the transcriptional regulator RosR in motility, synthesis of cell-surface components, and other cellular processes. BMC Genomics. 16

Ramachandran, V. K., East, A. K., Karunakaran, R., Downie, J. A., and Poole, P. S. 2011. Adaptation of Rhizobium leguminosarum to pea, alfalfa and sugar beet rhizospheres investigated by comparative transcriptomics. Genome Biol. 12

Ratcliff, W. C., and Denison, R. F. 2010. Individual-Level Bet Hedging in the Bacterium

Sinorhizobium meliloti. Curr. Biol. 20:1740–1744

Raun, W. R., and Johnson, G. V. 1999. Improving nitrogen use efficiency for cereal production. Agron. J. 91:357–363

Reid, C. J., and Poole, P. S. 1998. Roles of DctA and DctB in signal detection by the dicarboxylic acid transport system of Rhizobium leguminosarum. J. Bacteriol. 180:2660–2669

Resendis-Antonio, O., Reed, J. L., Encarnacion, S., Collado-Vides, J., and Palsson, B. O. 2007. Metabolic reconstruction and modeling of nitrogen fixation in Rhizobium etli. PLoS Comput Biol. 3:1887–1895

Rockstrom, J., Steffen, W., Noone, K., Persson, A., Chapin III, F. S., Lambin, E. F., Lenton, T. M., Scheffer, M., Folke, C., Schellnhuber, H. J., Nykvist, B., de Wit, C. A., Hughes, T., van der Leeuw, S., Rodhe, H., Sorlin, S., Snyder, P. K., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R. W., Fabry, V. J., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P., and Foley, J. A. 2009. A safe operating space for humanity. Nature. 461:472–475

Rogers, C., and Oldroyd, G. E. 2014. Synthetic biology approaches to engineering the nitrogen symbiosis in cereals. J Exp Bot. 65:1939–1946

Ronson, C. W., Astwood, P. M., and Downie, J. A. 1984. Molecular cloning and genetic organization of C4-dicarboxylate transport genes from Rhizobium leguminosarum. J. Bacteriol. 160:903–909

Roscher, A., Kruger, N. J., and Ratcliffe, R. G. 2000. Strategies for metabolic flux analysis in plants using isotope labelling. J. Biotechnol. 77:81–102

Roworth, J. D. 2017. The characterisation of root exudation and colonisation in the rhizosphere of land plants. D.Phil thesis, University of Oxford

Sabry, S. R. S., Saleh, S. A., Batchelor, C. A., Jones, J., Jotham, J., Webster, G., Kothari, S. L., Davey, M. R., and Cocking, E. C. 1997. Endophytic establishment of Azorhizobium caulinodans in wheat. Proc. R. Soc. B Biol. Sci. 264:341–346

Scalbert, A., Brennan, L., Fiehn, O., Hankemeier, T., Kristal, B. S., van Ommen, B., Pujos-Guillot, E., Verheij, E., Wishart, D., and Wopereis, S. 2009. Mass-spectrometry-based metabolomics: Limitations and recommendations for future progress with particular focus on nutrition research. Metabolomics. 5:435–458

Schafer, A., Tauch, A., Jager, W., Kalinowski, J., Thierbach, G., and Puhler, A. 1994. Small mobilizable multi-purpose cloning vectors derived from the Escherichia coli plasmids pK18 and pK19: selection of defined deletions in the chromosome of

Corynebacterium glutamicum. Gene. 145:69–73

Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270:467–470

Schlesinger, W. H. 2009. On the fate of anthropogenic nitrogen. Proc Natl Acad Sci U S A. 106:203–208

Schubert, K. R., and Evans, H. J. 1976. Hydrogen evolution: a major factor affecting the efficiency of nitrogen fixation in nodulated symbionts. Proc. Natl. Acad. Sci. U. S. A. 73:1207–1211

Schubert, P., Steinbüchel, A., and Schlegel, H. G. 1988. Cloning of the Alcaligenes eutrophus genes for synthesis of poly-beta-hydroxybutyric acid (PHB) and synthesis of PHB in Escherichia coli. J. Bacteriol. 170:5837–5847

Schümann, C., Michlmayr, H., del Hierro, A. M., Kulbe, K. D., Jiranek, V., Eder, R., and Nguyen, T. H. 2013. Malolactic enzyme from Oenococcus oeni: Heterologous expression in Escherichia coli and biochemical characterization. Bioengineered. 4

Schumpp, O., and Deakin, W. J. 2010. How inefficient rhizobia prolong their existence within nodules. Trends Plant Sci. 15:189–195

Scott, J. D., and Ludwig, R. A. 2004. Azorhizobium caulinodans electron-transferring flavoprotein N electrochemically couples pyruvate dehydrogenase complex activity to N-2 fixation. Microbiology-Sgm. 150:117–126

Sender, P. D., Martín, M. G., Peirú, S., and Magni, C. 2004. Characterization of an oxaloacetate decarboxylase that belongs to the malic enzyme family. FEBS Lett. 570:217–222

Senior, P. J., and Dawes, E. A. 1971. Poly- -hydroxybutyrate biosynthesis and the regulation of glucose metabolism in Azotobacter beijerinckii. Biochem. J. 125:55–66

Shah, R., and Emerich, D. W. 2006. Isocitrate dehydrogenase of Bradyrhizobium japonicum is not required for symbiotic nitrogen fixation with soybean. J. Bacteriol. 188:7600–7608

Sharma, M., and Dhingra, H. 2016. Poly-β-hydroxybutyrate: A biodegradable polyester, biosynthesis and biodegradation. Br. Microbiol. Res. J. 14:1–11

Shcherbak, I., Millar, N., and Robertson, G. P. 2014. Global metaanalysis of the nonlinear response of soil nitrous oxide (N2O) emissions to fertilizer nitrogen. Proc Natl Acad Sci U S A. 111:9199–9204

Stam, H., van Verseveld, H. W., de Vries, W., and Stouthamer, A. H. 1986. Utilization of

poly-β-hydroxybutyrate in free-living cultures of Rhizobium ORS571. FEMS Microbiol. Lett. 35:215–220

Stowers, M. 1985. Carbon metabolism in Rhizobium species. Annu. Rev. Microbiol. 39:89–108

Suzuki, S., Aono, T., Lee, K. B., Suzuki, T., Liu, C. T., Miwa, H., Wakao, S., Iki, T., and Oyaizu, H. 2007. Rhizobial factors required for stem nodule maturation and maintenance in Sesbania rostrata-Azorhizobium caulinodans ORS571 symbiosis. Appl Env. Microbiol. 73:6650–6659

Terpolilli, J. J., Hood, G. A., and Poole, P. S. 2012. What determines the efficiency of N_2 fixing rhizobium-legume symbioses? Adv. Microb. Physiol. Vol 60. 60:325–389

Terpolilli, J. J., Masakapalli, S. K., Karunakaran, R., Webb, I. U., Green, R., Watmough, N. J., Kruger, N. J., Ratcliffe, R. G., and Poole, P. S. 2016. Lipogenesis and redox balance in nitrogen-fixing pea bacteroids. J Bacteriol. 198:2864–2875

Tester, M., and Langridge, P. 2010. Breeding technologies to increase crop production in a changing world. Science (80-.). 327:818–822

Tilman, D., Balzer, C., Hill, J., and Befort, B. L. 2011. Global food demand and the sustainable intensification of agriculture. Proc Natl Acad Sci U S A. 108:20260–20264

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

Tombolini, R., and Nuti, M. P. 1989. Poly(ß-hydroxyalkanoate) biosynthesis and accumulation by different Rhizobium species. FEMS Microbiol. Lett. 60:299–304

Trainer, M. A., and Charles, T. C. 2006. The role of PHB metabolism in the symbiosis of rhizobia with legumes. Appl Microbiol Biotechnol. 71:377–386

Tsukada, S., Aono, T., Akiba, N., Lee, K. B., Liu, C. T., Toyazaki, H., and Oyaizu, H. 2009. Comparative genome-wide transcriptional profiling of Azorhizobium caulinodans ORS571 grown under free-living and symbiotic conditions. Appl. Environ. Microbiol. 75:5037–5046

Udvardi, M., and Poole, P. S. 2013. Transport and metabolism in legume-rhizobia symbioses. Annu. Rev. Plant Biol. Vol 64. 64:781–805

de Vries, W., Stam, H., and Stouthamer, A. H. 1984. Hydrogen oxidation and nitrogen fixation in rhizobia, with special attention focused on strain ORS571. Antonie Van Leeuwenhoek. 50:505–524

Walther, J. L., Metallo, C. M., Zhang, J., and Stephanopoulos, G. 2012. Optimization of ^{13}C isotopic tracers for metabolic flux analysis in mammalian cells. Metab. Eng. 144:167–174

Wang, C., Saldanha, M., Sheng, X., Shelswell, K. J., Walsh, K. T., Sobral, B. W. S., and Charles, T. C. 2007. Roles of poly-3-hydroxybutyrate (PHB) and glycogen in symbiosis of Sinorhizobium meliloti with Medicago sp. Microbiology. 153:388–398

Wang, Z., Gerstein, M., and Snyder, M. 2009. RNA-Seq: A revolutionary tool for transcriptomics. Nat. Rev. Genet. 10:57–63

Washburn, M. P., Wolters, D., and Yates, J. R. 2001. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat. Biotechnol. 19:242–247

Webb, I. 2016. Characterisation of the fixABCX operon in symbiotic nitrogen fixation. D.Phil thesis, John Innes Centre

Wetmore, K. M., Price, M. N., Waters, R. J., Lamson, J. S., He, J., Hoover, C. A., Blow, M. J., Bristow, J., Butland, G., Arkin, A. P., and Deutschbauer, A. 2015. Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. MBio. 6:1–15

Wheatley, R. 2018. Genome-scale characterisation of symbiotic fitness determinants of Rhizobium leguminosarum using INSeq. D.Phil thesis, University of Oxford

Wheatley, R. M., Ramachandran, V. K., Geddes, B. A., Perry, B. J., Yost, C. K., and Poole, P. S. 2017. Role of O_2 in the growth of Rhizobium leguminosarum bv. viciae 3841 on glucose and succinate. J. Bacteriol. 199

White, J., Prell, J., James, E. K., and Poole, P. 2007. Nutrient sharing between symbionts. Plant Physiol. 144:604–614

Williams, A., Wilkinson, A., Krehenbrink, M., Russo, D. M., Zorreguieta, A., and Downie, J. A. 2008. Glucomannan-mediated attachment of Rhizobium leguminosarum to pea root hairs is required for competitive nodule infection. J. Bacteriol. 190:4706–4715

Wong, P. P., and Evans, H. J. 1971. Poly-beta-hydroxybutyrate utilization by soybean (Glycine max Merr.) nodules and assessment of its role in maintenance of nitrogenase activity. Plant Physiol. 47:750–755

Wood, W. B. 1966. Host specificity of DNA produced by Escherichia coli: Bacterial mutations affecting the restriction and modification of DNA. J. Mol. Biol. 16:118–133

Xu, G., Fan, X., Miller, A. J., Fan X Fau - Miller, A. J., Miller, A. J., and Annu Rev Plant, B. 2012. Plant nitrogen assimilation and use efficiency. Annu. Rev. Plant Biol. 63:153–182

Yamada, M., Takahashi, S., Okahata, Y., Doi, Y., and Numata, K. 2013. Monitoring and kinetic analysis of the molecular interactions by which a repressor protein, PhaR, binds to target DNAs and poly[(R)-3-hydroxybutyrate]. AMB Express. 3:6

Young, J. D. 2013. Metabolic flux rewiring in mammalian cell cultures. Curr. Opin. Biotechnol. 24:1108–1115

Young, J. P. W., Crossman, L. C., Johnston, A. W. B., Thomson, N. R., Ghazoui, Z. F., Hull, K. H., Wexler, M., Curson, A. R. J., Todd, J. D., Poole, P. S., Mauchline, T. H., East, A. K., Quail, M. A., Churcher, C., Arrowsmith, C., Cherevach, I., Chillingworth, T., Clarke, K., Cronin, A., Davis, P., Fraser, A., Hance, Z., Hauser, H., Jagels, K., Moule, S., Mungall, K., Norbertczak, H., Rabbinowitsch, E., Sanders, M., Simmonds, M., Whitehead, S., and Parkhill, J. 2006. The genome of Rhizobium leguminosarum has recognizable core and accessory components. Genome Biol. 7

Zhang, Y., Aono, T., Poole, P., and Finan, T. M. 2012. NAD(P)$^{+}$-malic enzyme mutants of Sinorhizobium sp. strain ngr234, but not Azorhizobium caulinodans ORS571, maintain symbiotic N2 fixation capabilities. Appl. Environ. Microbiol. 78:2803–2812