Identification and characterization of pentose phosphate and Embden Meyerhof Parnas pathway mutants in *Sinorhizobium meliloti*

By

Sabhjeet Kaur

A Thesis Submitted to the Faculty of Graduate Studies of
The University of Manitoba
In partial Fulfillment of the Requirements of the Degree of

Doctor of Philosophy

Department of Microbiology
University of Manitoba
Winnipeg, Manitoba, Canada

Copyright © 2021 by Sabhjeet Kaur

The Great Tablecloth

Let us sit down to eat with all those

who haven't eaten,

let us spread great tablecloths, put salt in

the lakes of the world,

set up planetary bakeries, tables with

strawberries in snow,

and a plate like the moon itself

from which we can all eat.

For now I ask no more

than the justice of eating.

Pablo Neruda

Abstract

The interaction of *Sinorhizobium meliloti* and alfalfa is a well-studied model system for symbiotic establishment between rhizobia and legumes. The *S. meliloti* Rm1021 genome consists of a variety of genes involved in carbon metabolism and transport amongst others. Insertional inactivation of *tktA,* a gene encoding the major transketolase enzyme for the pentose phosphate pathway, resulted in a strain that was unable to establish an effective symbiosis with alfalfa, or produce symbiotically active exopolysaccharides (EPS) and showed conditional auxotrophy. Strains carrying secondary mutations were isolated that could suppress the effect of a mutation in *tktA.* The strains carrying the second site mutations were capable of partially restoring symbiosis, as well as reversing the conditional carbon phenotypes that were associated with *tktA.* Single nucleotide polymorphisms (SNPs) were identified in a Gnt-type negative regulator *SMc02340.* The suppression was attributed to upregulation of *tktB* due to mutations in *SMc02340. SMc02340* gene product was shown as a negative regulator of *tktB, SMc02341* and *SMc02339.* A unique symbiotic phenotype for *tktA* mutant strain was also identified by microscopic analysis. The *tktA* mutant strain is unable to make curled colonized root hairs (CCRH) on alfalfa roots. In *S. meliloti,* the catabolism of hexoses occurs using the Entner Doudoroff pathway, and the upper part of the Embden Meyerhof Parnas pathway is used in a gluconeogenic manner because of the lack of a gene encoding an ATP dependent phosphofructokinase. We characterized *SMc01852* (*pfp*), which encodes a putative pyrophosphate dependent phosphofructokinase. Overexpression and purification of Pfp showed that the protein had pyrophosphate dependent phosphofructokinase activity. We have also shown that the growth and symbiotic ability of the *pfp* mutant is dependent on *tal* by generating Δ*pfp*Δ*tal* strain. Overall, we have identified the role of *tktA* and *tktB* in the PP pathway and symbiosis as well as characterization of the missing phosphofructokinase.

Acknowledgments

I would like to acknowledge my supervisor Dr. Ivan Oresnik for believing in me and giving the opportunity to work with him. He has always encouraged me to become an independent thinker and have supported my research endeavours. He has created such an environment in the lab where a creative mind can thrive. I would also like to thank my advisory committee members Dr. Deb Court, Dr. Ayush Kumar and Dr. David Levin. They have provided me constructive feedback over the years and helped me navigate through the graduate program. I also want to thank the funding sources; NSERC, and University of Manitoba for UMGF and GETS.

My fellow lab members deserve a very special acknowledgement. Dr. Justin Hawkins for mentoring me at the beginning of my PhD and being my go-to person for asking all the research and life questions. Dr. MacLean Kohlmeier is also deserving of much praise for his support. Patricia, Derek, Gagan, Dr. Muntahi Mourin, Dr. Manu Singh and Dr. Brijesh Kumar have not only been my colleagues but my closest friends. Their friendship is invaluable to me. There are many people in the Department of Microbiology who have created a friendly and helping environment for graduate students, I want to thank everyone in the department for this special place.

I could not have been a researcher without the unconditional support and trust of my parents Sukhwinder and Harbhinder, and my brother Shahbaaz. My husband Upkar deserves a lot of credit for his support and love. He has been a great partner in life and grad school. I would also like to thank my family members Navdeep, Baljinder and my in-laws for always being there for me.

Dedication

This thesis is dedicated to my grandparents Sardarni Parmjit Kaur and Late Sardar Gurtej Singh Grewal (*Papa ji*). Papa ji, if you were alive today, you would have been so proud!

Table of Contents

Abstract i
Acknowledgments ii
Dedication iii
Table of Contents iv
List of Figures viii
List of Abbreviations x
Chapter 1: Literature Review 1
1.1 Nitrogen 2
1.1.1 Nitrogen is required for plant growth 2
1.1.2 Biological nitrogen fixation 2
1.2 Legume-Rhizobium symbiotic associations 3
1.2.1 *Sinorhizobium meliloti*-alfalfa symbiotic model system 4
1.2.2 Nod Factors 5
1.2.3 Exopolysaccharides 5
1.3 The process of symbiosis 8
1.3.1 Early stages of symbiosis 8
1.3.2 Late stages of symbiosis 9
1.4 Root exudates 10
1.4.1 Rhizosphere and 'the competition problem' 11
1.5 Carbon metabolism in *Sinorhizobium meliloti* 11
1.5.1 Pentose phosphate pathway 12
1.5.2 Shikimate pathway 15
1.5.3 Purine and Pyrimidine biosynthesis 16
1.6 Effect of carbon flow on symbiosis 19
1.6.1 Transketolase 21
1.6.2 Phosphofructokinase 21
1.6.3 Phosphoglucose isomerase 23
1.6.4 Phosphoenolpyruvte carboxylase 24
1.6.5 Pyruvate carboxylase 25

1.7 Metabolic modelling 26
1.7.1 *Sinorhizobium meliloti* metabolic model 28
1.8 Thesis goals and hypothesis 30
Chapter 2: *Sinorhizobium meliloti* strains carrying mutations in transketolase are unable to form curled colonized root hairs 33
2.1 Introduction 34
2.2 Materials and Methods 36
2.2.1 Bacterial strains, plasmids, and media. 36
2.2.2 Genetic manipulations and techniques 37
2.2.3 Metabolite analysis. 43
2.2.4 RNA extraction, cDNA synthesis and qRT-PCR. 44
2.2.5 Exopolysaccharide production. 45
2.2.6 Plant growth conditions and nodulation kinetics. 45
2.2.7 Microscopy. 46
2.3 Results 46
2.3.1 Symbiotic defects caused by the *tktA* mutation can be suppressed. 46
2.3.2 Strains carrying suppressor mutations do not completely restore central carbon pathways. 52
2.3.3 *stkR1* and *stkR3* are located in close proximity to *tktB*. 52
2.3.4 Phenotypic suppression is dependent on *SMc02340* and *tktB*. 56
2.3.5 Ten independent suppressors are all alleles of *SMc02340*. 62
2.3.6 Strains carrying *tktA* mutations do not form colonized curled root hairs. 65
2.4 Discussion 68
Chapter 3: Characterization of *tktB* locus in *Sinorhizobium meliloti*. 72
3.1 Introduction 73
3.2 Materials and methods 75
3.2.1 Bacterial strains, plasmids, and media 75
3.2.2 RNA extraction, cDNA synthesis and qRT-PCR 78
3.2.3 Genetic manipulations and techniques 78
3.2.4 Growth curves 82
3.2.5 Phenotypic microarray 83

3.3 Results 83
3.3.1 *SMc02341* has a growth phenotype on ribose in a suppressor background. 83
3.3.2 SMc02340 negatively regulates its flanking operons 90
3.3.3 Phenotypic microarray 92
3.3.4 Δ*SMc02341* strains do not display any symbiotic phenotypes. 99
3.4 Discussion 101
Chapter 4: *Sinorhizobium meliloti* contains a complete Embden Meyerhof Parnas pathway 106
4.1 Introduction 107
4.2 Materials and methods 109
4.2.1 Bacterial strains, plasmids, and media 109
4.2.2 Genetic manipulations and constructions 109
4.2.3 Plant growth conditions 113
4.2.4 Metabolite analysis 115
4.2.5 Growth curves 116
4.2.6 Phenotypic microarray 116
4.2.7 Protein overexpression and purification 117
4.2.8 Biochemical assays 118
4.3 Results 119
4.3.1 Identification and characterization of *pfp* by gene deletion 119
4.3.2 *SMc01852* encodes a pyrophosphate dependent phosphofructokinase (Pfp) 123
4.3.3 Metabolite analysis of Δ*pfp* strains 128
4.3.4 A Δ*pfp*Δ*tal* double mutation affects the utilization of multiple carbon substrates 132
4.3.5 Addition of glucose to a medium containing galactose leads to an inability of strain carrying Δ*pfp*Δ*tal* to grow 137
4.3.6 Δ*pfp*Δ*ta*l mutations cause accumulation of phosphorylated intermediates 139
4.3.7 The Δ*pfp*Δ*tal* strain of *S. meliloti* displays reduced symbiotic efficiency 143
4.4 Discussion 146
Chapter 5: Conclusions 151
5.1 Conclusions from the current work 152
5.2 Future work 158
Literature Cited 160

List of Tables

Table 1. Strains and plasmids list 38
Table 2. Primer sequences for insertional mutations and gene deletion constructs..... 41
Table 3. Metabolite concentrations..... 55
Table 4. PacBio sequencing parameters 57
Table 5. Carbon phenotypes associated with *tktA* suppression..... 58
Table 6. List of SNPs on *SMc02340* in the suppressor strains. 63
Table 7. Strains and plasmids list 76
Table 8. Primer sequences for insertional mutations and gene deletion constructs..... 79
Table 9. Growth phenotype on ribose in minimal and complex medium..... 86
Table 10. Strains and plasmids list 110
Table 11. Primer sequences for insertional mutations and gene deletion constructs..... 114
Table 12. Growth of *S. meliloti* strains on different carbon sources in minimal medium..... 122
Table 13. Kinetic properties of Pfp..... 126
Table 14. Central carbon metabolite concentrations found in wild type and Δ*pfp* strains. 130
Table 15. Central carbon metabolism phenotypes of the Δ*pfp*Δ*tal* strain 134
Table 16. Central carbon metabolite concentrations found in wild type and Δ*pfp*Δ*tal* strains .. 142

List of Figures

Figure 1. Network of central carbon metabolism pathways in *S. meliloti*. 14
Figure 2. Alfalfa plants inoculated with wild type, the *tktA* mutant strain (SRmD397) and a suppressor strain SRmD410 (*stkR1*). 48
Figure 3. Symbiotic phenotypes associated with *tktA* suppression..... 49
Figure 4. Exopolysaccharide production by the wild type, the *tktA* mutant strain and three suppressor strains; *stkR1* (SRmD410), *stkR2* (SRmD411), and *stkR3* (SRmD466)..... 51
Figure 5. Box plots represent concentrations of R5P/Ribu5P and S7P pools in the wild type, the *tktA* mutant strain and three suppressor strains; *stkR1* (SRmD410), *stkR2* (SRmD411), and *stkR3* (SRmD466) 53
Figure 6. Clustered heatmap of central carbon metabolite concentrations in the wild type, the *tktA* mutant strain and three suppressor strains; *stkR1* (SRmD410), *stkR2* (SRmD411), and *stkR3* (SRmD466) 54
Figure 7. Average shoot dry weight of alfalfa plants inoculated with the *S. meliloti* strains. 59
Figure 8. Mutations in *SMc02340* affect *tktB* expression. 61
Figure 9. Location of single nucleotide polymorphisms (SNPs) in *SMc02340* and the predicted open reading frame (ORF) in 10 suppressor strains: *stkR1-stkR10*. 64
Figure 10. Images of alfalfa root hairs and root nodules produced by GFP-labelled *S. meliloti* strains; the wild type, the *tktA* mutant strains, *stkR1*, Δ*trpE*, *exoR* and *exoY* mutant strains..... 67
Figure 11. Genetic map of *tktB* locus in *S. meliloti*. 85
Figure 12. Confirmation of *SMc02341* deletion. 88
Figure 13. Growth curve analysis of *S. meliloti* strains. 89
Figure 14. qRT-PCR analysis of *SMc02340* neighbouring genes expression in *S. meliloti* strains (wt, *tktA*, *stkR1*, *stkR3*, and *SMc02340*::pk-Gm *tktA*). 91
Figure 15. Phenotypic microarray analysis. 93
Figure 16. Phenotypic microarray analysis. 94
Figure 17. Phenotypic microarray analysis. 96
Figure 18. Phenotypic microarray analysis. 97
Figure 19. Growth curve analysis of *S. meliloti strains*. 98
Figure 20. Dry weight of alfalfa plants inoculated with *S. meliloti* strains with *tktB* and *SMc02341* deletions. 100
Figure 21. Agarose gel showing PCR fragments of unmarked and marked *pfp* deletion constructs of *S. meliloti*. 121
Figure 22. Western blot analysis of Pfp purification. 124
Figure 23. Kinetic analysis of Pfp using fructose-6P (top panel) or fructose-1,6biP (bottom panel) as a substrate. 125
Figure 24. Phosphofructokinase activity of purified Pfp using ATP, GTP or PPi as phosphate donor. 127
Figure 25. Clustered heatmap of central carbon metabolite concentrations in the wild type and Δ*pfp* strains grown on glucose or succinate as carbon source. 129

Figure 26. Metabolite concentrations 131
Figure 27. Phenotypic microarray analysis of wild type (black), Δ*tal* (blue), Δ*pfp* (red), Δ*pfp*Δ*tal* (orange) strains..... 135
Figure 28. Phenotypic microarray analysis of wild type (black), Δ*tal* (blue), Δ*pfp* (red), Δ*pfp*Δ*tal* (orange) strains..... 136
Figure 29. Bacterial growth curve analysis of the wild type (blue) and Δ*pfp*Δ*tal* (red) strains on galactose and glucose as carbon sources. 138
Figure 30. A Clustered heatmap of central carbon metabolite concentrations in the wild type and Δ*pfp*Δ*tal* strains grown either on galactose or galactose+glucose as carbon source. 140
Figure 31. A box plot representation of sedoheptulose-7P pools in the wild type and Δ*pfp*Δ*tal* strains. 141
Figure 32. Dry weights of alfalfa plants inoculated with *S. meliloti* strains. 144
Figure 33. Alfalfa plants grown in nitrogen deficient conditions inoculated with (from left) wild type, Δ*tal*, Δ*pfp*, Δ*pfp*Δ*tal* strains and uninoculated control. 145
Figure 34. Proposed role of Pfp in upper EMP and PP pathways. 147
Figure 35. Predicted structure of SMc02341 based on a crystal of gluconate kinase of *Lactobacillus acidophilus*. 155

List of Abbreviations

%	Percentage
2-PG	2-phosphoglycerate
3-PG	3- phosphoglycerate
6-PG	6-phosphogluconate
AICAR	4-amino-5-imidazole-carboxamide
aKG	Alpha ketoglutarate
ALE	Adaptive laboratory evolution
APRT	AMP pyrophosphate phosphoribosyl transferase
ATP	Adenosine triphosphate
BiGG	Biochemical, genetic, and genomic
BLAST	Basic local alignment sequence tool
BNF	Biological nitrogen fixation
C/g	Carbon per gram
C2	2 Carbon molecule
C3	3 Carbon molecule
C4	4 Carbon molecule
CBB	Celvin Bensen Bassham
CCRH	Curled colonized root hair
cDNA	Cyclic deoxyribose nucleic acid
CFU	Colony forming unit
Cm	Chloramphenicol
Cm^r	Chloramphenicol resistant
Da	Daltons
DHAP	Dihydroxyacetone phosphate
DME	NAD^+-dependent malic enzyme
ED	Entner Doudoroff
EMP	Embden Meyerhof Parnas
eno	Enolase
EPS	Exopolysaccharide
EPS I	Succinoglycan
EPS II	Galactoglucan
F6P	Fructose 6 phosphate
F1,6biP	Fructose-1,6-bisphosphate
FRT	Flippase recognition target
g	Grams
G1P	Glucose 1 phosphate
G6P	Glucose 6 phosphate
gdhA	Glutamine dehydrogenase
Gm	Gentamicin

GMP	Guanosine-5’-monophosphate
Gm^r	Gentamicin resistant
Gnd	6-phosphogluconate dehydrogenase
GTP	Guanosine triphosphate
h	Hours
HMW	High molecular weight
IMP	Inosine-5’-monophosphate
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
IRLC	Inverted repeat lacking clade
IT	Infection thread
kb	Kilobase
kDa	Kilodaltons
KDPG	2-Keto-3-deoxy-6-phosphogluconate
Km	Kanamycin
K_m	Michaelis constant
Km^r	Kanamycin resistant
L	Liter
LB	Luria Bertani
LMW	Low molecular weight
LPS	Lipopolysaccharide
M	Molar
Mb	Megabase
mg	Milligram
N	Nitrogen
N_2	Dinitrogen
Na	Sodium
NaCl	Sodium chloride
NAD	Nicotinamide adenine dinucleotide
NADH	Reduced nicotinamide adenine dinucleotide
NaOH	Sodium hydroxide
NADPH	Nicotinamide adenine dinucleotide phosphate
NCRs	Nodule like cystine rich polypeptides
ng	Nanogram
NF	Nod factor
Ni-NTA	Nickel Nitrilotriacetic acid
NH_3	Ammonia
NH_4^+	Ammonium
Nm^r	Neomycin resistant
OD	Optical density
OMP	Orotidine-5’-monophosphate

ORF	Open reading frame
pck	Phosphoenolpyruvate carboxylase
Pen^r	Penicillin resistant
pfk	Phosphofructoketolase
Pfp	Pyrophosphate dependent phosphofructose kinase
pgi	Phosphoglucose isomerase
PP	Pentose phosphate
PPDK	Pyruvate orthophosphate dikinase
PPi	Pyrophosphate
PRPP	5-phosphoribosylpyrophosphate
pyc	Pyruvate carboxylase
qRT-PCR	Quantitative real time polymerase chain reaction
R5P	Ribose 5 phosphate
Ribu5P	Ribulose 5 phosphate
RNA	Ribose nucleic acid
rpi	Ribose 5 phosphate isomerase
Rpm	Rounds per minute
S7P	Sedoheptulose 7 phosphate
Sm^r	Streptomycin resistant
SNF	Symbiotic nitrogen fixation
tal	Transaldolase
Tc	Tetracycline
TCA	Tri-carboxylic acid
TCP	1,2,3-trichloropropane
Tc^r	Tetracycline resistant
Tg	Trillion grams
tktA	Transketolase A
tktB	Transketolase B
ViNE	Virtual Nodule Environment
V_{max}	Maximum rate achieved by the system
VMM	Vincent's minimal medium
µl	Microlitre

Chapter 1:

Literature Review

1.1 Nitrogen

1.1.1 Nitrogen is required for plant growth

N_2 is abundant in the atmosphere in the form of di-nitrogen (N_2) making up 78% of the air in Earth's atmosphere. However, plants are unable to utilize the atmospheric N_2 in its gaseous form. N_2 must be converted into a plant utilizable form such as ammonia to fulfill its nitrogen needs (Cassman et al. 2002). In agriculture, a nitrogen deficiency is alleviated with the application of chemical fertilizers. The use of chemical fertilizers began in early 1900s with methods like the Haber-Bosch process that was invented to convert nitrogen gas to ammonia (Ertl 1991). The increased use of chemical fertilizers has led to an increase in food production; however, these amendments have drastically affected soil and human health over time (Oldroyd 2013). Plants often use only small amounts of applied nitrogen. The majority of the nitrogen runs off into receiving water bodies, leading to environmental pollution and eutrophication (Cassman et al. 2002). Due to the negative effects associated with the use of chemical fertilizers, the focus of research has shifted towards the development of more sustainable agricultural practices.

1.1.2 Biological nitrogen fixation

Biofertilizers can provide a sustainable alternative for nitrogen acquisition through biological nitrogen fixation (BNF). BNF can only be carried out by bacteria and can occur in free living, associative or symbiotic associations (Cassman et al. 2002). Symbiotic nitrogen fixation (SNF) is a process in which bacterial symbionts reduce atmospheric nitrogen to ammonia in return for carbon. The most important symbiotic association for nitrogen fixation happens between legumes and a group of bacteria found in the soil called 'rhizobia'. It should be noted that the process of symbiotic nitrogen fixation is limited to legumes except for *Parasponia andersonii* (Trinick 1973). According to recent global estimates, BNF contributes around 60 Tg

N per year in agricultural systems and most of it comes from legume-rhizobia symbiotic associations (Herridge et al. 2008; Fowler et al. 2013).

1.2 Legume-Rhizobium symbiotic associations

Rhizobia have been classified as α and β proteobacteria that have an innate ability to establish symbiotic associations with legumes resulting in the formation of root nodules. Atmospheric N_2 is reduced to ammonia by the enzyme nitrogenase in rhizobia. Nitrogenase is irreversibly inactivated by the presence of oxygen (Hoffman et al. 2014). Therefore, nitrogen fixation often happens under anoxic or nearly anoxic conditions.

The process of symbiotic nitrogen fixation takes place in plant-derived structures called nodules. Nodules are developed by the legume in response to a signal exchange between the rhizobium and its host legume (Gage 2004). Depending on the host plant, the nodules can be either determinate or indeterminate. Legumes growing in a temperate climate tend to form indeterminate nodules that are elongated in shape and possess a persistent apical meristem. This results in the new cells continually being infected by the bacteria that are present within the nodules. Within an indeterminate nodule there are distinct zones of development, with actively dividing cells near the apical meristem, and old senescent tissue proximal to the root. *Medicago sativa* (alfalfa), *Medicago truncatula* (sweet clover) and *Pisum sativum* (peas) are examples of some legumes that have indeterminate nodules.

Determinate nodules are round and lack a persistent meristem. Therefore, they lack the development gradient that is present in indeterminate nodules. Tropical climate legumes such as *Glycine max* (soybean), *Lotus japonicus* (lotus) and *Vicia faba* (beans) are some examples of legumes that make determinate nodules (Hirsch 1992).

1.2.1 Sinorhizobium meliloti-alfalfa symbiotic model system

The *S. meliloti*-alfalfa symbiotic association has long been studied as one of the model systems to understand symbiosis. *S. meliloti* is primarily studied for its interaction with alfalfa and sweet clover, but it can also enter a symbiotic relationship with trigonella. *S. meliloti* is an α-proteobacterium that belongs to the family *Rhizobiaceae*. *S. meliloti* has a dimorphic life cycle where it can live freely in the soil, or it can enter into a symbiotic association with its host legume. During the latter, the bacteria are transformed into bacteroids that are able to divide and fix nitrogen under low oxygen conditions (Geddes and Oresnik 2016).

S. meliloti Rm1021 is a laboratory strain that has been subjected to extensive genetic, biochemical, and metabolic research for decades. Rm1021 is a derivative of *S. meliloti* SU47 which was isolated in Australia by J. M. Vicent (Vincent 1941). The genome of *S. meliloti* Rm1021 has been sequenced and it consists of three replicons, a 3.6 Mb chromosome, a mega plasmid, pSymA (1.35Mb) and a chromid, pSymB (1.68 Mb) (Barnett et al. 2001; Capela et al. 2001; Finan et al. 2001; Galibert et al. 2001). pSymA contains the genes for nodule formation and nitrogen fixation like *nif*, *nod* and *fix* as well as a glutamine dehydrogenase (*gdhA*) and genes necessary for denitrification (Galibert et al. 2001). pSymA does not contain any essential genes, therefore it is called a megaplasmid. In the original *S. meliloti* 1021 sequencing papers pSymB was also classified as a megaplasmid, although the presence of putative essential genes was discussed (Galibert et al. 2001). pSymB contains essential genes like arginine tRNA $^{Arg}tRNA_{CCG}$ and *engA*, encoding for a putative GTP binding protein. The essential nature of these genes has been experimentally validated, therefore the classification of pSymB has been changed from a megaplasmid to a chromid (diCenzo and Finan 2017). pSymB also contains genes involved in small molecule transport and surface polysaccharide synthesis. Up to 12% of the genes present

on pSymB are dedicated towards polysaccharide synthesis like exopolysaccharide (EPS), lipopolysaccharide (LPS) and cyclic β-glucans (Galibert et al. 2001). In addition, 17.4% of the genes on pSymB are predicted to have an import function (Galibert et al. 2001). A great variety of carbon catabolism genes are present on the chromosome and pSymB, which explains the ability of *S. meliloti* to grow on many monosaccharides, polysaccharides, and organic acids.

1.2.2 Nod Factors

To establish a successful symbiosis, a series of complex signal exchanges must take place between rhizobium and its legume host (Jones et al. 2007a). Rhizobium responds to plant-derived flavonoids by activating genes that encode enzymes to produce a lipo-chito-oligosaccharide called Nod factor (Lerouge et al. 1990). Nod factors have a chitin backbone and a N-linked fatty acid moiety attached to a non-reducing oligosaccharide. Depending upon the species, the N-acetyl groups and decorations on the chitin backbone are different and contribute towards the host specificity. Nod factors act as a signal to induce the division of inner root cortical cells to develop nodule primordia (Timmers et al. 1999). Nod factor is recognized by specific LysR-type receptors present on plant roots (Luo et al. 2005).

1.2.3 Exopolysaccharides

In addition to Nod factors, exopolysaccharides also play a crucial role during symbiosis. *S. meliloti* produces two main types of exopolysaccharides: succinoglycan (EPS I), and galactoglucan (EPS II). Structurally, EPS I is a repeating polymer consisting of seven glucose and one galactose subunits, which can be modified by acetyl, succinyl and pyruvyl groups (Reinhold et al. 1994). EPS II is comprised of a repeating polymer of galactose and glucose, which can be decorated with acetyl and pyruvyl groups (Her et al. 1990). Both EPS I and EPS II are produced in either a high (HMW) or a low molecular weight (LMW) form. The HMW

fractions of both EPS I and EPS II consist of hundreds to thousands of the repeating units (Reinhold et al. 1994; Battisti et al. 1992).

S. meliloti strain Rm1021 produces only EPS I under most conditions. The production of EPS I has been shown to be critical for symbiotic establishment that plays a role in either stress tolerance or symbiotic signaling (Leigh and Lee 1988; Jones et al. 2007a; Hawkins et al. 2017a). Evidence that is consistent with the monomer, dimer or trimer fractions playing a critical role in early symbiotic interactions has been shown for LMW EPS I (Wang et al. 1999; González et al. 1998). As a result, it was generally believed that LMW EPS I plays an important role in bacterial infection of the plant tissues (Cheng and Walker 1998). However, a recent study has shown that LMW EPS is not essential as *S. meliloti* double mutants (*exoK exoH*) that only produced HMW EPS were able to establish symbiosis with host *M. truncatula* although with reduced efficiency than the wild type. The important step is succinylation of EPS as the *exoH* mutant strains that produced unsuccinylated EPS I failed to establish symbiosis (Mendis et al. 2016). This mutant strain only produces HMW EPS as the unsuccinylated EPS can not be cleaved by glycanases ExoK and ExsH (York and Walker 1998).

The production of EPS II was first found while characterizing of mucoid strains of Rm1021. Production of EPS II was found to be regulated by *mucR* and *expR*. Interestingly, EPS II that was produced in a *mucR* background could not rescue strains deficient in EPS I biosynthesis symbiotically, while the EPS II produced in an *expR*$^+$ background was symbiotically active (Keller et al. 1995). The differences between these two backgrounds seems to be the proportion of LMW EPS II that was produced (González et al. 1996). Further characterization of this demonstrated that EPS I deficient strains were able to form nodules when supplied with EPS II isolated from an *expR*$^+$ background or if small quantities of chemically

synthesized LMW EPS II were supplied to an infected root hair (González et al. 1996). Subsequent work showed that the *expR*$^{+}$ mutation was the result of a precise excision of an insertion element that is naturally found in the *expR* gene in Rm1021. This is often referred to as an *expR*$^{+}$ background. Of note, it has also been shown that EPS-II can be synthesized in Rm1021 background if the strain is grown under phosphate limiting conditions (Oresnik et al. 1994; Mendrygal and González 2000).

Inoculation of *M. truncatula* with succinoglycan deficient strains of *S. meliloti* resulted in the upregulation of genes responsible for plant defence responses as compared to plants inoculated with the wild type strain. This supported the hypothesis that the symbiotically active exopolysaccharides acted as signals for the plant host to initiate infection thread formation, and the absence of these signals would lead to the plant terminating the infection process, perhaps *via* a defense response (Jones et al. 2008). This hypothesis was further supported by the identification of Epr3, a receptor like kinase in *Lotus japonicus* which directly interacts with EPS. *Epr3* expression was shown to be inducible and dependent on host perception of bacterial nodulation (Nod) factors. A two-stage mechanism involving sequential receptor-mediated recognition of Nod factor and EPS signals was proposed for rhizobial access to the plant roots (Kawaharada et al. 2015). An *Epr3* orthologue *MtLYK10* (LysM-receptor like kinase) is present in *M. truncatula* and appears to play a symbiotic role. MtLYK10 mutant plants had severely reduced infection events, and there were few invaded nodules found on these plants. This suggests an important, but non-essential role of MtLYK10 during symbiosis in *M. truncatula* (Maillet et al. 2020).

1.3 The process of symbiosis

1.3.1 Early stages of symbiosis

The legume-Rhizobium symbiosis begins with the plant secreting signalling molecules including flavonoids to the rhizosphere that can be recognized by specific rhizobia (Oldroyd 2013). Alfalfa releases a variety of flavonoids in the soil but not all of them can induce the *nod* genes involved in Nod-factor production. Quercetin is a one such flavonoid that can enhance the growth of *S. meliloti in vitro* but is unable to induce *nodD* transcription (Hartwig et al. 1991). Luteolin, 3',4',5,7-tetrahydroxyflavone, was the first flavone identified to induce *nodABC* (Peters et al. 1986). In addition to luteolin, apigenin and eriodictyol are other related flavonoids that are also able to induce *nod* genes to relatively low levels of activity when present in low concentrations (Peters and Long 1988).

Flavonoids induce Nod factor production by inducing *nod* genes. Flavonoids are recognized by NodD. *S. meliloti* has three copies of *nodD*, and luteolin is recognized by NodD1. The genes *nodABC* are generic and found throughout all rhizobia, whereas *nodHPQ* and *nodL* are host-specific. The host specific *nod* are responsible for adding certain modifications to the Nod factor structure (Banfalvi and Kondorosi 1989; Fisher and Long 1993).

Once the Nod factors are released into the rhizosphere, they are perceived by specific LysR type receptors present on the plant cells (Amor et al. 2003; Radutoiu et al. 2003). This signal exchange leads to calcium spiking in root hair cells that is subsequently sensed by a calcium-calmodulin sensing kinase known as cCAMK. The activation of cCAMK leads in the division of inner cortical cells and development of nodule primordia (Ehrhardt et al. 1996; Shaw and Long 2003).

One of the earliest visible signs of this signal exchange is root hair curling (Mitra et al. 2004). The result of root hair curling is that the bacteria become entrapped within the curled root hair in an apoplastic compartment termed the curled colonized root hair (CCRH) (Cheng and Walker 1998). The bacteria entrapped in the CCRH continue to produce Nod factors, and symbiotically active exopolysaccharides that are necessary for the subsequent invasion of the plant. Isolates unable to produce succinogylcan, or only produce unsuccinylated succinoglycan, in *S. meliloti* are compromised in developing infection threads (IT) from the CCRH (Maillet et al. 2020). The CCRH has also been defined as an acidic compartment; however, it is unclear whether the acidity is plant derived, a product of bacterial metabolism, or a combination of both the plant and bacteria (Geddes et al. 2014). The bacteria exit the CCRH via an IT that leads towards the root cortex and the developing nodule primordium (Jones et al. 2007a). Once the bacteria enter into the IT, the bacteria at the extending tip are actively dividing while those distal become quiescent. The evidence suggests that the interior of the IT is filled with exopolysaccharides and plant cell material (Gage 2004).

1.3.2 Late stages of symbiosis

The bacteria that exit the IT are endocytosed by the plant cells in the nodule. These bacterial cells are surrounded by a plant derived membrane and the resulting structure is called a symbiosome. The bacterial cells continue to divide until the plant cells are packed with symbiosomes (Udvardi and Poole 2013). The bacteria are then transformed into a nitrogen fixing form termed bacteroids. The transformation into bacteroids is dictated by plant-derived nodule like cystine-rich polypeptides (NCRs) (Van de Velde et al. 2010). In some legumes, the bacteria can undergo multiple rounds endoreplication generating many copies of their genome within each cell. This event happens in legumes belonging to the inverted repeat lacking clade (IRLC),

a group of closely related legume species. The legumes in IRLC produce indeterminate nodules e.g. *M. sativa* and *M. truncatula.*

The chromosome copy number in a bacteroid is dictated by the legume host which has been shown by using recombinant rhizobial strains capable of nodulating legumes with indeterminate and determinate nodules (Mergaert et al. 2006). Bacteroids with multiple copies of the genome are larger in size and the cell membrane surrounding them is more permeable. Bacteroids also lose their ability to divide, and therefore are unable to return to their free-living stage if they are released from the nodules. In contrast, bacteroids in determinate nodules resemble their free-living stage and can return to their free-living state if released from a nodule (Mergaert et al. 2006).

1.4 Root exudates

Plants secrete a significant proportion of their photosynthate in the rhizosphere in the form of root exudates. These exudates contain carbohydrates, amino acids, and organic acids. Carbohydrates contribute to a larger proportion of these root exudates in most plants (Gunina and Kuzyakov 2015). Alfalfa releases 18.5 mg C/g root dry matter of water-soluble exudates and 46% of that is carbohydrates. Carbohydrates in root exudates are typically monosaccharides, disaccharides, oligosaccharides or polysaccharides (Hütsch et al. 2002). Glucose is the dominating sugar, whereas fructose, sucrose and ribose are present in smaller quantities. Ribose, arabinose and xylose are pentoses found in alfalfa root exudates along with hexoses (glucose, fructose and mannose), sugar alcohols (inositol) and a disaccharide maltose (Hamlen et al. 1972). A major portion of these exudates is used by the microorganisms present in the rhizosphere as food.

1.4.1 Rhizosphere and 'the competition problem'

The rhizosphere is termed as a zone in the soil which is directly influenced by root exudates. These exudates mainly consist of amino acids and carbohydrates that nourish the microbial population in the rhizosphere (Panda and Sarkar 2012). Rhizobia must compete with other microorganisms present in the soil to survive and establish a symbiosis with plant roots. A factor that limits the use of bacterial inocula as biofertilizers is their inability to compete with native rhizobia for nodule occupancy, creating what is called the 'rhizobial competition problem' (Triplett 1990). The ability to metabolize different sugars has been hypothesized to play a role in affecting competition for nodule occupancy (Triplett 1990). To address this hypothesis, several studies have focused on carbon catabolism by microorganisms and shown that mutants unable to catabolize sugars are less competitive for nodule occupancy (Jimenez-Zurdo 1995; Oresnik et al. 1998; Ding et al. 2012). In contrast to the premise that an inability to catabolize a compound leads to a decrease in its ability to compete, it has been found that the inability to utilize galactose (*dgoK* mutant strain) results in a strain that is more competitive for nodule occupancy in *S. meliloti*. Increased efficiency for nodule occupancy was linked to increased production of symbiotically active EPS by the mutant strains and its ability to acidify its growth medium (Geddes and Oresnik 2012b). Since the CCRH was shown to be an acidic compartment, acidification or the ability to tolerate lower pH was suggested to be an important determinant for the development of an effective nodule (Geddes and Oresnik 2012b; Geddes et al. 2014).

1.5 Carbon metabolism in *Sinorhizobium meliloti*

Carbon metabolism plays an important role at various stages of symbiosis. Nodule occupancy is affected by the ability of rhizobia to utilize different carbon sources. When rhizobia

leave the IT and transform into bacteroids, carbon sources like succinate and malate are taken up and used in bacteroid metabolism (Glenn et al. 2007; Finan et al. 1981).

The *S. meliloti* Rm1021 genome contains a variety of genes responsible for carbon metabolism and transport. Hexoses are metabolised using the Entner Doudoroff (ED) in *S. meliloti* Rm1021, whereas the Embden Meyerhof Parnas (EMP) has been suggested to play gluconeogenic role due to the lack of the enzyme phosphofructokinase in *S. meliloti* Rm1021 (Figure 1) (Stowers 1985). Complete pentose phosphate (PP) pathway and tricarboxylic acid (TCA) pathways are present in *S. meliloti* Rm1021 (Cerveñanský and Arias 1984; Stowers 1985; Dunn 1998). The PP pathway was not well characterized in *S. meliloti* Rm1021 until recently. It has been suggested to play a role in symbiosis but the exact mechanism is not known (Hawkins et al. 2018).

1.5.1 Pentose phosphate pathway

The PP pathway can be divided in to an oxidative and a non-oxidative branch (Figure 1). The oxidative branch has shared reactions with the ED pathway where glucose-6-phosphate is converted to 6-phosphogluconate (6PG). The enzyme 6PG dehydrogenase is used to convert 6-phosphogluconate to ribulose-5-phosphate, which enters the non-oxidative branch of PP pathway (Martínez-De Drets and Arias 1972). Presence of 6PG dehydrogenase is not ubiquitous among rhizobia, hence the oxidative branch. The oxidative branch generates NADPH by reducing $NADP^+$, which is used by the cell as an energy source. The non-oxidative part of PP pathway cycles pentose sugars and generates other phosphorylated sugars that can serve as precursors of nucleic acid, exopolysaccharide, and aromatic amino acid biosynthesis. Glyceraldehyde 3-P and fructose 6-P are the metabolites of the PP pathway that can be cycled back into central carbon metabolism in a gluconeogenic manner. The non-oxidative branch also serves as an entry point

for carbon sources such as erythritol, ribose and xylose into the central carbon metabolism (Geddes and Oresnik 2014). Transketolase and transaldolase are the two enzymes involved in interconversion of sugars in the non-oxidative branch of PP pathway (Cerveñanský and Arias 1984).

Transketolase is involved in using a ketose donor and aldose acceptor in two reversible reactions in the PP pathway. Five genes (*cbbT, tktA, tktB, SMc00269* & *SMc00270*) are annotated as putative transketolases in the *S. meliloti* Rm1021 genome database (Capela et al. 2001), and *tktA* has been shown to encode for a major transketolase. A plasmid overexpressing *tktB* was able to restore some carbon phenotypes associated with *tktA* mutation, therefore *tktB* was suggested to encode for an enzyme with minor transketolase activity (Hawkins et al. 2018). *E. coli* K-12 genome also contains a major (*tktA*) and a minor transketolase (*tktB*) (Josephson and Fraenkel 1969, 1974). Insertional inactivation of *tktA* had a pleiotropic effect including aromatic amino acid auxotrophy in *S. meliloti* (Hawkins et al. 2018). A similar observation was made in *E. coli*, where strains lacking major transketolase, *tktA,* were unable to grow on pentose sugars due to disruption of the shikimate pathway (Zhao and Winkler 1994). The metabolic pools of the intermediates associated with the PP pathway were altered in the *tktA* mutant strain as compared to the wild type in *S. meliloti* Rm1021. Significant reduction in amounts of sedoheptulose-7-phosphate were also observed in *tktA* mutant strain when grown on glucose in defined media, indicating blockage in the reversible reactions carried out by transketolase enzyme (see Figure 1 Hawkins et al. 2018).

Another putative transketolase *cbbT* is present with a cluster of other *cbb* genes that are involved in the Calvin-Benson-Bassham (CBB) cycle. The CBB cycle is necessary during

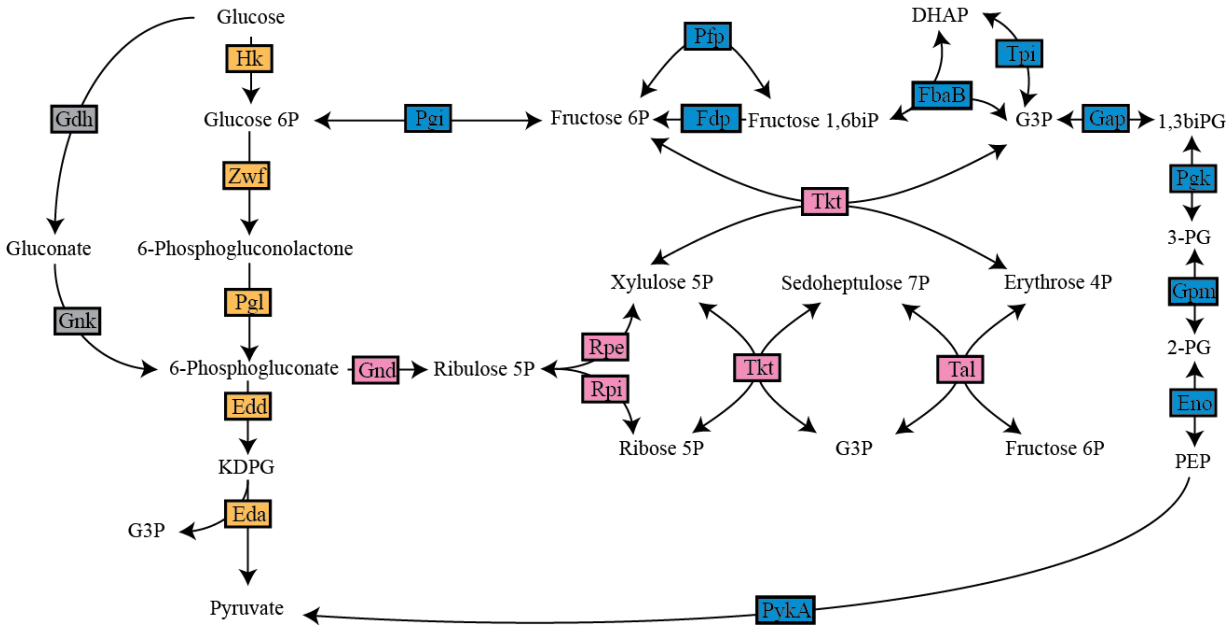


Figure 1. Network of central carbon metabolism pathways in *S. meliloti*.

Yellow, ED pathway; blue, EMP pathway; pink, PP pathway; and grey, auxiliary pathway. TCA cycle is not shown in this figure. Metabolites, DHAP, dihydroxyacetone phosphate; G3P, glyceraldehyde-3-phosphate; 1,3biPG, 1,3-biphosphoglycerate; 2-PG, 2-hosphoglycerate; 3-PG, 3-phosphoglycerate; PEP, phosphoenolpyruvate; KDPG, 2-dehydro-3-deoxy-D-gluconate 6-phosphate. Enzymes, Hk, hexokinase; Zwf, glucose-6-phosphate dehydrogenase; Pgl, 6-Phosphogluconolactonase; Edd, 6-Phosphogluconate dehydratase; Eda, 2-Keto-3-deoxy-6-phosphogluconate aldolase; Gnd, 6-Phosphogluconate dehydrogenase; Rpe, Ribulose-5-phosphate epimerase; Rpi, Ribose-5-phosphate isomerase; Tkt, transketolase; Tal, transaldolase; Pgi, Phosphoglucose isomerase; Fdp, Fructose bisphosphatase; Pfp, pyrophosphate dependent phosphofructokinase; FbaB, Fructose-bisphosphate aldolase; Tpi, Triose-phosphate isomerase; Gap, Glyceraldehyde-3-phosphate dehydrogenase; Pgk, Phosphoglycerate kinase; Gpm, Phosphoglycerate mutase; Eno, Enolase; PykA, Pyruvate kinase; Gdh, Glutamate dehydrogenase; Gnk, N-acetyl-D-glucosamine kinase

formate/bicarbonate dependent auxotrophic growth (Pickering and Oresnik 2008). Of note, a plasmid overexpressing *cbbT* was unable to rescue growth of *tktA* mutation on defined media, therefore was not considered to play a role in the PP pathway (Hawkins et al. 2018). The other two putative transketolases, *SMc00269* and *SMc00270,* are annotated as beta and alpha subunits of transketolase, respectively. A literature search revealed no published work on these two genes in *S. meliloti. Salmonella enterica* serovar Typhimurium contains three transketolases TktA, TktB and TktC. Interestingly, TktC is encoded by a set of two genes *STM14_2885* and *STM14_2886.* All three transketolases are reported to contribute for transketolase activity in PP pathway but TktA accounts for 88% of the total activity. The *S. enterica* serovar Typhimurium strain lacking all transketolases was avirulent (Shaw et al. 2018).

1.5.2 Shikimate pathway

Erythrose-4-phosphate is used in the synthesis of aromatic amino acids by making the aromatic compound shikimate. Erythritol has been shown to be converted to erythrose-4-phosphate which is a metabolite associated with the PP pathway (Figure *1*) (Barbier et al. 2014). Shikimate is converted to chorismate and then to prephenate. Prephenate, which serves as precursor for biosynthesis of anthranilate, 1-hydroxy-phenyl-pyruvate, and the aromatic amino acid phenylalanine (Barsomian et al. 1992). The other two aromatic amino acids, tryptophan, and tyrosine are made from anthranilate and 1-hydroxy-phenyl-pyruvate, respectively.

Previously, the inability to synthesize anthranilate in *S. meliloti* has been shown to lead to a defective symbiotic phenotype (Barsomian et al. 1992). The inability to synthesize anthranilate in *S. meliloti* due to a mutation in the first step of tryptophan biosynthesis (*trpE/G*) results in reduced symbiotic efficiency. However, mutants blocking the downstream steps of a *trpE(G)* mutation were effective in establishing symbiosis. The authors concluded that anthranilate

biosynthesis, not tryptophan biosynthesis is required for symbiosis (Barsomian et al. 1992). Anthranilate has been suggested to play a role in bacteroid maturation and nitrogen fixation in *S. meliloti* (Prasad et al. 2000). Histology of alfalfa nodules inoculated with *trpE(G)* mutant strains revealed the lack of a distinct nitrogen fixing zone in the nodules. Fewer bacteroids were observed in nodules that were formed with *trpE(G)* mutants as compared to wild type in a microscopic analysis. Additionally, transmission electron microscopy revealed the lack of condensed nuclear material in bacteroids, a typical trait of bacteroids, in *trpE(G)* mutants of *S. meliloti* (Prasad et al. 2000).

Apart from their direct role in symbiosis, aromatic compounds like protocatechuate have been shown to serve as carbon source for members of the family *Rhizobiaceae.* Aromatic compounds produced from the breakdown of plants cell materials such as lignin may serve as a carbon source for rhizobia present in the close vicinity of plant roots (Parke and Ornston 1984). The presence of aromatic catabolism genes in rhizobia might provide them an advantage over other soil bacteria to compete for their own survival in the rhizosphere and for nodule occupancy (Parke and Ornston 1984).

1.5.3 Purine and Pyrimidine biosynthesis

Purines and pyrimidines are some of the most important biomolecules in cellular metabolism. These biomolecules act as subunits to nucleic acids, part of the energy carrier molecules such as ATP and GTP, and precursors for NAD synthesis. The biosynthesis of purines and pyrimidines is directly connected to central carbon metabolism as the precursor for their biosynthesis is ribose-5P which is supplied through the PP pathway. Ribose-5P is converted into 5-phosphoribosylpyrophosphate (PRPP); the latter serves as an entry point for the purine and pyrimidine *de novo* biosynthesis. During purine biosynthesis, inosine-5’-monophosphate (IMP)

is synthesized which is later converted into either adenosine-5'monophosphate (AMP) or guanosine-5'-monophosphate (GMP). The conversion of PRPP to IMP usually requires 10 steps but the reactions can vary depending on the organism (Zhang et al. 2008). The key intermediate in pyrimidine biosynthesis is orotidine-5'-monophosphate (OMP), therefore pyrimidine biosynthetic pathway is also referred to as the "orotate pathway". The orotate pathway consists of six reactions that are usually common in plants, animals and microorganisms (Moffatt and Ashihara 2002).

Purine and pyrimidine auxotrophy causes symbiotic defects in *R. etli, R. leguminosarum* and *S. meliloti* (Pain 1979; Pankhurst and Schwinghamer 1974; Scherrer and Dénarié 1971). In *S. meliloti,* purine and pyrimidine mutant strains were able to grow on 4-amino-5-imidazole-carboxamide (AICAR), an intermediate of the purine biosynthetic pathway, but the nature of the mutations was not identified. However, growth on AICAR does indicate that the mutation should have caused a block before AICAR synthesis in the purine and pyrimidine biosynthesis pathways (Scherrer and Dénarié 1971). In another study, *S. meliloti* purine auxotrophs have been reported to form infection threads at a rate similar to the wild type, but the nodules that were formed were reported to be an irregular shape. Supplementation with AICAR did not restore the symbiotic phenotypes in these auxotrophs (Swamynathan and Singh 1992). Purine auxotrophs also have a pleiotropic effect on production of surface polysaccharides which might play a role in causing the symbiotic defects. It has been suggested that the changes in cell surface components affect the strains' abilities to interact with the host cells. The suggested rationale for the effect of purine auxotrophy on surface polysaccharides production is multifunctionality of purine biosynthetic pathway enzymes (Swamynathan and Singh 1995). As an example of the multifunctionality of these enzymes, the lack of AMP pyrophosphate phosphoribosyl transferase (APRT) in the purine

biosynthetic pathway has been linked to male sterility in *Arabidopsis* (Moffatt and Somerville 1988).

Purine and pyrimidine metabolic mutants of *R. etli* display a delay in nodulation. The nodules are small, white in colour, and lacked nitrogenase activity (Pain 1979). The formation of a nodule rather than a pseudonodule in the *R. etli* purine auxotroph was achieved by adding 4-aminoimidazole-5-carboxamide (AICA) riboside at early stages (day 2-6 from initial inoculation). Addition of AICA riboside later than day six had no effect on symbiosis (Newman et al. 1992). AICA riboside is a precursor of AICAR, and AICAR has been suggested to play a role at early stages of symbiosis, most likely in the development of infection threads (Newman et al. 1992). In later studies, AICAR and AICA riboside were shown to be negative effectors of the *fixNOQP* operon which encodes of cytochrome terminal oxidase cbb_3 in *R. etli* and *S. meliloti*, respectively (Soberón et al. 1997, 2001).

A *R. leguminosarum* strain that showed auxotrophy for adenine and thiamine was unable to establish symbiosis with pea plants. The symbiotic phenotype was restored by supplementing with adenine and/or thiamine. However, the effect of adenine deficiency was more severe on symbiosis than thiamine. In both cases, the rate of nitrogen fixation, measured by acetylene reduction, was significantly lower than the wild type strain. The auxotrophic strain also showed a block in early stages of symbiosis which could be partially overcome by continuous supply of adenine for the first six days from the day of inoculation (Pankhurst and Schwinghamer 1974).

Overall, purine or pyrimidine auxotrophy seem to affect strains forming both determinate and indeterminate nodules. The effect of purine auxotrophy on symbiosis is similar among these

rhizobia, as nodules formed on plants inoculated with these strains were small white bumps and were ineffective in nitrogen fixation.

1.6 Effect of carbon flow on symbiosis

Central carbon metabolism is a network of many metabolic pathways with an overall goal of shuffling intermediates for various biosynthetic functions as well as producing energy for the cell. Mutations in genes encoding enzymes at some key nodes in central carbon metabolism can have pleiotropic effects on these pathways as well as other anabolic pathway that depend on central carbon metabolism for substrates such as amino acid biosynthesis and purine and pyrimidine biosynthesis etc.

There are mainly three types of mutations affecting central carbon metabolism in relation to their effect on symbiosis. The first class of mutations does not have an effect on symbiosis. Although these mutations might lead to severe growth defect in growth medium, symbiosis is not affected. These mutations represent reactions that might produce metabolites that are either not necessary for symbiosis or can be produced by alternative ways. For example, a *tal* mutant strain (unable to produce transaldolase) of *S. meliloti* shows no significant decrease in plant dry weight when used as an inoculum on alfalfa plants, but it does show compromised growth using ribose as a sole carbon source (Hawkins et al. 2018). The *dct* genes encode determinants that are necessary for dicarboxylate transport in rhizobia. The *dct* locus in *S. meliloti* consists of three genes, *dctA, dctB,* and *dctD.* The genes *dctB* and *dctD* encode for a two-component system where *dctB* encodes a periplasmic protein and *dctD* encodes for a regulatory protein that interacts with σ^{54}. The transporter is encoded by *dctA*. Whereas all the components are necessary for the transport of succinate in free living bacteria, and it is generally agreed that dicarboxlyate

transport is essential for symbiosis, the strains with mutations in *dctD* have shown that dctD is not essential for symbiotic nitrogen fixation (Jiang et al. 1989; Watson 1990).

The second class of mutations in central carbon metabolism abolish symbiosis. The inability to establish symbiosis might result from the lack of a specific metabolite, or accumulation of phosphorylated intermediates that can be toxic for cell growth. *S. meliloti* strain with a mutation in *eno* does not produce enolase enzyme required during gluconeogenesis (Finan et al. 1991). The strain also shows less than 5% acetylene reduction (an indicator of nitrogen fixation) as compared to the wild type indicating severe symbiotic defects (Finan et al. 1991). Similarly, *S. meliloti* strains carrying mutations in *dctA*, which encodes for a C_4-dicarboxylate transport protein are unable to grow on C_4 substrates and display severe symbiotic defects (Engelke et al. 1989).

The third class of mutations results in a reduced symbiotic efficiency. Mutation of *dctB* in *S. meliloti* displays such a phenotype. Plants inoculated with *dctB* mutant strains show reduced symbiotic nitrogen fixation, whereas *dctD* mutant strains do not have any effect on nitrogen fixation. Other enzymes involved in succinate metabolism also show similar symbiotic phenotypes (Yurgel and Kahn 2004). For example, *eno* (encoding for enolase) mutant strains of *S. meliloti* form ineffective nodules on alfalfa, whereas *pck* (encoding for phosphoenolpyruvate carboxylase) mutant strains display a 50% reduction in SNF compared to the wild type (Finan et al. 1991).

Reduced nitrogen fixation shown by mutant strains in central carbon metabolism might arise from metabolic redundancy where in the absence of one enzyme a second enzyme can perform the same function with a reduced efficiency, or the amount of the enzyme produced is

not enough. To test the hypothesis that reduced carbon flow affects symbiosis, mutations in the central carbon metabolism with metabolic redundancy would be ideal candidates to study. These enzymes can be used as tools to study the effect of central carbon metabolism on symbiosis. Transketolase, phosphofructokinase, phosphoglucose isomerase and pyruvate carboxylase are some examples of such key nodes. These examples are reviewed in the next section.

1.6.1 Transketolase

A mutation in *tktA* has a pleiotropic effect on central carbon metabolism and symbiosis in *S meliloti* strains. The strains are unable to grow on ribose, xylose and erythritol. The *tktA* mutant strain produces less EPS and is severely affected symbiotically with alfalfa (Hawkins et al. 2018). The *tktA* mutation also affects the production of the PP pathway metabolites such as ribose -5P that may affect purine and pyrimidine biosynthesis. As discussed earlier in the review, once *S. meliloti* cells are transformed into bacteroids and they reach the symbiosome stage, they go through multiple rounds of endoreplication, reaching a chromosome copy number of up to 24 (Mergaert et al. 2006). The bacteria might be required to have an enhanced supply of purine and pyrimidines to go through this endoreplication process. The disruption of the PP pathway might affect chromosome copy number due to reduced carbon flow. To date there is no evidence to support this hypothesis. The carbon phenotypes associated with a *tktA* mutation in *S. meliloti* can be complemented by the introduction of the gene *tktB* on a plasmid (Hawkins et al. 2018). Interestingly the introduction of a plasmid containing *cbbT* did not complement the *tktA* mutation despite CbbT (60%) having higher sequence similarity to TktA than TktB (46%).

1.6.2 Phosphofructokinase

Glucose is catabolized through the ED pathway in *S. meliloti*. Earlier studies showed the production of pyruvate from 6-phosphogluconate in *S. meliloti* (Katznelson 1955). ^{13}C labelling

experiments done in *Bradyrhizobium japonicum* also supported the hypothesis that glucose is catabolized through the ED pathway in rhizobia (Keele et al. 1969). Later, the presence of the key enzymes with activities indicative of the ED pathway established it as the primary glycolytic pathway in *S. meliloti* and rhizobia in general (Martínez-De Drets and Arias 1972).

The EMP pathway has been thought to only play a gluconeogenic role in *S. meliloti* due to the lack of the key catabolic enzyme phosphofructokinase (Pfk, EC 2.7.1.11). Phosphofructokinase performs the irreversible reaction of adding a phosphate group in an ATP-dependent manner, converting fructose-6-phosphate (F6P) tofructose-1,6-bisphosphate (F1,6biP). Lack of phosphofructokinase activity was recorded in cell free extracts of *S. meliloti* L5-30 (Arias et al. 1979) and *S. meliloti* 102F51 (Irigoyen et al. 1990).

The phosphorylation of F6P can also be carried out by a pyrophosphate (PPi) dependent phosphofructokinase (EC 2.7.1.90) that uses inorganic PPi as a phosphate donor instead of ATP. The first PPi-dependent phosphofructokinase was found in *Entamoeba histolytica* (Reeves et al. 1974). These are also found in higher plants, animals, and some bacteria. PPi-dependent phosphofructokinase is considered the ancestral version of the ATP-dependent phosphofructokinase because PPi is believed to be an earlier source of energy than ATP (Ding et al. 2000; Baltscheffsky Herrick (Editor) 1996). PPi-dependent phosphofructokinase has been shown to be involved in the PP pathway in cellulolytic Clostridia lacking a transaldolase. Sedoheptulose-7P is phosphorylated to Sedoheptulose-1,7biP by PPi-dependent phosphofructokinase, which is later broken down to dihydroxy-acetonephosphate (DHAP) and erythrose-4P by a fructose bisphosphate aldolase (Koendjbiharie et al. 2020). Although no ATP dependent phosphofructokinase activity has been recorded in *S. meliloti*, fructose 1,6 bisphosphate aldolase activity is present (Irigoyen et al. 1990). *S. meliloti* does contain a

probable PPi- fructose 6-phosphate 1-phosphotransferase, encoded by *pfp* (*SMc01852*) (EC: 2.7.1.90). The non-oxidative PP pathway in *S. meliloti* contains only transketolase and transaldolase, but if there is a functional phosphofructokinase present, it can be thought of as an extended part of the PP pathway. Literature searches have revealed no reports of experimental data on *pfp* in *S. meliloti*. Weak phosphofructokinase (ATP dependent) activity was reported in *Rhizobium sp.* 32H1 but the activity was lower than *E. coli* PFK which served as a positive control. The low activity was only recorded when the strains were grown on glucose as a carbon source and not on succinate (Stowers 1985; Stowers and Elkan 1983).

1.6.3 Phosphoglucose isomerase

Glucose-6P is converted into fructose-6P by phosphoglucose isomerase (Pgi) in *S. meliloti*. The *pgi* mutant strain was first isolated by nitrosoguanidine mutagenesis. This strain was unable to grow on many carbon sources such as fructose, ribose, xylose, mannitol, sorbitol, or mannose, but it was able to grow on gluconate, and L-arabinose (Arias et al. 1979). The inability to grow on the listed carbon sources was explained by the fact that many of these sugars generate F6P that serves as an entry point for these sugars into central carbon metabolism (Geddes and Oresnik 2014). The strain was also severely affected in EPS production and nitrogen fixation. Reversion analysis of the *pgi* mutation demonstrated that revertants could be isolated that were able to restore Pgi activity, resulting in restoration of carbon and symbiotic phenotypes (Arias et al. 1979). Of note, the work also pointed out that a strain containing a *pgi* mutation was able to fix nitrogen, albeit at a greatly reduced rate that was only detectable 60 days after inoculation. Due to the positioning of Pgi in central carbon metabolism and the pleiotropic effect of the *pgi* mutation, and that it did have partial nitrogen fixation capabilities, it becomes a potential candidate to test the hypothesis that carbon flow affects symbiosis.

More recently a *pgi* knockout strain of *E. coli* was subjected to adaptive laboratory evolution (ALE) to understand the changes in metabolic flux and the rewiring of metabolism necessary to overcome growth defects associated with a *pgi* deletion. The ten strains isolated as a part of ALE showed mutations in transhydrogenase genes (*sthA* and *pntAB*), which help recover the NADPH and NADH production. Second site mutations in a phosphotransferase gene (*crr*) were also detected, and strains carrying this mutation showed an increased flux from pyruvate to PEP. The rewiring of metabolism did not occur by increased flux through the ED pathway or the glyoxylate shunt as one might expect, but through cofactor metabolism and global transcriptional regulation (Long et al. 2018).

1.6.4 Phosphoenolpyruvte carboxylase

Phosphoenolpyruvate carboxylase (Pck) catalyses the conversion of oxaloacetate to phosphoenolpyruvate. It plays a gluconeogenic role by converting a TCA cycle intermediate into something that can be used to make hexoses (Osterås et al. 1995). Rather than showing no growth, the strain with a *pck* mutation in *S. meliloti* grew slowly on gluconeogenic substrates such as succinate and malate. This indicated another pathway to compensate for the *pck* mutation and make phosphoenolpyruvate from oxaloacetate. This was shown by isolating suppressors to the *pck* mutation that can restore wildtype growth. The suppressors were mapped to the *pod* locus that encodes for pyruvate orthophosphate dikinase (PPDK). PPDK catalyses the conversion of pyruvate into phosphoenolpyruvate in an ATP dependent manner. The alternate route to convert oxaloacetate to phosphoenolpyruvate exists by converting oxaloacetate to malate by malate dehydrogenase, and malate to pyruvate by a NAD^+-dependent malic enzyme (DME) (Osterås et al. 1997).

S. meliloti strains lacking Pck and mutations that affect the lower EMP pathway (in glyceraldehyde 3P dehydrogenase, *gap*; enolase, *eno*; 3-phosphoglycerate kinase, *pgk*) display symbiotic deficiencies. Strains carrying mutations in genes involved in the lower part of the EMP pathway show a severe reduction in nitrogen fixation, whereas strains mutated in *pck* exhibit 60% the level of nitrogen fixation compared to wildtype (Finan et al. 1991). One possible explanation for the partial reduction of nitrogen fixation is DME is still functional and helps to bypass the metabolic block created by a mutation in *pck* in *S. meliloti.* The *S. meliloti* strain with mutation in *dme* display Fix^- phenotypes. This indicates that DME is essential for symbiosis (Driscoll and Finan 1993).

During symbiosis, bacteroids are dependent on the plant for carbon sources and in return provide fixed nitrogen. The carbon is supplied to the bacteroids in the form of C_4 dicarboxylates. These dicarboxylates are assimilated via the TCA cycle and gluconeogenesis. Genes that encode TCA cycle enzymes were also found to be upregulated in pea bacteroids in *R. leguminosarium* (Green et al. 2019). Strains carrying a mutation in *pck* show slower growth on C_4 substrates. Taken together, this data could indicate a reduction in carbon flow from C_4 substrates. Since carbon is supplied from plants to bacteroids as C_4, this reduction of carbon flow could translate into a reduction of symbiotic efficiency.

1.6.5 Pyruvate carboxylase

Pyruvate carboxylase (Pyc) is one of the enzymes used as an anaplerotic enzyme in *S. meliloti*. Pyruvate carboxylase is encoded by *pyc* present on the chromosome in *S. meliloti* (Dunn et al. 2001). Anaplerotic pathways are used to supply substrates of the TCA cycle that are used up during amino acid biosynthesis by bypassing TCA cycle reactions. The glyoxalate shunt is a

good example of an anaplerotic pathway that converts isocitrate into succinate and glyoxalate, and glyoxalate is converted into malate (Geddes and Oresnik 2014).

Like Pgi, Pyc also has a pleiotropic phenotype. A Δ*pyc* strain of *S. meliloti* is unable to grow on glycolytic carbon substrates like glucose, glycerol, erythritol, adonitol, L-arabitol and D-arabitol. The strain does grow on gluconeogenic substrates like succinate and arabinose and sugars that enter through PP pathway i.e., ribose and xylose. Interestingly, a Δ*pyc* mutation in combination with Δ*tpiA* Δ*tpiB* mutations, which should block both glycolysis and gluconeogenesis, still grows on ribose and xylose in *S. meliloti* (Geddes and Oresnik 2012a). This suggests that ribose and xylose catabolism can be carried out by an alternative route other than the known ED and EMP pathways. This alternative route is still unknown in *S. meliloti* (Geddes and Oresnik 2014).

1.7 Metabolic modelling

The development of sequencing technologies has led to the genomes of more and more organisms being sequenced (Heard et al. 2010; Rattray et al. 2018).The ‘omic’ techniques such as genomics, metagenomics, metabolomics and proteomics and other genome scale techniques such as transposon sequencing are providing large scale physiological knowledge about organisms. All this new information, along with the already available literature, is being used to create a common platform to understand the metabolic properties of a cell.

Metabolic modelling involves creating an *in silico* network of metabolic reactions under a defined set of conditions using biochemical, genetic and genomic (BiGG) knowledge of a given organism (Reed et al. 2006). This BiGG data can be used in a bottom-up approach to create mathematical models recreating cellular metabolism. This involves recreating all metabolic reactions in a cell, designating enzymes to catalyse those reactions, and the genes that encode the

enzymes (Thiele and Palsson 2010). These models are optimized by manual evaluation and experimental data. Metabolic models can run on a flux-based analysis where the model can operate under a given set of growth conditions and predict the reactions required to grow on a specific carbon source (diCenzo and Finan 2017). The accuracy of these predictions depends on the available data about that organism. The model can predict four types of outcomes: true positive, true negative, false positive and false negative. True positives and true negatives are generated when the model prediction and physiological data are in agreement. Having a higher percentage of true positives and true negatives represent the better quality of a model. This is very useful when models are used to maximize the production of a certain metabolite (Kay and Jewett 2020; Swayambhu et al. 2020). For example, a synthetic pathway to degrade 1,2,3-trichloropropane (TCP) was introduced in *E. coli*. The pathway was designed to use a metabolic model to maximize the enzyme efficiency and production of glycerol, while keeping toxic metabolite accumulation to a minimum (Kurumbang et al. 2014). This will only happen when the model is performing at a higher accuracy rate. False negatives represent a gap in knowledge with respect to a reaction in the model, but the reaction does take place *in vitro*. False negatives can be used in generating hypothesis and designing targeted experiments. Similarly, false positives occur when the reaction is carried out *in silico,* but no cell growth is recorded during the actual experiment. This means that the reaction was falsely included in the model, or there is no physiological evidence in the literature supporting its inclusion. However, it can also indicate metabolic redundancy where physiological knowledge is yet unavailable. (Reed et al. 2006; O'Brien et al. 2015). Second site suppressors or adaptive laboratory evolution can be used to test whether an alternative metabolic route exists (Fong et al. 2013). Both false positives and false negatives provide an opportunity to expand the knowledge base about the organism.

A range of metabolic models exist from a single cell models of prokaryotes to multi-tissue models in eukaryotes such as plants and human (Edwards and Palsson 2000; Robinson et al. 2020). Apart from metabolism, models for transcription, translation and regulation have been developed. A kinetic model of *E. coli* provides fluxomic data which gives a quantitative insight into metabolism (Khodayari et al. 2014). In contrast, stoichiometric models offer qualitative data which shows the presence or absence of a reaction but not the rate at which the reaction is carried out. These models are being using in synthetic biology approaches to engineer microorganisms to produce desired metabolites. This approach requires well studied chassis microorganisms such as *E. coli* to explore the full potential of metabolic modeling and synthetic biology.

1.7.1 Sinorhizobium meliloti metabolic model

S. meliloti is a model organism in the field of symbiotic nitrogen fixation and molecular genetics with a sizable volume of literature available about its genome and physiology. As a well studied microorganism, tools are easily and readily available for genetic manipulations of *S. meliloti*. Due to its unique position to establish symbiosis with its host legume and live freely in the soil, *S. meliloti* has a large potential to be developed as a chassis to study synthetic biology and metabolic modeling (diCenzo and Finan 2017).

The first metabolic model of *S. meliloti, i*HZ565, was published in 2012 and represented the bacteroid stage (Zhao et al. 2012). It is a relatively small network that consists of 503 metabolic reactions, 565 genes and 522 metabolites with a major portion of reactions focusing on amino acid metabolism, and co-factor and vitamin transport. In this model, the ‘objective function’ is symbiotic nitrogen fixation because *S. meliloti* cells stop dividing at bacteroid stage and there is no biomass production. The model finds central carbon metabolism pathways like the TCA cycle, PP pathway and glyoxylate shunt functioning in the bacteroids. In contrast, *R.*

etli model does not have a functioning PP pathway (Resendis-Antonio et al. 2007). It should be noted here that *R. etli* forms determinate nodules with its host plants and does not go through endoreplication. The model also predicts functioning gluconeogensis, which is important in utilizing C_4 dicarboxylic acid provided by the plant. In *S .meliloti,* deleting the *dme* gene encoding for NAD-malic enzyme from the model results in no symbiotic nitrogen fixation which is also in agreement with the experimental evidence (Driscoll and Finan 1993). These observations can be used to test the hypothesis that carbon flow affects symbiotic nitrogen fixation.

A genome scale metabolic model (*i*GD1575) representing the free living state and symbiotic state of *S. meliloti* has also been published (diCenzo *et al.* 2016). The goal of the model is to describe the importance of secondary replicons in adapting to varying environments. It has been established that *S. meliloti* has a complex life cycle. Therefore, the metabolic model takes this into account and works under three niches, i.e., the bulk soil, the rhizosphere, and the nodule. The *S. meliloti* genome contains genes that can help it adapt under different environments which is important for survival and competition for nodule occupancy. The model contains the data for carbon metabolism with growth of *S. meliloti* on 190 carbon sources and 1500 genes. The 'objective function' of this model is biomass production. It can accurately predict the growth rate and biomass production on glucose and succinate as well as ability to grow on various carbon and nitrogen sources with 85% accuracy (diCenzo *et al.* 2016). A metabolic model that functions under different environments has led to identifying genes that are necessary for environment specific fitness and has assigned function to a number of orphan genes. From an evolutionary perspective, the model can help understand multipartite genome and

the acquisition of the secondary replicons by predicting their role under different environments (diCenzo *et al.* 2016b).

The availability of these *S. meliloti* metabolic models makes it possible to ask questions about the effect of carbon flow on symbiosis. *In silico* deletions of key nodes in central carbon metabolism like *pgi*, *pck,* and *tktA* should predict the changes in flux caused by deleting these genes and possibly other effects that are still unknown. Those predictions can then be used to design targeted experiments.

Although the available models of *S. meliloti* address metabolic reactions in a free-living stage and bacteroid stage, the host metabolism is not considered in these models. Recently, a model representing the symbiosis between *S. meliloti* and *M. truncatula* was published (diCenzo et al. 2020). This model explores the interplay between rhizobia and its host in five different nodule environments representing spatially distinct development zones in a Virtual Nodule Environment (ViNE). The model comprises of the metabolism of both the bacteria and the plant. The availability of such model opens the opportunities to ask questions about central carbon metabolism and the role of carbon flux in symbiosis (diCenzo et al. 2020).

1.8 Thesis goals and hypothesis

In previous sections, different central carbon metabolic pathways and specific genes encoding key enzymes for these pathways were discussed for their role in metabolism and symbiosis. The overall hypothesis for this thesis was that carbon flow directly affects symbiosis in *S. meliloti.* To test this hypothesis, we utilized *S. meliloti* strains with mutations in *tktA, tktB* and, *pfp*.

A *S. meliloti* Rm1021 strain carrying a mutation in *tktA* forms an ineffective symbiotic association with alfalfa (Hawkins et al. 2018). It results in visually small, chlorotic plants when

grown under nitrogen deficient conditions. The nodules produced upon inoculation are white and cyst like structures instead of pink and cylindrical nodules as produced by the wild type Rm1021 strain (Hawkins et al. 2018). However, pink, and relatively cylindrical nodules were observed occasionally upon the inoculation with *tktA* mutant strain. Rhizobia that were isolated from these nodules still appeared to contain the Tn*5* insertion in the *tktA* gene. Such independent mutants were isolated, and these strains were used to re-inoculate alfalfa plants. They resulted in the plants that appeared phenotypically effective, indicating the symbiotic phenotypes associated with mutation of *tktA* can be suppressed by second site suppressor mutations. The suppressor strains can revert the conditional auxotrophic phenotype as well as restore the production of EPS to wild type levels. Therefore, the first goal of this thesis was to understand the role of transketolase in symbiosis by identification and characterization of second site suppressor mutations.

The second goal was to characterize the locus associated with *tktB* and the effect of a negative regulator *SMc02340* on *tktB* and other neighbouring genes. Deletion of *SMc02340* has led to the upregulation of *tktB* which plays a role in reinstating the PP pathway and restoring nodulation occupancy. This goal is explained in the third chapter of the thesis.

The third objective was to characterize *pfp* and identify its role in central carbon metabolism and symbiosis. *S. meliloti* does not contain an ATP dependent phosphofructokinase activity but the presence of *pfp* in the genome indicated a pyrophosphate dependent phosphofructokinase. A phylogenetic analysis has revealed that this *pfp* is also present in other rhizobia. Prior to this, no phosphofructokinase activity has been considered while explaining central carbon metabolism in *S. meliloti*, this chapter explains how we look at central carbon

metabolism or more specifically the EMP pathway by adding a pyrophosphate dependent phosphofructokinase in the scheme.

Chapter 2:

Sinorhizobium meliloti strains carrying mutations in transketolase are unable to form curled colonized root hairs

Authors: Sabhjeet Kaur, Justin P. Hawkins, and Ivan J. Oresnik. Submitted to Molecular Plant Microbe Interactions journal. This work was planned and executed by Sabhjeet Kaur. JH isolated the original suppressors and provided guidance early during the study.

2.1 Introduction

Nitrogen is often a limiting nutrient for agricultural crops (Vitousek et al. 2002). Legumes, which are grown agriculturally for human consumption or for forage, can enter into a symbiotic association with nitrogen fixing rhizobia (Poole et al. 2018). In an effort to realize the full potential of nitrogen fixation, it is important to understand the underlying genetic and biochemical mechanisms involved. This is often accomplished by having work focus on model systems. The interaction of *Sinorhizobium meliloti* with alfalfa is a model system studied for symbiosis (Geddes and Oresnik 2016).

To establish a successful symbiosis, a series of complex signal exchanges must take place (Jones et al. 2007b). The rhizobia respond to plant derived secreted flavonoids by activating genes that encode enzymes to produce Nod factors (Peters et al. 1986). Nod factor is recognized by a specific plant LysR-type receptor. This initiates a signal cascade in the plant root, which results in responses necessary for symbiosis (Oldroyd 2013). In addition to the production of Nod factor, exopolysaccharides (EPS) have also been shown to be essential for successful symbiotic establishment in *S. meliloti* Rm1021 (Jones et al. 2007b).

Root hair curling is one of the early events occurring during symbiosis. Rhizobia attached to plant root hairs during root hair curling can become trapped to form what is called the curled colonized root hair (CCRH) (Cheng and Walker 1998). Bacteria in the CCRH enter the root hair in a plant derived compartment termed the infection thread (IT). The IT penetrates towards the dividing inner cortical cells of the plant root. Upon reaching the target cells, rhizobia are endocytosed, surrounded by a symbiotic membrane, and become terminally differentiated (Van de Velde et al. 2010).

S. meliloti strain Rm1021 produces the exopolysaccharide EPS I (succinoglycan) under most conditions. The production of EPS I has been shown to be critical to symbiotic establishment, and it has been suggested that its production plays a role in either stress tolerance or symbiotic signaling (Leigh and Lee 1988; Jones et al. 2007b; Hawkins et al. 2018). Evidence that is consistent with the monomer, dimer or trimer fractions playing a critical role in early symbiotic interactions has been shown for LMW EPS I (Wang et al. 1999; González et al. 1998). As a result, it was generally believed that LMW EPS I plays an important role in bacterial infection of the plant tissues (Cheng and Walker 1998). More recently it has been shown that the succinylation of EPS I is important symbiotically and not whether it is LMW or HMW (Mendis et al. 2016).

Carbon metabolism plays important roles at various stages of symbiosis. Carbohydrates are exuded from roots and developing nodules and can induce genes within the rhizosphere (Ramachandran et al. 2011; Pini et al. 2017). More targeted approaches have shown that the inability to grow on some carbon sources can affect the ability of a rhizobial strain to compete with their isogenic wild type for nodule occupancy (Jimenez-Zurdo 1995; Oresnik et al. 1998; Ding et al. 2012; Geddes and Oresnik 2012b; Rosenblueth et al. 1998; Fry et al. 2001; Yost et al. 2006; Vanderlinde et al. 2014). However, the precise role each of these plays in the overall bacterial physiology within the rhizosphere and why it affects competition for nodule occupancy is not currently understood.

The majority of the genes that encode the enzymes necessary for the Entner Doudoroff (ED) pathway, the Embden Meyerhof Parnas (EMP) pathway, the pentose phosphate (PP) pathway, and the tricarboxylic acid (TCA) cycle have been identified (Geddes and Oresnik 2016). In addition, the role of gluconeogenesis in succinate metabolism and the use of

dicarboxylates as a carbon source to fuel nitrogen fixation is well understood and has been modeled (Finan et al. 1988, 1991; Udvardi and Poole 2013). The role or necessity of many of the reactions associated with central carbon metabolism are not known with respect to what occurs to the rhizobia during nodule development.

Mutations in the gene encoding transketolase (*tktA*), which is part of the non-oxidative PP pathway, were first isolated by screening for transposon mutants that were unable to acidify their growth medium (Hawkins et al. 2018). Initial characterization of the non-oxidative branch of the PP pathway showed that whereas mutations in the gene encoding transaldolase (*tal*) did not affect symbiosis, mutations in *tktA* were detrimental to symbiosis (Hawkins et al. 2018). Transketolase is necessary for the production of erythrose-4 phosphate which is also a part of the shikimate pathway. In *S. meliloti*, strains with mutations in the shikimate pathway display symbiotic defects on alfalfa. Specifically, inability to make anthranilate (mutations in *trpE*(*G*)) has been reported to play an important role in the *S. meliloti*-alfalfa symbiosis (Barsomian et al. 1992; Prasad et al. 2000).

This work describes the isolation, identification, and characterization of second site suppressor mutations that can suppress phenotypes associated with an *S. meliloti tktA* mutation. By understanding the nature of the suppression, we wish to gain insight into the symbiotic block associated with *tktA*.

2.2 Materials and Methods

2.2.1 Bacterial strains, plasmids, and media.

The bacterial strains and plasmids used in the study as listed in Table 1. *E. coli* strains were grown using Luria-Bertani (LB) as a complex medium at 37ºC (Davis et al. 1980). Antibiotics were added as necessary (see below). *S. meliloti* strains were routinely grown at 30ºC using

either LB (Davis et al. 1980), or yeast extract mannitol (YEM) (Vincent 1970) as complex media, and Vincent's minimal media (VMM) as a defined medium as previously described (Oresnik et al. 1998). For transductions, *S. meliloti* was grown using LB^{MC} as previously described (Finan et al. 1984). Carbon sources that were added to the defined medium were filter sterilized and generally used at a final concentration of 15 mM. Antibiotics were also filter sterilized and added to the growth medium when required at the following concentrations: streptomycin (Sm), 200 μg/ml; spectinomycin (Sp), 200 μg/ml; neomycin (Nm), 200 μg/ml; tetracycline (Tc), 5 μg/ml; gentamicin (Gm), 60 μg/ml or 20 μg/ml for *E. coli*. Aromatic amino acids were filter sterilized and used at final concentrations of tryptophan 25 μg/ml, tyrosine 20 μg/ml and phenylalanine 60 μg/ml. Constructs and strains were single colony purified 3 times prior to being used.

2.2.2 Genetic manipulations and techniques.

Conjugations and transductions were carried out as previously described (Finan et al. 1984, 1988). Standard conditions were used for DNA isolations, restriction digestions, ligations and transformations (Sambrook et al. 1989).

Mutations in *SMc02340* were constructed by PCR amplifying an internal fragment of *SMc02340* using primers 1 and 2 (Table 2), restricting the PCR product with *Bam*HI/*Hin*dIII and ligating it into pKnock-Gm as a fragment yielding pSK01. This was mobilized into both Rm1021 and SRmD410 yielding SRmD651 and SRmD652 respectively. The constructs were confirmed using primers 3 and 4 (Table 2). The PCR products were verified with Sanger sequencing.

Mutations in *tktB* were constructed by PCR amplifying an internal fragment of *tktB* using primers 5 and 6 (Table 2). The PCR fragment was restricted with *Bam*HI and *Xho*I and cloned into pKnock-Gm yielding pJH119. This plasmid was subsequently mobilized into Rm1021 and

Table 1. Strains and plasmids list

Strain or plasmid	Genotype or phenotype[a]	Reference
Strains		
S. meliloti		
Rm1021	SU47 *str-21* Sm^r	(Meade et al. 1982)
SRmD397	Rm1021, *tktA*::Tn*5* Sm^r Nm^r	(Hawkins et al. 2018)
SRmD557	Rm1021, Δ*tktA* Sm^r	This work
SRmD410	Rm1021, *tktA*::Tn*5, stkR1* Sm^r Nm^r	This work
SRmD411	Rm1021, *tktA*::Tn*5, stkR2* Sm^r Nm^r	This work
SRmD466	Rm1021, *tktA*::Tn*5, stkR3* Sm^r Nm^r	This work
SRmD515	Rm1021, *tktA*::Tn*5, stkR4* Sm^r Nm^r	This work
SRmD516	Rm1021, *tktA*::Tn*5, stkR5* Sm^r Nm^r	This work
SRmD517	Rm1021, *tktA*::Tn*5, stkR6* Sm^r Nm^r	This work
SRmD518	Rm1021, *tktA*::Tn*5, stkR7* Sm^r Nm^r	This work
SRmD570	Rm1021, Δ*tktA stkR8,* Sm^r	This work
SRmD571	Rm1021, Δ*tktA stkR9,* Sm^r	This work
SRmD572	Rm1021, Δ*tktA stkR10,* Sm^r	This work
SRmD485	*tktB*::pKnock-Gm	This work
SRmD647	*tktA*::Tn*5, stkR1, tktB*::pKnock-Gm	This work
SRmD651	*SMc02340*::pKnock-Gm	This work
SRmD652	*tktA*::Tn*5, SMc02340*::pKnock-Gm	This work
SRmD659	Rm1021, Δ*SMc02340* Sm^r	This work
SRmD660	SRmD397, Δ*SMc02340* Sm^r Nm^r	This work
SRmD661	Rm1021 Δ*tktB* Sm^r	This work
SRmD662	SRmD410 Δ*tktB* Sm^r Nm^r	This work
SRmD663	ΦSRmD397(*tktA*::Tn*5*) -> SRmD661(Δ*tktB*) Sm^r Nm^r	This work
SRmD668	Rm1021 Δ*trpE*(*G*) Sm^r	This work
E. coli		
MM294A	*pro-82 thi-1 hsdR17 supE44*	(Finan et al. 1986)
MT607	MM294A *recA56*	(Finan et al. 1986)
MT616	MT607(pRK600)	(Finan et al. 1986)
DH5α	λ^- *ϕ*80d*lacZ*ΔM15 Δ(*lacZYA-argF*)*U169 recA1 endA1 hsdR17*(r_k^- m_k^-) *supE44 thi-1 gyrA relA1*	(Hanahan 1983)
DH5αλpir	λ*pir* lysogen of DH5α	(House et al. 2004)
Plasmids		
pRK600	pRK2013 *npt*::TN*9* Cm^r	(Finan et al. 1985)
pRK602	pRK600ΩTn*5* Cm^r Nm^r	(Finan et al. 1985)
pRK607	pRK2013::Tn*5*-233	(Finan et al. 1986)
pRK7813	Broad-host-range cloning vector; Tc^r	(Jones and Gutterson 1987)
pKnock-Gm	Suicide vector for insertional mutagenesis; R6K ori RK4 *oriT* Gm^r	(Alexeyev 1999)
pCO37	pRK7813 containing *attB* sites; Gateway-compatible destination vector	(Jacob et al. 2008)

pMK2014	FRT-*ccdB*-Cm^r-FRT cassette Pen^r	(House et al. 2004)
pXINT129	*λint* and *xis* driven by P_{lac}; Km^r	(Platt et al. 2000)
pHC60	pHC41, S65T *gfp*	(Cheng and Walker 1998)
pJQ200SK	Gene replacement suicide vector Gm^r	Quandt & Hynes (1993)
pJH119	400-bp internal fragment of *tktB* cloned into pKnock-Gm, Gm^r	This work
pSK01	400-bp internal fragment of *SMc02340* cloned into pKnock-Gm, Gm^r	This work
pSK04	pJQ200SK/*trpE* flanking regions Gm^r	This work
pSK07	pJQ200SK/*SMc02340* flanking regions Gm^r	This work
pSK08	pJQ200SK/*tktB* flanking regions Gm^r	This work
pSK05	pCO37/*SMc02340* Tc^r	This work
pSK06	pCO37/*SMc02341* Tc^r	This work
pJH110	pCO37/*tktB* Tc^r	(Hawkins et al. 2018)

[a] Sm^r, streptomycin resistant; Nm^r, neomycin resistant; Gm^r, gentamicin resistant; Sp^r, spectinomycin resistant; Cm^r, chloramphenicol resistant; Tc^r, tetracycline resistant; Pen^r, penicillin resistant; Km^r, kanamycin resistant; FRT, flippase recognition target.

SRmD410 yielding SRmD485 and SRmD647 respectively. The constructs were confirmed by nucleotide sequencing the PCR product generated by primers 7 and 8.

Plasmids for unmarked deletions were constructed using Gibson assembly. Primers were designed using the NEBuilder Assembly tool (http://nebuilder.neb.com/#!/ , New England Biolabs, Inc.). The primers were designed by taking the flanking regions of gene of interest and pJQ200SK as a plasmid vector (Quandt and Hynes 1993). The PCR reactions were carried out using Q5 polymerase and the assembly protocol was carried out as suggested (New England Biolabs). The resultant plasmids were then mobilized into *S. meliloti*. Allelic exchanges were isolated as previously described (Geddes and Oresnik 2012b). To construct SRmD668, primers 9 to 12 were used to generate the construct. The final construct was verified by using primers 13 and 14 to generate a PCR fragment to amplify across the deletion junction. The PCR fragment was nucleotide sequenced. Similarly, a *tktB* deletion was generated using primers 15 to 18 (Table 2). The resultant plasmid, pSK08 was used to construct SRmD661, SRmD662, and SRmD663. The introduction of this deletion was confirmed using primers 19 and 20 (Table 2). The PCR fragment generated in each case was verified by nucleotide sequencing. A deletion of *SMc02340* was generated using primers 21-24 (Table 2). The resultant plasmid was used to generate SRmD659 and SRmD660 by using Rm1021 and SRmD397 as recipients. Deletions were confirmed using primers 25 and 26 (Table 2). The PCR fragment that was generated in each case was verified by Sanger sequencing.

To genetically map suppressors of *tktA*, an initial genetic approach was used to localize the second site mutation mutations (Oresnik et al. 1994). Briefly, a random mutagenesis of Rm5000 was carried out using pRK607 which contained Tn*5*-233 (Gm^RSp^R). Approximately 3,000 colonies were pooled. An aliquot of this was grown in LB^{MC} and infected with ΦM12. This

Table 2. Primer sequences for insertional mutations and gene deletion constructs

Primer #	Primer sequence 5'-3'	Primer name
	***tktB*:: pKnock-Gm**	
1	ATTC**GGATCC**ATCTGCTGCAAGACGATCATC	pKnock_tktB_BamHI_F
2	ATTC**CTCGAG**CGAACCACCGATCACCTCC	pKnock _tktB_XhoI_R
3	GTCTATGACGACAACAGGATC	pKnock _tktB_confir_F
4	CATCATTGCCGACATGCCGA	pKnock _tktB_confir_R
	***SMc02340*:: pKnock-Gm**	
5	ATTC**GGATCC**TCCGGCAGGCGCTTCAGTCG	pKnock_SMc02340_BamHI_F
6	ATTC**AAGCTT**TGCCGGTCGTCGCCACGGC	pKnock _SMc02340_HindIII_R
7	GAACGCAGTTCACCCATC	pKnock _SMc02340_confir2_F
8	GGATGGTGCCAGGACTCG	pKnock _SMc02340_confir2_R
	Δ*tktB*	
9	tcctgcagcccggggAACGGCGTGATCGCGGCG	ΔtktB_SacI_SMc02343_F
10	tcatcgtgATGAAGCCTCCAAATATCTCTTCA	
	TCCTATAATGTTGTAATTACAACATG	ΔtktB_SacI_SMc02343_R
11	ggcttcatCACGATGATCCAATGATGGG	ΔtktB_BamHI_SMc02341_F
12	gaacaaaagctggagctCCGCTGACAACGGGATTG	ΔtktB_BamHI_SMc02341_R
13	TCCTTTGCATCGCGGCCATAG	ΔtktB_conf_F
14	TTGTGCCCACCAGTCCGCG	ΔtktB_conf_R
	Δ*SMc02340*	
15	tcctgcagcccggggGCCGCGTCCTCCTCCGGT	ΔSMc02340_left_400_BamHI_F
16	caccggtcGCGAATCTCTGCCTTTCGTACCGTG	ΔSMc02340_left_400_BamHI_R
17	agattcgcGACCGGTGCGGCATTCTC	ΔSMc02340_right_400_SacI_F
18	gaacaaaagctggagctTCTACTCGATGGCCGCCC	ΔSMc02340_right_400_SacI_R
19	TACTCCCGGATTTGATGTCCC	ΔSMc02340_confir_F
20	CTCGGCAATGCCACAATCGG	ΔSMc02340_confir_R
	Δ*trpE*	
21	tcctgcagcccggggCCTCCTGCGCGTCGGCTG	trpE _left_400bp_F
22	gccttgaatCCTTGTGATGCCCGAAGGGG	trpE _left_400bp_R
23	tcacaaggATTCAAGGCGATCATCGG	trpE _right_600bp_F

24	gaacaaaagctggagctTCGTCTCGCTCCGTAAAAG	trpE _right_600bp_R
25	CCGCTTGCGCGCCATCGCAA	ΔtrpE _conf_F
26	CGCAGGATGGTTCTATTGTCGG	ΔtrpE _conf_R
	qRT-PCR primers	
27	AAACGATGCGGTGTTTCTGG	SMc02340_qRT_F
28	CTCCTTCGACTGCAACAGGT	SMc02340_qRT_R
29	GGCTGGCAATGGAACTTCTT	SMc02343_qRT_F
30	AGAAAGCCCACGAGATCGAA	SMc02343_qRT_R
31	AGATCGGCCTTGTCCTTGAA	tktB_qRT_F
32	AGCACGCTCTCCTGATAGTC	tktB_qRT_R
33	CACGAGCTCGAGATCGCCAT	SMc00128_qRT_F
34	AAATGGACAGCTTCTTTCCGC	SMc00128_qRT_R

Restriction sites are shown in bold; lower case letters represent the overlapping regions of the Gibson assembly plasmid.

lysate was then used as a donor for transductions into strains SRmD410 and SRmD466 which carried *stkR1* and *stkR3* alleles, respectively. Insertions linked in transductions were identified by screening the resultant transductants for the loss of their ability to utilize ribose. The position of these insertions was determined using arbitrary PCR followed by nucleotide sequencing of the resultant product as previously described (Poysti et al. 2007). The genetic distance to the suppressor locus was determined by transducing the inserts back into either SRmD410 or SRmD466. The genetic distance was converted to a physical distance as previously described (Charles and Finan 1990), and was used to calculate the linkage of transposon insertion to a specific locus (Finan et al. 1984).

To sequence strains SRmD410 and SRmD466, genomic DNA was extracted using an UltraClean® Microbial DNA Isolation Kit (MO BIO Laboratories Inc., Carlsbad, CA) following the manufacturer's instructions. The genomic DNA was sequenced using Pacific Biosciences SMRT sequencing platform at McGill University and Genome Quebec Innovation Centre (Montreal, QC, Canada). Sequence alignment was done using RS_Bridgemapper pipeline available in the SMRT portal. The sequence alignment was done against *S. meliloti* 1021 reference sequence. The localized area in the genome was analyzed for SNPs using NCBI BLASTn. BioEdit software was used to generate the predicted open reading frames for the SNP variants in the suppressor strains. SNPs were confirmed by nucleotide sequencing.

2.2.3 Metabolite analysis.

Metabolite analysis was carried out as described previously (Hawkins et al., 2018). Briefly, bacterial cells were grown in LB overnight and sub-cultured into 500 ml VMM containing 15 mM glucose as carbon source and aromatic amino acids (phenylalanine, tyrosine and tryptophan). The cultures were grown overnight with shaking at 200 rpm. 100 ml cultures were

harvested by centrifugation at an OD_{600} of 0.5. The cell pellets were washed with double distilled water, pelleted, weighed, and flash frozen in liquid N_2. Metabolite concentrations were normalized to pellet weight. Metabolite extraction and LC-MS was performed as described previously (Hawkins et al., 2018) at Metabolomics Innovation Resource (MIR) - McGill University. Statistical analysis was performed using MetaboAnalyst 4.0 (https://www.metaboanalyst.ca) (Chong et al. 2019). The metabolite concentrations were normalized by log-transformation. Principal component analysis, analysis of variance and clustering analysis were done as described previously after normalization (Hawkins et al. 2018). Significant difference was determined by performing Welch's two sample *t* test with a *P*-value criterion of < 0.05. False discovery rate (< 0.1) between two samples was determined for significance as well.

2.2.4 RNA extraction, cDNA synthesis and qRT-PCR.

RNA extraction and cDNA synthesis were performed as described previously (Geddes and Oresnik 2012a). Bacterial cultures were grown in LB overnight and 1 ml cultures were harvested at OD_{600}~ 0.8. The cell pellet was resuspended in TE (Tris 10 mM, EDTA 1 mM, pH 8) buffer with lysozyme (0.4 mg/ml). RNA extraction was carried out using Qiagen RNeasy isolation kits and the samples were treated with RNase free DNase to remove DNA contamination. RNA concentration was determined spectrophotometrically using a NanoDrop (ThermoFisher Scientific). One μg RNA was used to synthesize cDNA using a VILO® cDNA synthesis kit from Invitrogen. Approximately 200 ng of cDNA was used to carry out the qRT-PCR reaction using SYBR green RT-PCR kit from Invitrogen. The quality of cDNA was determined by gel electrophoresis and quantified spectrophotometrically. Primer pairs that were unique to this study

are listed in Table 2. The data from qRT-PCR was analysed by $2^{\wedge-(\Delta\Delta Ct)}$ method (Livak and Schmittgen 2001).

2.2.5 Exopolysaccharide production.

Succinoglycan production was visualized by the characteristic fluorescence emitted by cultures under UV light that were grown on plates containing Calcofluor as previously described (Finan et al. 1985). To quantify exopolysaccharide production, the EPS was precipitated from the growth medium as previously described (Hawkins et al. 2017a), and its concentration was determined using an anthrone assay (Morris 1948). The data were normalized to the OD_{600} of the cultures. Statistical significance was determined by conducting Student's *t* test with a *P*-value criterion of < 0.05.

2.2.6 Plant growth conditions and nodulation kinetics.

Plant dry weights were determined as previously described (Poysti et al. 2007). Briefly, surface sterilized alfalfa seeds were germinated for 2 days on water agar plates. Approximately 10 seedlings were aseptically transplanted into autoclaved Leonard jars that contained a mix of sand and vermiculite (approximately 1:1 ratio) and were soaked with 200 ml of nitrogen free Jensen's plant growth medium (Glazebrook and Walker 1991). The pots were inoculated with approximately 10^7 cfu of rhizobia in a volume of 10 ml sterile water. Plants were harvested 4 weeks after inoculation.

Nodulation kinetics assays were carried out as previously described (Geddes et al. 2014; Hawkins et al. 2018). Briefly, germinated alfalfa seedlings were placed on Jensen's agar slants and were inoculated with 100 µl of 10^7 CFU/ml rhizobial cultures. The appearance of nodules was scored every other day for 3 weeks. Each experiment consisted of at least 5 biological replicates with at least 5 plants per replicate.

2.2.7 Microscopy.

The microscopic analysis was performed as previously described (Gage et al. 1996; Cheng and Walker 1998; Geddes et al. 2014). *S. meliloti* strains were labelled with green fluorescent protein (GFP) by introducing pHC60 into each strain to be assayed (Cheng and Walker 1998). Alfalfa seedlings were grown on glass slides containing an even layer of solidified Jensen's agar that was covered by a dialysis membrane to keep the seedling in place. Each seedling was inoculated with 100 µl of 10^7 CFU/ml rhizobial culture and the slide was placed in a 50 ml tube containing 20 ml sterile Jensen's broth media. The tube was loosely covered with plastic cling wrap. Visualization of curled colonized root hair, infection thread formation and nodule primordia were done from day 7 post inoculation. The plants were observed until the appearance of fully developed nodules. Each experiment consisted of three biological replicates that contained at least 10 plants. Each experiment was repeated three times.

The plants were visualized using a Zeiss Axio Imager.Z1 microscope equipped with an AxioCamMR digital camera. GFP expression was visualized using Zeiss filter set 38 (eGFP). Autofluorescence was decreased in merged images by using Zeiss filter set 20 (Rhodamine). Images were generated using the software ZENBlue 2.

2.3 Results

2.3.1 Symbiotic defects caused by the tktA mutation can be suppressed.

S. meliloti strains carrying a *dgoK* mutation are unable to utilize galactose as a sole carbon source (Geddes and Oresnik 2012a). If these strains are grown in the presence of galactose and a second carbon source which they are able to utilize, it results in an acidification of the growth medium (Geddes et al. 2014). Mutations in *tktA* were first identified in screens designed to determine what genes contributed to the acidification of the medium (Hawkins et al. 2017b). In

addition to providing evidence that *tktA* was essential for the PP pathway in *S. meliloti*, it was also found that a strain carrying a mutation in *tktA* was symbiotically compromised (Hawkins et al. 2018).

When strain SRmD397, which carries a transposon (Tn*5*) insertional mutation in *tktA*, was inoculated onto alfalfa plants and grown using nitrogen deficient conditions, the resultant plants were small and chlorotic. The nodules that were produced were cyst-like white nodules (Figure 2). However, it was noted that a small number of plants turned green, and they appeared to have an occasional effective (Fix^+) nodule. When bacteria were isolated from one of these Fix^+ nodules, it was found that they still contained the Tn*5* insertion within *tktA*. The Tn*5* marker was transduced into Rm1021 and the resulting strains showed the same plant phenotypes as the parental strain (SRmD397). The original isolate was designated SRmD410. When it was reinoculated onto alfalfa seedlings, the resultant plants were visibly green and had nodules that were similar to those made by the wild type (Figure 2). This suggested that a second site suppressor mutation may have arisen in SRmD397 that was greater than 150 kilobases distant from the original mutation. Over the course of several experiments two other independent suppressor mutations were isolated. The suppressor alleles were designated *stkR1, stkR2,* and *stkR3* (suppressor of transketolase) in strains SRmD410, SRmD411, and SRmD466 respectively.

When these strains were used to re-inoculate alfalfa plants, they resulted in plants that appeared phenotypically effective with nodules that resembled that of the wild type; however, it appeared that the plant biomass was lower (Figure 2). To address this observation, strains carrying each of the three suppressor alleles were used to inoculate alfalfa seedlings. In addition, nodule development was also observed. The results showed that although the suppressor strains

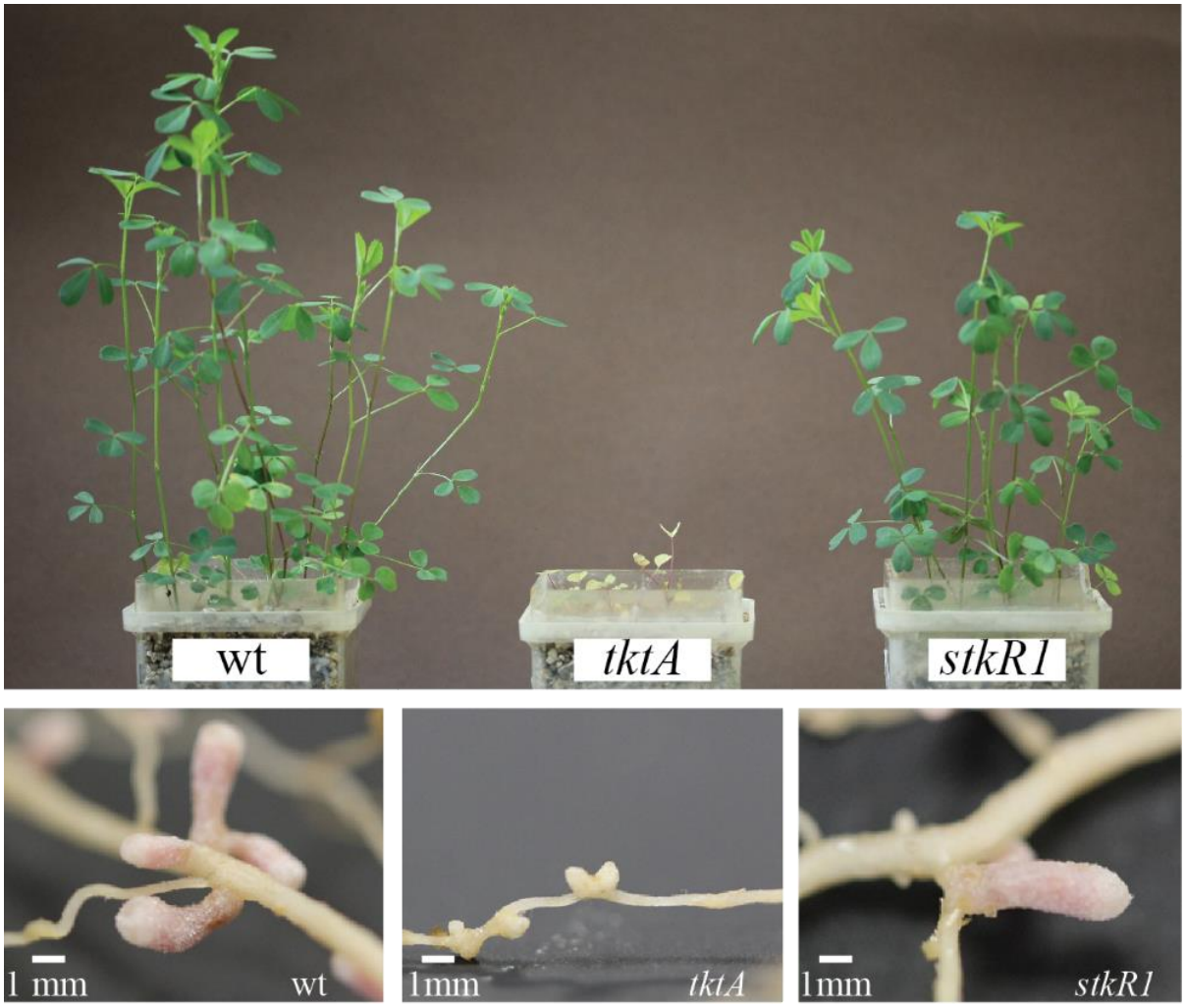


Figure 2. Alfalfa plants inoculated with wild type, the *tktA* mutant strain (SRmD397) and a suppressor strain SRmD410 (*stkR1*).

Top picture shows alfalfa plants grown in N-free Jensen broth. Sterile sand and vermiculite mixture was used in the pots for physical support. The bottom panel shows the nodules produced by each strain. The scale bar represents 1mm.

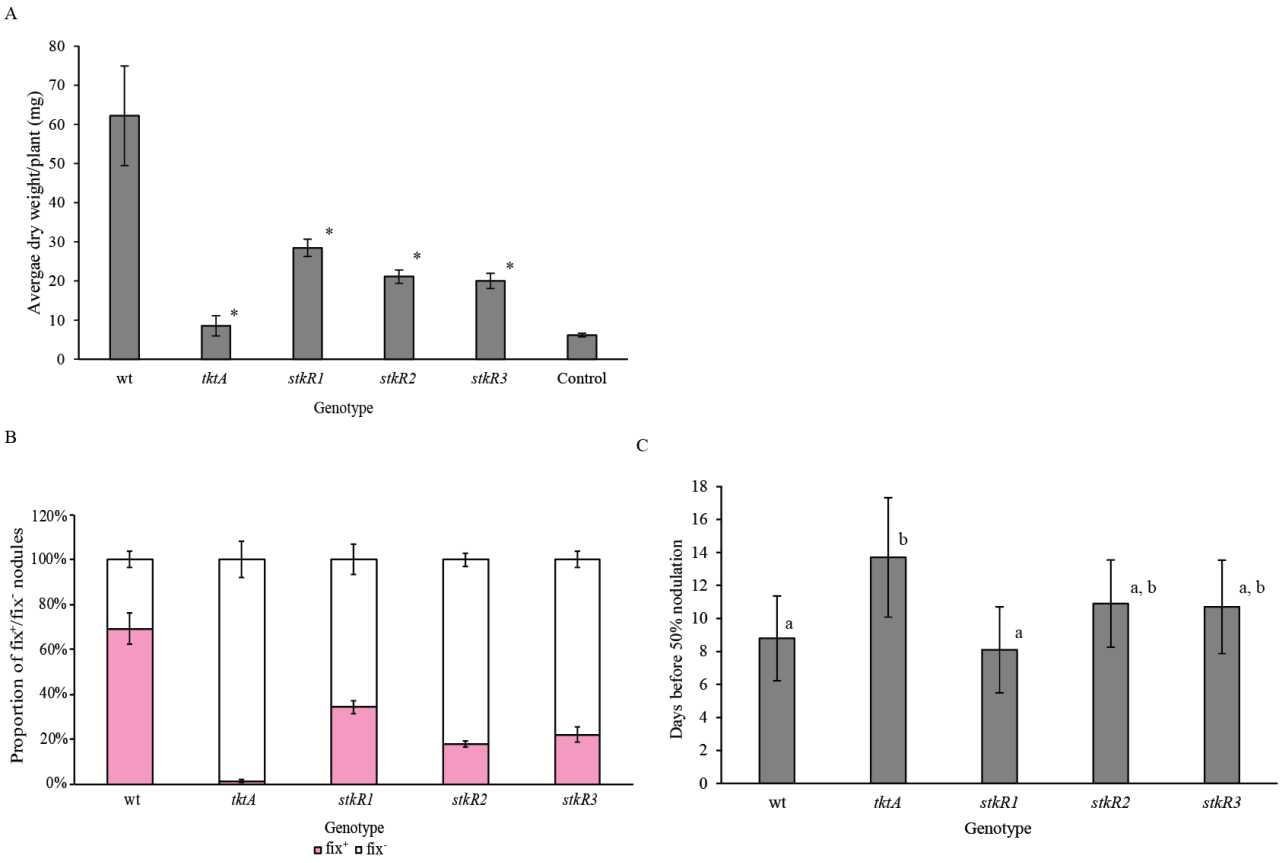


Figure 3. Symbiotic phenotypes associated with *tktA* suppression.

A: Average shoot dry weight of alfalfa plants inoculated with the wild type, the *tktA* mutant strain (SRmD397) and three suppressor strains; *stkR1* (SRmD410), *stkR2* (SRmD411), and *stkR3* (SRmD466). * Represents statistical significance calculated by Student's t test with $P < 0.05$, error bars represent standard deviation of three biological replicates. Each bar the average shoot dry weight of 30 alfalfa plants. B: Proportion of effective (pink) and ineffective (white) nodules on alfalfa plants in a nodulation kinetics study. The error bar represents standard deviation of 5 biological replicates each containing at least 5 plants. C: Bar graph representing nodulation kinetics shown as days before 50% of the plants had nodules. The value for each bar is an average from two experiments with 10 biological replicates and 5 plants per replicate. The error bars represent standard deviation. Statistical significance was calculated by single factor analysis of variance (ANOVA) and Tukey's HSD with $P < 0.05$. $P = 0.0008$, a, no significant difference between wt, *stkR1*, *stkR2*, and *stkR3*; b, no significant difference between *tktA*, *stkR2*, and *stkR3*

gave rise to effective nodules, the accumulated shoot dry weight was significantly lower than that of plants that were inoculated with the wild type (Figure 3A). When the nodules were closely examined, it was found that although there were effective nodules on all the root systems of plants inoculated with the suppressors, the proportion of visibly pink effective nodules was lower than that found in the wild type (Figure 3B, 3C). We note a single effective nodule was observed on the root system of a plant that was inoculated with strain carrying a *tktA* mutation (Figure 3B); presumably this was due to a suppression event.

In addition to effective nitrogen fixation, *tktA* mutations were also correlated with lower exopolysaccharide production, slower nodulation kinetics, and an inability to grow on defined medium in the absence of aromatic amino acids (Hawkins et al. 2018). Strains SRmD410, SRmD411, and SRmD466 had nodulation kinetics similar to that of the wild type, had a calcofluor bright phenotype, increased exopolysaccharide synthesis relative to the parental strain which carried the *tktA* mutation, and no longer exhibited an amino acid auxotrophy (Figure 4).

Since the amino acid auxotrophy was associated with *tktA* mutation, it was of interest to determine whether this phenotype could be exploited to determine the rate of second site mutations which led to the phenotypic suppression. Five independent overnight cultures were plated on defined medium containing succinate. The rate of phenotypic reversion was determined to be $1.34 \times 10^{-8} \pm 3.3 \times 10^{-7}$. Strains isolated from these experiments were designated as carrying alleles *stkR4-stkR7* in strains SRmD515-518 respectively (Table 1). Strains SRmD570-572 were subsequently isolated on defined medium with ribose as a sole carbon source in a *tktA* deletion strain (Table 1). All of these suppressor alleles had phenotypes similar to that described for *stkR1-stkR3*.

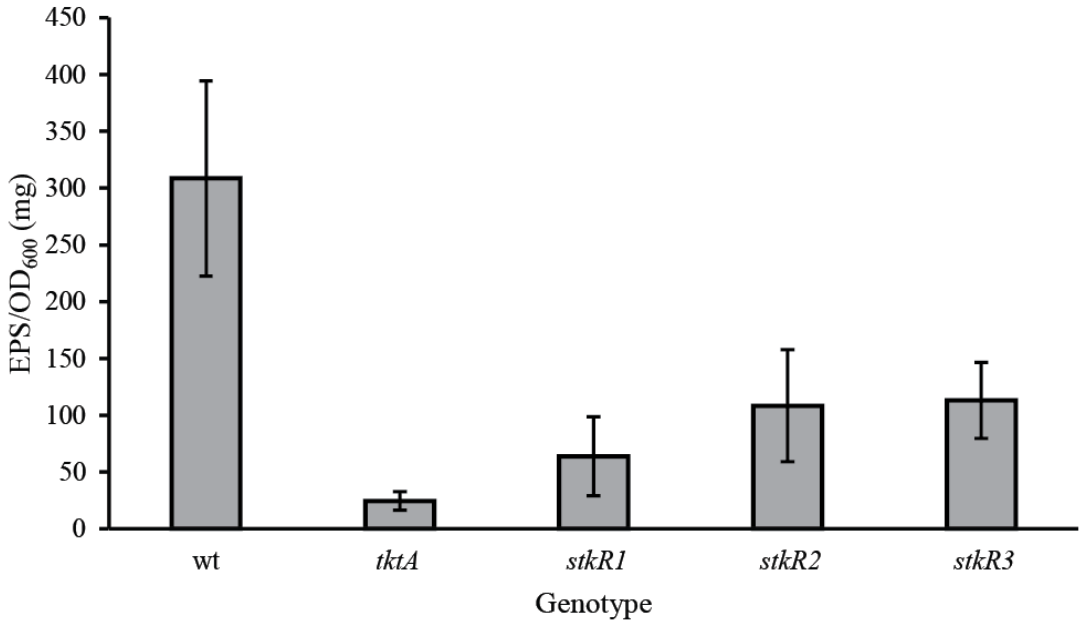


Figure 4. Exopolysaccharide production by the wild type, the *tktA* mutant strain and three suppressor strains; *stkR1* (SRmD410), *stkR2* (SRmD411), and *stkR3* (SRmD466).

The error bars represent standard deviation of three biological replicates. EPS was quantified using Anthrone assay.

2.3.2 Strains carrying suppressor mutations do not completely restore central carbon pathways.

Metabolomic analysis of the strain carrying a *tktA* mutation had previously shown that when it was grown on defined medium with glucose there was increased pentose phosphate intermediates and a corresponding decrease in sedoheptulose-7-phosphate pools (Hawkins et al. 2018). It was hypothesized that since symbiotic as well as other growth phenotypes were suppressed, the nature of the suppressor mutation must be metabolic in nature. To test this hypothesis a metabolite analysis focusing on the central carbon metabolites was carried out by growing the strains on glucose as a sole carbon source. The analysis revealed that although the levels of pentose phosphates in the suppressor strains were decreased compared to the *tktA* mutant strain, and that the level of sedoheptulose-7-phosphate increased, they were not comparable to the wild type (Figure 5). A wider comparison using hierarchical clustering based on a broader spectrum of central carbon metabolites showed that the three strains carrying the suppressor mutations were different from both the wild type and SRmD397 (Figure 6, Table 3).

2.3.3 stkR1 and stkR3 are located in close proximity to tktB.

To broadly identify what region of the genome the mutations that lead to the suppression of *tktA* associated phenotypes, a genetic approach of finding transposons linked in transduction to the allele was employed (Oresnik et al. 1994). This strategy was employed to determine if the *stk* mutations behaved as point mutations, whether multiple loci were responsible for the suppression events, as well as to provide guidance for subsequent SNP analysis of the strains carrying suppressors following whole genome sequencing.

Two Tn*5*-233 insertions linked in transduction were isolated and found to be 77% and 80% linked to the *stkR1* and *stkR3* alleles respectively. Correlating genetic linkage to physical

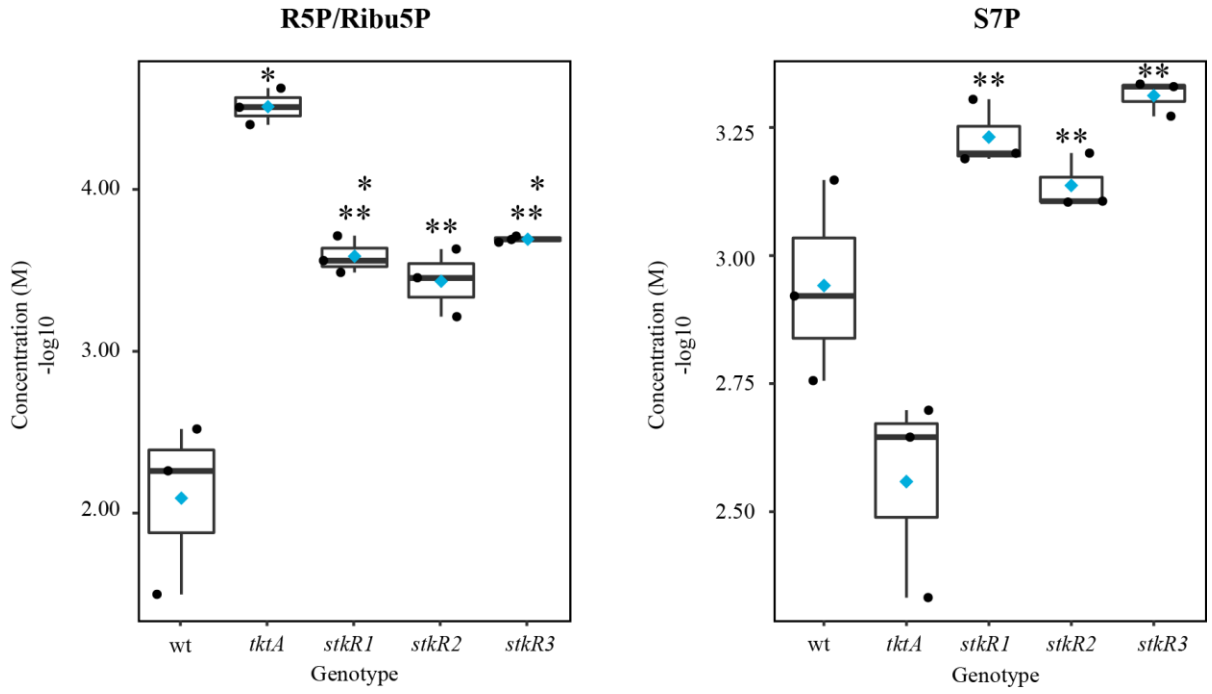


Figure 5. Box plots represent concentrations of R5P/Ribu5P and S7P pools in the wild type, the *tktA* mutant strain and three suppressor strains; *stkR1* (SRmD410), *stkR2* (SRmD411), and *stkR3* (SRmD466)

The y-axis shows the concentration on a -log10 scale and x-axis has the strains. * Significant difference samples compared to the wt. ** Significant difference samples compared to the *tktA* mutant strain. Significance was calculated using a Student's *t*-test with significance defined as $P < 0.05$. A blue diamond indicates the mean value of replicates, whereas a black dot indicates the value of each individual replicate.

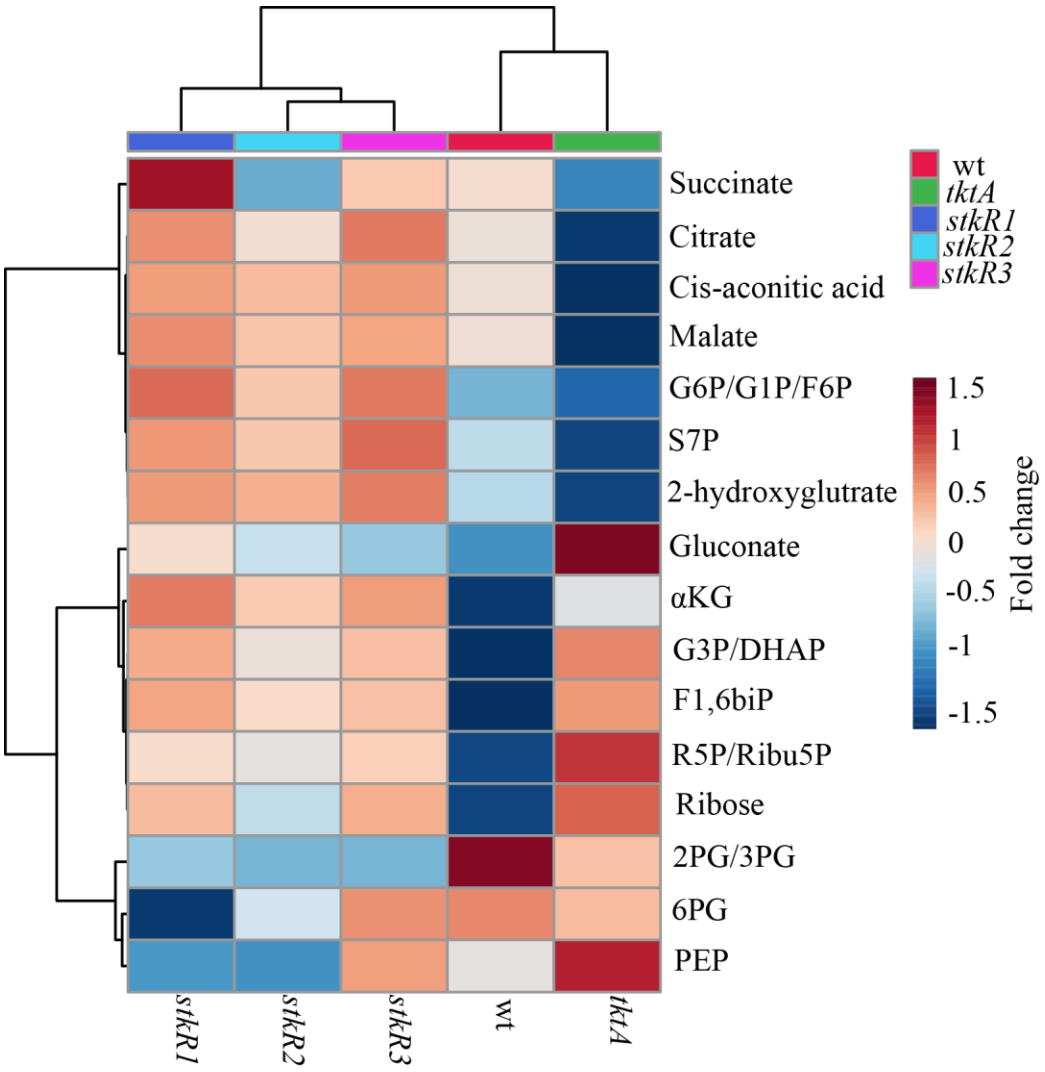


Figure 6. Clustered heatmap of central carbon metabolite concentrations in the wild type, the *tktA* mutant strain and three suppressor strains; *stkR1* (SRmD410), *stkR2* (SRmD411), and *stkR3* (SRmD466)

The cultures were grown on glucose and supplemented with aromatic amino acids tryptophan, phenylalanine, and tyrosine. Classes represent *S. meliloti* strains as: red, wild type; green, *tktA*; dark blue, *stkR1*; light blue, *stkR2;* pink, *stkR3*. Metabolites are represented by each row, while mean of three replicates for each strain is represented by a column. The dendrogram and colored boxes were generated with MetaboAnalyst 5.0. The color scale represents fold changes with respect to the average value of each metabolite. Colors indicate fold increases or decreases in concentrations, as indicated in the key. G6P, glucose-6-phosphate; G1P, glucose-1-phosphate; F6P, fructose-6-phosphate; S7P, sedoheptulose-7-phosphate; aKG, a-ketoglutarate; G3P, glyceraldehyde-3-phosphate; DHAP, dihydroxyacetone phosphate; F1,6biP, fructose-1,6-bisphosphate; R5P, ribose-5-phosphate; Ribu5P, ribulose-5-phosphate 2-PG, 2-phosphoglycerate; 3-PG, 3-phosphoglycerate; 6PG, 6-phosphogluconate; PEP, phosphoenolpyruvate.

Table 3. Metabolite concentrations

Metabolite (nM/mg)	**wt**	***tktA***	***stkR1***	***stkR2***	***stkR3***
ED/EMP					
G6P/F6P/G1P	10469	5425	26589*	18349*	25096
DHAP/G3P	1121	13418*	10129*	5859*	8520*
Gluconate	1191	3786	1924	1642*	1414
6P-Gluconic acid	191	71	37	61	88
2P-Glyceric acid/3P-Glyceric acid	679	293*	156*	142*	145*
Phosphoenolpyruvate	500	588	187	186	446
F1,6 biP	65	566	455*	314*	380*
PPP					
Ribose 5P/Ribulose 5P	190	34807*	4147*	3054	5159*
Sedoheptulose-7P	936	385	1716*	1378	2057
E4P	ND	ND	ND	ND	ND
PRPP (5P-Ribose 1diP)	ND	304	ND	ND	ND
Ribose	74211	143627	124544	104229*	128053
TCA					
Citrate	2461	431	3566	1991	3625
Cis-Aconitic acid	586	222*	775	694	780
α-Ketoglutarate	1616	2735	3693*	3280	3429*
2-Hydroxyglutarate	73	36*	124	114	136*
Succinate	10036	7587	11985	8324	9761
Malate	1031	631	1167	1069	1124

*Significant differences in metabolite concentrations of glucose grown cells when compared to the wild type. Significance ($P < 0.05$) was determined by Student's T-test. $n = 3$

ND, not detected; G6P, glucose-6-phosphate; F6P, fructose-6-phosphate; G1P, glucose-1-phosphate; DHAP, dihydroxyacetone phosphate; G3P, glyceraldehyde-3-phosphate.

distance suggests that these insertions were approximately 10-15 kb from each the *stk* alleles. Using arbitrary PCR, followed by the sequencing of these products it was found that these insertions were within *gldA* and *SMc02344* respectively. Taken together, this suggests that the *stk* alleles were flanked by these insertions. We note that *tktB* (*SMc02342*), which has previously been shown to be able to complement a *tktA* mutation *in trans*, is found in this region of the chromosome (Hawkins et al. 2018).

Following whole genome sequencing of strains carrying *stkR1* and *stkR3* (Table *4*) it was found that each carried a mutation in *SMc02340*, which encodes a GntR family negative regulator. The mutation in *stkR1* results in a frameshift that affects the C-terminal end of SMc02340 leading to no identity to the wild type over its terminal 25 amino acids as well as having an extension of 14 amino acids. The mutation in *stkR3* led to an amino acid substitution such that the isoleucine at position 9 was now an asparagine.

2.3.4 Phenotypic suppression is dependent on SMc02340 and tktB.

Some GntR family regulators are negative regulators that consist of a helix-turn-helix DNA binding domain at the N-terminus, which is linked to a effector binding domain found at the C-terminus (Suvorova et al. 2015). Since it had been shown that the introduction of *tktB* on a plasmid could complement the carbon utilization phenotypes associated with a strain carrying a *tktA* mutation (Hawkins et al. 2018), it was hypothesized that *SMc02340* must be a negative regulator of *tktB* and that the *stkR1* and *stkR3* alleles must result in non-functional variants of *SMc02340*.

To test this hypothesis an insertional mutation in *SMc02340* was constructed in both the wild type (SRmD651) as well as the *tktA* backgrounds (SRmD652) (Table 5). The results

Table 4. PacBio sequencing parameters

Sequencing Parameters (bp)	***stkR1***	***stkR3***
Average read length	11,786	11,745
Average read quality	84	84
Subreads	170,851	174,657
Mean length of subreads	6943	6975
Circular consensus sequence (CCS) read count	5039	5103
Mean CCSs length	7944	8305
Total bases of CCSs	40,030,832	42,380,448
N50	8143	8316
Assembly parameters (bp)		
Coverage	165	170
Contigs	5	3
Total bases in contigs	6,746,180	6,745,689
Minimum contig length	8936	1,371,559
Maximum contig length	3,667,978	3,675,645

Table 5. Carbon phenotypes associated with *tktA* suppression.

Strain	**Relevant Genotype**	**Ery**	**Ery^{FWY}**	**Rib**	**Rib^{FWY}**	**Glc**	**Glc^{FWY}**
Rm1021	wild type	++	++	++	++	++	++
SRmD397	*tktA*::Tn*5*	-	++	-	++	-	++
SRmD410	*tktA*::Tn*5*, *stkR1*	+	+	++	++	++	++
SRmD651	*SMc02340*::pKnock-Gm	++	++	++	++	++	++
SRmD485	*tktB*::pKnock-Gm	++	++	++	++	++	++
SRmD652	*tktA*::Tn*5*, *SMc02340*::pKnock-Gm	++	++	++	++	++	++
SRmD647	*tktA*::Tn*5*, *stkR1*, *tktB*::pKnock-Gm	-	++	-	++	-	++
SRmD647 (pJH110)	*tktA*::Tn*5*, *stkR1*, *tktB*::pKnock-Gm, (*tktB*$^{+}$)	++	++	++	++	++	++

Glc: Glucose; Rib: Ribose; Ery: Erythritol, F: Phenylalanine; W: Tryptophan; Y: Tyrosine, *++ wild type growth; + weak growth; - no growth

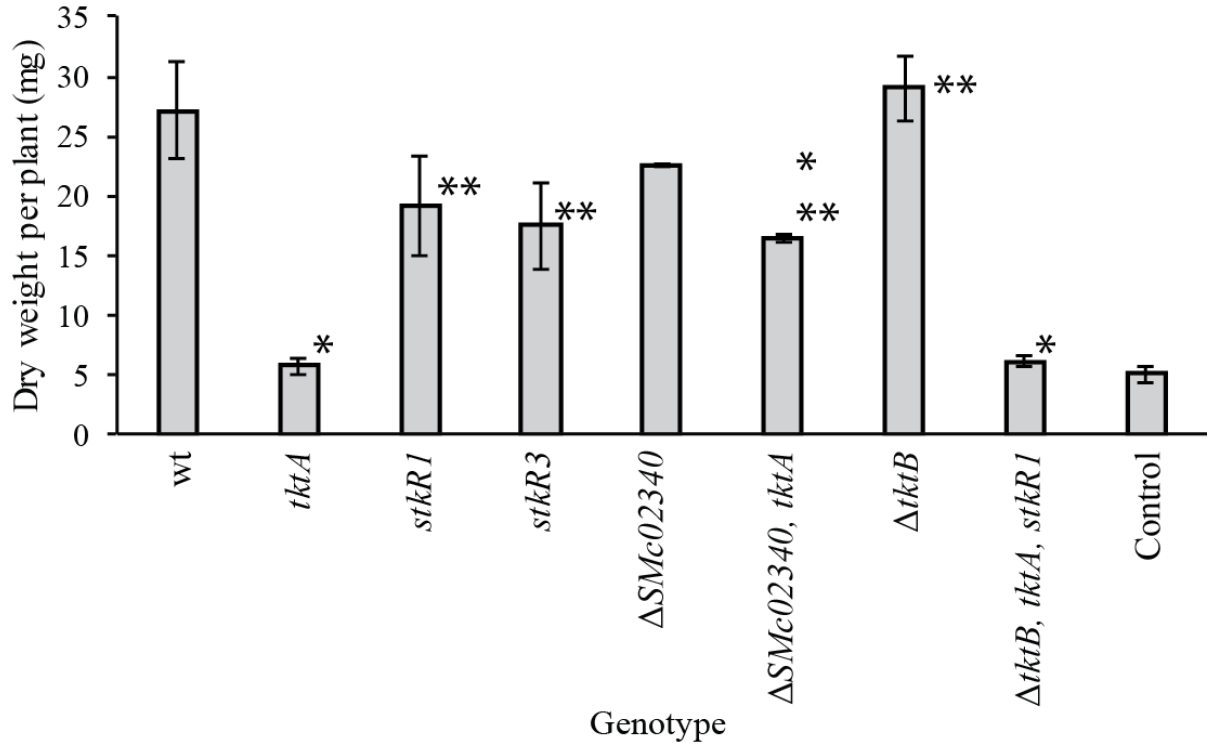


Figure 7. Average shoot dry weight of alfalfa plants inoculated with the *S. meliloti* strains.

Seedlings were inoculated with strains carrying the indicated genotypes and the plants were harvested 28 days after inoculation. Each bar represents the mean of 3 biological replicates. Each replicate consisted of 10 plants. Error bars represents standard deviation.

* Represents statistical significance calculated using a Student's *t*-test with significance defined as $P < 0.05$ in comparison to the wt, ** represents significance in comparison to the *tktA* mutant strain.

showed that whereas a mutation of *SMc02340* in the wild type background did not alter the carbon utilization phenotype, the introduction of this mutation into a SRmD397, which carries a *tktA* mutation, resulted in the suppression of *tktA* associated auxotrophy (Table 5). To show that the suppression was due to *tktB*, strains carrying mutations in *tktB* in both the wild type as well as the *stkR1* background were constructed. The introduction of *tktB* into the wild type did not have any discernable phenotype, whereas the introduction of this mutation to SRmD647, which carries a *tktA* and a *stkR1* mutation, resulted in a phenotype identical to that of the original strain carrying a *tktA* mutation (Table 5). Addition of *tktB* on a plasmid to this strain resulted in alleviation of the *tktA* associated auxotrophy (Table 5).

Since the single cross-over mutations can be unstable when used for nodulation experiments as the plants can select revertants of the mutation, a series of non-reverting deletion alleles of *SMc02340* and *tktB* were constructed in a number of backgrounds to corroborate the carbon utilization phenotypes (Figure 7). The results showed that the deletion of *SMc02340* did not have any discernable effect on nodulation and nitrogen fixation in a wild type background; however when introduced into SRmD397, it resulted in suppression of the symbiotic phenotype associated with *tktA* (Figure 7). We note that although there was a suppression of the symbiotic phenotype, the dry matter accumulation was significantly lower from the wild type (Figure 7). Similar to what was seen with the carbon utilization phenotypes (Table 5), the introduction of a *tktB* deletion into a suppressor strain reverted the phenotype to that which was observed with the *tktA1* allele (Figure 7).

To directly demonstrate that the mechanism of suppression was due to an up-regulation of *tktB* because of a loss of a negative regulator, qRT-PCR experiments were carried out (Figure 8). The gene encoding the regulator SMc02340 is located in close proximity to *tktB* and oriented

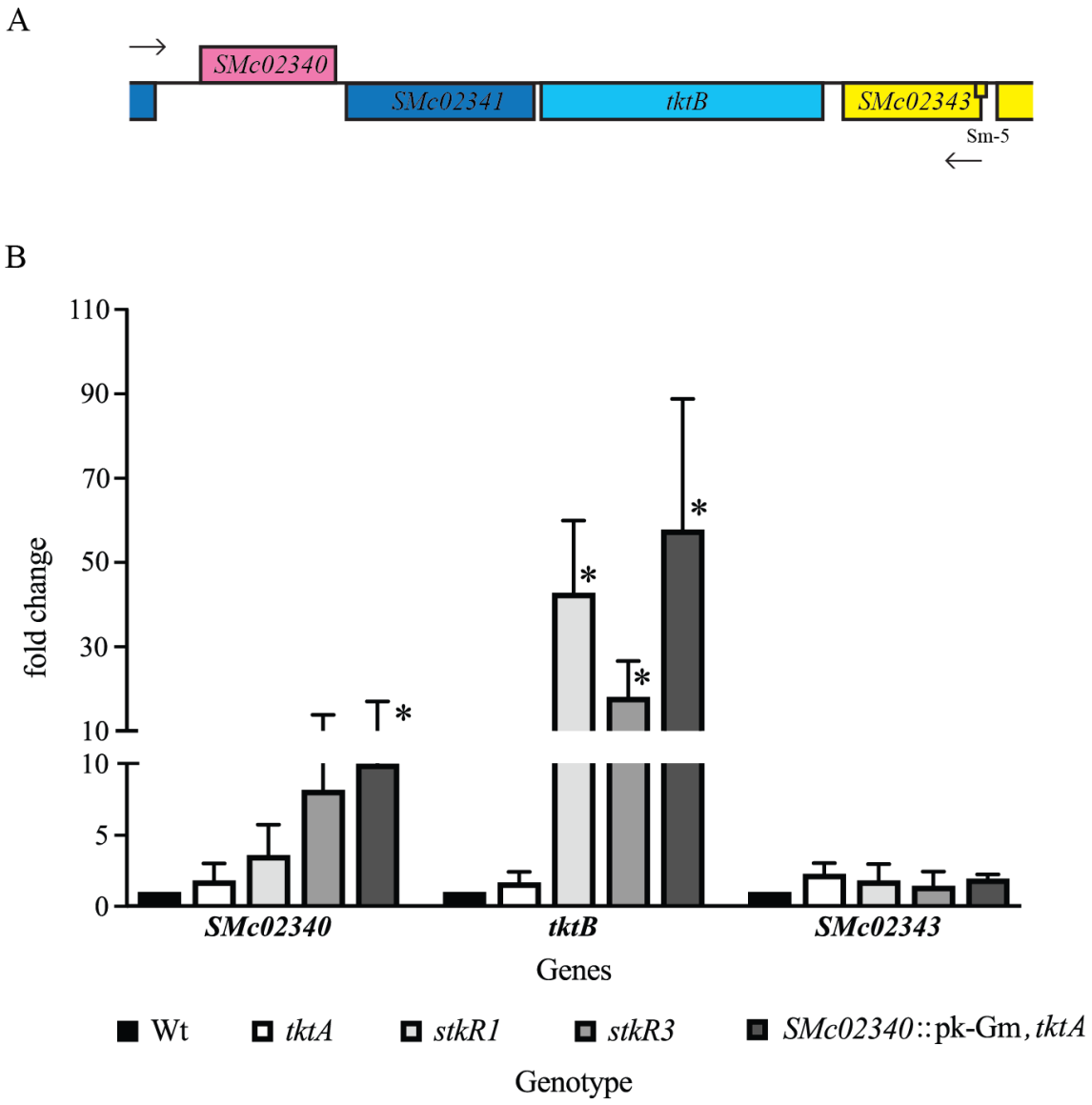


Figure 8. Mutations in *SMc02340* affect *tktB* expression.

A: Genetic locus representing *tktB* operon in *S. meliloti.* Arrows indicate the direction of transcription (ie open reading frames depicted on top are transcribed left to right; open reading frames depicted below are transcribed right to left), the colors indicate the predicted functional category of the gene according to the *S. meliloti* genome database (https://iant.toulouse.inra.fr/bacteria/annotation/cgi/rhime.cgi). Yellow; cell processes, dark blue; small molecule metabolism, light blue; central intermediary metabolism, pink; miscellaneous. B: qRT-PCR analysis of *SMc02340, tktB* and *SMc02343* expression in *S. meliloti* strains (wt, *tktA stkR1*, *stkR3*, and *SMc02340*::pk-Gm *tktA*).* Represents a significant difference from the wt strain with a $P < 0.05$ using a Student's *t*-test. Error bars represent standard deviation between 3 biological replicates. The figure was created using GraphPad Prism software.

such that the transcription of *tktB* and *SMc02340* are convergent (Figure 8A). In addition, a gap of approximately 90 bp exists between the stop codon of *SMc02343* and the start codon of *tktB*. Since this intergenic region is sufficient to contain a promoter, we also wanted to determine if *SMc02343* was co-transcribed with *tktB* (Figure 8A).

Whereas there were very low levels of transcription of either *tktB* or *SMc02340* in the wild type, the transcription of *tktB* was significantly increased in either the *stkR1* or *stkR3* backgrounds. Consistent with the hypothesis that *stk* alleles led to non-functional variants, a pKnock mutation of *SMc02340* also led to significant increases of *tktB* transcription. Lack of any changes of the transcription of *SMc02343* are consistent with the hypothesis that the regulation of the genes upstream of *tktB* are not dependent on *SMc02340* (Figure 8).

2.3.5 Ten independent suppressors are all alleles of SMc02340.

Through the course of the study a number of independent suppressor alleles were isolated (Table 6). It was hypothesized that some of these might be found in *SMc02340*. However, some of the mutations might also be in the promoter region of *SMc02340*, or possibly in the promoter region of *tktB*. In addition, it was probable that a mutation not found in any of these locations might identify other genetic determinants capable of suppressing a *tktA* mutation.

A series of PCR primers were designed such that the upstream region of *SMc02340* and the entire coding region of *SMc02340* could be amplified. In addition, primers to the region upstream of *tktB* were also designed. Following amplification of these regions and sequencing both strands it was found that all of the mutations were within *SMc02340* (Table 6, Figure 9). Five of these mutations were due to single base pair changes that resulted in changes to a single amino acid (Table 6). Three of the mutations were either an insertion or deletion which led to a

Table 6. List of SNPs on *SMc02340* in the suppressor strains.

Allele	**Mutation[a]**	**Predicted change**
stkR1	621, insertion of TCCT	Divergent amino acid sequence from amino acid 209 and results in a 259 amino acid protein.
stkR2	62, G to C	A21P
stkR3	26, T to A	I9N
stkR4	539, A to C	T180P
stkR5	689, G to A	Generation of stop codon resulting in a truncated 228 amino acid protein
stkR6	276, deletion A	Frame shift, leads to divergent sequence from amino acid 93 and results in a 100 amino acid protein.
stkR7	203, C to G	R68G
stkR8	616, deletion T	Frame shift, leads to divergent sequence from amino acid 206 and results in a 236 amino acid protein.
stkR9	441, T to A	L147Q
stkR10	48, Insertion G	Frame shift, leads to divergent sequence from amino acid 18 and results in a 92 amino acid protein.

SMc02340 open reading frame consists of 735 bp that encode a protein that is predicted to contain 244 amino acids. [a] Mutations are denoted such that the number represents the bp at which the mutation was identified, followed by the nature of the mutation was found.

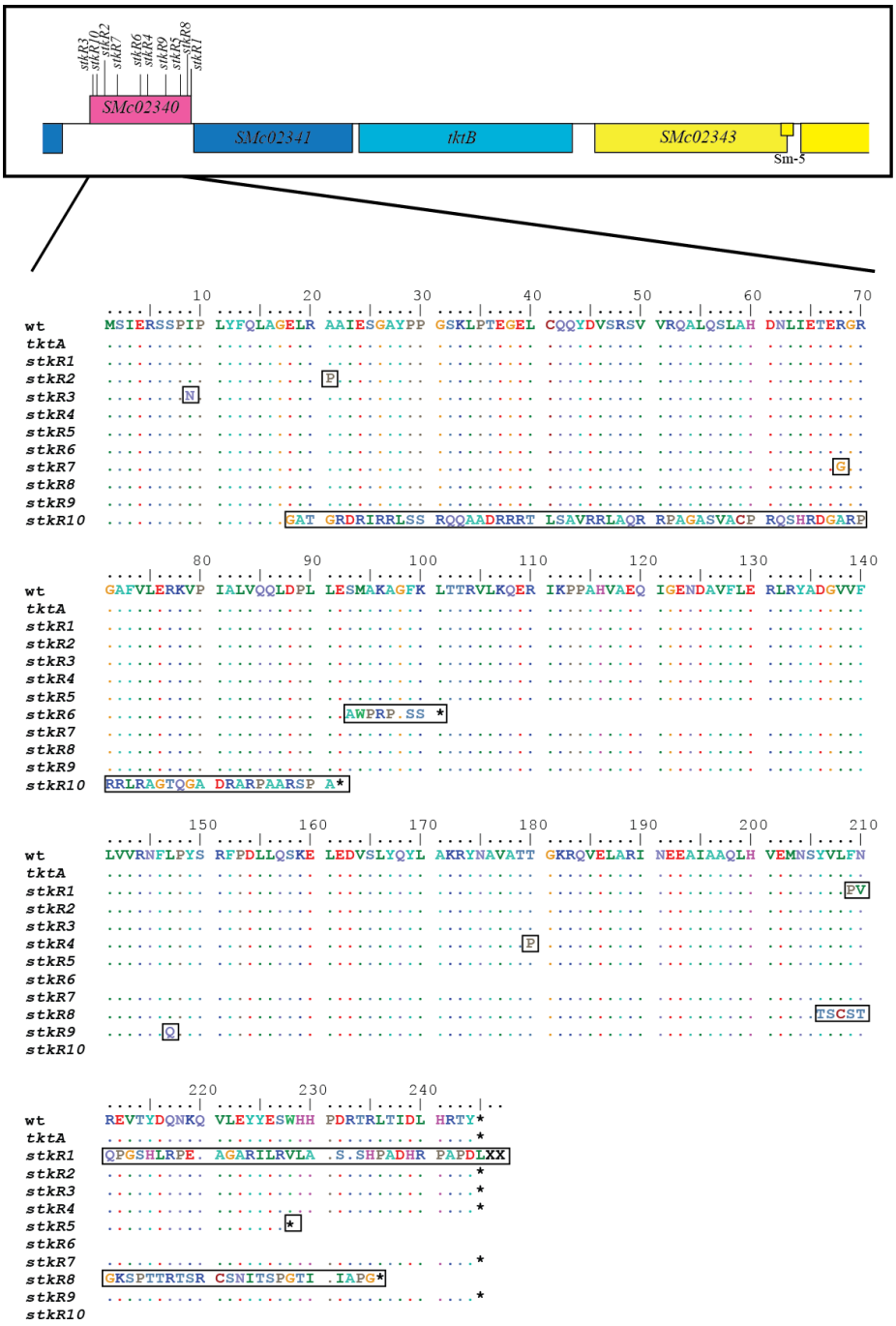


Figure 9. Location of single nucleotide polymorphisms (SNPs) in *SMc02340* and the predicted open reading frame (ORF) in 10 suppressor strains: *stkR1-stkR10*.

The top is the genetic locus representing *tktB* operon in *S. meliloti* with an approximate positioning of SNPs represented by vertical lines in each suppressor strain. Below is a *SMc02340* amino acid sequence alignment of the predicted ORFs of the wild type, the *tktA* mutant strain and 10 suppressor strains. Positions with residues identical to the wild-type are represented as dots. Rectangles highlight the position and nature of the SNP. * represents the stop codon.

missense protein. One of the mutations was a single base pair change that resulted in a premature stop codon. Three of these mutations found in *stkR2, stkR3,* and *stkR7* were found in the helix turn helix region. Of these, R68 is one of the 19 conserved amino acids that defines the GntR family of transcriptional regulators (Suvorova et al. 2015). The remaining two mutations that led to single amino acid changes, *stkR4* and *stkR9*, are found within the C- terminal UTRA (**U**biC **t**ranscription **r**egulator **a**ssociated) domain which is characterized as a ligand binding domain that modulates transcription factor activity (Fillenberg et al. 2016).

2.3.6 Strains carrying tktA mutations do not form colonized curled root hairs.

Previous characterization of SRmD397, which carries a transketolase mutation, had shown that it caused a delayed in nodulation, formed fewer nodules per plant, and was unable to fix nitrogen when inoculated on alfalfa (Hawkins et al. 2018). Since the transketolase reaction is a central step in the PP pathway, the absence of this step can affect multiple physiological processes. For example, SRmD397 is unable to synthesize aromatic amino acids because of the inability to produce chorismate (Hawkins et al. 2018), and it is known that lesions in *trpE(G)* lead to an altered symbiotic association (Barsomian et al. 1992; Prasad et al. 2000). Transketolase mutants were also found to have altered succinoglycan production (Hawkins et al. 2018), and many exopolysaccharide mutants are unable to form a symbiotic association (Jones et al. 2007b).

To further characterize early interactions, plants inoculated with strains carrying the transketolase mutation, or the *tktA* suppressor strains were compared to plants inoculated with strains carrying mutations in *trpE*(*G*), *exoR*, and *exoY*. The strain with mutation in *trpE*(*G*) has been shown to be capable of nitrogen fixation and is affected early in bacteroid development (Barsomian et al. 1992; Prasad et al. 2000). An *exoR* mutation was isolated as succinoglycan

overproducing and has been characterized as having reduced nodulation efficiency and reduced root hair colonization (Doherty et al. 1988; Yao et al. 2004). *exoY* encodes for a galactosyl-transferase which is necessary for the initiation of succinoglycan biosynthesis. An *exoY* mutant strain is able to form curled colonized root hairs but forms predominantly aborted infection threads on alfalfa (Maillet et al. 2020).

Plants inoculated with the wild type typically developed between 2-10 nodules per plant. Over 10 plants, 80 nodules were found. In the vicinity of the nodule primordia, 73 CCRH were found and of these 63 had developed infection threads (Figure 10). The phenotypes of *exoR, exoY,* and *trpE(G)* mutant strains are in agreement to what has been already reported (Barsomian et al. 1992; Cheng and Walker 1998; Yao et al. 2004). Strains carrying either *exoR* or *exoY* mutations were severely affected in the number of curled colonized root hairs and the number of infection threads they produced in comparison to the wild type. Over 10 plants inoculated with Rm7055 which carries *exoY55*, 85 nodule primordia with 29 curled colonized root hairs were found (Figure 10) and a total of 10 aborted infection threads were detected. A total of 76 nodule primordia were detected over 10 plants inoculated with Rm7095 which carries the *exoR95* allele. Similar to what has been previously observed (Yao et al. 2004), only 5 CCRH were detected, and in each case an infection thread was visible. Plants inoculated with the *trpE(G)* allele formed what appeared to be normal looking nodules. Over 10 plants a total of 41 nodules were examined. In each case a CCRH and extended infection threads were observed. Plants inoculated with SRmD397 typically produced misshapen nodules (Figure 10). Curled root hairs could be found microscopically in the vicinity of these nodule masses or along the plant root; however, examining well over 100 nodules in multiple experiments a single colonized curled root hair or

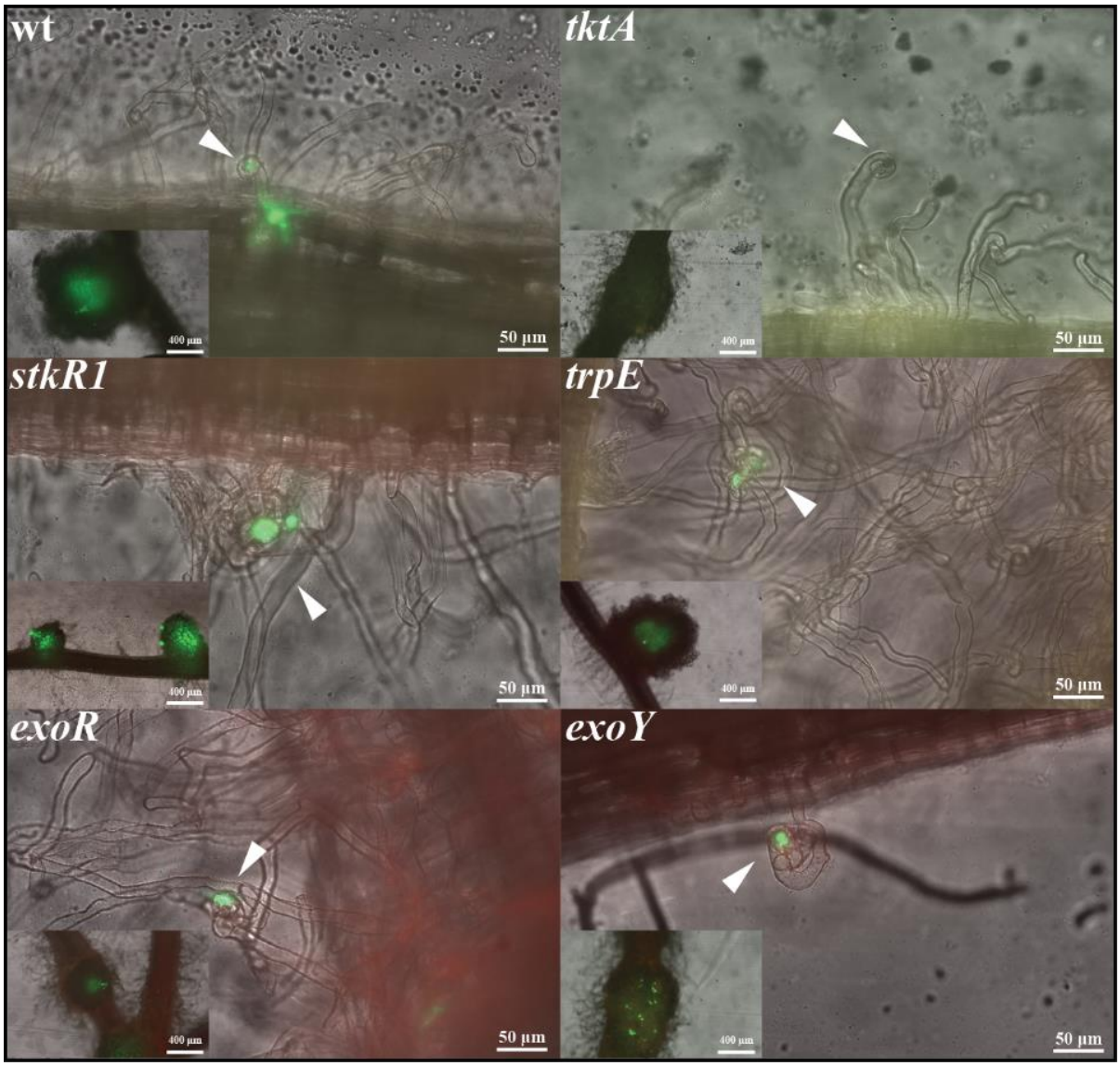


Figure 10. Images of alfalfa root hairs and root nodules produced by GFP-labelled *S. meliloti* strains; the wild type, the *tktA* mutant strains, *stkR1*, *ΔtrpE*, *exoR* and *exoY* mutant strains.

White arrow heads indicate the formation of curled root hair with bacteria entrapped in some of them. The merged images were created from differential interface of contrast (DIC), eGFP and Rhodamine filter images using Zeiss ZENBlue 2.0 software.

the presence of a single infection thread was never observed (Figure 10). In contrast, curled colonized root hairs and infection threads were apparent on plants inoculated with the suppressor strains (Figure 10). The numbers of these were similar to those found with the wild type (Figure 10).

We observed that nodules formed by plants inoculated with the *tktA* allele appeared dark, suggesting that nodules were devoid of bacteria (Figure 10). To confirm this, 10 nodules from plants inoculated with either the wild type or the transketolase mutant were picked, surface sterilized, crushed, and plated. Whereas bacteria could readily be isolated from the wild type and suppressor nodules, no bacteria were isolated from nodules formed by the transketolase mutant strain. Collectively, these results suggest that the symbiotic block is early, and it does not resemble the phenotype of either an exopolysaccharide or tryptophan auxotroph.

2.4 Discussion

The loss of transketolase presents a major block to central carbon metabolism that has been quantified metabolomically (Hawkins et al. 2018). This lesion affects the ability of *S. meliloti* to acidify its medium (Hawkins et al. 2017b), synthesize aromatic amino acids (Hawkins et al. 2018), reduces its ability to compete for iron acquisition (Hawkins et al. 2018), reduces succinoglycan production, as well as causes a severe symbiotic defect (Hawkins et al. 2018).

The data show that suppressors to *tktA* can be readily isolated and that they revert all of the tested phenotypes by the upregulation of *tktB* (Figure 3, Figure 4, Figure 6, Figure 7, Table 5). The *S. meliloti* genome contains four putative genes encoding transketolase activity (Capela et al. 2001). These include *tktA* (*SMc03978*), *tktB* (*SMc02342*), *cbbT* (*SMb20200*), as well as putative transketolase containing an alpha and beta chain encoded by *SMc00270* and *SMc00269* respectively. With respect to sequence identity, it is clear that all of these can encode proteins

that contain transketolase domains, but that SMc00269 and SMc00270 only have partial identity to TktA, TktB, and CbbT. These contain putative dehydrogenase domains suggesting that these proteins may carry out a transketolase-type reaction, but their primary activity is not likely to use xylulose-5-phosphate as a keto-donor and either ribose-5-phosphate or erythrose-4-phosphate as keto-acceptor substrates. Of the remaining transketolases, TktB has been shown to complement *tktA* (Hawkins et al. 2018), or in the case of CbbT to play a functional role during autotrophic growth (Pickering and Oresnik 2008); both presumably are capable of phenotypically suppressing the loss of *tktA*.

It is striking that all ten independent suppressors mapped to the open reading frame of *stkR* (*SMc02340*). Five of these mutations led to single amino acid substitutions (Table 6). The *stkR2* (A21P) and *stkR3* (I9N) mutations are within the helix-turn helix motif and most likely lead to an inability to bind DNA, thus leading to up-regulation of *tktB*. The *stkR4* (T180P) and *stkR9* L147Q alleles are in the C-terminal domain associated with substrate binding, and these mutations led to changes within the protein that presumably cause an inability to bind a consensus sequence on a substrate. We did not find any mutations in the putative promoter region between *SMc02343* and *tktB*, upstream of *stkR*, or that activate the operon containing the genes necessary for Calvin Benson Bassam cycle which includes *cbbT* that is found on pSymB (Finan et al. 2001).

The phenotypic suppression that was found with the *stkR* alleles never returned the strain to being physiologically equivalent to the wild type. Clearly in Rm1021, *tktB* is not expressed under standard growth conditions (Figure 8). However, once it is up-regulated, the suppressor strains appear to grow normally, yet their metabolic profile is distinct (Figure 6), and the ability to produce exopolysaccharide, or to attain full symbiotic capabilities is lower than the wild type

(Figure 2, Figure 3, Figure 4, Figure 10). Collectively these data suggest that although TktB can bring about phenotypic suppression, it is not as efficient as TktA.

In *E. coli*, triose phosphate mutants that were originally isolated still retained 5% of their triose phosphate isomerase activity, and with this still had good growth on some defined media (Anderson and Cooper 1969), suggesting that full enzyme activity may not be necessary in all circumstances. Alternately in *S. meliloti* it was found that two genes *tpiA* and *eryH* (formerly *tpiB*) encoded proteins that were incapable of providing triose-phosphate isomerase activity although the true activity of EryH is to convert 3-tetrulose-4-phosphate to erythrulose 4-phosphate (Poysti et al. 2007; Barbier et al. 2014). Although the true nature of the kinetic activity of TktB is not known, either of these scenarios can possibly explain why TktB may not be able to completely substitute for TktA.

Our understanding of early *S. meliloti* interactions with its host have mostly been derived from the characterization of mutants affected in nodulation or exopolysaccharide production (Jones et al. 2007b; Gibson et al. 2008; Zipfel and Oldroyd 2017). A *tkt* mutant can produce exopolysaccharide and is clearly able to elicit curled root hairs as well as nodule-like structures suggesting that it also has the ability to produce Nod factor. However, we were unable to find a colonized curled root hair. An *exoR* mutation was originally published as being unable to produce colonized curled root hairs; this however has been shown to occur, albeit at a low frequency (Cheng and Walker 1998; Yao et al. 2004). In our hands we were unable to observe this event with a *tkt* mutation whereas we were able to find this using an *exoR* mutant, implying that if a *tkt* mutant strain is able to colonize a curled root hair it is at a lower frequency. We currently do not know whether the *tkt* mutant strains reach the apoplastic space provided by the curled root hair, or that they become non-viable upon entering the space. We do note that we

were able to observe *tkt* mutants on the roots suggesting that it is not likely the inability to colonize the root surface. Recently *cbbT*, the *tkt* orthologue in *R. leguminosarum*, was identified as being essential for the progression of nodule development through an INSeq analysis, suggesting that this phenotype is not solely restricted to the *S. meliloti*-alfalfa symbiosis (Wheatley et al. 2020).

The isolation of suppressors to characterize a mutant has proven to be successful strategy to provide insight into the nature of a mutation (Yarosh et al. 1989; Nagpal et al. 1992; Oresnik et al. 1994; Osterås et al. 1997; diCenzo et al. 2017). Our isolation of suppressors that have not completely restored symbiosis suggests that central carbon metabolism plays an uncharacterized role during the progression of nodule formation. We are currently using these suppressors as a genetic platform with which we can elucidate the role that *tktA* and carbon metabolism play in the overall physiology and progression of nodule development.

Chapter 3:

Characterization of *tktB* locus in *Sinorhizobium meliloti.*

Authors: Sabhjeet Kaur and Ivan J. Oresnik. This work was planned and executed by Sabhjeet Kaur.

3.1 Introduction

Central carbon metabolism in *Sinorhizobium meliloti* consists of a few major pathways. Hexose sugars are catabolised through the Entner Doudoroff (ED) pathway which is the predominant glycolytic pathway in *S. meliloti* (Stowers 1985) and gluconeogenesis is carried out through Embden Meyerhof Parnas (EMP) pathway (Martínez-De Drets and Arias 1972; Irigoyen et al. 1990; Finan et al. 1988). The presence of a functional tri-carboxylic acid (TCA) cycle has been established as well (Dunn 1998; Stowers 1985).

Pentose sugars are metabolised primarily through a fully functional pentose phosphate (PP) pathway in *S. meliloti.* The PP pathway is divided into two branches; an oxidative and a non-oxidative branch (Geddes and Oresnik 2014). The oxidative branch consists of reactions where glucose-6-phosphate is converted in to 6-phosphoglucononate via the formation of 6-phopshogluconolactone and is shared with the ED pathway. These reactions play a role in energy production by generating NADPH. The PP pathway is separated from the ED pathway by the first dedicated enzyme 6-phosphogluconate dehydrogenase (Gnd) that converts 6-phosphogluconate to ribulose 5-phosphate and CO_2. The generation of ribulose 5-phosphate serves as an entry point to the non-oxidative branch (Martínez-De Drets and Arias 1972). The role of the non-oxidative branch is to cycle phosphorylated sugars through central carbon metabolism and to generate substrates for other anabolic processes. These biosynthetic reactions are purine and pyrimidine synthesis, aromatic amino acid biosynthesis, and the generation of a number of co-factors and vitamins (Geddes and Oresnik 2014; Hawkins et al. 2018).

The two enzymes that play a key role in the interconversion of the phosphorylated sugars in the PP pathway are transketolase and transaldolase. Genes encoding for these two enzymes have been characterized for their impact on central carbon metabolism in *S. meliloti* Rm1021

(Hawkins et al. 2018). Mutations in *tktA*, which encodes the transketolase for *S. meliloti*, has severe pleiotropic effects on metabolism as well as its ability to establish a symbiotic association with its host plant *Medicago sativa* (alfalfa) (Hawkins et al. 2018).

Hierarchical clustering of metabolomic data has shown that the central carbon metabolome of a *tal* (encodes for transaldolase) mutant more closely resembles the wild type than that of a *tktA* mutant strain (Hawkins et al. 2018). However, it was noted a *tal* mutant was slightly impaired for its ability to utilize ribose as a sole carbon source on defined medium, but this mutation did not affect symbiosis (Hawkins et al. 2018). Ribose, xylose as well as polyols like L-arabitol, adonitol and erythritol, enter central carbon metabolism through the PP pathway (Geddes and Oresnik 2014; Geddes et al. 2010). The locus responsible for L-arabitol, adonitol and erythritol metabolism was identified and characterized in *S. meliloti* (Geddes et al. 2010; Geddes and Oresnik 2012a). Ribose and xylose must be phosphorylated before entering the PP pathway. Ribose kinase and xylose kinase mutants were isolated in *S. meliloti* L5-30 that were unable to grow on ribose and xylose, respectively (Duncan 1981). The mutant strains also lacked kinase activity in cell free extracts (Duncan 1981). A gene encoding for a xylulose kinase (*xylB*) was identified and characterized along with a xylose isomerase gene (*xylA*) in *S. meliloti* (Geddes and Oresnik 2012a); however, the gene encoding a ribose kinase has not been functionally confirmed. In the *S. meliloti* genome database, there is a gene *rbsK* (*SMc01103*) that is annotated as a putative ribokinase. In addition, a putative ribose 5P isomerase (*rpi*) has been identified. These genes have not been functionally verified (Galibert et al. 2001; Capela et al. 2001).

An alternate "secondary" metabolic route has also been suggested for ribose and xylose metabolism (Geddes and Oresnik 2012b). Ribose kinase and xylose isomerase mutants of *S. meliloti* were able to incorporate ^{14}C-labelled ribose and xylose at wild type levels (Duncan

1981). Pyruvate carboxylase (*pyc*) mutants of *S. meliloti* are unable to grow on sole carbon sources that are metabolized to pyruvate, meaning these mutants cannot affect glycolysis (Geddes and Oresnik 2012b). Similarly, *S. meliloti* strains that carry lesions in triosephosphate isomerase (*tpiA tpiB/eryH*) cannot grow on compounds that are needed to carry out gluconeogenesis. However, it has been shown that in *S. meliloti*, a *pyc tpiA tpiB* triple mutant strain is able to grow on ribose and xylose (Geddes and Oresnik 2012a). Collectively these data support the hypothesis of an alternate pathway linking the PP pathway and the TCA cycle.

The only other *S. meliloti* strains unable to grow on ribose and xylose are ones with mutations in *tktA* (Hawkins et al. 2018). Second site suppressor strains of the *tktA* mutant were isolated that can grow on ribose and xylose. Identification and characterization of these suppressor strains revealed that an upregulation of *tktB* (a minor transketolase) is responsible for the suppressor phenotype (Chapter 2). The expression of *tktB* is regulated by a negative regulator, *SMc02340.* We have already established that *tktB* plays a role in partially restoring the PP pathway. Interestingly, the *tktB* locus also contains an uncharacterized carbohydrate kinase gene (*SMc02341*) just downstream of *tktB* which seems to be a part of the *tktB* operon. In this chapter, it is hypothesized that *tktB* and *SMc02341* may encode a metabolic bypass that plays a role in pentose catabolism. We describe the phenotypes associated with *tktB* operon and other neighbouring genes.

3.2 Materials and methods

3.2.1 Bacterial strains, plasmids, and media

The bacterial strains and plasmids used in the study as listed in Table 7. *E. coli* strains were grown on Luria-Bertani (LB) as complex medium at 37ºC with appropriate antibiotics

Table 7. Strains and plasmids list

Strain or plasmid	Genotype and relevant phenotype[a]	Reference
Strains		
S. meliloti		
Rm1021	SU47 *str-21* Sm^r	(Meade et al. 1982)
SRmD397	Rm1021, *tktA*::Tn*5* Sm^r Nm^r	(Hawkins et al. 2018)
SRmD410	Rm1021, *tktA*::Tn*5, stkR1* Sm^r Nm^r	Chapter 2
SRmD466	Rm1021, *tktA*::Tn*5, stkR3* Sm^r Nm^r	Chapter 2
SRmD660	SRmD397, Δ*SMc02340* Sm^r Nm^r	Chapter 2
SRmD661	Rm1021 Δ*tktB* Sm^r	Chapter 2
SRmD662	SRmD410 Δ*tktB* Sm^r Nm^r	Chapter 2
SRmD652	SRmD397, *SMc02340*::pKnock-Gm	Chapter 2
SRmD655	Rm1021 pKnock-Gm::*SMc02341* Sm^r Gm^r	This work
SRmD656	SRmD466 pKnock-Gm::*SMc02341* Sm^r Nm^r Gm^r	This work
SRmD681	SRmD466 Δ*SMc02341* Sm^r Nm^r	This work
SRmD690	Rm1021 Δ*SMc02341* Sm^r Nm^r	This work
SRmD689	SRmD410 pKnock-Gm::*SMc02343* Sm^r Nm^r Gm^r	This work
SRmD702	SRmD410 pKnock-Gm::*SMc02339* Sm^r Nm^r Gm^r	This work
SRmD703	Rm1021 pKnock-Gm::*SMc02339* Sm^r Gm^r	This work
E. coli		
MM294A	*pro-82 thi-1 hsdR17 supE44*	(Finan et al. 1986)
MT607	MM294A *recA56*	(Finan et al. 1986)
MT616	MT607(pRK600)	(Finan et al. 1986)
DH5α	λ^- *ϕ*80d*lacZ*ΔM15 Δ(*lacZYA-argF*)*U169 recA1 endA1 hsdR17*(r_k^- m_k^-) *supE44 thi-1 gyrA relA1*	(Hanahan 1983)
DH5αλpir	λ*pir* lysogen of DH5α	(House et al. 2004)
Plasmids		
pRK7813	Broad-host-range cloning vector; Tc^r	(Jones and Gutterson 1987)
pCO37	pRK7813 containing *attB* sites; Gateway-compatible destination vector	(Jacob et al. 2008)

pMK2014	FRT-*ccdB*-Cm^r-FRT cassette Pen^r	(House et al. 2004)
pXINT129	*λint* and *xis* driven by P_{lac}; Km^r	(Platt et al. 2000)
pJQ200SK	Gene replacement suicide vector Gm^r	(Quandt and Hynes 1993)
pKnock-Gm	Suicide vector for insertional mutagenesis; R6K *ori* RK4 *oriT* Gm^r	(Alexeyev 1999)
pSK03	400bp internal fragment of *SMc02341* cloned in pKNOCK-Gm, Gm^r	This work
pSK09	pJQ200SK/*SMc02341* flanking regions Gm^r	This work
pSK13	400bp internal fragment of *SMc02343* cloned in pKNOCK-Gm, Gm^r	This work
pSK17	400bp internal fragment of *SMc02339* cloned in pKNOCK-Gm, Gm^r	This work
pSK06	pCO37/*SMc02341* Tc^r	This work
pJH109	pCO37/*tktB* Tc^r	(Hawkins et al. 2018)

[a] Sm^r, streptomycin resistant; Nm^r, neomycin resistant; Gm^r, gentamicin resistant; Cm^r, chloramphenicol resistant; Tc^r, tetracycline resistant; Pen^r, penicillin resistant; Km^r, kanamycin resistant; FRT, flippase recognition target.

(Davis et al. 1980). *S. meliloti* strains were grown at 30ºC on LB as complex and Vincent's minimal media (VMM) as a defined medium (Vincent 1970). The carbon sources were filter sterilized and used at a final concentration of 15 mM. Antibiotics were filter sterilized prior to use at following concentrations: streptomycin (Sm), 200 μg/ml; neomycin (Nm), 200 μg/ml; kanamycin (Km), 200 μg/ml; chloramphenicol (Cm), 20 μg/ml; tetracycline (Tc), 5 μg/ml; gentamicin (Gm), 60 μg/ml *S. meliloti* and 20 μg/ml for *E. coli.*

3.2.2 RNA extraction, cDNA synthesis and qRT-PCR

RNA extraction and cDNA synthesis were performed as described previously (Geddes and Oresnik 2012b). Bacterial cultures were grown in LB overnight and 1 ml cultures were harvested at OD_{600}~ 0.8. The cell pellet was resuspended in 200 μl of Tris-HCl pH 8 buffer with 4 mg/ml lysozyme. RNA extraction was carried out using Qiagen RNeasy isolation kits and the samples were treated with RNase free DNase to remove DNA contamination. The RNA concentration was determined spectrophotometrically. One μg RNA was used to synthesize cDNA using VILO® cDNA synthesis kit from Invitrogen. Approximately 200 ng of cDNA was used to carry out the qRT-PCR using the SYBR green RT-PCR kit from Invitrogen. The quality of cDNA was determined spectrophotometrically. The data from qRT-PCR analysed by 2^-($^{\Delta\Delta}$Ct) method. The gene encoding a hypothetical protein, *SMc00128,* was used as an internal control (Krol and Becker 2004; Geddes and Oresnik 2012b).

3.2.3 Genetic manipulations and techniques

Plasmid isolation, restriction digestion, ligation, PCR amplification and gel electrophoresis were carried out using standard protocols (Sambrook et al. 1989). Conjugations and the genetic manipulations were carried out as described previously (Finan et al. 1988). Primers used are listed in Table 8. Mutations in *SMc02341* were generated by inserting a suicide plasmid vector

Table 8. Primer sequences for insertional mutations and gene deletion constructs

Primer #	Primer sequence 5'-3'	Primer name
	***SMc02341*:: pKnock-Gm**	
1	ACTG**GGATCC**GCTCGCCAATCCCGTTGTC	pKnock_SMc02341_BamHI_F
2	ACTG**AAGCTT**GATGGTGGAGCCGGGCTTC	pKnock_SMc02341_HindIII_R
3	TCGCGGCATCGGCCTCTCG	pKnock _SMc02341_confir_F
4	CGAAGCCGGAGCGGTACATT	pKnock _SMc02341_confir_R
	***SMc02339*:: pKnock-Gm**	
5	ACTG**GGATCC**CATGCGATCCTTCAGGCGAA	pKnock_SMc02339_BamHI_F
6	ACTG**AAGCTT**TCGCGCCCACCTACATCGA	pKnock_SMc02339_HindIII_R
7	TTGCCATCATCGAACGTGATC	pKnock_SMc02339_confir2_F
8	GTTGAGCGGCGTTTCGATGT	pKnock_SMc02339_confir2_R
	***SMc02343*:: pKnock-Gm**	
9	ACTG**GGATCC**CTGGATCACCACTCACAGCG	pKnock_SMc02343_BamHI_F
10	ACTG**AAGCTT**AACAGCAGCGGCAGCAGGAA	pKnock_SMc02343_HindIII_R
11	ATCACCTGGCCGGCCCTG	pKnock _SMc02343_confir_F
12	CAAGACGAGCGACCCTGTC	pKnock _SMc02343_confir_R
	Δ*SMc02341*	
13	tcctgcagcccggggtagggataacagggtaatGAATGCC GCACCGGTCTC	Δ*SMc02341*_left_ISceI_F
14	gctgatgcGCGGAGCAGATCGGCGAA	Δ*SMc02341*_left_R
15	tgctccgcGCATCAGCCGTGCGAGCA	Δ*SMc02341*_right_F
16	gaacaaaagctggagctGCTTTCCCCATCATTGG ATCATCG	Δ*SMc02341*_right_R
17	CAGCGCAGGACCGTGACTAT	Δ*SMc02341*_confir2_F
18	ACGTCTCCCTCTATCAGTATCT	Δ*SMc02341*_confir2_R
	qRT-PCR Primers	
19	AAACGATGCGGTGTTTCTGG	SMc02340_qRT_F
20	CTCCTTCGACTGCAACAGGT	SMc02340_qRT_R
21	AGATCGGCCTTGTCCTTGAA	tktB_qRT_F
22	AGCACGCTCTCCTGATAGTC	tktB_qRT_R

23	GCCCATCGTCGAATCCCATT	SMc02341_F_qRT
24	AAGCCGTATCGGAAAGACCT	SMc02341_R_qRT
25	CACGAGCTCGAGATCGCCAT	SMc00128_qRT_F
26	AAATGGACAGCTTCTTTCCGC	SMc00128_qRT_F
27	CGAAGGGCTACGACGTCTC	SMc02339_qRT_F
28	GCATTGTTGACGAGGATGGA	SMc02339_qRT_R

Restriction sites are shown in bold; lower case letters represent the overlapping regions of the Gibson assembly plasmid; the I-SceI endonuclease site is underlined in primer 13.

pKnock-Gm in the genome by single crossover (Alexeyev 1999). The plasmid vector construct was made by amplifying a 400-bp region from the middle of the *SMc02341* sequence using PCR and primers 1 and 2 (Table 8). The primer 1 contained BamHI and primer 2 contained HindIII restriction sites on the 5' ends. These endonucleases were used to restrict the pKnock-Gm plasmid. The amplified DNA fragment was gel isolated, restricted, and ligated with T4 DNA ligase to pKnock-Gm yielding pSK03. The insertional mutations were made by conjugating the plasmid to the appropriate *S. meliloti* strain and selecting for Gm resistance. pSK03 was recombined in Rm1021 and SRmD466, yielding strains SRmD655 and SRmD656, respectively. The insertions were confirmed by PCR using primers 3 and 4 (Table 8). PCR product was sequenced to confirm the construct.

An insertional mutation in *SMc02339* was similarly constructed. Briefly, a 400-bp long internal fragment of the *SMc02339* was amplified by PCR using primers 5 and 6 (Table 8). The amplified DNA fragment was cloned into the pKnock-Gm plasmid vector yielding pSK17. pSK17 was recombined into Rm1021 and SRmD410 generating strains SRmD703 and SRmD702, respectively. The mutations were confirmed by PCR amplification using primers 7 and 8. The PCR products were confirmed by sequencing. The *SMc02343*::pKnock-Gm construct was made using primers 9 and 10 to amplify a 400-bp internal fragment. This was cloned into pKnock-Gm yielding plasmid pSK13. pSK13 was recombined in to SRmD410 producing strain SRmD689. The mutation was confirmed by generating a PCR fragment using primers 11 and 12 (Table 8). The fragment was subsequently sequenced to confirm proper construction.

Gene deletions were made using Gibson assembly constructs designed using NEBuilder Assembly tool (https://nebuilderv1.neb.com/) (New England Biolabs, Inc.). The primers were designed by taking flanking regions of the gene of interest and using pJQ200SK as a plasmid

vector backbone (Quandt and Hynes 1993). This plasmid contains an antibiotic marker (Gm) as well as a *sacB* marker to allow for screening of putative deletion events in the chromosome.

Briefly, for the Δ*SMc02341* construct a 600-bp upstream region of *SMc02341* was amplified using primers 15 and 16 (Table 8). To amplify 400-bp of the downstream region primers 13 and 14 were used where primer 13 included an I-SceI endonuclease site which could be used to induce double stranded breaks in DNA to increase the probability of a deletion event. All the PCR products were amplified using Q5 DNA polymerase (New England Biolabs, Inc.). The plasmid construct was assembled using the Gibson assembly master mix (New England Biolabs, Inc.) following the manufacturer's instructions. The assembled plasmid was integrated into the target genome via homologous recombination and putative deletions were screened using sucrose counterselection. The deletions were confirmed by PCR using primers 17 and 18 and subsequently sequenced.

Plasmids overexpressing *SMc02340, SMc02341* and *tktB* were constructed using the *S. meliloti* ORFeome library as described previously (House et al. 2004). The open reading frame (ORF) for the gene of interest was recombined into pCO37, a gateway compatible destination vector. Plasmids overexpressing *SMc02340, SMc02341* and *tktB* were named pSK05, pSK06, and pJH109, respectively.

3.2.4 Growth curves

Growth curve experiments were conducted in complex (LB) or defined (VMM) medium supplemented with the appropriate carbon source as described previously (Kohlmeier et al. 2019). Briefly, bacterial cultures were grown overnight in LB, washed with saline and sub-cultured into fresh medium. Aliquots of 200 µl of the diluted cultures were dispensed into the wells of a 96-well plate. The inside of the plate lid was coated with 0.05% TritonX in 20%

ethanol solution to avoid condensation (Brewster 2003). The experiment was conducted at 28 ºC with continuous orbital shaking. Optical density was recorded at 600 nm wavelength after every hour for 72 h on a BioTek Synergy 2 plate reader (BioTek Instruments, Inc., Winooski, VT, USA).

3.2.5 Phenotypic microarray

Rhizobial strains were subjected to phenotypic microarray study using PM01 and PM02A Biolog plates (Biolog Inc. CA, USA). Bacterial cultures were grown overnight in complex medium (LB), washed with saline and sub-cultured into fresh defined medium (VMM). Aliquots (100 µl) of the sub-cultured medium were dispensed into the wells the Biolog plate. The inside of the plate lid was coated with 0.05% TritonX in 20% ethanol solution to avoid condensation (Brewster 2003). The rhizobial strains were tested on 190 carbon sources for growth over 72 h using a BioTek Synergy 2 plate reader. Absorbance was recorded at 600 nm wavelength every hour. The data were analysed and plotted using the program DuctApe (Galardini et al. 2014). The phenome analysis in DuctApe was completed by using the commands provided by the authors in the DuctApe manual (https://combogenomics.github.io/DuctApe/howto.html).

3.3 Results

3.3.1 SMc02341 has a growth phenotype on ribose in a suppressor background.

Previous work suggested that *tktA* encodes for the major transketolase activity in *S. meliloti*, whereas *tktB* plays a minor role (Hawkins et al. 2018). In the last chapter it was shown that the second site suppressor mutations partially alleviate the metabolic and symbiotic block associated with a mutation in *tktA.* The symbiotic and metabolic phenotypes displayed by the suppressor strains are associated with an overexpression of *tktB* caused by point mutations in its

negative regulator *SMc02340.* Ten independently isolated second site suppressor mutant strains were all shown to contain single nucleotide polymorphisms (SNPs) in *SMc02340.* The genomic arrangement of *tktB* and *SMc02340* shown in Figure 11. The genes *tktB* and *SMc02341* appear to be arranged as an operon and these genes are convergently arranged with respect to the GntR type regulator *SMc02340* (Galibert et al. 2001; Capela et al. 2001).

Due to its proximity to *tktB* and predicted function as a carbohydrate kinase, it is plausible that *SMc02341* plays a role in metabolising pentoses along side *tktB.* It should also be noted that a gene encoding for ribose kinase activity is yet to be identified in *S. meliloti*, despite *rbsK* being annotated as a putative ribokinase (Duncan 1981; Geddes and Oresnik 2014). To test if *SMc02341* functions as a pentose kinase, insertional mutations were created in *SMc02341* in wild type background yielding a strain SRmD655 (Table 7). SRmD655 was tested for its growth phenotype on minimal medium containing ribose as a sole carbon source. However, SRmD655 did not display any growth defects in ribose. We have established in the prior chapter that SRmD661, a strain with Δ*tktB* in wild type background, did not show any growth defects on ribose as a sole carbon source in minimal medium. However, a strain with Δ*tktB* in a suppressor background (SRmD662) completely lost its ability to grow on ribose (Table 9).

To test if *SMc02341* is functioning in a similar fashion as *tktB* such that a carbon phenotype for *SMc02341* can only be identified in the suppressor background, a *SMc02341* insertional mutant was generated in a suppressor background. The resulting strain was named SRmD656 and when tested on ribose for growth, the mutant strain showed slow growth on minimal medium plates as compared to its parent suppressor strain SRmD466 (*stkR3*) (Table 9). Unlike the *tktB* mutant strain that showed no growth on ribose in the suppressor background, the

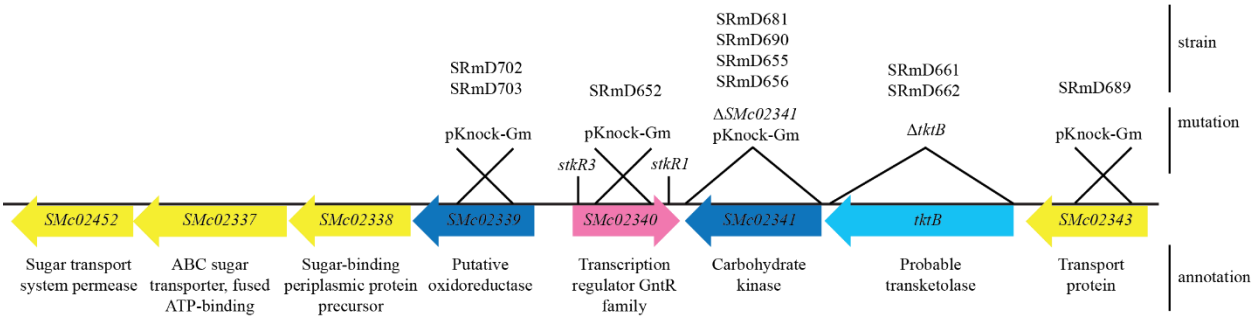


Figure 11. Genetic map of *tktB* locus in *S. meliloti*.

Arrows indicate ORFs and the direction of transcription, the colors indicate the predicted functional category of the gene according to the *S. meliloti* genome database (https://iant.toulouse.inra.fr/bacteria/annotation/cgi/rhime.cgi). Yellow, cell processes; dark blue, small molecule metabolism; light blue, central intermediary metabolism; pink, miscellaneous. Type of mutations created in different genes are described above the locus with strain numbers that contained these mutations. Crosses represent insertions; straight vertical lines represent point mutations; triangles represent deletions.

Table 9. Growth phenotype on ribose in minimal and complex medium

Strain	Genotype	Growth on defined medium		Growth on complex medium
		Glucose	Ribose	LB
Rm1021	Wild type	+	+	+
SRmD397	*tktA*::Tn*5*	-	-	+
SRmD410	*tktA*::Tn*5 stkR1*	+	+	+
SRmD466	*tktA*::Tn*5 stkR3*	+	+	+
SRmD661	Δ*tktB*	+	+	+
SRmD662	*tktA*::Tn*5 stkR1* Δ*tktB*	-	-	+
SRmD662 (pJH109)	*tktA*::Tn*5 stkR1* Δ*tktB* (*tktB*$^{+}$)	+	+	+
SRmD655	*SMc02341*::pKnock-Gm	+	+	+
SRmD656	*tktA*::Tn*5 stkR3 SMc02341*::pKnock-Gm	+	+/-	+
SRmD656 (pSK06)	*tktA*::Tn*5 stkR3 SMc02341*::pKnock-Gm (*SMc02341*$^{+}$)	+	+	+

Growth phenotypes were as follows: +, wild type growth; +/-, slow growth; -, no growth.

mutations in *SMc02341* slowed the growth on ribose. These data suggest that in a suppressor background, the auxotrophy associated with the mutations in *tktB* can be uncoupled by generating mutations in *SMc02341*. Also, that *SMc02341* does have a role in ribose metabolism but not an essential one as the ribose phenotype can only be recorded in a suppressor background. This is in agreement with the previous chapter where we have established that *tktB* does not play an essential role in the wild type, but it is able to partially rescue the phenotypes associated with mutations in *tktA.*

To generate more stable mutations in *SMc02341,* unmarked Δ*SMc02341* strains were created using a suicide vector pJQ200SK with the Gibson assembly approach in wild type and suppressor (*stkR3*) backgrounds, yielding strains SRmD690 and SRmD681, respectively. The deletions were confirmed by sequencing, which revealed that *SMc02341* was completely deleted but a small section (~280 bp) of the suicide plasmid vector was still integrated in the genome (Figure 12). As the downstream gene *SMc02340* is transcribed from the opposite strand and the growth phenotype of the deletion strain on ribose plates matches with the pKnock-Gm insertion strain SRmD656, no further changes were made to SRmD690 and SRmD681. These stable mutations were generated for plant dry weight assays and were also used to run bacterial growth curves in complex and minimal medium to further investigate the ribose phenotype associated with *SMc02341.*

To quantify the ribose phenotype associated with the *SMc02341,* the deletion strains were subjected to a growth curve assay in complex (LB) and minimal medium supplemented with ribose. The data in Figure 13 show the growth curve of wild type, a suppressor strain (*stkR3),* the *tktA* mutant strain and the strains with Δ*SMc02341* in wild type and suppressor backgrounds. The strain with Δ*SMc02341* in the suppressor background (SRmD681) showed significantly slower

ATCATCGTGTCAGCGCAGGACCNTNACTATCAGGAGAGCGTGCTGGCTTTCGG
CACGCCGACGGTATTTGTCGAGGCGGCCTCTCCGCTTTATTGGTACCAGTTCCT
CAAGGGACCGGGTACGGTCGTCGGCGTCGACCGTTTCGGCGAATCCGCGCCAG
GACCGGAAATCTACCGCGTCATGAACCTGACGGCTGAGCGGATCGTATCCGAG
GCCGGTCAGTTCCTGTGCTGACACGATGATCCAATGATGGGGAAAGCAGCTCC
AGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGC
TGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAA
GCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGT
TGCGCTCACTGCACTATAGGGCGAATTGGGTACCGGGCCCCCCCTCGAGGTCGACG
GTATCGATAAGCTTGATATCGAATTCCTGCAGCCCGGGTAGGGATAACAGGGTAAT
GAATGCCGCACCGGTCTCAATAGGTCCGGTGCAGGTCGATGGTCAGCCGGGTG
CGATCGGGATGGTGCCAGGACTCGTAATATTCGAGCACCTGCTTGTTCTGGTC
GTAGGTGACTTCCCGGTTGAACAGGACGTAACTGTTCATTTCCACATGTAACTG
GGCGGCGATCGCCTCTTCATTGATGCGGGCGAGCTCGACCTGCCGTTGCC

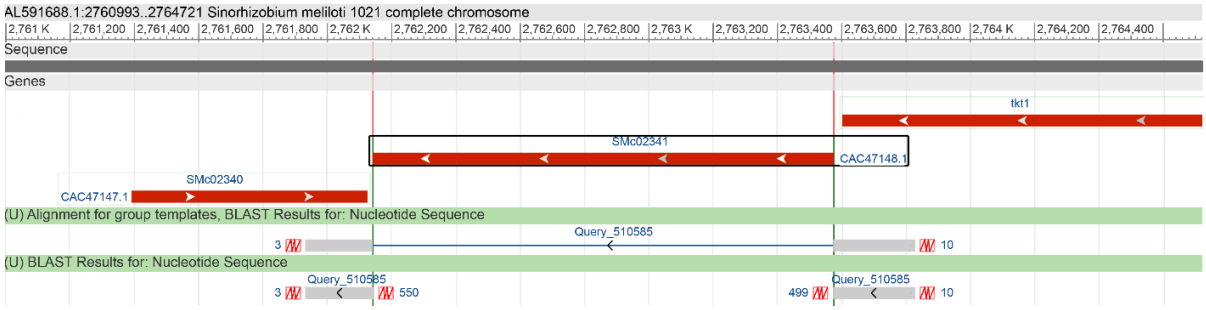


Figure 12. Confirmation of *SMc02341* deletion.

On the top panel is the nucleotide sequence of the *SMc02341* flanking region. The underlined sequence represents the parts of suicide vector that were left behind during homologous recombination. The bold sequence part represents the nucleotide sequence that is found flanking *SMc02341*. indicating the entire gene was successfully deleted. The bottom panel represents a Basic Local Alignment Search Tool (BLASTn) analysis of the above sequence to *S. meliloti* reference sequence (AL591688). Horizontal red bars indicate the genomic arrangement of *SMc02340, SMc02341* and *tktB* genes. Grey horizontal bars represent the flanking sequence of Δ*SMc02341* strain successfully aligned to the reference genome. The thin blue horizontal line indicates the missing sequence which indicates that SMc02341 was deleted.

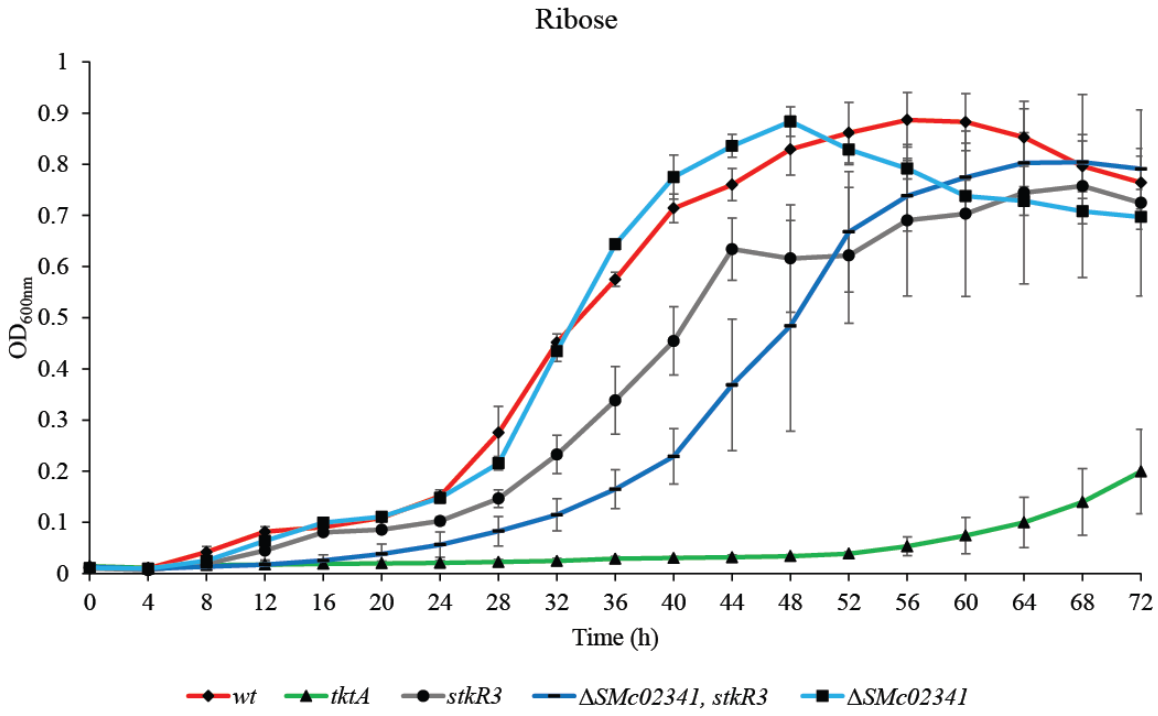


Figure 13. Growth curve analysis of *S. meliloti* strains.

wt (red), *tktA* (green), *stkR3* (grey), Δ*SMc02341*, *stkR3* (dark blue) and Δ*SMc02341* (light blue) with ribose as sole carbon source in Vincent's minimal medium. The error bars represent standard deviation of 3 biological replicates.

growth on ribose during the log phase as compared to the suppressor strain or the wild type. Growth rate for all the strains was calculated by taking log phase into account. The average log phase growth rate for SRmD681 was 0.3631/h whereas the suppressor strain had a growth rate of 0.6248/h and the wild type with 0.7428/h. Evident from the growth curve (Figure 13), SRmD681 does catch up to the wild type after 60 hrs. The strain with Δ*SMc02341* in wild type background (SRmD690) did not show any growth defects during the assay and displayed a growth rate of 0.822/h. Slow growth on ribose supported the evidence that *SMc02341* does play a role in ribose metabolism but only in the suppressor background and not wild type. As published previously, the *tktA* mutant strain was unable to grow on ribose (Hawkins et al. 2018).

3.3.2 SMc02340 negatively regulates its flanking operons

The regulation of *stkR* has been shown to affect transcription of *tktB* (Chapter 2). It is not uncommon for a regulator to regulate divergently transcribed genes (Richardson et al. 2004). To determine if *SMc02340* regulates divergently transcribed genes, transcription of *SMc02339* was measured using qRT-PCR in various backgrounds (Figure 14). In addition, it was also of interest to determine if *tktB* and *SMc02341* were co-ordinately regulated. Expression of these genes was tested in wild type, the *tktA* mutant strain and three suppressor strains carrying mutations in *SMc02340.* The results show that both *SMc02339* and *SMc02341* were significantly upregulated in all three suppressor strains carrying mutations in *SMc02340.* The *tktA* mutant strain however did not show any upregulation of these genes. A 40-to-100-fold increase in the expression of *SMc02339* in the suppressor strains was recorded. Similarly, *SMc02341* expression was increased by 7-to-26-fold in the suppressor strains. Although the expression of *SMc02340* is increased by a maximum of 10-fold in the suppressor strains, the values were only significant in

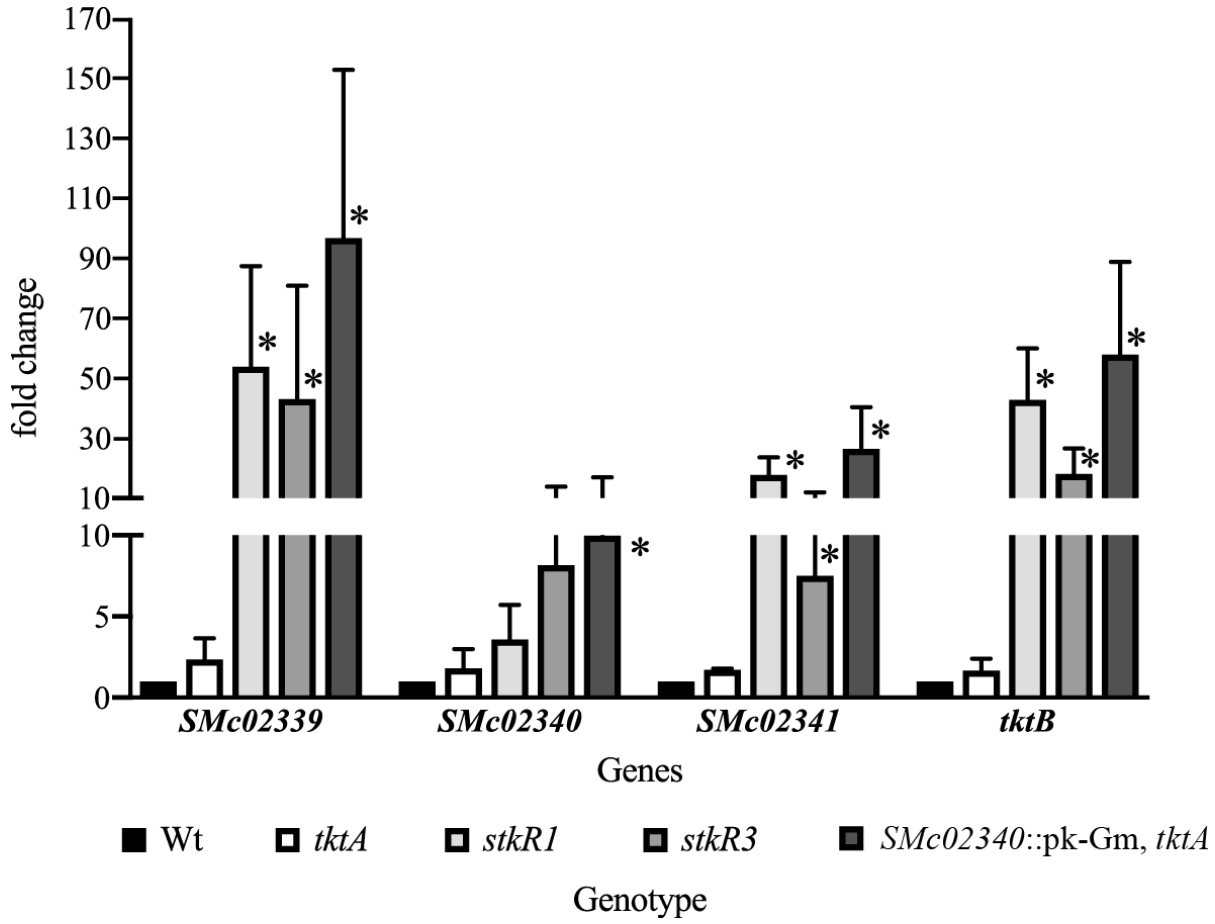


Figure 14. qRT-PCR analysis of *SMc02340* neighbouring genes expression in *S. meliloti* strains (wt, *tktA*, *stkR1*, *stkR3*, and *SMc02340*::pk-Gm *tktA*).

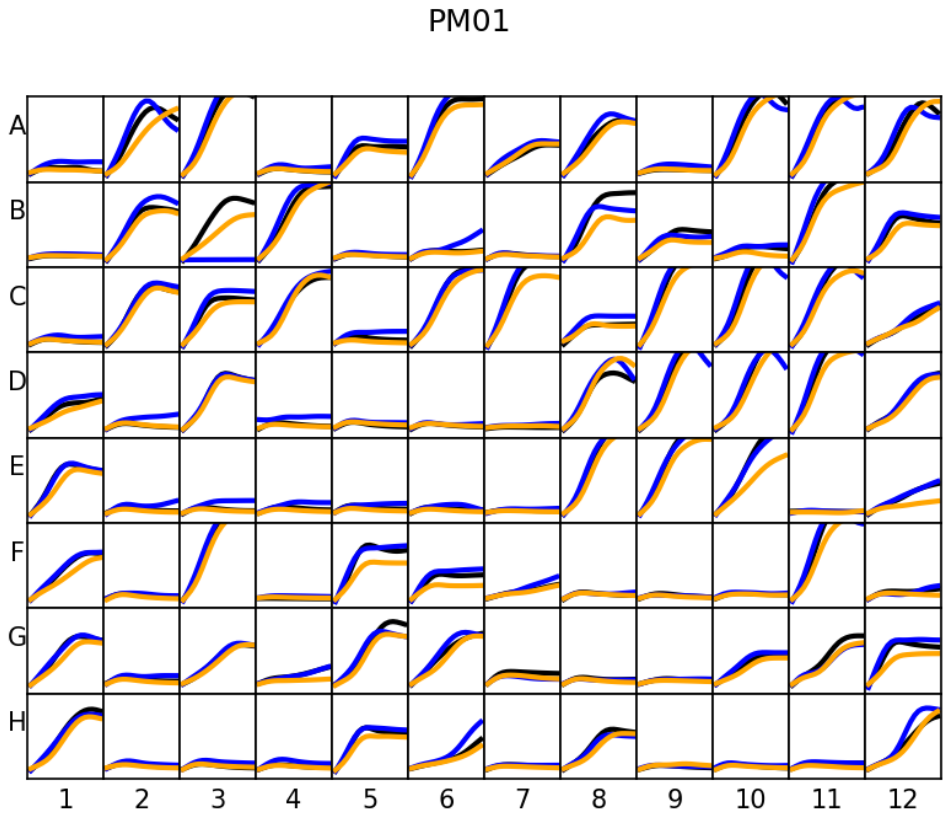
* Represents significant difference from the wt strain with a $P < 0.05$ with Student's *t*-test. The represented is an average of three biological replicates. Figure was created using GraphPad Prism software.

a suppressor strain that contains a pKnock-Gm insertion in *SMc02340.* The other two suppressor strains contain point mutations in *SMc02340.* In all three suppressor strains, the expression of *SMc02340* is higher than the wild type and the *tktA* mutant strain. Despite the statistical insignificance in two of the suppressor strains, *SMc02340* seems to autoregulate its expression. As expected, the expression of *tktB* increased by 18-to-58-fold. It has been already shown that the expression of *SMc02343,* which is present upstream of *tktB* is not affected in the suppressor strains. With the fact that *SMc02343* was not affected by *stkR,* these data suggest that *SMc02340* regulates the expression of only two of its convergently encoded genes i.e., *tktB* and *SMc02341* as well as the divergently transcribed *SMc02339,* and possibly other genes beyond *SMc02339.*

3.3.3 Phenotypic microarray

According to the *S. meliloti* genome database, *SMc02339* is a putative oxidoreductase. Three downstream genes of *SMc02339* (*SMc02452, SMc02337* and *SMc02338*) are annotated as parts of a putative sugar transporter (Capela et al. 2001). The genomic arrangement of these genes indicates that *SMc02339* is part of the same operon as its downstream genes. Assuming these genes are transcribed together, insertions in *SMc02339* should have a polar effect on the downstream genes. A pKnock-Gm suicide vector was integrated in *SMc02339* to disrupt its expression. The mutations were created in *SMc02339* in the wild type and a suppressor background, yielding strains SRmD703 and RmD702, respectively (Table 7).

To identify a potential substrate targets of *SMc02339,* strains SRmD702 and SRmD703 were subjected to a large-scale screening by growing on 190 carbon sources using PM01 and PM02A Biolog plates (Biolog Inc.). The growth patterns of *SMc02339* mutant strains were

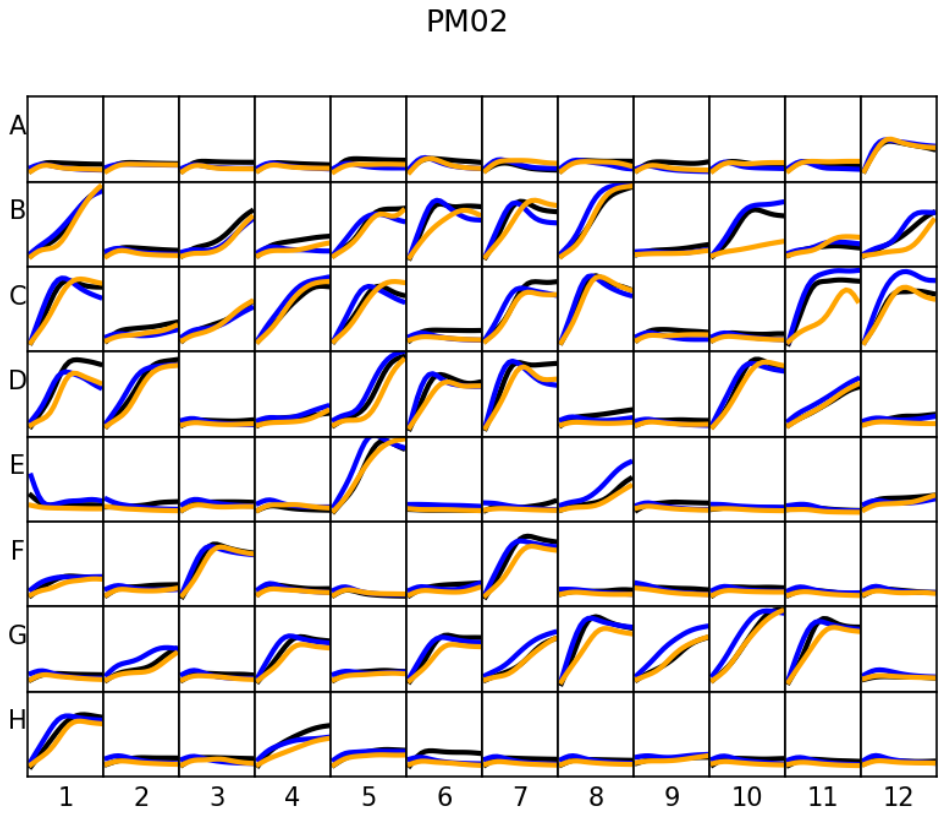


PM1 MicroPlate™ Carbon Sources

A1 Negative Control	A2 L-Arabinose	A3 N-Acetyl-D-Glucosamine	A4 D-Saccharic Acid	A5 Succinic Acid	A6 D-Galactose	A7 L-Aspartic Acid	A8 L-Proline	A9 D-Alanine	A10 D-Trehalose	A11 D-Mannose	A12 Dulcitol
B1 D-Serine	B2 D-Sorbitol	B3 Glycerol	B4 L-Fucose	B5 D-Glucuronic Acid	B6 D-Gluconic Acid	B7 D,L-α-Glycerol-Phosphate	B8 D-Xylose	B9 L-Lactic Acid	B10 Formic Acid	B11 D-Mannitol	B12 L-Glutamic Acid
C1 D-Glucose-6-Phosphate	C2 D-Galactonic Acid-γ-Lactone	C3 D,L-Malic Acid	C4 D-Ribose	C5 Tween 20	C6 L-Rhamnose	C7 D-Fructose	C8 Acetic Acid	C9 α-D-Glucose	C10 Maltose	C11 D-Melibiose	C12 Thymidine
D-1 L-Asparagine	D2 D-Aspartic Acid	D3 D-Glucosaminic Acid	D4 1,2-Propanediol	D5 Tween 40	D6 α-Keto-Glutaric Acid	D7 α-Keto-Butyric Acid	D8 α-Methyl-D-Galactoside	D9 α-D-Lactose	D10 Lactulose	D11 Sucrose	D12 Uridine
E1 L-Glutamine	E2 m-Tartaric Acid	E3 D-Glucose-1-Phosphate	E4 D-Fructose-6-Phosphate	E5 Tween 80	E6 α-Hydroxy Glutaric Acid-γ-Lactone	E7 α-Hydroxy Butyric Acid	E8 β-Methyl-D-Glucoside	E9 Adonitol	E10 Maltotriose	E11 2-Deoxy Adenosine	E12 Adenosine
F1 Glycyl-L-Aspartic Acid	F2 Citric Acid	F3 myo-Inositol	F4 D-Threonine	F5 Fumaric Acid	F6 Bromo Succinic Acid	F7 Propionic Acid	F8 Mucic Acid	F9 Glycolic Acid	F10 Glyoxylic Acid	F11 D-Cellobiose	F12 Inosine
G1 Glycyl-L-Glutamic Acid	G2 Tricarballylic Acid	G3 L-Serine	G4 L-Threonine	G5 L-Alanine	G6 L-Alanyl-Glycine	G7 Acetoacetic Acid	G8 N-Acetyl-β-D-Mannosamine	G9 Mono Methyl Succinate	G10 Methyl Pyruvate	G11 D-Malic Acid	G12 L-Malic Acid
H1 Glycyl-L-Proline	H2 p-Hydroxy Phenyl Acetic Acid	H3 m-Hydroxy Phenyl Acetic Acid	H4 Tyramine	H5 D-Psicose	H6 L-Lyxose	H7 Glucuronamide	H8 Pyruvic Acid	H9 L-Galactonic Acid-γ-Lactone	H10 D-Galacturonic Acid	H11 Phenylethyl-amine	H12 2-Aminoethanol

Figure 15. Phenotypic microarray analysis.

Rm1021 (wt; black); SRmD702 (*SMc02339*::pk-Gm, *stkR1*; orange), and SRmD703 (*SMc02339*::pk-Gm; blue) strains of *S. meliloti* on PM01 and Biolog plates. The grid at the bottom represents the location of each substrate as it is laid out on the 96 well plate. The bottom figure is adapted from the Biolog Inc. website.



PM2A MicroPlate™ Carbon Sources

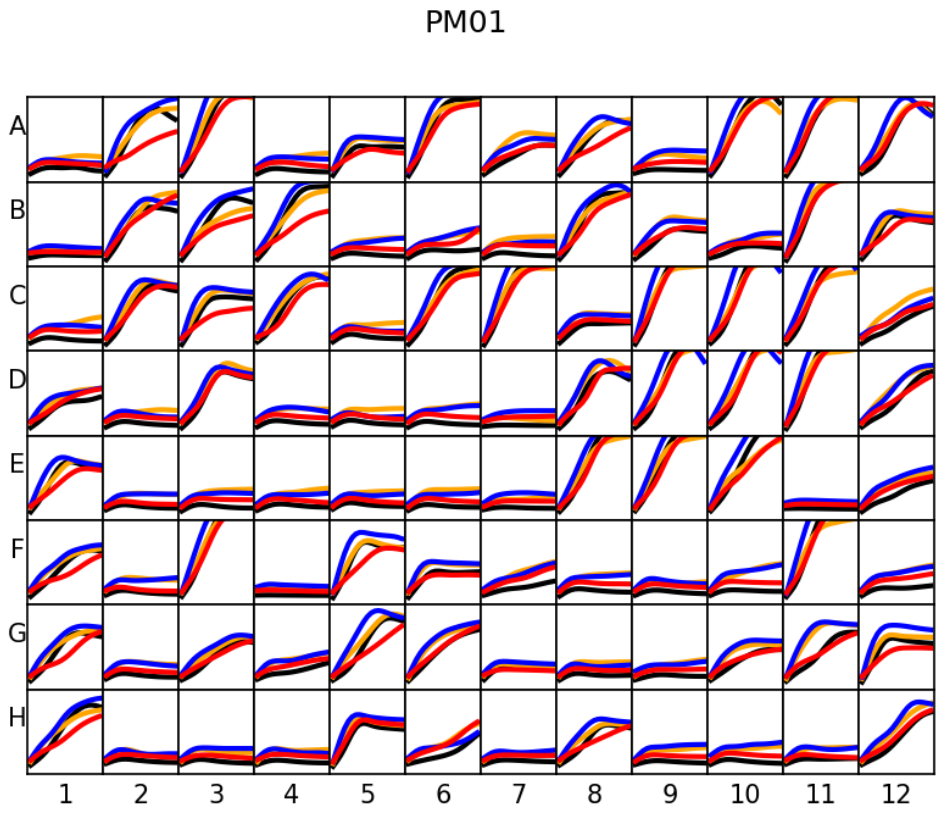
A1 Negative Control	A2 Chondroitin Sulfate C	A3 α-Cyclodextrin	A4 β-Cyclodextrin	A5 γ-Cyclodextrin	A6 Dextrin	A7 Gelatin	A8 Glycogen	A9 Inulin	A10 Laminarin	A11 Mannan	A12 Pectin
B1 N-Acetyl-D-Galactosamine	B2 N-Acetyl-Neuraminic Acid	B3 β-D-Allose	B4 Amygdalin	B5 D-Arabinose	B6 D-Arabitol	B7 L-Arabitol	B8 Arbutin	B9 2-Deoxy-D-Ribose	B10 i-Erythritol	B11 D-Fucose	B12 3-O-β-D-Galactopyranosyl-D-Arabinose
C1 Gentiobiose	C2 L-Glucose	C3 Lactitol	C4 D-Melezitose	C5 Maltitol	C6 α-Methyl-D-Glucoside	C7 β-Methyl-D-Galactoside	C8 3-Methyl Glucose	C9 β-Methyl-D-Glucuronic Acid	C10 α-Methyl-D-Mannoside	C11 β-Methyl-D-Xyloside	C12 Palatinose
D1 D-Raffinose	D2 Salicin	D3 Sedoheptulosan	D4 L-Sorbose	D5 Stachyose	D6 D-Tagatose	D7 Turanose	D8 Xylitol	D9 N-Acetyl-D-Glucosaminitol	D10 γ-Amino Butyric Acid	D11 δ-Amino Valeric Acid	D12 Butyric Acid
E1 Capric Acid	E2 Caproic Acid	E3 Citraconic Acid	E4 Citramalic Acid	E5 D-Glucosamine	E6 2-Hydroxy Benzoic Acid	E7 4-Hydroxy Benzoic Acid	E8 β-Hydroxy Butyric Acid	E9 Glycolic Acid	E10 α-Keto-Valeric Acid	E11 Itaconic Acid	E12 5-Keto-D-Gluconic Acid
F1 D-Lactic Acid Methyl Ester	F2 Malonic Acid	F3 Melibionic Acid	F4 Oxalic Acid	F5 Oxalomalic Acid	F6 Quinic Acid	F7 D-Ribono-1,4-Lactone	F8 Sebacic Acid	F9 Sorbic Acid	F10 Succinamic Acid	F11 D-Tartaric Acid	F12 L-Tartaric Acid
G1 Acetamide	G2 L-Alaninamide	G3 N-Acetyl-L-Glutamic Acid	G4 L-Arginine	G5 Glycine	G6 L-Histidine	G7 L-Homoserine	G8 Hydroxy-LProline	G9 L-Isoleucine	G10 L-Leucine	G11 L-Lysine	G12 L-Methionine
H1 L-Ornithine	H2 L-Phenylalanine	H3 L-Pyroglutamic Acid	H4 L-Valine	H5 D,L-Carnitine	H6 Sec-Butylamine	H7 D,L-Octopamine	H8 Putrescine	H9 Dihydroxy Acetone	H10 2,3-Butanediol	H11 2,3-Butanedione	H12 3-Hydroxy-2-Butanone

Figure 16. Phenotypic microarray analysis.

Rm1021 (wt; black); SRmD702 (*SMc02339*::pk-Gm, *stkR1*; orange), and SRmD703 (*SMc02339*::pk-Gm; blue) strains of *S. meliloti* on PM02A Biolog plates. The grid at the bottom represents the location of each substrate as it is laid out on the 96 well plate. The bottom figure is adapted from the Biolog Inc. website.

compared with the wild type strain (Figure 15 and Figure 16). Of all the carbon sources that were tested, the strain with mutations in *SMc02339* in wild type background (SRmD703, blue) and the wild type (black) showed similar growth patterns on all the carbon sources except for glycerol (B3, PM01) (Figure 15). SRmD703 showed no growth on glycerol on the Biolog plate; however, the phenotype could not be confirmed when the experiment was repeated (Figure 19). The strain with mutations in *SMc02339* in the suppressor background (SRmD702, orange) showed similar growth patterns to the wild type on almost all the carbon sources as well except erythritol. SRmD702 was unable to grow on erythritol (B10, PM02A) (Figure 16). When the experiment was repeated in the form of a growth curve analysis using erythritol as a sole carbon sources in minimal medium, the *ery* $^{-}$ phenotype was confirmed. However, a suppressor strain, SRmD410 (*stkR1*) that was used as a control did not grow on erythritol as well (Figure 19). All the strains were also tested for growth on minimal medium plates with erythritol as a sole carbon source. The strain with mutations in *SMc02339* in the suppressor background and a suppressor strain showed weak growth, whereas the strain with mutations in *SMc02339* in the wild type background showed the same growth pattern as the wild type strain. Taken together, the erythritol phenotype appears to be caused by *stkR* and not by mutations in *SMc02339*.

A similar analysis of *SMc02341* mutations was done using strains with Δ*SMc02341* in different backgrounds along with a suppressor strain and the wild type strain. Growth curves comparing all four strains are shown in Figure 17 and Figure 18. The results show that all the strains were able to grow on most of 190 carbon sources on PM01 and PM02A Biolog plates. The Δ*SMc02341* strain in a suppressor background showed slow growth initially on L-arabinose (A2, PM01) as compared to the other three strains, but after 72 hrs it reached a similar optical density as the other strains. As observed previously, the suppressor strain was unable to grow on

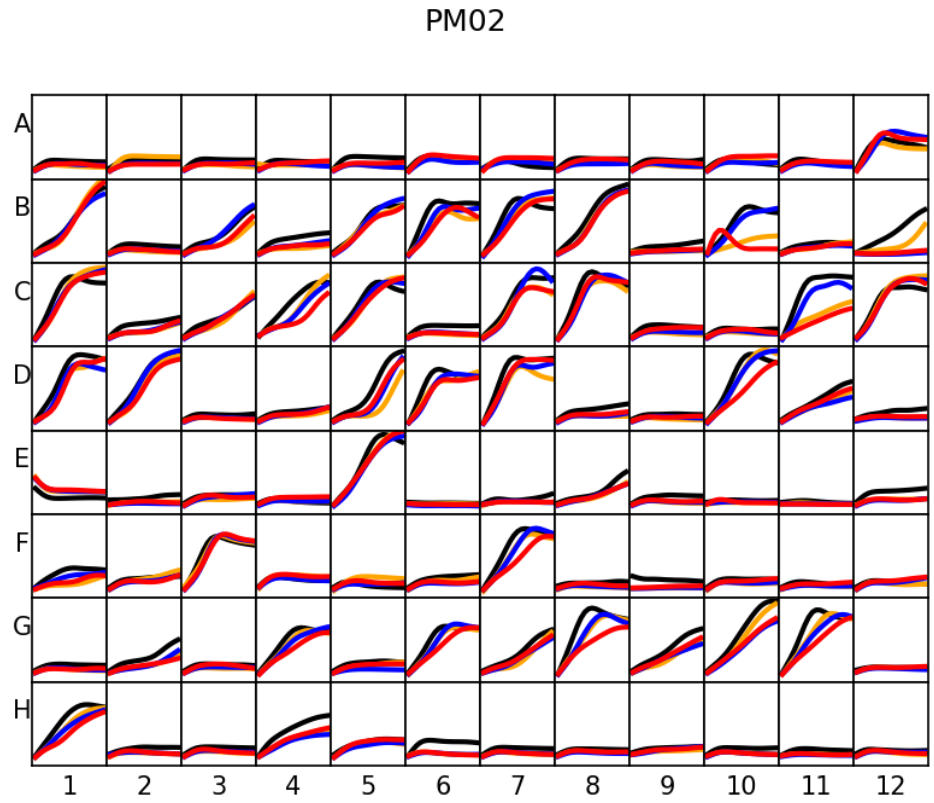


PM1 MicroPlate™ Carbon Sources

A1 Negative Control	A2 L-Arabinose	A3 N-Acetyl-D-Glucosamine	A4 D-Saccharic Acid	A5 Succinic Acid	A6 D-Galactose	A7 L-Aspartic Acid	A8 L-Proline	A9 D-Alanine	A10 D-Trehalose	A11 D-Mannose	A12 Dulcitol
B1 D-Serine	B2 D-Sorbitol	B3 Glycerol	B4 L-Fucose	B5 D-Glucuronic Acid	B6 D-Gluconic Acid	B7 D,L-α-Glycerol-Phosphate	B8 D-Xylose	B9 L-Lactic Acid	B10 Formic Acid	B11 D-Mannitol	B12 L-Glutamic Acid
C1 D-Glucose-6-Phosphate	C2 D-Galactonic Acid-γ-Lactone	C3 D,L-Malic Acid	C4 D-Ribose	C5 Tween 20	C6 L-Rhamnose	C7 D-Fructose	C8 Acetic Acid	C9 α-D-Glucose	C10 Maltose	C11 D-Melibiose	C12 Thymidine
D-1 L-Asparagine	D2 D-Aspartic Acid	D3 D-Glucosaminic Acid	D4 1,2-Propanediol	D5 Tween 40	D6 α-Keto-Glutaric Acid	D7 α-Keto-Butyric Acid	D8 α-Methyl-D-Galactoside	D9 α-D-Lactose	D10 Lactulose	D11 Sucrose	D12 Uridine
E1 L-Glutamine	E2 m-Tartaric Acid	E3 D-Glucose-1-Phosphate	E4 D-Fructose-6-Phosphate	E5 Tween 80	E6 α-Hydroxy Glutaric Acid-γ-Lactone	E7 α-Hydroxy Butyric Acid	E8 β-Methyl-D-Glucoside	E9 Adonitol	E10 Maltotriose	E11 2-Deoxy Adenosine	E12 Adenosine
F1 Glycyl-L-Aspartic Acid	F2 Citric Acid	F3 myo-Inositol	F4 D-Threonine	F5 Fumaric Acid	F6 Bromo Succinic Acid	F7 Propionic Acid	F8 Mucic Acid	F9 Glycolic Acid	F10 Glyoxylic Acid	F11 D-Cellobiose	F12 Inosine
G1 Glycyl-L-Glutamic Acid	G2 Tricarballylic Acid	G3 L-Serine	G4 L-Threonine	G5 L-Alanine	G6 L-Alanyl-Glycine	G7 Acetoacetic Acid	G8 N-Acetyl-β-D-Mannosamine	G9 Mono Methyl Succinate	G10 Methyl Pyruvate	G11 D-Malic Acid	G12 L-Malic Acid
H1 Glycyl-L-Proline	H2 p-Hydroxy Phenyl Acetic Acid	H3 m-Hydroxy Phenyl Acetic Acid	H4 Tyramine	H5 D-Psicose	H6 L-Lyxose	H7 Glucuronamide	H8 Pyruvic Acid	H9 L-Galactonic Acid-γ-Lactone	H10 D-Galacturonic Acid	H11 Phenylethyl-amine	H12 2-Aminoethanol

Figure 17. Phenotypic microarray analysis.

Rm1021(wt; black); SRmD681 (Δ*SMc02341, stkR3*; red); SRmD690 (Δ*SMc02341*; blue), and SRmD410 (*stkR1*; orange) strains of *S. meliloti* on a PM01 Biolog plate. The grid at the bottom represents the location of each substrate as it is laid out on the 96 well plate. The bottom figure is adapted from the Biolog Inc. website.



PM2A MicroPlate™ Carbon Sources

A1 Negative Control	A2 Chondroitin Sulfate C	A3 α-Cyclodextrin	A4 β-Cyclodextrin	A5 γ-Cyclodextrin	A6 Dextrin	A7 Gelatin	A8 Glycogen	A9 Inulin	A10 Laminarin	A11 Mannan	A12 Pectin
B1 N-Acetyl-D-Galactosamine	B2 N-Acetyl-Neuraminic Acid	B3 β-D-Allose	B4 Amygdalin	B5 D-Arabinose	B6 D-Arabitol	B7 L-Arabitol	B8 Arbutin	B9 2-Deoxy-D-Ribose	B10 i-Erythritol	B11 D-Fucose	B12 3-O-β-D-Galactopyranosyl-D-Arabinose
C1 Gentiobiose	C2 L-Glucose	C3 Lactitol	C4 D-Melezitose	C5 Maltitol	C6 α-Methyl-D-Glucoside	C7 β-Methyl-D-Galactoside	C8 3-Methyl Glucose	C9 β-Methyl-D-Glucuronic Acid	C10 α-Methyl-D-Mannoside	C11 β-Methyl-D-Xyloside	C12 Palatinose
D1 D-Raffinose	D2 Salicin	D3 Sedoheptulosan	D4 L-Sorbose	D5 Stachyose	D6 D-Tagatose	D7 Turanose	D8 Xylitol	D9 N-Acetyl-D-Glucosaminitol	D10 γ-Amino Butyric Acid	D11 δ-Amino Valeric Acid	D12 Butyric Acid
E1 Capric Acid	E2 Caproic Acid	E3 Citraconic Acid	E4 Citramalic Acid	E5 D-Glucosamine	E6 2-Hydroxy Benzoic Acid	E7 4-Hydroxy Benzoic Acid	E8 β-Hydroxy Butyric Acid	E9 Glycolic Acid	E10 α-Keto-Valeric Acid	E11 Itaconic Acid	E12 5-Keto-D-Gluconic Acid
F1 D-Lactic Acid Methyl Ester	F2 Malonic Acid	F3 Melibionic Acid	F4 Oxalic Acid	F5 Oxalomalic Acid	F6 Quinic Acid	F7 D-Ribono-1,4-Lactone	F8 Sebacic Acid	F9 Sorbic Acid	F10 Succinamic Acid	F11 D-Tartaric Acid	F12 L-Tartaric Acid
G1 Acetamide	G2 L-Alaninamide	G3 N-Acetyl-L-Glutamic Acid	G4 L-Arginine	G5 Glycine	G6 L-Histidine	G7 L-Homoserine	G8 Hydroxy-LProline	G9 L-Isoleucine	G10 L-Leucine	G11 L-Lysine	G12 L-Methionine
H1 L-Ornithine	H2 L-Phenylalanine	H3 L-Pyroglutamic Acid	H4 L-Valine	H5 D,L-Carnitine	H6 Sec-Butylamine	H7 D,L-Octopamine	H8 Putrescine	H9 Dihydroxy Acetone	H10 2,3-Butanediol	H11 2,3-Butanedione	H12 3-Hydroxy-2-Butanone

Figure 18. Phenotypic microarray analysis.

Rm1021(wt; black); SRmD681 (Δ*SMc02341, stkR3*; red); SRmD690 (Δ*SMc02341*; blue), and SRmD410 (*stkR1*; orange) strains of *S. meliloti* on a PM02 Biolog plate. The grid at the bottom represents the location of each substrate as it is laid out on the 96 well plate. The bottom figure is adapted from the Biolog Inc. website.

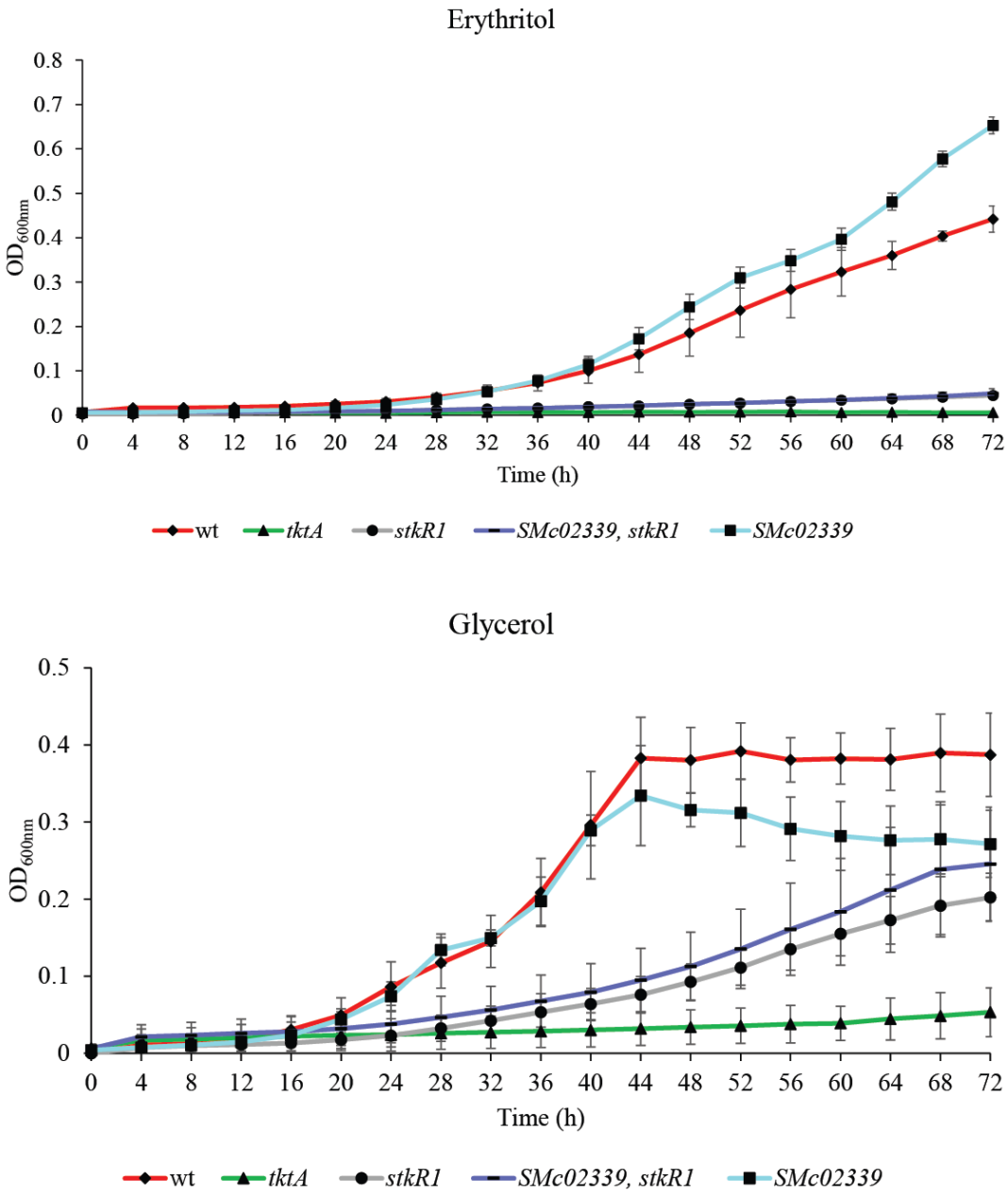


Figure 19. Growth curve analysis of *S. meliloti* strains.

wt (red), *tktA* (green), *stkR1* (grey), *SMc02339*::pKnock-Gm, *stkR3* (purple) and *SMc02339*::pKnock-Gm (light blue) on erythritol and glycerol in Vincent's minimal medium. The error bars represent standard deviation of 3 biological replicates.

erythritol and so was the strain with Δ*SMc02341* in the suppressor background. The strain with Δ*SMc02341* in wild type background did not show *ery* ⁻ phenotype. The same trend was seen for b-methly-D-xyloside (C11, PM02A) where the slow growth phenotype seems to be caused by *stkR* and not by the deletion of *SMc02341* (Figure 18).

3-O-b-D-galactopyranosyl-D-arabinose (B12, PM02A) was the only carbon source where both strains containing Δ*SMc02341* showed no growth and the suppressor strain showed weak growth (Figure 18). 3-O-b-D-galactopyranosyl-D-arabinose, also know as galactosylarabinose, is a disaccharide that contains galactose and arabinose subunits.

3.3.4 ΔSMc02341 strains do not display any symbiotic phenotypes.

Mutations in *tktA* cause severe symbiotic defects in alfalfa plants in *S. meliloti* (Hawkins et al. 2018), and the suppressor strains can partially rescue symbiosis and fix nitrogen. The ability to revert the symbiotic phenotype in the suppressor strains has been attributed to the overexpression of *tktB.* To investigate if *SMc02341* contributes to the symbiotic phenotypes associated with the suppressor strains alongside *tktB,* unmarked Δ*SMc02341* strains, SRmD690 and SRmD681 were tested for symbiosis. Alfalfa plants were grown under nitrogen-deficient conditions in Jensen's media and were inoculated with the wild type, the *tktA* mutant strain, a suppressor strain and the Δ*SMc02341* strains. Shoot dry weight of alfalfa plants was recorded as an indicator of nitrogen fixation (Figure 20). The plants inoculated with Δ*SMc02341* strains were green and healthy looking with no visible defects seen on root nodules. The average dry weight of plants inoculated with the Δ*SMc02341* strain in suppressor background was significantly lower than the wild type but not different from its parent suppressor strain (*stkR3*). The dry weight of plants inoculated with the Δ*SMc02341* strain in the wild type background was not

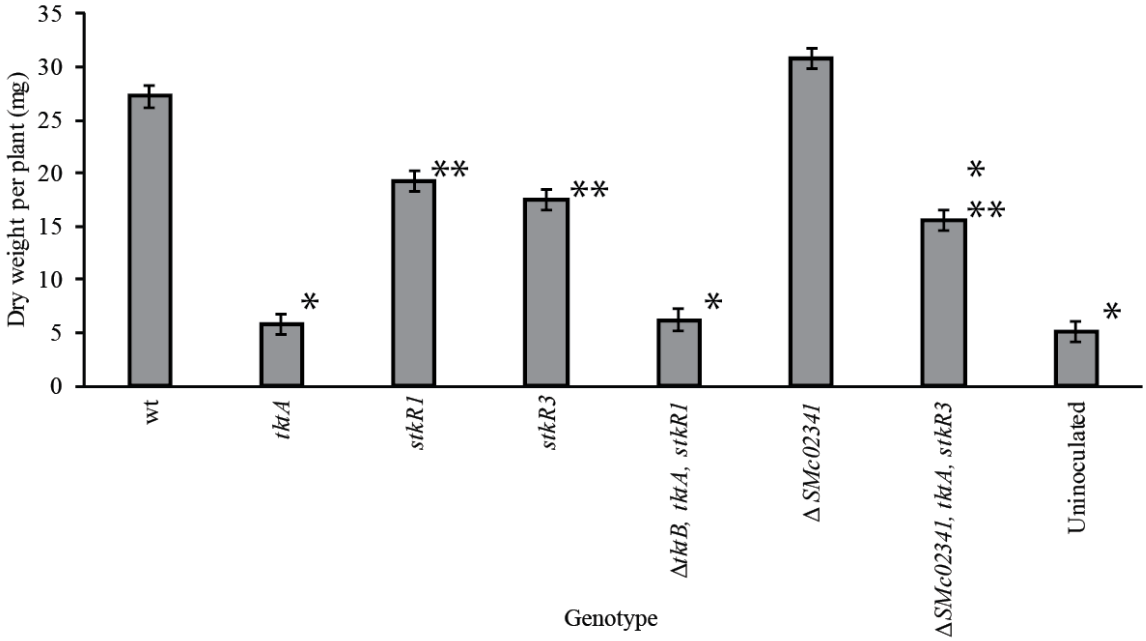


Figure 20. Dry weight of alfalfa plants inoculated with *S. meliloti* strains with *tktB* and *SMc02341* deletions.

Grey bars represent average shoot dry weight of alfalfa plants from 3 biological replicates with 10 plants per replicate. * Represents significant difference from the dry weight recorded for the wt strain and ** represents significant difference from the dry weight recorded for the *tktA* mutant strain with a $P < 0.05$ with Student's *t*-test.

significantly different from the wild type. Both Δ*SMc02341* strains produced higher dry weight than the *tktA* mutant strain and the uninoculated control. As reported previously, the *tktA* mutant strain produced small chlorotic plants that did not fix nitrogen and the dry weight of these plants was comparable to the uninoculated control. These data indicate that mutations in *SMc02341* do not cause any symbiotic defect in either wild type or suppressor backgrounds, suggesting the contribution of *SMc02341* is not symbiotic but a metabolic one.

3.4 Discussion

In this study we characterize the locus associated with *tktB* in *S. meliloti*. The genome of *S. meliloti* has fives genes that are annotated as transketolases (*tktA*, *tktB, cbbT, SMc00269,* and *SMc00270*); however only *tktA* and *tktB* seem to play a role in the non-oxidative PP pathway. We have previously shown that in the absence of a functional *tktA*, upregulation of *tktB* is the preferred metabolic route by isolating 10 independent second site suppressor strains to the *tktA* mutant strain. All suppressor strains contain point mutations in *SMc02340,* a negative regulator of *tktB* and its neighbouring genes.

Transketolase catabolises a reversible reaction where it cleaves the C2-C3 bond from a keto sugar and transfers the C2 to an aldo sugar. The C2 is a glycolaldehyde residue which is transferred to the aldose and the C3 product is released into the medium. Recently, it was shown that it is possible to carry out only a half reaction in the absence of an aldose sugar (Solovjeva et al. 2020). Using either hydroxypyruvate or xylulose-5P as a substrate for the half reaction, erythrulose (C4) can be produced. The mechanism showed that a glycolaldehyde (C2) is formed and remain attached to the enzyme while the second glycolaldehyde is produced and these are combined to form a C4 (Solovjeva 2021).

Glycolaldehyde can be converted to glycolate. An aldehyde oxidase (ALOD) has been functionally characterized in *Burkholderia* sp. AIU129 that can carry out this biochemical reaction (Yamada et al. 2015). The ALOD contains three α, β and γ subunits and the N-terminal sequence contains highly sequence similarity to the subunits of a xanthine dehydrogenase (Yamada et al. 2015). The pSymB chromid of *S. meliloti* contains two genes *SMb21338* and *SMb21339* that are annotated as probable subunits of an oxidoreductase like aldehyde oxidase with highly similar N-terminal sequences to a xanthine dehydrogenase (Barnett et al. 2001). Both *SMb21338* and *SMb21339* are functionally uncharacterized in *S. meliloti.* Evidence for glycolate oxidation to glyoxylate exists in *E. coli* where *glc* operon plays a key role (Pellicer et al. 1996). Glyoxylate can be converted to malate which is a TCA cycle intermediate. *S. meliloti* genome consists of a similar set of *glc* genes that are functionally uncharacterized (Capela et al. 2001).

The genetic components required for converting glycoaldehyde to glyoxylate do exist in *S. meliloti.* Therefore, we propose that the secondary metabolic bypass in *S. meliloti* is carried out by a transketolase half reaction (using *tktB*) and the glycolaldehyde produced as a result of this half reaction is converted into glyoxylate using the *SMb21338, SMb21339* and the *glc* operon. The glyoxylate is then converted to malate which enters the TCA cycle.

The putative carbohydrate kinase, *SMc02341,* is present directly downstream of *tktB* and the gap between the two genes is too small for a promotor (Figure 11). We have found that both *tktB* and *SMc02341* are co-ordinately regulated which substantiates that these two genes are likely part of the same operon. Furthermore, it is possible that *SMc02341* works in conjunction with *tktB* to phosphorylate pentoses that act as substrates for TktB activity. It is evident from the lack of *SMc02343* overexpression in the suppressor strains that it may not be part of the *tktB*

operon. Additionally, no ribose phenotype was identified in the strains with mutations in *SMc02343.*

rhaR is a DeoR type negative regulator of the genetic locus necessary of rhamnose uptake and catabolism in *S. meliloti.* The *rhaR* is flanked by other *rha* genes on the chromosome and it has been shown to divergently regulate the expression of *rha* operon (Richardson et al. 2004). Insertional and point mutations in *SMc02340* cause an upregulation of neighbouring genes that might be controlled by divergent promotors. We have also seen some overexpression of *SMc02340* in the strains that carry mutations in this gene but the change in expression seems to depend on the type of mutation. Two suppressor strains (*stkR1* and *stkR3*) with point mutations in *SMc02340* showed higher transcription of *SMc02340,* but the values were not statistically significant. However, a strain with a pKnock-Gm insertion in *SMc02340* displayed a significant change in its expression (Figure 14). In any case, the transcriptional data do suggest autoregulation by *SMc02340*. A similar phenomenon was reported in *E. coli.* The *crp* gene encodes for a cyclic AMP receptor protein in *E. coli.* CRP-cAMP was shown to reciprocally and coordinately regulate *crp* and divergent promotors (Hanamura and Aiba 1991). In *Pseudomonas aerugnosa,* expression of genes involved in the ED pathway is regulated by PtxS, HexR, GtrS/GltR systems. However, an additional GntR (PA2320) was identified to also regulate the expression of those gene as well as repress its own expression. GntR was able to bind to it own promotor as well as a promotor for *gntP* encoding for a gluconate permease (Daddaoua et al. 2017).

The mutations in *tktB* or *SMc02341* in the suppressor background can cause growth defects on ribose. The introduction of a Δ*tktB* in suppressor background results in no growth on ribose and the strain is auxotrophic, but Δ*SMc02341* on the other hand, has only reduced growth

phenotype on ribose in suppressor backgrounds (Figure 13). This means that the auxotrophy can be uncoupled from the ribose phenotype, and ribose might be a putative substrate of the determinants encoded by this operon. These phenotypes were only identified in the suppressor background that has mutations in *SMc02340*, which suggests that the *tktB* operon does not play an essential role in the wild type. However, in the absence of a functional *tktA*, *tktB* can contribute transketolase activity. This can be considered as a classic example of metabolic redundancy which is common in rhizobia and many other bacteria. In *E. coli tktB* expression is regulated by ppGpp and RpoS. Under wild type conditions, *tktB* is "buffered" by *tktA* but in a *tktA* mutant strain in a $ppGpp^0$ background *tktB* is transcribed (Harinarayanan et al. 2008; Jung et al. 2005). Similarly, three transketolases were identified in *Salmonella* where TktA contributes approximately 88% of enzymatic activity in the wild type cells and TktB has a minor role. However, the kinetic ability of TktA and TktB were comparable. TktC on the other hand showed 10-100 fold less catalytic efficiency (Shaw et al. 2018).

Phenotypic screening of *SMc02339* and *SMc02341* mutant strains did not reveal any striking growth defects on any carbon sources but some growth impairments were recorded in the *SMc02339* mutant strains. The strain with mutations in *SMc02339* in the suppressor background showed subtle growth defects on L-arabinose, glycerol, D-xylose, and adenosine as compared to the wild type and the parent suppressor strain. All these carbon substrates are one or two reactions away from entering the PP pathway. Although these phenotypes are not lethal, further investigation should reveal potential targets for *SMc02339* and its downstream genes. A kinase assay should be done to identify a substrate for SMc02341 activity. The biochemical characterization will also further our understanding for the role of *tktB* operon in carbon

metabolism and the metabolic redundancy. Additionally, a transketolase assay is required to identify the contribution of TktA and TktB towards transketolase activity in *S. meliloti*.

The *glc* operon should also be functionally characterized to determine the role it may play in glyoxylate metabolism. Existence of a secondary metabolic bypass can be determined by generating mutations in the *glc* operon.

Chapter 4:

Sinorhizobium meliloti contains a complete Embden Meyerhof Parnas pathway

Authors: Sabhjeet Kaur, MacLean G. Kohlmeier, Justin P. Hawkins, and Ivan J. Oresnik. This work is part of a larger manuscript that will be written in collaboration with Dr George C. diCenzo. The collaborator's work is not shown or discussed here. This work was planned and executed by Sabhjeet Kaur. MK designed the original Pfp overexpression construct and set up the protein purification protocol. JH designed and constructed the *tal* deletion as well as provided guidance for protein purification and enzyme assays.

4.1 Introduction

The breakdown of glucose is carried out via the Entner Doudoroff (ED) pathway in *Sinorhizobium meliloti* (Fuhrer et al. 2005). This pathway is often considered an alternate pathway to the Embden Meyerhof Parnas (EMP) pathway that is found in bacteria like *E. coli* and *Bacillus subtilus*. Although it is often not as well known as the EMP pathway, the ED pathway can be found in many bacteria, archaea and higher organisms (Chen et al. 2016; Siebers and Schönheit 2005; Vegge et al. 2016).

The pathway consists of four reactions. Starting with glucose-6-phosphate, the first step consists of it being oxidized to 6-phosphogluconolactone by the action of glucose-6-phosphate dehydrogenase. The resultant product is linearized by the addition of H_2O by 6-phosphogluconolactonase to yield 6-phosphogalactone (6-PG). The 6-PG is in turn dehydrated by 6-phosphogluconate dehydratase to yield 2-keto-3-deoxy-phosphogluconate (KDPG). The KDPG is a substrate for 2-keto-3-deoxy-6-phosphonate aldolase which carries out an aldol cleavage to yield glyceraldehyde 3-phosphate and pyruvate (Katznelson 1955). The presence of the enzyme activities for this pathway has been established in *S. meliloti* (Martínez-De Drets and Arias 1972). In addition, ^{13}C labelling studies have shown that glucose is catabolised through the ED pathway in *Bradyrhizobium japonicum* and *S. meliloti* (Keele et al. 1969; Fuhrer et al. 2005).

All the enzymes that play a role in the EMP pathway have been shown to be present in *S. meliloti* except for phosphofructokinase (Arias et al. 1979). *S. meliloti* was found to be lacking an ATP-dependent phosphofructokinase activity that should carry out the physiologically irreversible reaction of converting fructose-6-phosphate to fructose-1,6-bisphosphate (Arias et al. 1979; Irigoyen et al. 1990). Therefore, it has long been assumed that the EMP pathway only plays a gluconeogenic role in *S. meliloti* and rhizobia in general (Geddes and Oresnik 2014).

The genome of *S. meliloti* 1021 is annotated to contain a probable PP_i- fructose 6-phosphate 1-phosphotransferase, encoded for by *pfp* (*SMc01852*) (Capela et al. 2001). This enzyme has been shown to utilize inorganic pyrophosphate (PP_i) instead of ATP as a phosphate donor, and to carry out a reversible reaction of converting fructose-6-phosphate to fructose-1,6-bisphosphate in some bacteria (Ding et al. 2000). The first PP_i-dependent phosphofructokinase was found in *Entamoeba histolytica*, subsequently this activity was also found in higher plants, animals, and some bacteria (Reeves et al. 1974). PP_i-dependent phosphofructokinase is considered to be an ancestral version of the ATP-dependent phosphofructokinase because PP_i is believed to be a more ancient source of energy than ATP (Ding et al. 2000; Baltscheffsky Herrick (Editor) 1996). In addition to playing a role in the EMP pathway, PP_i-dependent phosphofructokinase has also been shown to be involved in the pentose phosphate (PP) pathway in microorganisms like cellulolytic Clostridia. In these organisms, sedoheptulose-7-phosphate, which is a PP pathway intermediate, is phosphorylated to sedoheptulose-1,7-bisphosphate by a PP_i-dependent phosphofructokinase. This 7 carbon intermediate is subsequently broken down to dihydroxy-acetone phosphate (DHAP) and erythrose-4-phosphate by a fructose bisphosphate aldolase (Koendjbiharie et al. 2020), thus bypassing a transketolase reaction.

In this chapter we hypothesized that *S. meliloti*, contrary to published literature (Geddes and Oresnik 2014), has the ability to carry out a phosphofructokinase reaction. However, this reaction is predicted to be carried out by a PP_i dependent phosphofructokinase encoded by *SMc01852* which is annotated as *pfp*.

4.2 Materials and methods

4.2.1 Bacterial strains, plasmids, and media

The bacterial strains and plasmids used in the study are listed in Table 10. *E. coli* strains were grown on Luria-Bertani (LB) as complex medium at 37 ºC. *S. meliloti* strains were grown at 30 ºC on LB as a complex medium, and Vincent's minimal medium (VMM) as a defined medium (Vincent 1970). Carbon sources were filter sterilized and used at a final concentration of 15 mM unless otherwise noted. Antibiotics were filter sterilized and were used as necessary for selection of genetic markers. When used, the concentration of antibiotics were as follow: ampicillin (Amp), 100 μg/ml; streptomycin (Sm), 200 μg/ml; kanamycin (Km), 200 μg/ml; chloramphenicol (Cm), 20 μg/ml; neomycin (Nm), 200 or 50 μg/ml; tetracycline (Tc), 5 μg/ml, gentamicin (Gm), 60 μg/ml for *S. meliloti* and 20 μg/ml for *E. coli*.

4.2.2 Genetic manipulations and constructions

Basic genetic manipulations such as transductions, transformations, and conjugations were carried out as previously described (Finan et al. 1988, 1984). The plasmids expressing *pfp* (pMK66) and *tal* (pJH118) were constructed by using ORFeome library designed for *S. meliloti* (House et al. 2004) and using the Gateway compatible vector pCO37 (Jacob et al. 2008) as previously described (Geddes and Oresnik 2012a).

Genetic deletions of specific genes were constructed using Gibson assembly. Constructs were designed using NEBuilder Assembly tool (https://nebuilderv1.neb.com/) (New England Biolabs, Inc.). The primers were designed by using flanking regions from the gene of interest and using pJQ200SK as an integration vector to create chromosomal deletions (Quandt and Hynes 1993). All PCR amplifications using this method were carried out using Q5 DNA polymerase (New England Biolabs, Inc.)

Table 10. Strains and plasmids list

Strain or plasmid	Genotype and relevant phenotype[a]	Reference
Strains		
S. meliloti		
Rm1021	SU47 *str-21* Sm^r	(Meade et al. 1982)
SRmD574	Rm1021, Δ*tal* Sm^r	This work
SRmD673	Rm1021, Δ*pfp* Sm^r	This work
SRmD674	Rm1021, Δ*pfp::*Nm Sm^r Nm^r	This work
SRmD688	Rm1021, Δ*pfp::*Nm Δ*tal* Sm^r Nm^r	This work
E. coli		
MM294A	*pro-82 thi-1 hsdR17 supE44*	(Finan et al. 1986)
MT607	MM294A *recA56*	(Finan et al. 1986)
MT616	MT607(pRK600)	(Finan et al. 1986)
DH5α	λ^- *ϕ*80d*lacZ*ΔM15 Δ(*lacZYA-argF*)*U169 recA1 endA1 hsdR17*(r_k^- m_k^-) *supE44 thi-1 gyrA relA1*	(Hanahan 1983)
DH5αRif	Rifampin-resistant variant of DH5α	(House et al. 2004)
DH5αλpir	λ*pir* lysogen of DH5α	(House et al. 2004)
JM109	*end*A1, *rec*A1, *gyr*A96, *thi, hsd*R17 (r_k^-, m_k^+), *rel*A1, *sup*E44, Δ(*lac-pro*AB), [F´ *tra*D36, *pro*AB, *laq*I^qZΔM15]	(Hanahan and Glover 1985)
Plasmids		
pRK600	pRK2013 *npt*::Tn*9* Cm^r	(Finan et al. 1986)
pRK7813	Broad-host-range cloning vector; Tc^r	(Jones and Gutterson 1987)
pCO37	pRK7813 containing *attB* sites; Gateway-compatible destination vector	(Jacob et al. 2008)
pMK2014	FRT-*ccdB*-Cm^r-FRT cassette Pen^r	(House et al. 2004)
pXINT129	*λint* and *xis* driven by P_{lac}; Km^r	(Platt et al. 2000)
pJQ200SK	Gene replacement suicide vector Gm^r	Quandt & Hynes (1993)
pSK11	pJQ200SK/*pfp* flanking regions with Nm marker Gm^r Nm^r	This work
pDK36	pJQ200SK/*pfp* flanking regions Gm^r	This work

pJH123	pBS/*tal*±500bp Amp^r	This work
pJH125	pBS/Δ*tal* Amp^r	This work
pJH118	pCO37/*tal* Tc^r	(Hawkins et al. 2018)
pMK66	pCO37/*pfp* Tc^r	This work
pSK16	pQE9/*pfp* codon optimized	This work

[a] Sm^r, streptomycin resistant; Nm^r, neomycin resistant; Gm^r, gentamicin resistant; Sp^r, spectinomycin resistant; Cm^r, chloramphenicol resistant; Tc^r, tetracycline resistant; Pen^r, penicillin resistant; Km^r, kanamycin resistant; FRT, flippase recognition target.

To create a marked mutation in *pfp* (*SMc01852*), the flanking regions were amplified using primer pairs 1-2 and 3-4, respectively (Table 11). using Rm1021 genomic DNA as a template. The Nm marker cassette was amplified with primers 5 and 6 using pOGG008 as template DNA (Table 11) (Geddes et al. 2018). The plasmid pJQ200SK was linearized with BamHI and SacI (Invitrogen) and all components were assembled using a Gibson assembly master mix (New England Biolabs, Inc.) by incubating at 50 ºC for 1 h in a thermocycler to generate pSK11 (Table 10). This plasmid was transformed into DH5α, and subsequently mobilized into Rm1021. Single cross over recombinants were selected using Sm and Gm. Single colonies of these putative recombinants were tested to ensure sucrose sensitivity, and then streaked onto LB containing 10% sucrose. The resultant sucrose resistant colonies were then screened for Gm and Nm resistance. Colonies that were Nm resistant and Gm sensitive were putatively the result of an allelic exchange via putative double crossover. The mutation was confirmed by PCR amplifying using primers 7 and 8, which flank the target gene. The PCR product was sequenced to verify correct assembly and integration. The resultant strain was designated as SRmD674 (Table 10).

A similar strategy was used to construct an unmarked *pfp* mutation. Flanking regions were amplified using primers 9-10 and primers 11-12 (Table 11). The integration plasmid pJQ200SK was linearized using BamHI and SacI, and the plasmid pDK36 was generated using a Gibson assembly master mix (New England Biolabs, Inc.) following the manufacturer's instructions. This construct was mobilized into Rm1021 and recombinants were selected using Gm. Some of the resultant colonies were streaked onto LB containing 10% sucrose and the resultant colonies were screened for the loss of Gm resistance. These putative Δ*pfp* constructs were then screened by PCR using primers 7 and 8. Deletions were finally confirmed by sequencing the PCR

products to ensure proper integration and loss of *pfp*. The resultant strain was designated as SRmD673.

The *tal* deletion in Rm1021 background was constructed by using an inverse PCR strategy. The *tal* ORF including 400-500 bp of its flanking regions was amplified as a single product using Rm1021 genomic DNA template using primers 13 and 14 (Table 11). The PCR product was gel isolated, restricted with BamHI and XbaI and ligated into the vector pBS SK^+ that was restricted with the same enzymes yielding pJH123 (Table 10). A second set of primers, 15 and 16, were designed to amplify the flanking area of *tal* in an outwards direction such that PCR product would contain the flanking regions and the entirety of the plasmid backbone, but not the gene of interest. The resultant plasmid was designated pJH125 (Table 10). The plasmid pJH125 was restricted with BamHI and XbaI to remove the insert DNA and ligated into pJQ200SK and subsequently mobilized into Rm1021. Generation of the Δ*tal* mutation was carried out as described above. The resultant strain was designated SRmD574.

The carbon phenotypes associated with the deletions were confirmed by complementing the deletion with a plasmid overexpressing the target gene on a gateway compatible vector pCO37.

4.2.3 Plant growth conditions

Quantification of the nitrogen fixation was carried out by using plant dry weights following their growth in nitrogen deficient medium as described previously (Poysti et al. 2007). Briefly, surface sterilized alfalfa seeds were germinated for 2 days on water agar plates. The seedlings were aseptically transplanted into Magenta jars that were modified to work as Leonard jars.

Table 11. Primer sequences for insertional mutations and gene deletion constructs

Primer #	Primer sequence 5'-3'	Primer name
	Δ*pfp*-Nm	
1	cgaattcctgcagcccggggCCGTAAGCGCGTTACCAG	pfp_L_BamHI_F
2	tatggctcatGGGGTCCTCCTGGGGATTTC	pfp_L_BamHI_R
3	aatttttctaaTGAACGACGAGGAACGGTTC	pfp_R_SacI_F
4	agggaacaaaagctggagctTCTCGAAGATCGCCATCTG	pfp_R_SacI_R
5	ggaggaccccATGAGCCATATTCAGCGTG	Neo_BamSac_F
6	tcctcgtcgttcaTTAGAAAAATTCATCCAGCATC	Neo_BamSac_R
7	TTCCTGGACGCCACCGTGC	del_pfp_confir_F
8	GAATGGCAAGCAGCACCGCA	del_pfp_confir_R
	Δ*pfp*	
9	tcctgcagcccgggggatcATTGTAAAGCCTGGCGACG	Left_pfp_um_F
10	gtcgttcaGGGGTCCTCCTGGGGATTTC	Left_pfp_um_R
11	aggaccccTGAACGACGAGGAACGGTTC	Right_pfp_um_F
12	gaacaaaagctggagctcATGCCGAGCCAGACGAGATAG	Right_pfp_um_F
	Δ*tal*	
13	AGTC**GGATCC**GTCAGCGTATAGGCGGCC	tal_BamHI_FW
14	AGTC**TCTAGA**CTGCTCTTAAGCGTTTATCGACC	tal_XbaI_RV
15	GGATGTCCTTCACATCGGC	talinv_RV
16	CCAGAAGATCGCCTGACTTT	talinv_FW
	***pfp* codon optimized**	
17	GATC**GGATCC**GCAAAAAAGAAAGTAGCTAT	Pfp_Bam_FW
18	GATC**AAGCTT**TTAGTCAGCGGTCTGCCACT	Pfp_Hind_RV

Restriction sites are shown in bold; lower case letters represent the overlapping regions of the Gibson assembly plasmid.

These contained a sterile mix of sand and vermiculite in 1:1 ratio and were soaked with 200 ml of Jensen's medium. The pots were inoculated with appropriate strains at a titre of approximately 10^7 CFU/ml in 10 ml sterile water. Plants were grown in a growth chamber with 16 h/day light cycle. The plants were harvested 4 weeks after inoculation. Dry weight numbers were taken from an average of 10 plants/pot. Typically, each strain was assayed using at least 3 biological replicates.

4.2.4 Metabolite analysis

Sample preparation for metabolite analysis

Metabolite analysis was carried out as described previously (Hawkins et al. 2018). Briefly, bacterial cells were grown in LB overnight and sub-cultured into 500 ml VMM containing either 15 mM glucose or 15 mM succinate. These cultures were grown overnight at 30 ºC with shaking (300 rpm). One hundred ml of the cultures were harvested by centrifugation after reaching an OD_{600} of 0.5. The cell pellet was washed with double distilled water and flash frozen in liquid N_2. The weight of the pellet was taken before freezing to allow normalization of the metabolite concentrations. Metabolite extraction and LC-MS was performed as described previously (Hawkins et al. 2018) at the Metabolomics Innovation Resource (MIR) - McGill University. The data represent an average of at least 3 biological replicates.

Metabolite analysis

Statistical analysis was performed using MetaboAnalyst 5.0 (Chong et al. 2019) (https://www.metaboanalyst.ca). The metabolite concentrations were normalized by log-transformation. Principal component analysis, analysis of variance and clustering analysis were done as described previously (Hawkins et al. 2018). Significant difference was determined by

preforming Welch's two sample *t* test with a *P*-value criterion of <0.05. False discovery (below 0.1) rate between two samples was determined for significance as well.

4.2.5 Growth curves

Bacterial growth curve experiments were conducted in complex (LB) and defined (VMM) media that was supplemented with an appropriate carbon source as described previously (Kohlmeier et al. 2019). The bacterial cultures were grown overnight in LB, washed and sub-cultured into fresh medium. Two hundred microlitre aliquots of the sub-cultured medium were dispensed into the wells of a 96 well plate. A final concentration of 15 mM of a carbon source was used in VMM unless otherwise stated. To decrease condensation on the inside of the plate lid, the lid was coated with 0.05% TritonX in 20% ethanol solution (Brewster 2003). The experiment was conducted at 28 ºC under continuous orbital shaking and the optical density was recorded at 600 nm wavelength every hour for 72 h on a BioTek Synergy 2 plate reader (BioTek Instruments, Inc., Winooski, VT, USA).

4.2.6 Phenotypic microarray

Rhizobial strains were subjected to phenotypic microarray study using PM01 and PM02A Biolog plates (Biolog Inc. CA, USA). Bacterial cultures were grown overnight in complex medium (LB), washed and sub-cultured into fresh defined medium (VMM). One hundred microliter aliquots of the sub-cultured medium were dispensed into the wells of the Biolog plate. The inside of the plate lid was coated with 0.05% TritonX in 20% ethanol solution to avoid condensation (Brewster 2003). The rhizobial strains were tested on 190 carbon sources for growth in a 72 h growth curve on a BioTek Synergy 2 plate reader. Optical density was recorded at 600 nm wavelength every hour. The data were analysed and plotted using DuctApe (Galardini

et al. 2014). The phenome analysis was completed by following commands provided by the authors in the DuctApe manual (https://combogenomics.github.io/DuctApe/howto.html).

4.2.7 Protein overexpression and purification

The *pfp* open reading frame was codon optimized for overexpression in *E. coli* and delivered as an insert in pUC57. This open reading frame was subsequently amplified using primers 17 and 18 (Table 11), restricted with BamHI and HindIII, and ligated into pQE-9 such that it contained an in-frame N-terminal RGS-His6X tag. This plasmid was designated as pSK16 (Table 10). To overexpress *pfp* pSK16 was transformed into JM109. Overexpression and purification was carried out similarly to what was previously described (Kohlmeier et al. 2021). A 5 ml culture of pSK16 was grown overnight at 28 ºC and used to inoculate 500 ml of LB broth. The culture was grown at 28 ºC with shaking conditions (200 rpm) to an OD_{600} of 0.6-0.8 at which time the cells were harvested by centrifugation at 4000 rpm for 1 h at 4 ºC. The cell pellet was then stored at -80 ºC until it was needed. To lyse the cells the pellet was resuspended in lysis buffer (20 mM Na_2PO_4, pH 8, containing 500 mM NaCl, 20 mM imidazole,) and the cells were lysed using an Emulsiflex. The lysed cells were centrifuged at 12,000 rpm for 1 h at 4 ºC to clear the cell debris. The cell-free lysate was applied to a nickel-nitrilotriacetic acid (Ni-NTA) column and washed with 3 column volumes of the lysis buffer. This was subsequently washed with 2 column volumes lysis buffer that contained 100 mM imidazole to remove non-specific proteins. To elute the protein the column was washed with 3 column volumes of elution buffer (20 mM Na_2PO_4, pH 8, containing 500 mM NaCl, 500 mM imidazole) and collected as 1.5 ml fractions. The eluted protein was dialyzed overnight at 4 ºC in a dialysis buffer (50 mM Tris pH 8, 0.1 mM EDTA, 1mM DTT). Protein fractions were routinely analyzed using SDS PAGE (Laemmli 1970) as well as by Western blotting. For Western blots the proteins were first

separated using SDS-PAGE and subsequently transferred to nitrocellulose and detected using His6 monoclonal antibody that was subsequently detected colorimetrically (Richardson et al. 2008).

4.2.8 Biochemical assays

The PP_i dependent phosphofructokinase assays were performed as described previously (Reshetnikov et al. 2008; Zhou et al. 2013). Briefly, assays were performed at 30 ºC by measuring optical density changes due to the oxidation of NADH, or the reduction of $NADP^+$ at 340 nm using a BioTek Synergy 2 plate reader. The forward reaction was performed by measuring NADH oxidation in a coupled reaction. A 200 µL reaction mix contained 100 mM MOPS (pH 8), 5 mM $MgCl_2$, 0.2 mM NADH, 2 mM PPi (or ATP or GTP), 2 mM fructose-6P, 5U/ml aldolase, 5U/ml triose phosphate isomerase, 5U/ml α-glycerophosphate dehydrogenase. The reaction was initiated by the addition of PP_i. Absorbance measurements were taken every 50 seconds for 10 minutes.

The reverse reaction mix contained 100 mM MOPS (pH 8), 5 mM $MgCl_2$, 0.3 mM NADP, 15 mM NaH_2PO_4, 2 mM fructose-1,6biP, 0.24 U/ml phosphoglucose isomerase, 0.12 U/ml glucose 6-phosphate dehydrogenase. The assays were carried out in a 200 µL volume. The reaction was initiated by adding the addition of the substrate fructose-1,6biP. Absorbance measurements were taken every 50 seconds for 10 minutes.

In all cases enzyme activity was proportional to the volume of extract that was added to the reaction. Activities were calculated from the linear portion of the curve and the specific activities were calculated by normalizing to the protein concentration. Protein concentrations were measured using a Nanodrop to measure the absorbance of the samples at 280 nm. The data were

analyzed using GraphPad Prism. The data were fitted to Michaelis-Menton model to calculate V_{max} and K_m.

4.3 Results

4.3.1 Identification and characterization of pfp by gene deletion

The annotation of the *S. meliloti* genome clearly showed that it did not contain an open reading frame encoding an ATP-dependent phosphofructokinase (PFK). Interestingly, it was noted that the genome did have the potential to encode a pyrophosphate dependent phosphofructokinase (*SMc01852*), but at the time of publication these were not known to be found within the alpha-proteobacteria, and no follow-up work was carried out (Capela et al. 2001). Although a genome scale metabolic model for this organism did include this gene and its potential activity, it was not recognized as a core gene necessary for biomass production and was included as an uncurated reaction (diCenzo et al. 2016). The reactants for phosphofructokinase are fructose 6-phosphate and fructose 1,6, bisphosphate. Since the non-oxidative PP pathway generates metabolites that are found on either side of this reaction, we were intrigued at the possibility that *S. meliloti* might contain this activity.

A microbial fitness browser was developed that displays genome-wide fitness data on 42 diverse bacteria (https://fit.genomics.lbl.gov/cgi-bin/myFrontPage.cgi) (Price 2018). *S. meliloti* is one of the organisms that was included. The browser data were generated from using pools of random bar-coded transposon mutants that had been subjected to a number of different growth conditions. Within the collection, *S. meliloti* Rm1021 was grown on 35 different carbon sources. To determine whether *SMc01852* should be pursued within the scope of this thesis, the browser was used to query whether mutations in this gene might result in a loss of fitness on any carbon sources.

Fitness values are defined as $\log_2$ ratios that describe the changes in insertions on a given carbon source as compared to that in the starting pool of inserts (Price et al. 2018). The strongest phenotypes are defined as those with an absolute fitness change of greater to or equal to 1.6. The query of *SMc01852* showed that mutations in this gene led to a decrease of fitness on 10 different carbons sources. These were: glycerol, -3.7; L-rhamnose, -3.0; m-inositol, -2.9; D-galactose, -2.6; D, L-lactate, -2.5; D-tagatose, -2.1; D-cellobiose, -2.0; L-glutamine, -1.7; L-histidine, -1.7; and D-trehalose, -1.7. Although it is difficult to directly assess how *SMc01852* directly affects all of these conditions, it strongly suggests that this gene encodes an activity involved in general carbon metabolism.

To further characterize *SMc01852* a genetic strategy was employed. Two mutations were constructed and verified (Figure 21). These were a deletion of the gene that was marked with a neomycin resistance cassette (SRmD674), as well as an unmarked deletion of the open reading frame (SRmD673).

SRmD674 grew in defined medium on sugars such as glucose, succinate, glycerol, as well as galactose. When compared to the wild type, we noted that growth appeared to be markedly slower (Table 12). This growth phenotype was more pronounced when the strains were grown in broth and more nuanced on agar. This indicates that *pfp* is not essential for growth on these sugars and the deletion of *pfp* did not make the strains auxotrophic (Table 12). Interestingly, SRmD674 was unable to grow using glycerol as a carbon source if M9 was used as a defined medium. The major difference between these two defined media is the use of NH_4Cl_2 as a N-

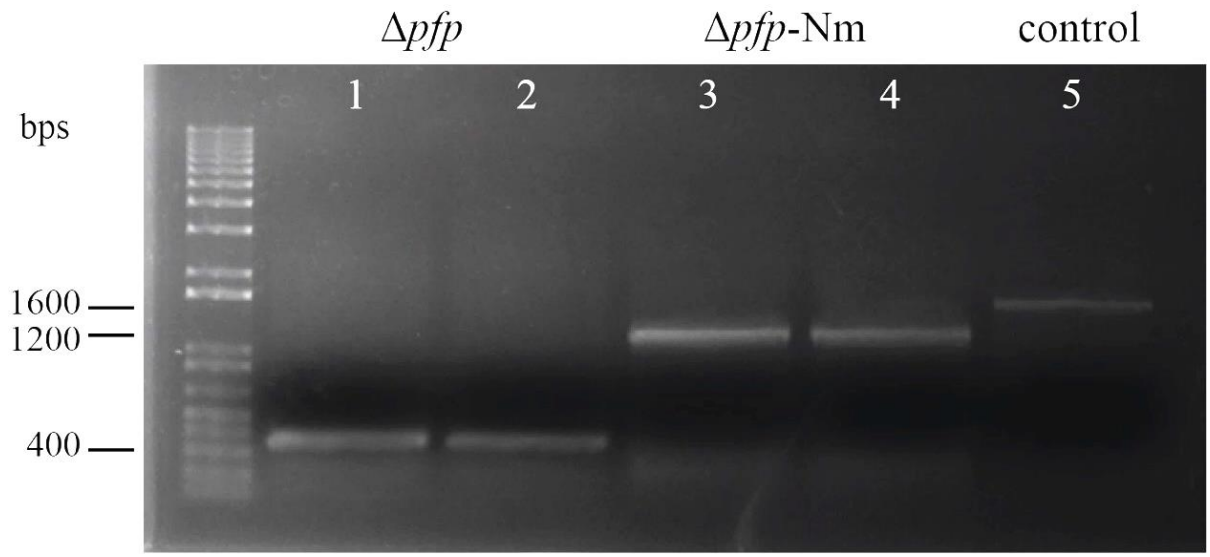


Figure 21. Agarose gel showing PCR fragments of unmarked and marked *pfp* deletion constructs of *S. meliloti*.

Lane 1-2, unmarked *pfp* deletion; lane 3-4, *pfp* deletion marked by a neomycin cassette; lane 5, wild type control of *pfp*.

Table 12. Growth of *S. meliloti* strains on different carbon sources in minimal medium

Strain	**Genotype**	**Growth on VMM agar**				
		Glucose	**Succinate**	**Glycerol**	**Galactose**	**LB**
Rm1021	Wild type	++	++	++	++	++
SRmD574	Δ*tal*	++	++	++	++	++
SRmD674	Δ*pfp*	++	++	+	++	++
SRmD688	Δ*tal*Δ*pfp*	-	-	-	++	++
SRmD688 (pMK66)	Δ*tal*Δ*pfp* (*pfp*$^+$)	++	++	++	++	++

Growth phenotypes are as follows: ++, wild type growth; +, moderate growth; and -, no growth; VMM, Vincent's minimal medium; LB, Luria-Bertani medium

sources in M9 medium, whereas VMM uses KNO_3. In addition, we also noted a growth defect on M9 medium plates with succinate as a carbon source and NH_4 as a N-source. When the N-source in M9 was replaced with NO_3, growth defects were not observed. These results indicate that the growth phenotypes seen on glycerol and succinate on M9 medium might not have been caused by the C-sources but a preference of NO_3 as N-source over NH_4. This was not pursued further.

4.3.2 SMc01852 encodes a pyrophosphate dependent phosphofructokinase (Pfp)

To determine if *SMc01852* encodes Pfp we also employed a biochemical approach. A codon optimized *pfp* with an RGS 6-His tag was cloned into a pQE-9 plasmid expression vector and transformed to *E. coli* JM109 cells for overexpression and protein purification. Protein purification was carried out using the Ni-NTA column and 4 elutions of the purified protein sample were collected. A Western blot was done to ensure purity and size of the isolated protein (Figure 22). A band just above the 40 kDa marker was observed which is close to the predicted size of 43.7 kDa for Pfp. Presence of the correct band size was seen in the crude extract, elution 1, 2 and 3 (Figure 22; lanes 1, 5, 6 & 7). Since the strongest band was found in elution 2, it was dialyzed and used to perform the enzyme assays. Protein from elution 2 was quantified by measuring absorbance at 280 nm wavelength and the purification yielded 0.203 mg/ml of protein in the preparation.

The forward and reverse enzymes reactions were carried out using purified Pfp as described previously with some modifications (Reshetnikov et al. 2008; Zhou et al. 2013). For the forward reaction, fructose-6P was used as a substrate and the rate of NADH oxidation was measured as an indicator of enzyme activity. Fructose-1,6 biP was used as a substrate for the reverse reaction and NADP reduction was measured. Saturation curves were generated with varying substrate

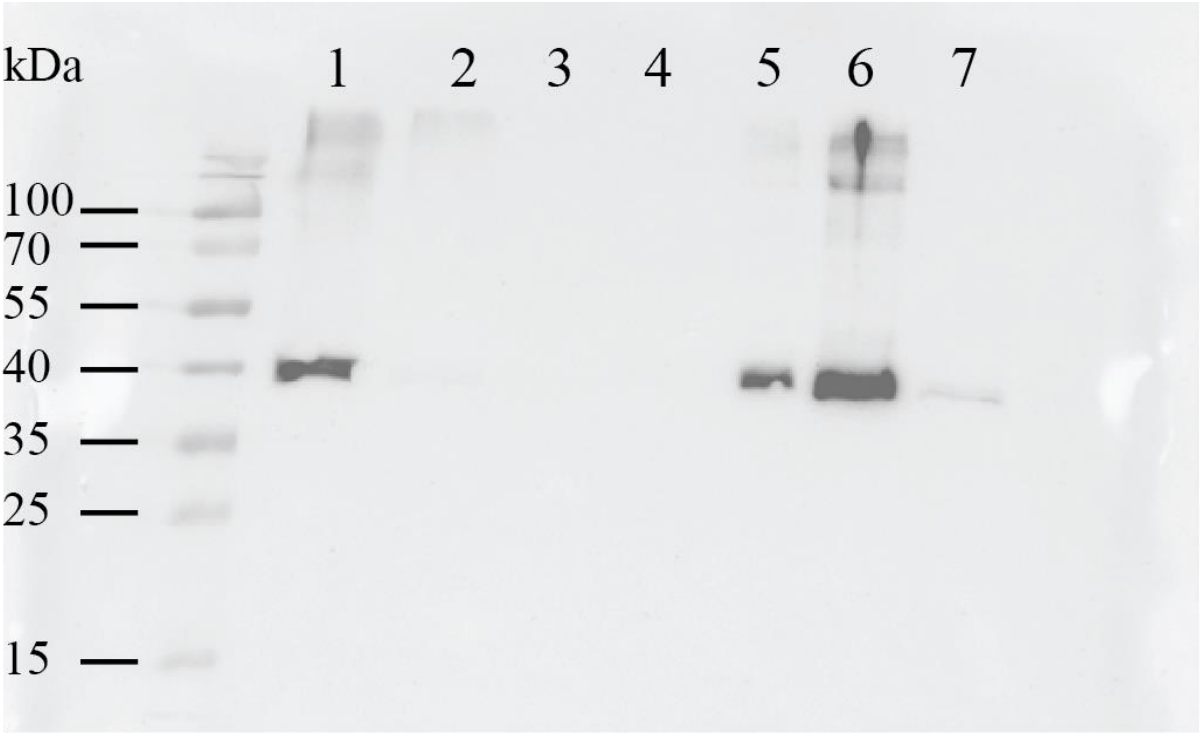


Figure 22. Western blot analysis of Pfp purification.

Protein ladder shows the approximate size of proteins in kDa. Lane 1, cleared lysate; lane 2, flow through; lane 3, binding buffer wash; lane 4, imidazole wash; lane 5, elution 1; lane 6, elution 2; lane 7, elution 3.

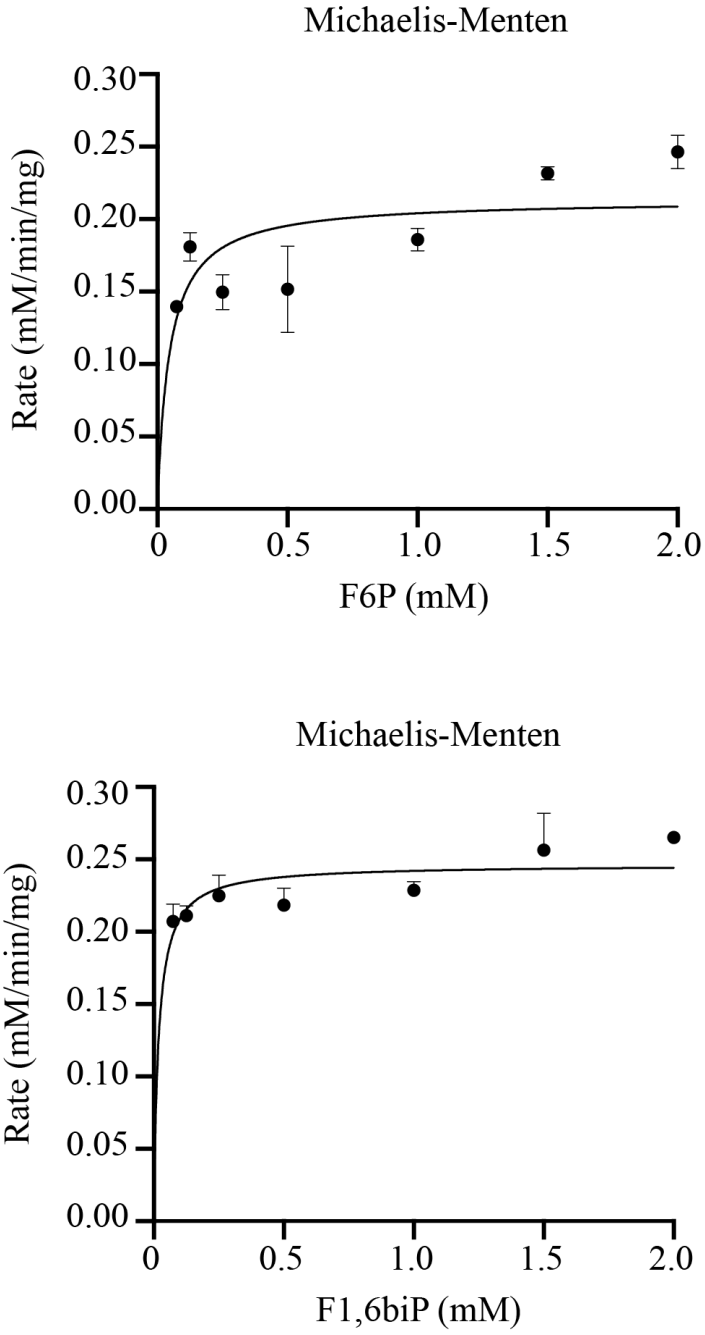


Figure 23. Kinetic analysis of Pfp using fructose-6P (top panel) or fructose-1,6biP (bottom panel) as a substrate.

The data were fitted to a Michaelis-Menten model using GraphPad Prism. The data represented is an average of 3 technical replicated with error bars representing standard deviation.

Table 13. Kinetic properties of Pfp

Substrate	K_m **(mM)**	V_{max} **(mM min^{-1} mg^{-1})**	V_{max}/K_m **(min^{-1})**
Fructose-6P	0.018	0.2464	13.69
Fructose-1, 6biP	0.046	0.2135	4.64

K_m, reaction constant; V_{max}, maximum rate achieved by the reaction.

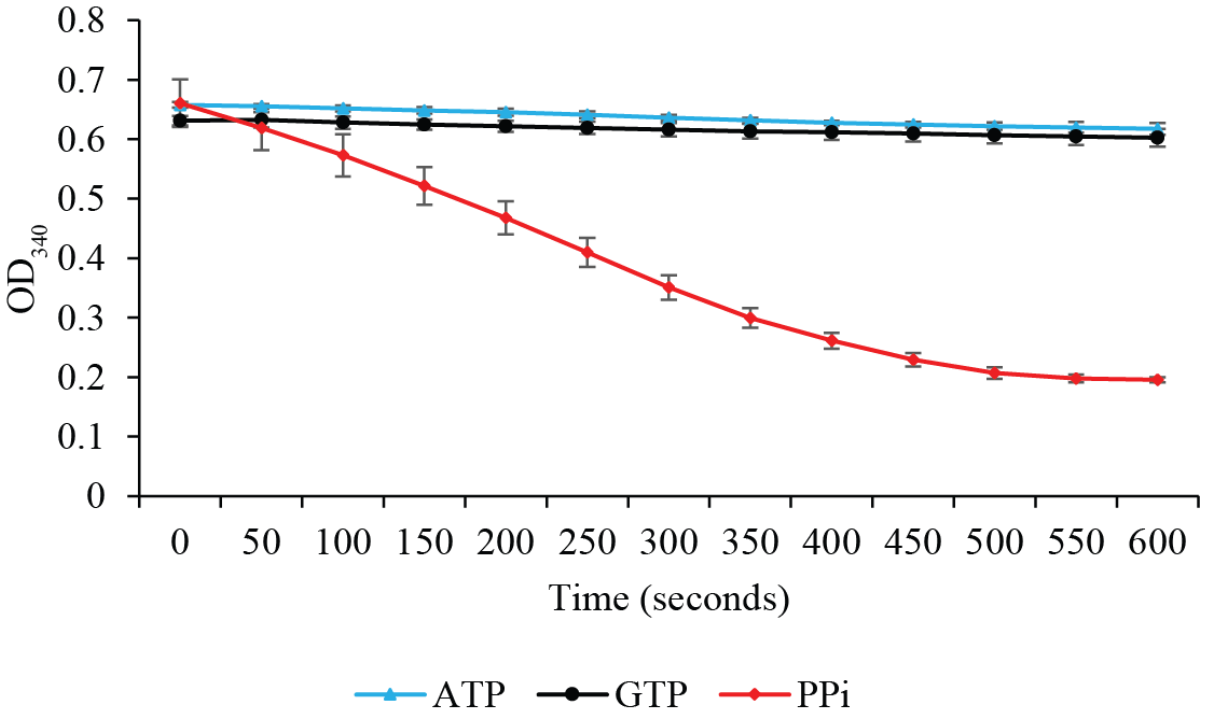


Figure 24. Phosphofructokinase activity of purified Pfp using ATP, GTP or PPi as phosphate donor.

The activity was coupled to the production of NAD^+ by glycerol-3P dehydrogenase, triosephosphate isomerase and fructose 1, 6bisP aldolase at 340 nm for 10 minutes. Error bars represent standard deviation of 3 technical replicates.

concentrations to measure Pfp activity and calculate the reaction constants (Figure 23). Pfp was able to catalyse both forward and reverse reactions, meaning a fully functional Pfp is present in *S. meliloti*. The data from enzyme kinetics were fitted to Michelis-Menton equation to calculate the rate and maximum velocity of the reactions (Figure 23). K_m for the forward reaction was recorded to be 0.018 mM whereas the reverse reaction has a K_m of 0.046 mM. The maximum velocity (V_{max}) for fructose-6P was recorded at 0.2464 mM/mg/min whereas for fructose-1,6 biP it was 0.2135 mM/mg/min. Taking these values into account the overall reaction efficiency (V_{max}/ K_m (min^{-1})) of the forward reaction (13.69) is more than the reverse reaction (4.64) (Table 13). We note that the forward and reverse reactions were both carried out on two independent purifications. The kinetics, however, were only carried out as a single experiment with the appropriate replicates and were not repeated due to time constraints.

To test if the reaction can be carried out using other phosphate donors ATP and GTP were used in place of PP_i. Pfp was unable to carry out the conversion of fructose-6P to fructose-1,6biP using ATP or GTP, suggesting Pfp is strictly uses inorganic PP_i as a phosphate donor in *S. meliloti* (Figure 24).

4.3.3 Metabolite analysis of Δpfp strains

The genetic and biochemical data together suggested that *S. meliloti* contains a functional Pfp. To understand the broader effect of Pfp on central carbon metabolite pools, a metabolite analysis was performed. To carry out this experiment SRmD674 which carries a Δ*pfp* and the wild type were grown in defined medium with either glucose as a glycolytic carbon source, or succinate as a gluconeogenic carbon source. The individual metabolite concentrations as well as a hierarchical clustering of the data are presented (Table 14, Figure 25).

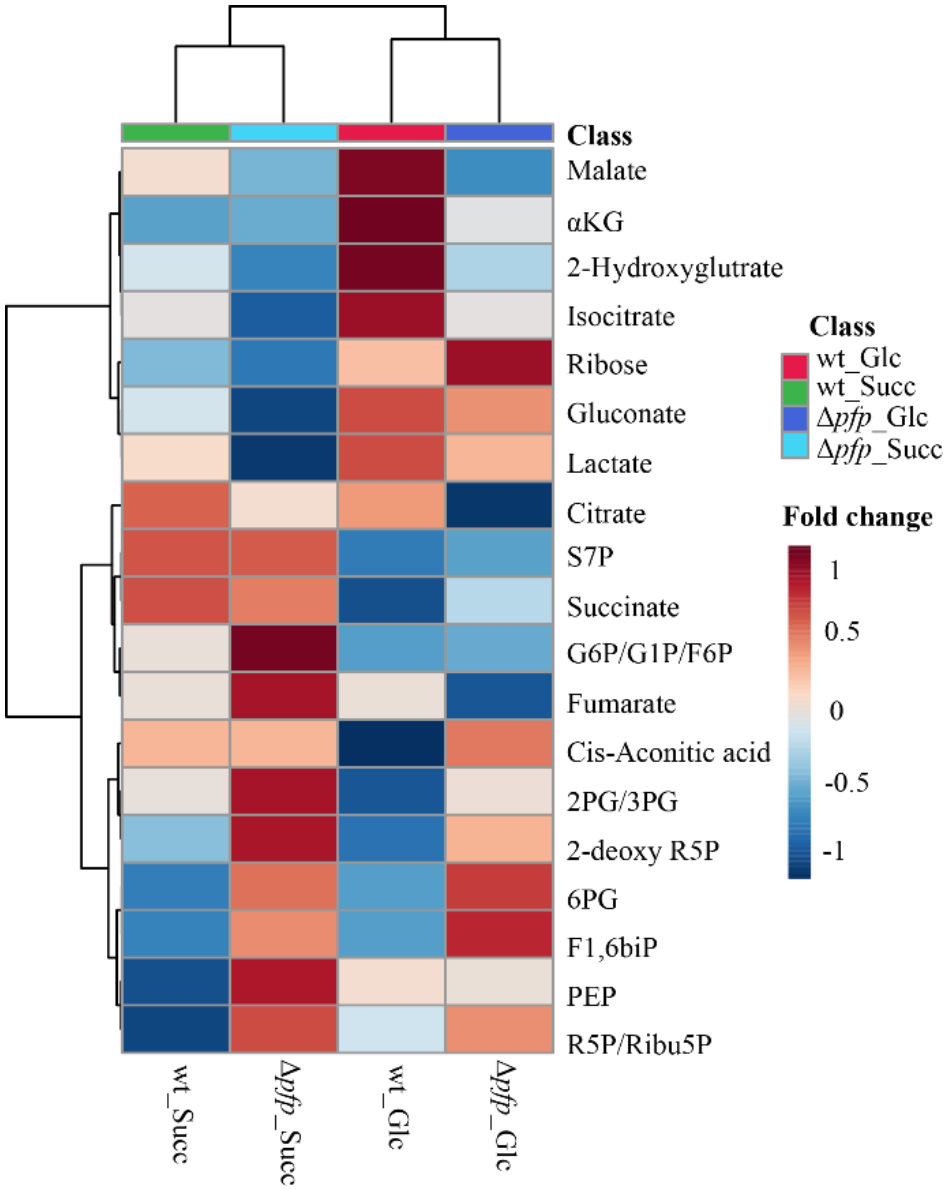


Figure 25. Clustered heatmap of central carbon metabolite concentrations in the wild type and Δ*pfp* strains grown on glucose or succinate as carbon source.

Classes represent *S. meliloti* strains as: red, wild type (glucose); green, wild type (succinate), dark blue, Δ*pfp* (glucose), light blue, Δ*pfp* (succinate). Metabolites are represented by each row, while the mean of three replicates for each strain is represented by a column. The dendrogram and colored boxes were generated with MetaboAnalyst 5.0. The color scale represents fold changes with respect to the average value of each metabolite. Colors indicate fold increases or decreases in concentrations, as indicated in the key. aKG, a-ketoglutarate; S7P, sedoheptulose-7-phosphate; G6P, glucose-6-phosphate; G1P, glucose-1-phosphate; F6P, fructose-6-phosphate; 2-PG, 2-hosphoglycerate; 3-PG, 3-phosphoglycerate; 6PG, 6-phosphogluconate; F1,6biP, fructose-1,6-bisphosphate PEP, phosphoenolpyruvate; R5P, ribose-5-phosphate; Ribu5P, ribulose-5-phosphate.

Table 14. Central carbon metabolite concentrations found in wild type and Δ*pfp* strains.

Metabolite (nM/mg)	**wt glucose**	**wt succinate**	**Δ*pfp* glucose**	**Δ*pfp* succinate**
ED/EMP				
G6P/G6P/G1P	1943	2772	2027	5153
Gluconate	212	68	154	19**
6P-Gluconic acid	321	304	718*	643
2P-Glyceric acid/3P-Glyceric acid	95	386	398	1434**
Phosphoenolpyruvate	298	258	290	374 **
F1,6 biP	126	95	2011*	979**
PPP				
Ribose 5P/Ribulose 5P	731	581	882	918
Sedoheptulose-7P	441	1971	562	1871
2-Deoxy-Ribose-5P	82	74	189	426
Ribose	19682	15279	27800	13394**
TCA				
Isocitrate	154	86	86*	50
Fumarate	732	703	526	1015
Cis-aconitic acid	14	12	15	12
α-ketoglutrate	73	29	41	30
2-Hydroxyglutarate	27	15	14*	12
Succinate	1295	47393	3831	29963**
Lactate	458	369	386	228
Malate	304540	234833	195822	206870

*Significant differences in metabolite concentrations of glucose grown cells when compared to the glucose grown wild type. ** Significant differences in metabolite concentrations of succinate grown cells when compared to the succinate grown wild type. Significance ($P < 0.05$) was determined by Student's *t*-test. n = 3

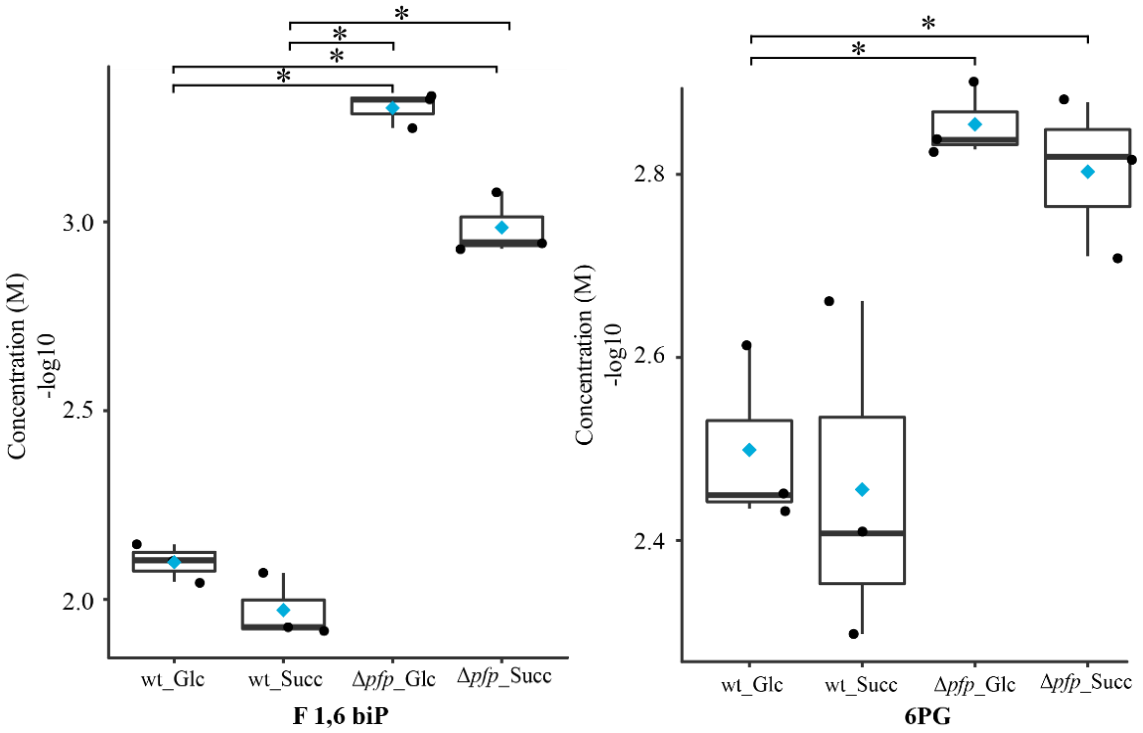


Figure 26. Metabolite concentrations

Box plot representation of metabolite concentrations of F1, 6biP (panel A) and 6PG (panel B) pools in the wild type and Δ*pfp* strains. The Y-axis shows the concentration on a -log10 scale and whereas the strains are labelled on the x-axis.

* Significant difference between two samples. Significance was calculated by Student's *t*-test with $P < 0.05$. The blue diamond indicates the means value of replicates, the black dots indicate the values of each individual replicate.

By analyzing the data using a hierarchical clustering of central carbon metabolites in wild type and Δ*pfp* strains it was found that both strains clustered together based on the carbon source they were grown. However, there are differences among the strains in some specific metabolites.

The Δ*pfp* strain shows an accumulation of some ED, EMP and PP pathway metabolites such as 6-phosphogluconate, fructose-1,6 biP and ribose/ribulose-P, whereas the TCA cycle metabolite pools have minor differences amongst both strains (Table 14). The most striking differences can be seen for fructose 1,6biP (Figure 26). In this case the Δ*pfp* strain accumulation of fructose-1,6biP was 10-fold higher than that of the wild type. This value is significantly higher than the wild type when it is grown on either glucose or succinate. The accumulation of fructose-1,6 biP should also have some effect on the PP pathway metabolites which can be seen in increased levels of ribose/ribulose-5P pools in the Δ*pfp* strain. Similarly, a 2-fold increase in 6PG pool was recorded in the Δ*pfp* strain when grown on glucose; this value is significantly higher than the wild type (Figure 26). Since glucose is assimilated in a glycolytic fashion mainly via the ED pathway, the Δ*pfp* strain should have a blocked EMP pathway, resulting in an increase in flux through the ED pathway and causing the accumulation of 6PG. Collectively these data corroborate the biochemical assays that show PP_i dependent Pfk activity.

4.3.4 A Δpfp Δtal double mutation affects the utilization of multiple carbon substrates

A Δ*pfp*Δ*tal* double mutant strain of *S. meliloti* was constructed to determine if *pfp* can be linked to the PP pathway. The *tal* mutation was chosen over *tktA* (encoding for transketolase) since mutations in *tktA* are auxotrophic and severely affected in symbiosis, whereas the strains with mutations in *tal* do not show such stark phenotypes (Hawkins et al. 2018). Our preliminary screening of this mutant showed that the strain carrying both mutations was more compromised on glucose, succinate, and glycerol but not on galactose (Table 12). The reintroduction of *pfp* on

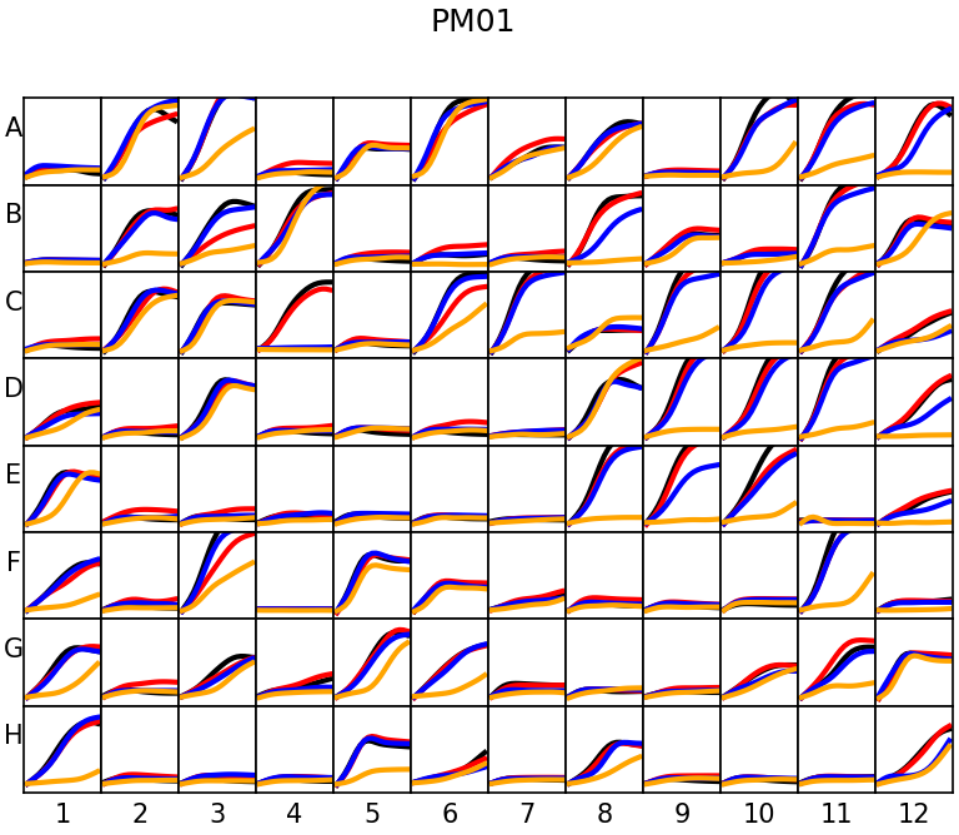
a plasmid reversed these phenotypes implying that these phenotypes were a result of the addition of a *pfp* mutation to a *tal* mutation. The *pfp tal* double mutant strain was grown on phenotypic microarray plates (Biolog plates PM01 and PM02A) to get a better idea of carbon utilization profile of this strain.

Four *S. meliloti* strains, wild type, Δ*pfp*, Δ*tal* and, Δ*pfp*Δ*tal* were tested for growth on 190 carbon sources. The data are shown in Figure 27 and Figure 28 which show a composite grid of growth curves representing these strains on all carbon sources. Table 15 contains the list of sugars on which Δ*pfp*Δ*tal* strain was unable to grow.

The Δ*pfp*Δ*tal* strain was unable to grow on sugars that are incorporated through the ED, EMP or PP pathways. More interestingly, the Δ*pfp*Δ*tal* strain could not grow on glucose but showed wild type levels of growth on galactose. Although both sugars are stereoisomers, their entry into central carbon metabolism in *S. meliloti* is at different points (Geddes and Oresnik 2014). Glucose is converted to glucose-6P and enters mainly glycolysis via the ED pathway, galactose on the other hand is incorporated through the De Ley-Doudoroff pathway and gets broken down to glyceraldehyde-3P and pyruvate (Gosselin et al. 2001; Arias and Cerveñansky 1986). In addition to galactose, the Δ*pfp*Δ*tal* strain was able to grow on L-fucose (B4) and myo-inositol (F3) (Figure 27). Both sugars are incorporated via pyruvate, therefore avoiding the ED and EMP pathways (Milunovic et al. 2014; Jacob et al. 2008; Geddes and Oresnik 2014; Kohler et al. 2010). It was recently shown that L-fucose is converted to 2, 4- diketo-3-deoxy-L-fuconate by a series of reactions and later broken down into pyruvate and lactate (Kohlmeier 2019).

Table 15. Central carbon metabolism phenotypes of the Δ*pfp*Δ*tal* strain

Growth of Δ*pfp*Δ*tal*	**Carbon source**
Growth	galactose, L-fucose, myo-inositol
No Growth	glucose, mannitol, sorbitol, fructose, D-arabitol, mannose, glycerol, adonitol, L-arabitol, ribose, xylose, erythritol, arabinose, dulcitol, maltose, D-melibiose, L-lactose, sucrose, maltotriose, D-cellobiose, D-psicose, maltitol

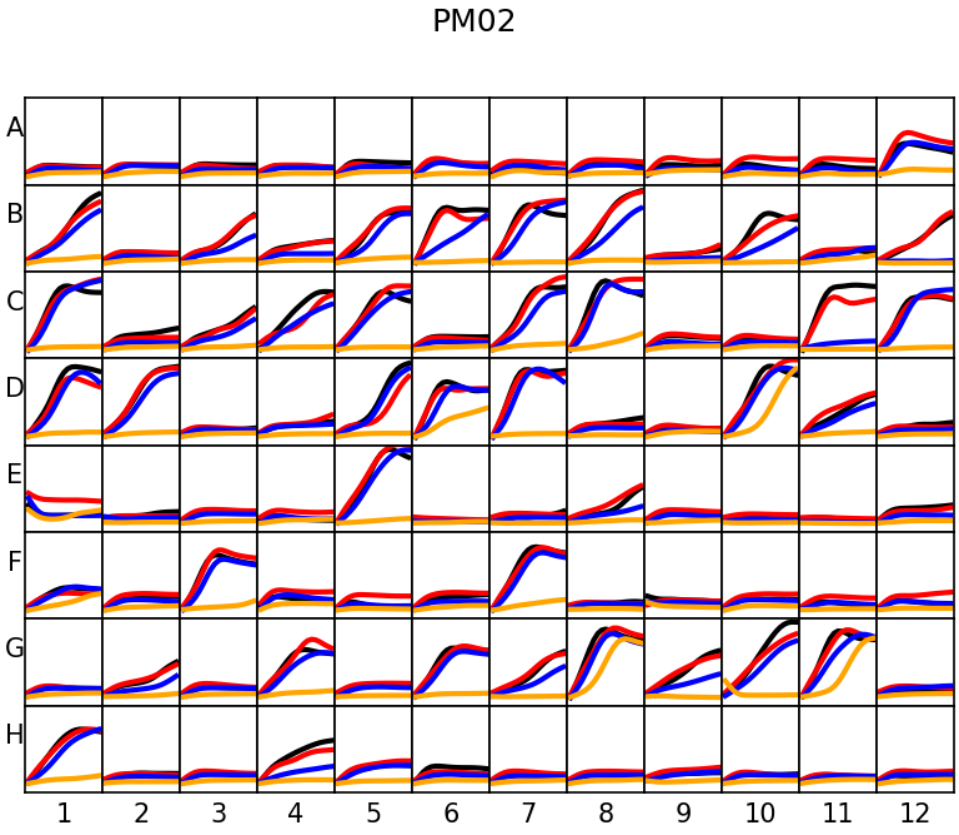


PM1 MicroPlate™ Carbon Sources

A1 Negative Control	A2 L-Arabinose	A3 N-Acetyl-D-Glucosamine	A4 D-Saccharic Acid	A5 Succinic Acid	A6 D-Galactose	A7 L-Aspartic Acid	A8 L-Proline	A9 D-Alanine	A10 D-Trehalose	A11 D-Mannose	A12 Dulcitol
B1 D-Serine	B2 D-Sorbitol	B3 Glycerol	B4 L-Fucose	B5 D-Glucuronic Acid	B6 D-Gluconic Acid	B7 D,L-α-Glycerol-Phosphate	B8 D-Xylose	B9 L-Lactic Acid	B10 Formic Acid	B11 D-Mannitol	B12 L-Glutamic Acid
C1 D-Glucose-6-Phosphate	C2 D-Galactonic Acid-γ-Lactone	C3 D,L-Malic Acid	C4 D-Ribose	C5 Tween 20	C6 L-Rhamnose	C7 D-Fructose	C8 Acetic Acid	C9 α-D-Glucose	C10 Maltose	C11 D-Melibiose	C12 Thymidine
D-1 L-Asparagine	D2 D-Aspartic Acid	D3 D-Glucosaminic Acid	D4 1,2-Propanediol	D5 Tween 40	D6 α-Keto-Glutaric Acid	D7 α-Keto-Butyric Acid	D8 α-Methyl-D-Galactoside	D9 α-D-Lactose	D10 Lactulose	D11 Sucrose	D12 Uridine
E1 L-Glutamine	E2 m-Tartaric Acid	E3 D-Glucose-1-Phosphate	E4 D-Fructose-6-Phosphate	E5 Tween 80	E6 α-Hydroxy Glutaric Acid-γ-Lactone	E7 α-Hydroxy Butyric Acid	E8 ß-Methyl-D-Glucoside	E9 Adonitol	E10 Maltotriose	E11 2-Deoxy Adenosine	E12 Adenosine
F1 Glycyl-L-Aspartic Acid	F2 Citric Acid	F3 myo-Inositol	F4 D-Threonine	F5 Fumaric Acid	F6 Bromo Succinic Acid	F7 Propionic Acid	F8 Mucic Acid	F9 Glycolic Acid	F10 Glyoxylic Acid	F11 D-Cellobiose	F12 Inosine
G1 Glycyl-L-Glutamic Acid	G2 Tricarballylic Acid	G3 L-Serine	G4 L-Threonine	G5 L-Alanine	G6 L-Alanyl-Glycine	G7 Acetoacetic Acid	G8 N-Acetyl-ß-D-Mannosamine	G9 Mono Methyl Succinate	G10 Methyl Pyruvate	G11 D-Malic Acid	G12 L-Malic Acid
H1 Glycyl-L-Proline	H2 p-Hydroxy Phenyl Acetic Acid	H3 m-Hydroxy Phenyl Acetic Acid	H4 Tyramine	H5 D-Psicose	H6 L-Lyxose	H7 Glucuronamide	H8 Pyruvic Acid	H9 L-Galactonic Acid-γ-Lactone	H10 D-Galacturonic Acid	H11 Phenylethyl-amine	H12 2-Aminoethanol

Figure 27. Phenotypic microarray analysis of wild type (black), Δ*tal* (blue), Δ*pfp* (red), Δ*pfp*Δ*tal* (orange) strains.

The growth curves were generated using DuctApe. The grid at the bottom represents the location of each substrate as it is laid out on the 96 well plate. The bottom figure is adapted from the Biolog Inc. website.



PM2A MicroPlate™ Carbon Sources

A1 Negative Control	A2 Chondroitin Sulfate C	A3 α-Cyclodextrin	A4 ß-Cyclodextrin	A5 γ-Cyclodextrin	A6 Dextrin	A7 Gelatin	A8 Glycogen	A9 Inulin	A10 Laminarin	A11 Mannan	A12 Pectin
B1 N-Acetyl-D-Galactosamine	B2 N-Acetyl-Neuraminic Acid	B3 ß-D-Allose	B4 Amygdalin	B5 D-Arabinose	B6 D-Arabitol	B7 L-Arabitol	B8 Arbutin	B9 2-Deoxy-D-Ribose	B10 i-Erythritol	B11 D-Fucose	B12 3-0-ß-D-Galactopyranosyl-D-Arabinose
C1 Gentiobiose	C2 L-Glucose	C3 Lactitol	C4 D-Melezitose	C5 Maltitol	C6 α-Methyl-D-Glucoside	C7 ß-Methyl-D-Galactoside	C8 3-Methyl Glucose	C9 ß-Methyl-D-Glucuronic Acid	C10 α-Methyl-D-Mannoside	C11 ß-Methyl-D-Xyloside	C12 Palatinose
D1 D-Raffinose	D2 Salicin	D3 Sedoheptulosan	D4 L-Sorbose	D5 Stachyose	D6 D-Tagatose	D7 Turanose	D8 Xylitol	D9 N-Acetyl-D-Glucosaminitol	D10 γ-Amino Butyric Acid	D11 δ-Amino Valeric Acid	D12 Butyric Acid
E1 Capric Acid	E2 Caproic Acid	E3 Citraconic Acid	E4 Citramalic Acid	E5 D-Glucosamine	E6 2-Hydroxy Benzoic Acid	E7 4-Hydroxy Benzoic Acid	E8 ß-Hydroxy Butyric Acid	E9 Glycolic Acid	E10 α-Keto-Valeric Acid	E11 Itaconic Acid	E12 5-Keto-D-Gluconic Acid
F1 D-Lactic Acid Methyl Ester	F2 Malonic Acid	F3 Melibionic Acid	F4 Oxalic Acid	F5 Oxalomalic Acid	F6 Quinic Acid	F7 D-Ribono-1,4-Lactone	F8 Sebacic Acid	F9 Sorbic Acid	F10 Succinamic Acid	F11 D-Tartaric Acid	F12 L-Tartaric Acid
G1 Acetamide	G2 L-Alaninamide	G3 N-Acetyl-L-Glutamic Acid	G4 L-Arginine	G5 Glycine	G6 L-Histidine	G7 L-Homoserine	G8 Hydroxy-L-Proline	G9 L-Isoleucine	G10 L-Leucine	G11 L-Lysine	G12 L-Methionine
H1 L-Ornithine	H2 L-Phenylalanine	H3 L-Pyroglutamic Acid	H4 L-Valine	H5 D,L-Carnitine	H6 Sec-Butylamine	H7 D,L-Octopamine	H8 Putrescine	H9 Dihydroxy Acetone	H10 2,3-Butanediol	H11 2,3-Butanedione	H12 3-Hydroxy-2-Butanone

Figure 28. Phenotypic microarray analysis of wild type (black), Δ*tal* (blue), Δ*pfp* (red), Δ*pfp*Δ*tal* (orange) strains.

The growth curves were generated using DuctApe. The grid at the bottom represents the location of each substrate as it is laid out on the 96 well plate. The bottom figure is adapted from the Biolog Inc. website.

The Δ*tal* strain showed slow growth on xylose and did not grow on ribose in liquid medium; however, it can grow slowly on VMM agar plates with ribose as sole carbon source (Hawkins et al. 2018). The discrepancy in growth pattern might be caused by more access to ribose in liquid creating a toxic effect as compared to solid surface. In addition, it was found that the Δ*tal* strain was unable to grow on 3-O-β-D-galactopyranosyl-D-arabinose and β-methyl-D-xyloside (B12) (Figure 28).

4.3.5 Addition of glucose to a medium containing galactose leads to an inability of strain carrying ΔpfpΔtal to grow

The phenotypic microarray analysis revealed that the Δ*pfp*Δ*tal* strain is unable to grow on glucose as a sole carbon source but grows well on galactose. It was hypothesized that this phenotype could be due to the accumulation of phosphorylated intermediates that are generated when glucose, or another substrate that feeds into the EMP, ED, of PP pathways are used as carbons sources. To test this hypothesis, a growth curve of the Δ*pfp*Δ*tal* mutant strain was carried out on galactose as well as galactose that contained increasing concentrations of glucose.

The results showed that whereas the wild type strain can grow under all of the tested conditions, the Δ*pfp*Δ*tal* strain showed varying growth (Figure 29). As expected, the Δ*pfp*Δ*tal* strain showed wild type levels of growth on galactose and no growth on glucose. When 1, 2.5 or 5 mM glucose was added to the medium containing 15 mM galactose, the Δ*pfp*Δ*tal* strain still showed significant growth, albeit the growth of the strain decreased with increasing concentrations of glucose (Figure 29). However, when the concentration of glucose was increased to 7.5 or 15 mM there was essentially no growth.

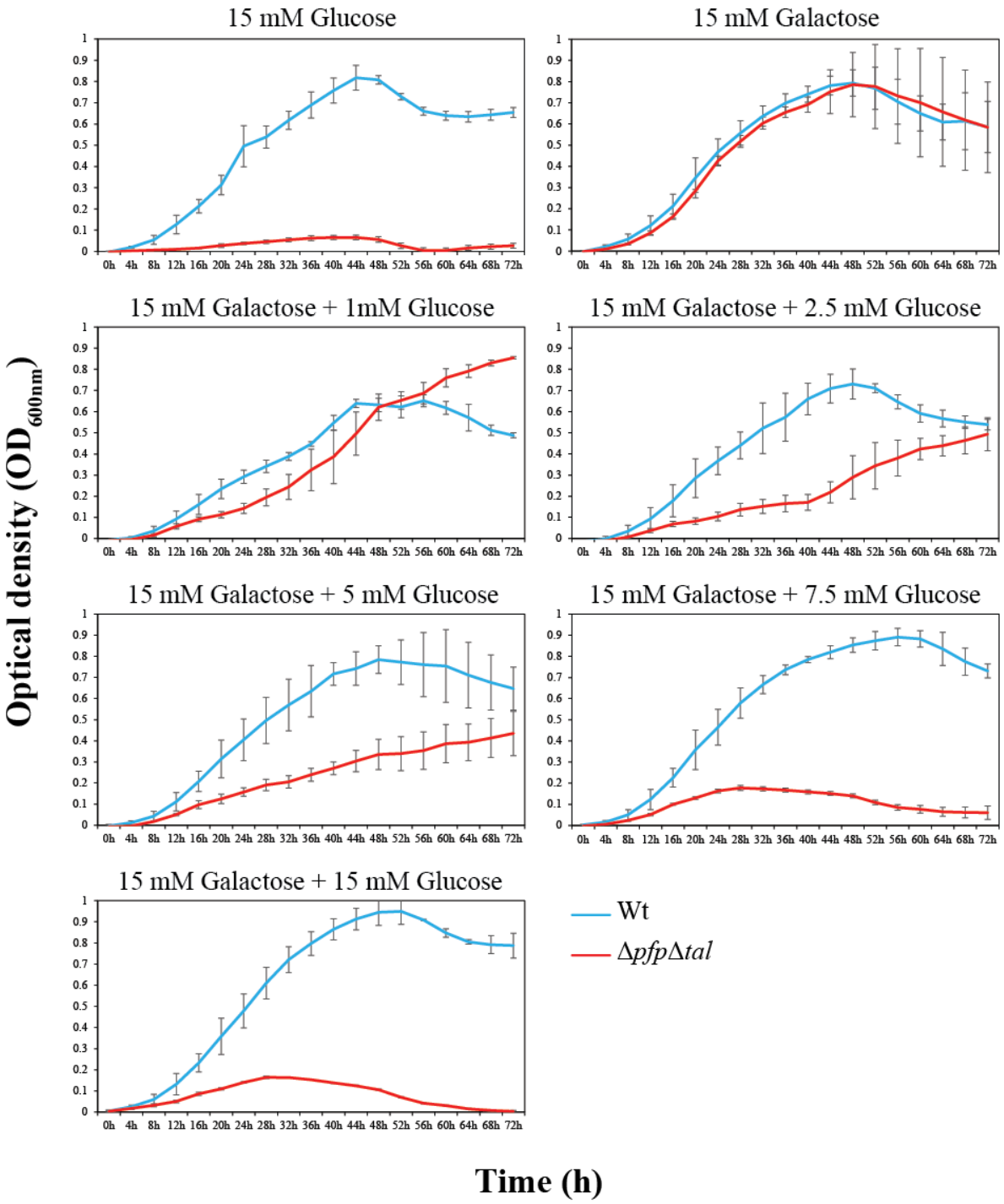


Figure 29. Bacterial growth curve analysis of the wild type (blue) and Δ*pfp*Δ*tal* (red) strains on galactose and glucose as carbon sources.

Error bars represent standard deviation of three biological replicates.

These data indicate that the Δ*pfp*Δ*tal* strain can assimilate glucose at lower concentrations but as the concentration increases it becomes toxic to the cells. We note that growth of the Δ*pfp*Δ*tal* strain on glucose can be rescued by complementing the mutation with a plasmid overexpressing *pfp* (Table 12).

4.3.6 ΔpfpΔtal mutations cause accumulation of phosphorylated intermediates

To determine if higher concentrations of phosphorylated compounds could be correlated to exposure of glucose, a metabolite analysis was performed on the wild type and the Δ*pfp*Δ*tal* strain to study the impact of these mutations on the pools of metabolites involved in major pathways of central carbon metabolism. The *S. meliloti* cells were grown in defined medium using either galactose as sole carbon source or galactose grown cells that were given 5 mM glucose, then allowed to grow for an additional 4 hours prior to harvesting.

A hierarchical clustering analysis of the 19 metabolites assayed shows that the Δ*pfp*Δ*tal* and the wild type strain cluster together despite the growth conditions (Figure 30). The metabolites glucose-6P/glucose-1P/fructose-6P, 2P/3P-glyceric acid, ribose-5P/ribulose-5P, sedoheptulose-7P and ribose, that represent the upper EMP pathway and parts of the PP pathway are clustered together, and these have increased levels in the Δ*pfp*Δ*tal* strain. A significant increase in 6-phosphogluconic acid was recorded in galactose+glucose treatment. 2P/3P-glyceric acid pools were significantly increased for Δ*pfp*Δ*tal* strain grown under both conditions as compared to the wild type (Table 16). As well, significantly higher amounts of ribose were detected in the Δ*pfp*Δ*tal* strain (Table 16). The origin of this free ribose in the cell is unclear. Most of the TCA cycle intermediate pools were relatively unchanged in the Δ*pfp*Δ*tal* strain. Exceptions to this are citrate, α-ketoglutarate levels, that were significantly more reduced in both

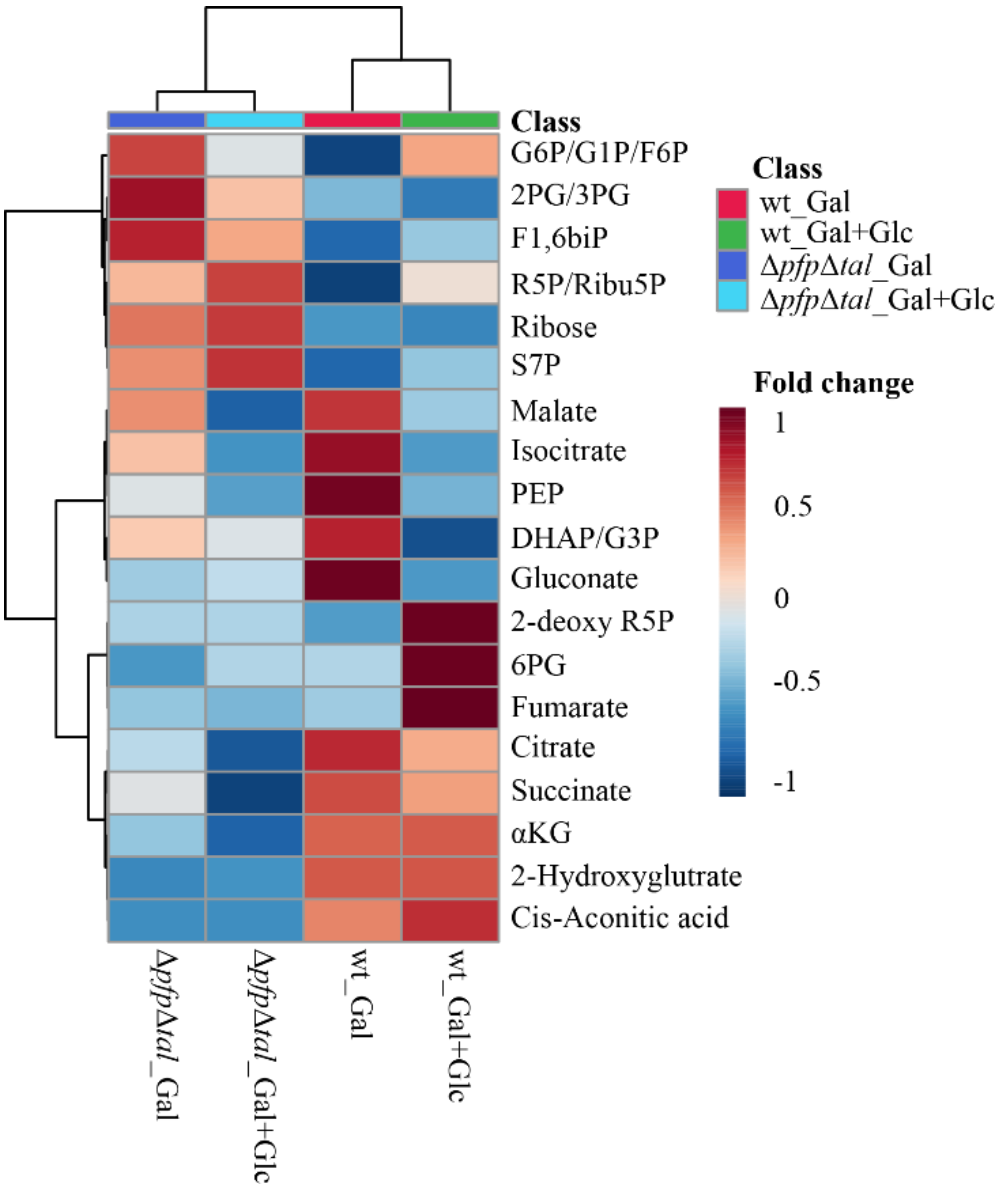


Figure 30. A Clustered heatmap of central carbon metabolite concentrations in the wild type and Δ*pfp*Δ*tal* strains grown either on galactose or galactose+glucose as carbon source.

Classes represent *S. meliloti* strains as: red, wild type (galactose); green, wild type (galactose+glucose), dark blue, Δ*pfp* (galactose), light blue, Δ*pfp* (galactose+glucose). Metabolites are represented by each row, while mean of four replicates for each strain is represented by a column. The dendrogram and colored boxes were generated with MetaboAnalyst 5.0. The color scale represents fold changes with respect to the average value of each metabolite. Colors indicate fold increases or decreases in concentrations, as indicated in the key. G6P, glucose-6-phosphate; G1P, glucose-1-phosphate; F6P, fructose-6-phosphate; 2-PG, 2-hosphoglycerate; 3-PG, 3-phosphoglycerate; F1,6biP, fructose-1,6-bisphosphate; R5P, ribose-5-phosphate; Ribu5P, ribulose-5-phosphate; S7P, sedoheptulose-7-phosphate; PEP, phosphoenolpyruvate; DHAP, dihydroxyacetone phosphate; G3P, glyceraldehyde-3-phosphate; 6PG, 6-phosphogluconate; aKG, a-ketoglutarate.

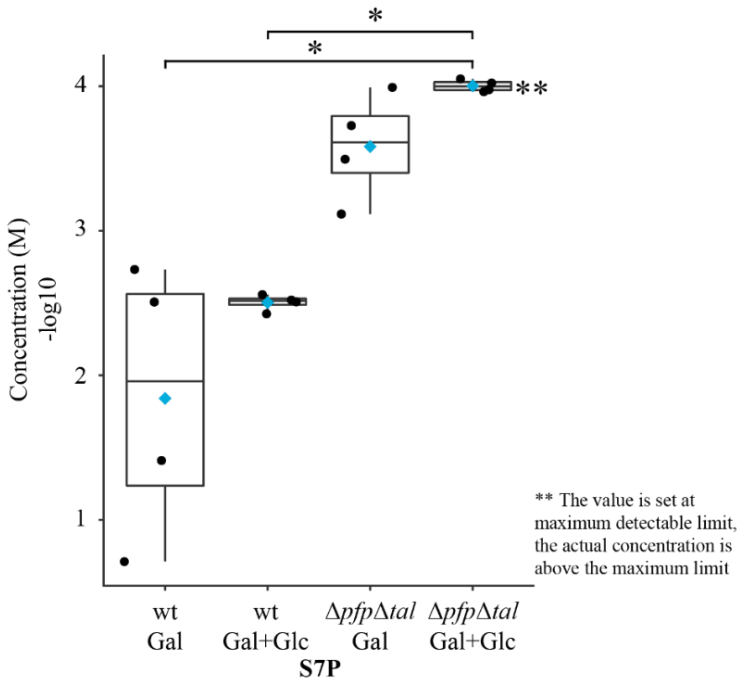


Figure 31. A box plot representation of sedoheptulose-7P pools in the wild type and Δ*pfp*Δ*tal* strains.

The Y-axis shows the concentration on a -$\log_{10}$ scale and the strains and conditions are labelled on the X-axis. *Significant difference from the wild type grown under same conditions; ** The value of sedoheptulose-7P was set to the maximum detectable limit of the instrument. Significance was calculated by Student's *t*-test with $P < 0.05$. Blue diamonds represent the mean value of the replicates. The black dot indicates the value of each of the individual replicates.

Table 16. Central carbon metabolite concentrations found in wild type and Δ*pfp*Δ*tal* strains

Metabolite (nM/mg)	**wt Galactose**	**wt Galactose+Glucose**	**Δ*pfp*Δ*tal* Galactose**	**Δ*pfp*Δ*tal* Galactose+Glucose**
ED/EMP				
G6P/G6P/G1P	940	1600	2697*	1114
DHAP/G3P	297	168	243	219
Gluconate	897	448	552	634
6P-Gluconic acid	277	352	76	118**
2P-Glyceric acid/3P-Glyceric acid	180	142	522*	306**
Phosphoenolpyruvate	163	113	125	113
F1,6 biP	81	74	1516	290
PPP				
Ribose 5P/Ribulose 5P	365	425	1053	1782
Sedoheptulose-7P	296	317	6100	9723**
2-Deoxy-Ribose-5P	92	151	33	27
Ribose	9139	7698	62123*	85653**
TCA				
Isocitrate	692	304	477	309
Fumarate	634	944	543	583
Citrate	741	383	191*	68**
Cis-aconitic acid	67	49	ND	ND
α-ketoglutrate	217	213	79*	43**
2-Hydroxyglutarate	312	314	126*	122
Succinate	4566	4024	3495	2413**
Malate	275255	226055	260422	210256

*Significant differences in metabolite concentrations of galactose grown cells when compared to the galactose grown wild type. ** Significant differences in metabolite concentrations of galactose grown and induced with glucose cells when compared to the galactose grown and induced with glucose wild type. Significance ($P < 0.05$) was determined by Student's *t*-test, n = 4.

conditions in the mutant than the wild type (Figure 30). Succinate was the only other metabolite that had a significant reduction, but only in the presence of glucose (Table 16).

The changes seen in sedoheptulose-7P are the largest in this data set (Figure 31). Although the concentration of sedoheptulose-7P was higher in the Δ*pfp*Δ*tal* strain under all conditions, the sedoheptulose-7P levels were not significantly higher than the wild type when the Δ*pfp*Δ*tal* strain was grown on just galactose (Figure 31). However, with the addition of glucose, the levels of sedoheptulose-7P levels were consistently above the top limit of detection of the instrument. To provide an approximate estimate of the amount of the metabolite, it was assumed that the value was at the limit of detection (Figure 31). With this assumption, the concentration of sedoheptulose 7-phospahte is nearly two orders of magnitude greater than that measured in the wild type (Figure 31). This is consistent with the hypothesis that the accumulation of phosphorylated intermediates occurs in the *ΔpfpΔtal* when it is exposed to glucose, and by extension to any carbon compound that enters through the ED, EMP, or PP pathway.

4.3.7 The ΔpfpΔtal strain of S. meliloti displays reduced symbiotic efficiency

To determine if the *pfp* mutation affected the symbiotic process, the strain SRmD674 which contains just the *Δpfp* mutation as well as SRmD688 which contains the *Δpfp* mutation as well as the Δ*tal* mutation were inoculated onto alfalfa. After 4 weeks all plants looked healthy and green. However, the plants inoculated with the Δ*pfp*Δ*tal* strain were shorter than the wild type, the Δ*pfp* mutant and Δ*tal* mutant strains (Figure 33). The uninoculated plants were small, chlorotic and with plate yellow leaves. Upon inspection of the root nodules, all plants contained elongated pink nodules except the ones inoculated with the Δ*pfp*Δ*tal* strain which produced a mix of elongated and short nodules. Generally, the nodules induced by the Δ*pfp*Δ*tal* strain were small relative to the wild type. The shoot dry weight of these plants is

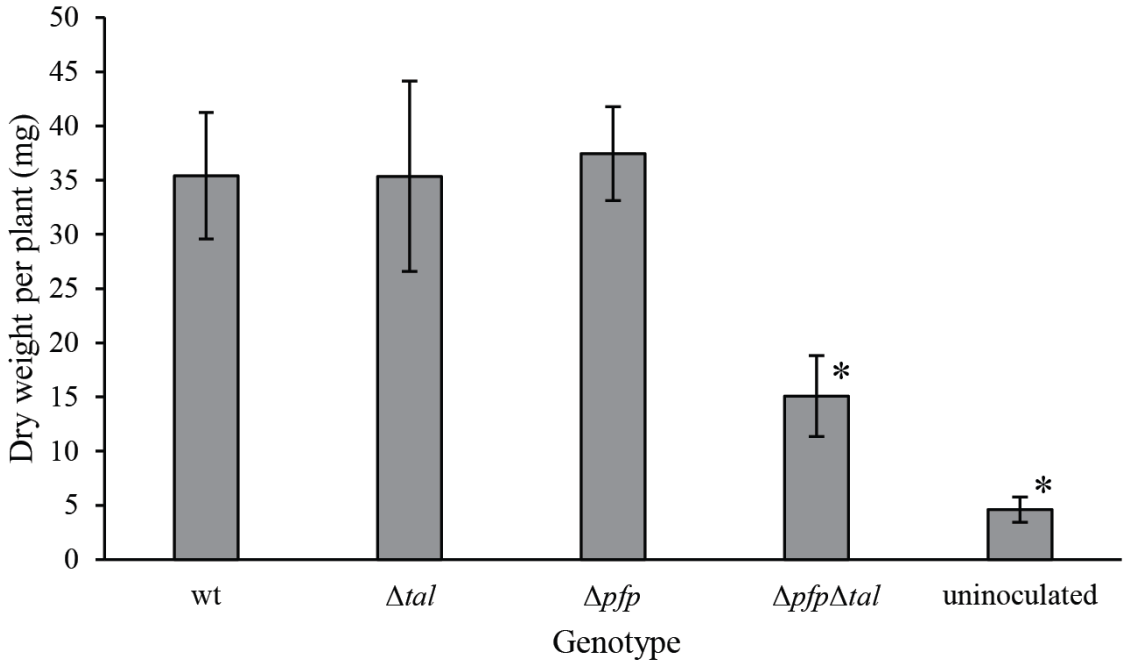


Figure 32. Dry weights of alfalfa plants inoculated with *S. meliloti* strains.

Plants were grown in sterile sand-vermiculite in nitrogen deficient conditions. The relevant genotypes are on the x-axis. Plants were harvested 4-weeks after inoculation and were dried at 60ºC for 24h to calculate dry weight. Each bar represents data from 3 biological replicates each containing 10 plants. *Significant difference from wild type inoculated plants calculated by Student's *t*-test with $P < 0.05$.



Figure 33. Alfalfa plants grown in nitrogen deficient conditions inoculated with (from left) wild type, Δ*tal*, Δ*pfp*, Δ*pfp*Δ*tal* strains and uninoculated control.

presented in Figure 32. The plants inoculated with the Δ*pfp* or Δ*tal* strains produced wild type levels of shoot dry weight and were significantly different from the uninoculated control. The Δ*pfp*Δ*tal* strain on the other hand showed significant reduction in shoot dry weight in comparison to wild type. The Δ*pfp*Δ*tal* strain produced 15 mg/plant dry weight whereas the wild type plants produced 35 mg/plant of shoot dry weight.

4.4 Discussion

The phosphofructokinase reaction is often considered the first committed step of the EMP pathway (Mulukutla et al. 2014; Hofmann 1978). In many organisms the ATP dependent Pfk reaction is physiologically essentially irreversible, and its activity is highly sensitive to allosteric inhibition (Babul 1978). The lack of this activity in *S. meliloti* in conjunction with a number of ^{13}C-labelling studies that reported C-flux through 6-phosphogluconate and the ED pathway have been used as evidence to support the lack of this biochemical activity (Fuhrer et al. 2005, Arias et al. 1979). In this chapter we show unambiguously that *S. meliloti* Rm1021 contains a gene that is capable of carrying out a pyrophosphate dependent phosphofructokinase reaction (Figure 23, Figure 24, Table 13). Moreover, this reaction does play a role in central carbon metabolism that becomes evident by the changes associated with metabolite pools that can be seen in a strain that carries a Δ*pfp* mutation (Figure 25, Figure 26, Table 14).

A major difference between Pfk and Pfp is that the latter can carry out both forward and reverse reactions. It can be argued that the ability to perform a reversible reaction may provide a metabolic advantage since it allows the EMP pathway to play a glycolytic and a gluconeogenic role. Pfp has been used as an example to show the energetic advantage of using PP_i instead of ATP. The free energy (ΔG^o) for PP_i hydrolysis is lower than that for ATP under a range of conditions that were modeled to mimic a cytosol. (Davies et al. 1993). A quantitative assessment

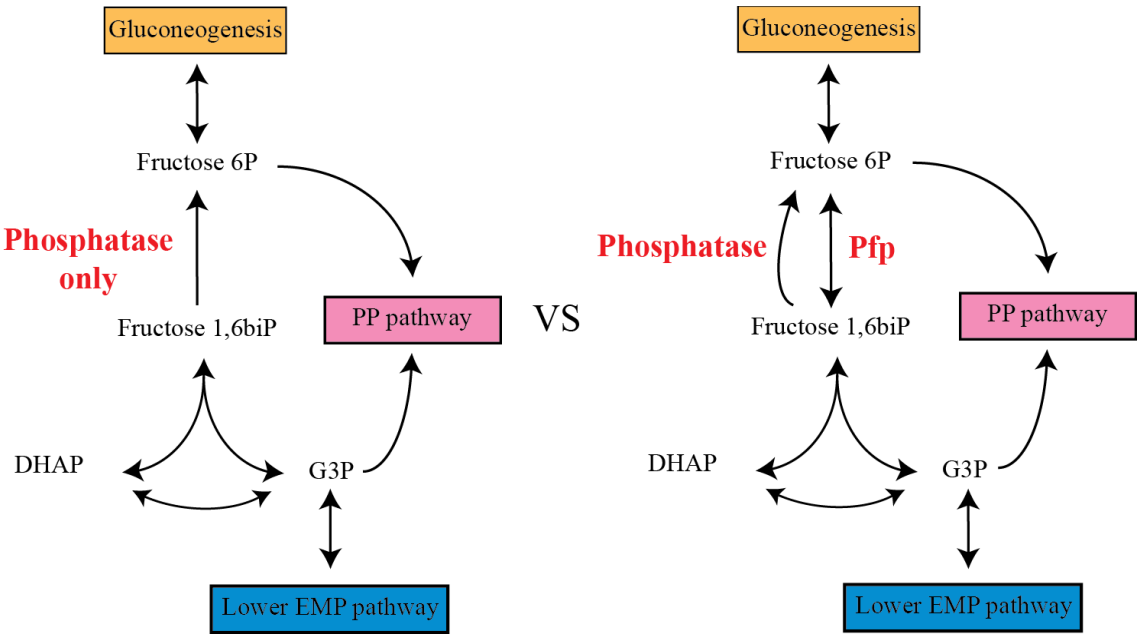


Figure 34. Proposed role of Pfp in upper EMP and PP pathways.

Pfp, pyrophosphate dependent phosphofructokinase; Fructose 6P, fructose-6-phosphate; Fructose 1,6biP, fructose-1,6-bisphosphate; DHAP, dihydroxyacetone phosphate; G3P, glyceraldehyde-3-phosphate; EMP, Embden-Meyerhof-Parnas pathway; PP, pentose phosphate pathway

of *ΔG* in glycolysis was done using radioisotope labelling in the anerobic bacterium *Clostridium cellulolyticum* (Park et al. 2019). In this organism glycolysis operates at near equilibrium with an overall change in the *ΔG* ~ -3 Kj/mol from glucose-6P to phosphoenolpyruvate. The reason for this small change in *C. cellulolyticum* is that it uses a PP_i-dependent Pfp which has a small driving force to produce fructose-1,6biP as compared to Pfk (ATP to ADP *ΔG'*= -10 Kj/mol; PPi to Pi *ΔG'*= -1 Kj/mol). This leads to a lower flux rate through glycolysis following the Pfp reaction. An advantage of this slower glycolysis comes in the form of extra ATP equivalents produced during glycolysis (2-3 ATP equivalents from 1 glucose) as compared to the canonical glycolytic pathway (Park et al. 2019). At present we do not know what bioenergetic advantage might be gained by *S. meliloti* by having this enzyme.

The data clearly show that in a strain that is carrying a *pfp* mutation there is an increase of fructose-1,6biP (Figure 26), suggesting that the lesion does lead to a metabolic block. Surprisingly, the strain can grow in a defined medium using either glucose or succinate (Table 12). The ability to utilize glucose as a sole carbon source is predictable based on the use of the ED pathway. The ability to grow on a gluconeogenic substrate suggests that a metabolic bypass must exist that can bypass this reaction. It has been previously predicted that this reaction could be carried out by a phosphatase reaction (Finan et al. 1988) (Figure 34). It is noteworthy that within the genome there is an annotated fructose 1,6-biP aldolase (*cbbF*) encoded in the *cbb* operon that encodes the enzymes for the Calvin-Benson-Bassham pathway (Geddes et al. 2014). Alternately a more complex bypass may also exist that might utilize the reactions of the non-oxidative PP pathway.

Mutations in *pfp* do not affect symbiosis with alfalfa (Figure 32, Figure 33). A recent study using insertion sequencing (In-seq) to study the lifestyle stages of symbiosis in *R.*

leguminosarum-pea symbiotic association has been carried out and genes were classified as either growth essential, growth defective, or growth neutral at various stages of symbiosis (Wheatley et al. 2020). Within their analysis the *R. leguminosarum* orthologue of *pfp* (RL3322), was classified as being neutral with respect to either affecting growth in the rhizosphere, root colonization, or bacteroid development (Wheatley et al. 2020). Collectively, this suggests that *pfp* does not play an essential role in symbiosis in either *R. leguminosarum* or *S. meliloti*. The gene however was classified as growth defective in bacteria isolated from developed nodules (Wheatley et al. 2020). It is unclear whether this suggests that the Pfp may be utilized within the infection thread and prior to terminally differentiating into a nitrogen fixing bacteroid, or if it reflects the ability of the bacteria to survive once released from the nodule (Wheatley et al. 2020).

The phenotype of a Δ*pfp*Δ*tal* strain encompasses all aspects of its physiology. This strain is compromised in its ability to utilize a number of carbon sources, and when it enters into a symbiotic relationship it is capable of fixing nitrogen, but at reduced levels. The ability of this strain to develop nitrogen fixing bacteroids suggests that while in the curled colonized root hair and the infection thread it does not encounter high concentrations of carbon sources that are catabolized via the EMP or PP pathways since these can lead to a build-up of phosphorylated intermediates. In addition, it is not clear why the amount of nitrogen that was fixed by this strain is lower. It could be a result of slower growth that leads to slower development. Alternately it may also be due to a reduced capacity to reduce nitrogen. Since the primary carbon source to the actively fixing bacteroid is a dicarboxylate (Udvardi and Poole 2013; Yurgel and Kahn 2004), it is difficult to envision a direct role of Pfp and Tal in contributing to the energetics directly

involved with nitrogenase. This suggests that the role of these enzymes must occur prior to bacteroid differentiation. This is consistent with the recent In-seq analysis (Wheatley et al. 2020).

Chapter 5:

Conclusions

5.1 Conclusions from the current work

The goal of this thesis was to use *S. meliloti* strains with mutations in *tktA, tktB* and *pfp* to assess their role in central carbon metabolism and symbiosis in the *S. meliloti*-alfalfa model system. In chapter 2 and 3 we have used *tktA* and *tktB* mutants to characterize their role in the pentose phosphate pathway physiologically as well as their role in symbiosis. Chapter 4 addresses the missing phosphofructokinase reaction in the EMP pathway of *S. meliloti*. Overall, we tested the hypothesis that central carbon metabolism plays a role in symbiosis by controlling the flow of carbon during symbiosis. The conclusions and main observations from the thesis are discussed below.

In the second chapter we have described that the metabolic and symbiotic block caused by mutations in *tktA* can be partially alleviated by suppression. Ten independent second site suppressor strains of *tktA* were isolated and all were found to have point mutations in *SMc02340*, a negative regulator in close proximity of *tktB*. A qRT-PCR assay demonstrated that *tktB* is upregulated in the suppressor strains. This is supported by previous work that showed that a plasmid overexpressing *tktB* could complement the carbon phenotypes associated with a *tktA* mutant strain (Hawkins et al. 2018). Surprisingly, characterization of some of these suppressors demonstrated that there was not a full restoration of metabolism or symbiosis as shown by metabolomic analysis and the plant dry weights assays.

A possible reason for incomplete suppression might be that the overall amount of transketolase produced by *tktB* may be less than what is normally produced by *tktA*. *Salmonella* possess more than one transketolase and enzymatic assays have shown that *tktA* contributes 88% of the total transketolase activity (Shaw et al. 2018). We have noted that Δ*tktB* mutation in a wild type background does not affect either growth on minimal medium or symbiosis. This implies

that in the wild type *tktB* either does not play a role or plays a very minor role in the pentose phosphate pathway that we are currently unable to quantify. A transketolase assay using the wild type, a *tktA* mutant, and the suppressor strains could reveal the contribution of each transketolase. Having two functional transketolases does provide the evidence of metabolic redundancy in *S. meliloti*. Similar observations have been reported previously where both *tpiA* and *eryH* (previously annotated as *tpiB*) exhibit triosephosphate isomerase (Tpi) activity in *S. meliloti*. However, *eryH* also plays a role in erythritol metabolism (Geddes and Oresnik 2012a). The *tpiA* mutations can be complemented by *eryH*; however, erythritol phenotypes associated with *eryH* mutation cannot be complemented by *tpiA* (Poysti and Oresnik 2007).

A microscopic analysis revealed that the *tktA* mutant strain is unable to form curled colonized root hairs (CCRH) and infection threads (IT) on alfalfa. The *tktA* mutant strain is affected in EPS production and displays aromatic amino acid auxotrophy (Hawkins et al. 2018), but the symbiotic phenotype of the *tktA* mutant strain is different from published EPS mutants like *exoR* and *exoY* (Doherty et al. 1988; Jones 2012; Cheng and Walker 1998). In addition, it does not resemble a *trpE* mutant which is incapable of anthranilate biosynthesis. Anthranilate is an important intermediate of aromatic amino acid biosynthesis in the shikimate pathway (Barsomian et al. 1992). According to the literature and our experimental observations *exoR*, *exoY*, and *trpE* mutant strains are able to form CCRH and ITs to varying degrees but no CCRH or IT could be identified in the *tktA* mutant strains (Barsomian et al. 1992; Prasad et al. 2000; Cheng and Walker 1998; Leigh and Lee 1988). Moreover, the cyst like nodules produced by the *tktA* mutant strains do not contain viable bacterial cells. The microscopic analysis that was carried out did provide evidence of root hair deformation and clumps of *tktA* mutant could be

found on the plant roots. This suggests that the *tktA* mutant can colonize roots and does have a limited interaction with the root hairs which causes them to curl, but not colonize the curled root.

In the third chapter, the *tktB* locus characterized by generating mutations in *tktB* and neighbouring genes. Using a qRT-PCR analysis, it was established that *SMc02340* is a negative regulator of *tktB, SMc02341*, and *SMc02339. SMc02341* is a putative carbohydrate kinase that has a growth phenotype using ribose as a sole carbon source. It is noteworthy that it does grow on defined media and is therefore not an aromatic amino acid auxotroph. The auxotrophy therefore is completely associated with the loss of transketolase activity which can be provided by either TktA in the wild type, or by TktB in the suppressor strains. An amino acid sequence analysis of *SMc02341* using Phyre2 predicted a 3D structure with 100% confidence based on a gluconate kinase from *Lactobacillus acidophilus* which has a 24% amino acid sequence identity to *SMc02341* (Figure 35). The secondary structures are predicted to contain 43% alpha helix and 20% beta strands. The membrane topology has both N and C-terminals in the cytoplasm with two transmembrane helices (aa 227-256 and 285-300) (Kelley et al. 2015). Other similar proteins include *E. coli* xylulose kinase with 35% amino acid sequence identity. The sequence of *SMc02341* does show 35% sequence similarities to the predicted xylulose kinase (*xylB*) of *S. meliloti* 1021. Collectively these data strongly suggest that SMc02341 is a pentulose kinase, the nature of its true substrate is not presently known and difficult to predict.

It has also been determined that *tktB* and *SMc02341* are coordinately expressed. The co-ordinated expression would indicate that both *tktB* and *SMc02341* may be part of a catabolic pathway and work in conjunction with each other. One possible explanation could be that the true role of tktB may be a secondary route to catabolize ribose. A metabolic bypass in *S. meliloti* has been proposed since there is evidence to show that ribose and xylose can be metabolised in

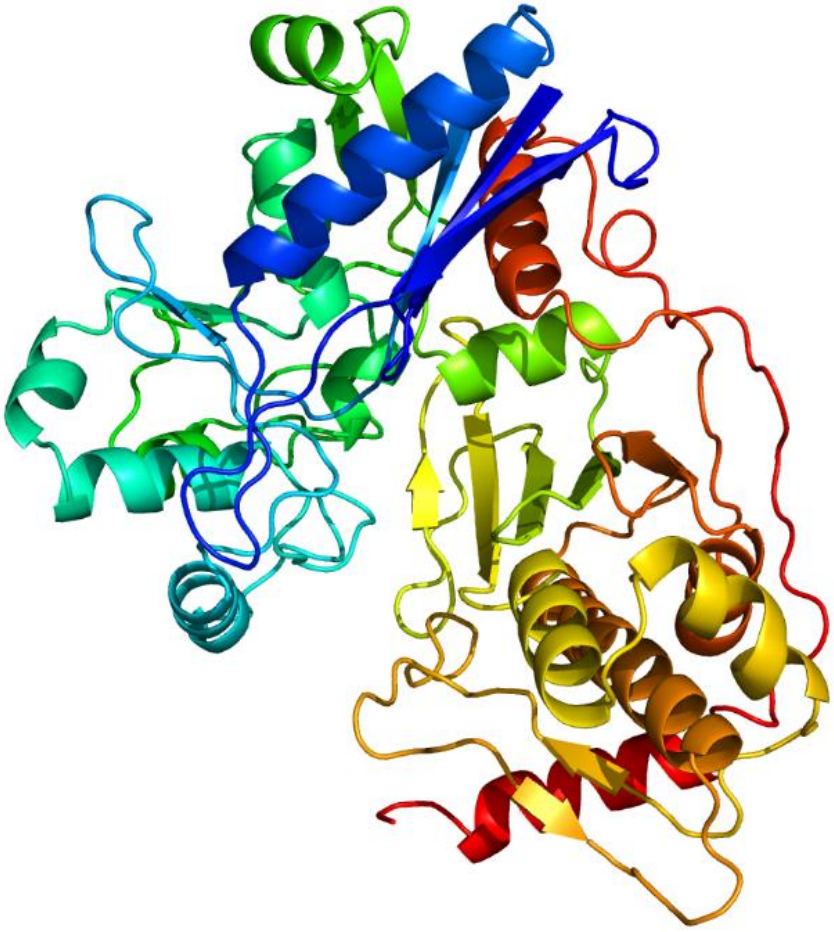


Figure 35. Predicted structure of SMc02341 based on a crystal of gluconate kinase of *Lactobacillus acidophilus*.

Model dimensions Aº: X, 70.552; Y, 50.838; Z, 59.470, Image colored by rainbow, N to C terminus. The template used to construct the model was c3gbtA_ with 100% confidence and 100% coverage. PDB header, transferase; PDB molecule, gluconate kinase; chain, A and 476 aligned residues. The model was constructed using Phyre2.

strains that have both incomplete gluconeogenic and glycolytic pathways because of mutations in *pyc*, *tpiA*, and *eryH* (*tpiB*).

Transketolase enzymes act by transferring a 2C molecule from a keto-sugar to an aldo-acceptor. There is evidence in the literature that transketolases are capable of carrying out only half reactions in the absence of an aldose sugar acceptor. These reactions can use xylulose-5P as a substrate and produce a 2C glycolaldehyde (Solovjeva et al. 2020). As discussed in chapter 3, *S. meliloti* carries the genetic potential to convert glycolaldehyde to malate via glyoxylate. We propose that *tktB* is capable of carrying out the half transketolase reaction as a part of the secondary metabolic bypass for ribose and xylose catabolism and *SMc02341* plays a role by phosphorylating pentoses.

In chapter 4, we provide evidence for a functional pyrophosphate dependent phosphofructokinase (Pfp) as a part of the EMP pathway. It is generally accepted that the EMP pathway in *S. meliloti* only functions in a gluconeogenic manner due to the lack of an ATP-dependent phosphofructokinase (Pfk) (Arias et al. 1979). Although the presence of *pfp* was pointed out when the complete sequence of *S. meliloti* was published, it had not been characterized and was not recognized either in review articles, or within the published metabolic models (Capela et al. 2001; Geddes and Oresnik 2014). In the current work, it was shown that Δ*pfp* mutation has a significant impact on central carbon metabolite pools. Accumulation of fructose-1,6biP in the Δ*pfp* mutant strain grown on either glucose or succinate strongly suggests that *pfp* functions as a part of the EMP pathway. Moreover, overexpression of the *pfp* open reading frame and purification of the encoded product followed by enzyme assays provided direct evidence that Pfp was able to carry out both forward and reverse reactions for converting fructose-6P to fructose-1,6biP and vice versa. A functional Pfp changes the way we look at

central carbon metabolism in *S. meliloti.* This would mean that the EMP pathway can play a glycolytic and gluconeogenic role; however, this does not change the fact that the ED pathway is the primary route for glycolysis. Furthermore, the Δ*pfp* strain was able to grow on both gluconeogenic and glycolytic sugars, suggesting that a metabolic bypass of this reaction exists. It was also discovered that the growth of the *pfp* mutant is dependent on transaldolase. The Δ*pfp*Δ*tal* strain was unable to grow on carbon substrates that are metabolised through the PP or ED pathway. We also see reduced nitrogen fixation when both *pfp* and *tal* are deleted, whereas the strains carrying either *pfp* or *tal* mutations do not affect symbiosis. The Δ*pfp*Δ*tal* strain demonstrates a link between Pfp and the PP pathway. In some cellulolytic Clostridia that lack a transaldolase, Pfp has been shown to act on sedoheptuloase-7P by adding one more phosphate to it. Then a fructose bisphosphate aldolase can break the sedoheptulose-1,7biP into 4C and 3C erythrose-4P and dihydroxy-acetone phosphate (DHAP), respectively (Koendjbiharie et al. 2020).

S. meliloti has a complex life cycle as it is exposed to different growth conditions and stresses. The *S. meliloti* genome contains a diversity of genes for carbohydrate catabolism and transport (Geddes and Oresnik 2014). There are genes that exhibit functional redundancy where the loss of one gene does not have drastic effect on growth. The *tktA* and *tktB* genes would fall under that category as loss of *tktB* does not affect the growth or symbiosis and loss of *tktA* can be partially complemented by upregulating *tktB.* The mutations in *pfp* or *tal* can also be thought as metabolically redundant as the Δ*pfp* or Δ*tal* strains can still grow on almost all carbon sources. But when more than one of these mutations are combined, they can exert a greater effect on the cell physiology. Perhaps these mutations impact the carbon flow to an extent that it affects nitrogen fixation. It has been suggested previously that reduced carbon flow to polysaccharide

biosynthesis can impact symbiosis (Hawkins et al. 2018), however, EPS mutations generally lead to an inability of the bacteria to develop into nitrogen fixing bacteroids, whereas the phenotype of the Δ*pfp*/Δ*tal* strain results in lower amounts of total nitrogen fixed based on plant dry weight data.

5.2 Future work

The aim of the second and third chapters was to identify the role of *tktA* and *tktB* in central carbon metabolism and symbiosis. Although both genes were metabolically and symbiotically characterized, we still do not know the biochemical contribution of *tktA* and *tktB* towards transketolase activity. A kinetic study of the biochemistry of these two transketolases could reveal significant information about their biochemical roles in *S. meliloti.* Additionally, the inability to make CCRH in the *tktA* mutant is a very interesting phenotype that to our knowledge has not been previously reported. More work in characterizing this very early symbiotic block could reveal new information about the initial interactions between rhizobia and its host.

A gene encoding for ribose kinase has not been identified in *S. meliloti*, although a putative ribokinase (*rbsK*) does exist. Biochemical characterization of *SMc02341* should reveal the substrate for this putative kinase. That information would be very useful in identifying the role of *tktB* operon in pentose metabolism and the potential secondary metabolic bypass. In addition, the *glc* operon (*SMc00832, SMc00833, and SMc00926*) has the genetic potential to convert glycolaldehyde to glyoxylate which can be subsequently used to synthesize malate. However, no functional characterization of this operon has been carried out yet. Therefore, characterization of the *glc* operon should reveal if a secondary metabolic bypass can connect the PP pathway to the TCA cycle.

We have provided evidence of a major missing link in the EMP pathway by characterizing the *pfp* in *S. meliloti.* Another gene that is important in completing the EMP pathway is a phosphatase that can convert fructose-1,6biP to fructose-6P. The gene *cbbF* has been putatively identified as encoding a phosphatase that plays this role in the CBB cycle. The operon containing all of the genes encoding for the enzymes of the CBB pathway however has not been reported to be expressed under normal growth conditions. The *S. meliloti* genome contains many putative phosphatases that are uncharacterized. The identity of the fructose 1,6 bisphosphatase is tentatively the only gluconeogenic enzyme which has not been identified. To truly understand central metabolic pathways in this organism future work should focus on identifying this phosphatase.

Literature Cited

Alexeyev, M. F. 1999. The pKNOCK series of broad-host-range mobilizable suicide vectors for gene knockout and targeted DNA insertion into the chromosome of gram-negative bacteria. Biotechniques. 26:824–6, 828

Amor, B. Ben, Shaw, S. L., Oldroyd, G. E. D., Maillet, F., Penmetsa, R. V., Cook, D., Long, S. R., Dénarié, J., and Gough, C. 2003. The NFP locus of *Medicago truncatula* controls an early step of Nod factor signal transduction upstream of a rapid calcium flux and root hair deformation. Plant J. 34:495–506

Anderson, A., and Cooper, R. A. 1969. Gluconeogenesis in *Escherichia coli* the role of triose phosphate isomerase. FEBS Lett. 4:19–20

Arias, A., and Cerveñansky, C. 1986. Galactose metabolism in *Rhizobium meliloti* L5-30. J. Bacteriol. 167:1092–4

Arias, A., Cerveńansky, C., Gardiol, A., and Martínez-Drets, G. 1979. Phosphoglucose isomerase mutant of *Rhizobium meliloti*. J. Bacteriol. 137:409–14

Baltscheffsky Herrick (Editor). 1996. *Origin and evolution of biological energy conversion*. New York, N.Y. : VCH.

Banfalvi, Z., and Kondorosi, A. 1989. Production of root hair deformation factors by *Rhizobium meliloti* nodulation genes in *Escherichia coli*: *HsnD* (*NodH*) is involved in the plant host-specific modification of the *NodABC* factor. Plant Mol. Biol. 13:1–12

Barbier, T., Collard, F., Zúñiga-Ripa, A., Moriyón, I., Godard, T., Becker, J., Wittmann, C., Van Schaftingen, E., and Letesson, J.-J. 2014. Erythritol feeds the pentose phosphate pathway via three new isomerases leading to D-erythrose-4-phosphate in *Brucella*. Proc. Natl. Acad. Sci. 111:17815–17820

Barnett, M. J., Fisher, R. F., Jones, T., Komp, C., Abola, A. P., Barloy-Hubler, F., Bowser, L., Capela, D., Galibert, F., Gouzy, J., Gurjal, M., Hong, A., Huizar, L., Hyman, R. W., Kahn, D., Kahn, M. L., Kalman, S., Keating, D. H., Palm, C., Peck, M. C., Surzycki, R., Wells, D. H., Yeh, K. C., Davis, R. W., Federspiel, N. A., and Long, S. R. 2001. Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti* pSymA megaplasmid. Proc. Natl. Acad. Sci. U. S. A. 98:9883–8

Barsomian, G. D., Urzainqui, A., Lohman, K., and Walker, G. C. 1992. *Rhizobium meliloti* mutants unable to synthesize anthranilate display a novel symbiotic phenotype. J. Bacteriol. 174:4416–26

Battisti, L., Lara, J. C., and Leigh, J. A. 1992. Specific oligosaccharide form of the *Rhizobium meliloti* exopolysaccharide promotes nodule invasion in alfalfa. Proc. Natl. Acad. Sci. U. S. A. 89:5625–5629

Brewster, J. D. 2003. A simple micro-growth assay for enumerating bacteria. J. Microbiol. Methods. 53:77–86

Capela, D., Barloy-Hubler, F., Gouzy, J., Bothe, G., Ampe, F., Batut, J., Boistard, P., Becker, A., Boutry, M., Cadieu, E., Dréano, S., Gloux, S., Godrie, T., Goffeau, A., Kahn, D., Kiss, E., Lelaure, V., Masuy, D., Pohl, T., Portetelle, D., Pühler, A., Purnelle, B., Ramsperger, U., Renard, C., Thébault, P., Vandenbol, M., Weidner, S., and Galibert, F. 2001. Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. Proc. Natl. Acad. Sci. U. S. A. 98:9877–82

Cassman, K. G., Dobermann, A., and Walters, D. T. 2002. Agroecosystems, nitrogen-use efficiency, and nitrogen management. MBio. 31:132–40

Cerveñanský, C., and Arias, A. 1984. Glucose-6-phosphate dehydrogenase deficiency in pleiotropic carbohydrate-negative mutant strains of *Rhizobium meliloti*. J. Bacteriol. 160:1027–30

Charles, T. C., and Finan, T. M. 1990. Genetic map of *Rhizobium meliloti* megaplasmid pRmeSU47b. J. Bacteriol. 172:2469–76

Chen, X., Schreiber, K., Appel, J., Makowka, A., Fähnrich, B., Roettger, M., Hajirezaei, M. R., Sönnichsen, F. D., Schönheit, P., Martin, W. F., and Gutekunst, K. 2016. The Entner–Doudoroff pathway is an overlooked glycolytic route in cyanobacteria and plants. Proc. Natl. Acad. Sci. 113:5441–5446

Cheng, H. P., and Walker, G. C. 1998. Succinoglycan is required for initiation and elongation of infection threads during nodulation of alfalfa by *Rhizobium meliloti*. J. Bacteriol. 180:5183–91

Chong, J., Wishart, D. S., and Xia, J. 2019. Using MetaboAnalyst 4.0 for Comprehensive and Integrative Metabolomics Data Analysis. Curr. Protoc. Bioinforma. 68

Daddaoua, A., Corral-Lugo, A., Ramos, J.-L., and Krell, T. 2017. Identification of GntR as regulator of the glucose metabolism in *Pseudomonas aeruginosa*. Environ. Microbiol. 19:3721–3733

Davies, J. M., Poole, R. J., and Sanders, D. 1993. The computed free energy change of hydrolysis of inorganic pyrophosphate and ATP: apparent significance. for inorganic-pyrophosphate-driven reactions of intermediary metabolism. Biochim. Biophys. Acta - Bioenerg. 1141:29–36

Davis, R., Roth, J., and Botstein, D. 1980. Advanced Bacterial Genetics: A Manual for Genetic Engineering.

diCenzo, G. C., Checcucci, A., Bazzicalupo, M., Mengoni, A., Viti, C., Dziewit, L., Finan, T. M., Galardini, M., and Fondi, M. 2016. Metabolic modelling reveals the specialization of secondary replicons for niche adaptation in *Sinorhizobium meliloti*. Nat. Commun. 7:12219

diCenzo, G. C., and Finan, T. M. 2017. The Divided Bacterial Genome: Structure, Function, and Evolution. Microbiol. Mol. Biol. Rev. 81:e00019-17

diCenzo, G. C., Muhammed, Z., Østerås, M., O'Brien, S. A. P., and Finan, T. M. 2017. A key regulator of the glycolytic and gluconeogenic central metabolic pathways in *Sinorhizobium*

meliloti. Genetics. 207:961–974

diCenzo, G. C., Tesi, M., Pfau, T., Mengoni, A., and Fondi, M. 2020. Genome-scale metabolic reconstruction of the symbiosis between a leguminous plant and a nitrogen-fixing bacterium. Nat. Commun. 11:2574

Ding, H., Yip, C. B., Geddes, B. A., Oresnik, I. J., and Hynes, M. F. 2012. Glycerol utilization by *Rhizobium leguminosarum* requires an ABC transporter and affects competition for nodulation. Microbiology. 158:1369–78

Ding, Y.-H. R., Ronimus, R. S., and Morgan, H. W. 2000. Sequencing, cloning, and high-level expression of the *pfp* gene, encoding a PPi-dependent phosphofructokinase from the extremely thermophilic eubacterium *Dictyoglomus thermophilum*. J. Bacteriol. 182:4661–4666

Doherty, D., Leigh, J. A., Glazebrook, J., and Walker, G. C. 1988. *Rhizobium meliloti* mutants that overproduce the *R. meliloti* acidic calcofluor-binding exopolysaccharide. J. Bacteriol. 170:4249–4256

Driscoll, B. T., and Finan, T. M. 1993. NAD(+)-dependent malic enzyme of *Rhizobium meliloti* is required for symbiotic nitrogen fixation. Mol. Microbiol. 7:865–73

Duncan, M. J. 1981. Properties of Tn*5*-induced carbohydrate mutants in *Rhizobium meliloti*. Microbiology. 122:61–67

Dunn, M. F. 1998. Tricarboxylic acid cycle and anaplerotic enzymes in rhizobia. FEMS Microbiol. Rev. 22:105–23

Dunn, M. F., Araíza, G., and Finan, T. M. 2001. Cloning and characterization of the pyruvate carboxylase from *Sinorhizobium meliloti* Rm1021. Arch. Microbiol. 176:355–63

Edwards, J. S., and Palsson, B. O. 2000. The *Escherichia coli* MG1655 *in silico* metabolic genotype: its definition, characteristics, and capabilities. Proc. Natl. Acad. Sci. U. S. A. 97:5528–33

Ehrhardt, D. W., Wais, R., and Long, S. R. 1996. Calcium spiking in plant root hairs responding to *Rhizobium* nodulation signals. Cell. 85:673–81

Engelke, T., Jording, D., Kapp, D., and Pühler, A. 1989. Identification and sequence analysis of the *Rhizobium meliloti dctA* gene encoding the C4-dicarboxylate carrier. J. Bacteriol. 171:5551–5560

Ertl, G. 1991. Elementary Steps in Ammonia Synthesis. Pages 109–132 in: Catalytic Ammonia Synthesis: Fundamentals and Practice, J.R. Jennings, ed. Springer US, Boston, MA.

Fillenberg, S. B., Friess, M. D., Körner, S., Böckmann, R. A., and Muller, Y. A. 2016. Crystal structures of the global regulator DasR from Streptomyces coelicolor: Implications for the allosteric regulation of GntR/HutC repressors. PLoS One. 11:e0157691

Finan, T. M., Elizabeth, M., Brian, D., and Robert, W. J. 1991. Complex symbiotic phenotypes

result from gluconeogenic mutations in *Rhizobium meliloti*. Mol. Plant-Microbe Interact. 4:386

Finan, T. M., Hartweig, E., LeMieux, K., Bergman, K., Walker, G. C., and Signer, E. R. 1984. General transduction in *Rhizobium meliloti*. J. Bacteriol. 159:120–124

Finan, T. M., Hirsch, A. M., Leigh, J. A., Johansen, E., Kuldau, G. A., Deegan, S., Walker, G. C., and Signer, E. R. 1985. Symbiotic mutants of *Rhizobium meliloti* that uncouple plant from bacterial differentiation. Cell. 40:869–877

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

Finan, T. M., Oresnik, I., and Bottacin, A. 1988. Mutants of *Rhizobium meliloti* defective in succinate metabolism. J. Bacteriol. 170:3396–403

Finan, T. M., Weidner, S., Wong, K., Buhrmester, J., Chain, P., Vorholter, F. J., Hernandez-Lucas, I., Becker, A., Cowie, A., Gouzy, J., Golding, B., and Puhler, A. 2001. The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endosymbiont *Sinorhizobium meliloti*. Proc. Natl. Acad. Sci. 98:9889–9894

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

Fisher, R. F., and Long, S. R. 1993. Interactions of NodD at the nod Box: NodD binds to two distinct sites on the same face of the helix and induces a bend in the DNA. J. Mol. Biol. 233:336–48

Fong, N. L., Lerman, J. A., Lam, I., Palsson, B. O., and Charusanti, P. 2013. Reconciling a *Salmonella enterica* metabolic model with experimental data confirms that overexpression of the glyoxylate shunt can rescue a lethal *ppc* deletion mutant. FEMS Microbiol. Lett. 342:62–9

Fowler, D., Coyle, M., Skiba, U., Sutton, M. a, Cape, J. N., Reis, S., Sheppard, L. J., Jenkins, A., Grizzetti, B., Galloway, J. N., Vitousek, P., Leach, A., Bouwman, A. F., Butterbach-Bahl, K., Dentener, F., Stevenson, D., Amann, M., and Voss, M. 2013. The global nitrogen cycle in the twenty-first century. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 368:20130164

Fry, J., Wood, M., and Poole, P. S. 2001. Investigation of myo -inositol catabolism in *Rhizobium leguminosarum* bv. *viciae* and its effect on nodulation competitiveness. Mol. Plant-Microbe Interact. 14:1016–1025

Fuhrer, T., Fischer, E., and Sauer, U. 2005. Experimental identification and quantification of glucose metabolism in seven bacterial species. J. Bacteriol. 187:1581–90

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

Gage, D. J., Bobo, T., and Long, S. R. 1996. Use of green fluorescent protein to visualize the

early events of symbiosis between *Rhizobium meliloti* and alfalfa (*Medicago sativa*). J. Bacteriol. 178:7159–66

Galardini, M., Mengoni, A., Biondi, E. G., Semeraro, R., Florio, A., Bazzicalupo, M., Benedetti, A., and Mocali, S. 2014. DuctApe: A suite for the analysis and correlation of genomic and OmniLog™ Phenotype Microarray data. Genomics. 103:1–10

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, J., 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., Thebault, P., Vandenbol, M., Vorholter, F. J., Weidner, S., Wells, D. H., Wong, K., Yeh, K. C., and Batut, J. 2001. The composite genome of the legume symbiont *Sinorhizobium meliloti*. Science. 293:668–72

Geddes, B. A., González, J. E., and Oresnik, I. J. 2014. Exopolysaccharide production in response to medium acidification is correlated with an increase in competition for nodule occupancy. Mol. Plant. Microbe. Interact. 27:1307–17

Geddes, B. A., Mendoza-Suárez, M. A., and Poole, P. S. 2018. A bacterial expression vector archive (BEVA) for flexible modular assembly of golden gate-compatible vectors. Front. Microbiol. 9:3345

Geddes, B. A., and Oresnik, I. J. 2012a. Genetic characterization of a complex locus necessary for the transport and catabolism of erythritol, adonitol and L-arabitol in *Sinorhizobium meliloti*. Microbiology. 158:2180–91

Geddes, B. A., and Oresnik, I. J. 2012b. Inability to catabolize galactose leads to increased ability to compete for nodule occupancy in *Sinorhizobium meliloti*. J. Bacteriol. 194:5044–5053

Geddes, B. A., and Oresnik, I. J. 2014. Physiology, genetics, and biochemistry of carbon metabolism in the alphaproteobacterium *Sinorhizobium meliloti*. Can. J. Microbiol. 60:491–507

Geddes, B. A., and Oresnik, I. J. 2016. The Mechanism of Symbiotic Nitrogen Fixation. Pages 69–97 in: The Mechanistic Benefits of Microbial Symbionts, C.J. Hurst, ed. Springer International Publishing, Cham.

Geddes, B. A., Pickering, B. S., Poysti, N. J., Collins, H., Yudistira, H., and Oresnik, I. J. 2010. A locus necessary for the transport and catabolism of erythritol in *Sinorhizobium meliloti*. Microbiology. 156:2970–81

Gibson, K. E., Kobayashi, H., and Walker, G. C. 2008. Molecular determinants of a symbiotic chronic infection. Annu. Rev. Genet. 42:413–41

Glazebrook, J., and Walker, G. C. 1991. Genetic techniques in *Rhizobium meliloti*. Methods Enzymol. 204:398–418

Glenn, S. A., Gurich, N., Feeney, M. A., and González, J. E. 2007. The *ExpR/Sin* quorum-sensing system controls succinoglycan production in *Sinorhizobium meliloti*. J. Bacteriol. 189:7077–7088

González, J. E., Reuhs, B. L., and Walker, G. C. 1996. Low molecular weight EPS II of *Rhizobium meliloti* allows nodule invasion in *Medicago sativa*. Proc. Natl. Acad. Sci. U. S. A. 93:8636–41

González, J. E., Semino, C. E., Wang, L. X., Castellano-Torres, L. E., and Walker, G. C. 1998. Biosynthetic control of molecular weight in the polymerization of the octasaccharide subunits of succinoglycan, a symbiotically important exopolysaccharide of *Rhizobium meliloti*. Proc. Natl. Acad. Sci. U. S. A. 95:13477–82

Gosselin, I., Wattraint, O., Riboul, D., Barbotin, J. N., and Portais, J. C. 2001. A deeper investigation on carbohydrate cycling in *Sinorhizobium meliloti*. FEBS Lett. 499:45–49

Green, R. T., East, A. K., Karunakaran, R., Downie, J. A., and Poole, P. S. 2019. Transcriptomic analysis of *Rhizobium leguminosarum* bacteroids in determinate and indeterminate nodules. Microb. Genomics. 5

Gunina, A., and Kuzyakov, Y. 2015. Sugars in soil and sweets for microorganisms: Review of origin, content, composition and fate. Soil Biol. Biochem. 90:87–100

Hamlen, R. A., Lukezic, F. L., and Bloom, J. R. 1972. Influence of age and stage of development on the neutral carbohydrate components in the root exudates from alfalfa plants grown in a gnotobiotic environment. Can. J. Plant Sci. 52:633–642

Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. 166:557–80

Hanahan, D., and Glover, D. M. 1985. *DNA cloning: a practical approach*.

Hanamura, A., and Aiba, H. 1991. Molecular mechanism of negative autoregulation of *Escherichia coli crp* gene. Nucleic Acids Res. 19:4413–4419

Harinarayanan, R., Murphy, H., and Cashel, M. 2008. Synthetic growth phenotypes of *Escherichia coli* lacking ppGpp and transketolase A (*tktA*) are due to ppGpp-mediated transcriptional regulation of *tktB*. Mol. Microbiol. 69:882–894

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

Hawkins, J. P., Geddes, B. A., and Oresnik, I. J. 2017a. Common dyes used to determine bacterial polysaccharides on agar are affected by medium acidification. Can. J. Microbiol. 63:559–562

Hawkins, J. P., Geddes, B. A., and Oresnik, I. J. 2017b. Succinoglycan production contributes to

acidic pH tolerance in *Sinorhizobium meliloti* Rm1021. Mol. Plant. Microbe. Interact. 30:1009–1019

Hawkins, J. P., Ordonez, P. A., and Oresnik, I. J. 2018. Characterization of mutations that affect the nonoxidative pentose phosphate pathway in *Sinorhizobium meliloti*. J. Bacteriol. 200

Heard, E., Tishkoff, S., Todd, J. A., Vidal, M., Wagner, G. P., Wang, J., Weigel, D., and Young, R. 2010. Ten years of genetics and genomics: what have we achieved and where are we heading? Nat. Rev. Genet. 11:723–33

Her, G. R., Glazebrook, J., Walker, G. C., and Reinhold, V. N. 1990. Structural studies of a novel exopolysaccharide produced by a mutant of *Rhizobium meliloti* strain Rm1021. Carbohydr. Res. 198:305–12

Herridge, D. F., Peoples, M. B., and Boddey, R. M. 2008. Global inputs of biological nitrogen fixation in agricultural systems. Plant Soil. 311:1–18

Hirsch, A. 1992. Developmental biology of legume nodulation. New Phytol. 122:211–237

Hoffman, B. M., Lukoyanov, D., Yang, Z.-Y., Dean, D. R., and Seefeldt, L. C. 2014. Mechanism of Nitrogen Fixation by Nitrogenase: The Next Stage. Chem. Rev. 114:4041–4062

Hofmann, E. 1978. Phosphofructokinase—a favourite of enzymologists and of students of metabolic regulation. Trends Biochem. Sci. 3:145–147

House, B. L., Mortimer, M. W., and Kahn, M. L. 2004. New Recombination Methods for *Sinorhizobium meliloti* Genetics. Appl. Environ. Microbiol. 70:2806–2815

Hütsch, B. W., Augustin, J., and Merbach, W. 2002. Plant rhizodeposition — an important source for carbon turnover in soils. J. Plant Nutr. Soil Sci. 165:397–407

Irigoyen, J. J., Sanchez-Diaz, M., and Emerich, D. W. 1990. Carbon metabolism enzymes of *Rhizobium meliloti* cultures and bacteroids and their distribution within alfalfa nodules. Appl. Environ. Microbiol. 56:2587–2589

Jacob, A. I., Adham, S. A. I., Capstick, D. S., Clark, S. R. D., Spence, T., and Charles, T. C. 2008. Mutational analysis of the *Sinorhizobium meliloti* short-chain dehydrogenase/reductase family reveals substantial contribution to symbiosis and catabolic diversity. Mol. Plant. Microbe. Interact. 21:979–87

Jiang, J., Gu, B. H., Albright, L. M., and Nixon, B. T. 1989. Conservation between coding and regulatory elements of *Rhizobium meliloti* and *Rhizobium leguminosarum dct* genes. J. Bacteriol. 171:5244–5253

Jimenez-Zurdo, J. I. 1995. Characterization of a *Rhizobium meliloti* proline dehydrogenase mutant altered in nodulation efficiency and competitiveness on alfalfa roots. Mol. Plant-Microbe Interact. 8:492

Jones, J. D., and Gutterson, N. 1987. An efficient mobilizable cosmid vector, pRK7813, and its use in a rapid method for marker exchange in *Pseudomonas fluorescens* strain HV37a.

Gene. 61:299–306

Jones, K. M. 2012. Increased production of the exopolysaccharide succinoglycan enhances *Sinorhizobium meliloti* 1021 symbiosis with the host plant *Medicago truncatula*. J. Bacteriol. 194:4322–4331

Jones, K. M., Kobayashi, H., Davies, B. W., Taga, M. E., and Walker, G. C. 2007a. How rhizobial symbionts invade plants: the *Sinorhizobium-Medicago* model. Nat. Rev. Microbiol. 5:619–33

Jones, K. M., Kobayashi, H., Davies, B. W., Taga, M. E., and Walker, G. C. 2007b. How rhizobial symbionts invade plants: the *Sinorhizobium–Medicago* model. Nat. Rev. Microbiol. 5:619–633

Jones, K. M., Sharopova, N., Lohar, D. P., Zhang, J. Q., VandenBosch, K. A., and Walker, G. C. 2008. Differential response of the plant *Medicago truncatula* to its symbiont *Sinorhizobium meliloti* or an exopolysaccharide-deficient mutant. Proc. Natl. Acad. Sci. 105:704–709

Josephson, B. L., and Fraenkel, D. G. 1974. Sugar metabolism in transketolase mutants of *Escherichia coli*. J. Bacteriol. 118:1082–9

Josephson, B. L., and Fraenkel, D. G. 1969. Transketolase mutants of *Escherichia coli*. J. Bacteriol. 100:1289–95

Jung, I. L., Phyo, K. H., and Kim, I. G. 2005. RpoS-mediated growth-dependent expression of the *Escherichia coli tkt* genes encoding transketolases isoenzymes. Curr. Microbiol. 50:314–318

Katznelson, H. 1955. Production of pyruvate from 6-Phospho-gluconate by bacterial plant pathogens and legume bacteria. Nature. 175:551–552

Kawaharada, Y., Kelly, S., Nielsen, M. W., Hjuler, C. T., Gysel, K., Muszyński, A., Carlson, R. W., Thygesen, M. B., Sandal, N., Asmussen, M. H., Vinther, M., Andersen, S. U., Krusell, L., Thirup, S., Jensen, K. J., Ronson, C. W., Blaise, M., Radutoiu, S., and Stougaard, J. 2015. Receptor-mediated exopolysaccharide perception controls bacterial infection. Nature. 523:308–12

Kay, J. E., and Jewett, M. C. 2020. A cell-free system for production of 2,3-butanediol is robust to growth-toxic compounds. Metab. Eng. Commun. 10:e00114

Keele, B. B., Hamilton, P. B., and Elkan, G. H. 1969. Glucose catabolism in *Rhizobium japonicum*. J. Bacteriol. 97:1184–91

Keller, M., Roxlau, A., Weng, W. M., Schmidt, M., Quandt, J., Niehaus, K., Jording, D., Arnold, W., and Pühler, A. 1995. Molecular analysis of the *Rhizobium meliloti mucR* gene regulating the biosynthesis of the exopolysaccharides succinoglycan and galactoglucan. Mol. Plant. Microbe. Interact. 8:267–77

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. 10:845–858

Khodayari, A., Zomorrodi, A. R., Liao, J. C., and Maranas, C. D. 2014. A kinetic model of *Escherichia coli* core metabolism satisfying multiple sets of mutant flux data. Metab. Eng. 25:50–62

Koendjbiharie, J. G., Hon, S., Pabst, M., Hooftman, R., Stevenson, D. M., Cui, J., Amador-Noguez, D., Lynd, L. R., Olson, D. G., and van Kranenburg, R. 2020. The pentose phosphate pathway of cellulolytic clostridia relies on 6-phosphofructokinase instead of transaldolase. J. Biol. Chem. 295:1867–1878

Kohler, P. R. A., Zheng, J. Y., Schoffers, E., and Rossbach, S. 2010. Inositol catabolism, a key pathway in *Sinorhizobium meliloti* for competitive host nodulation. Appl. Environ. Microbiol. 76:7972–80

Kohlmeier, M. G. 2019. Sugar alcohol metabolism in the legume endosymbiont *Sinorhizobium meliloti*.

Kohlmeier, M. G., Bailey-Elkin, B. A., Mark, B. L., and Oresnik, I. J. 2021. Characterization of the sorbitol dehydrogenase SmoS from *Sinorhizobium meliloti* 1021. Acta Crystallogr. Sect. D Struct. Biol. 77:380–390

Kohlmeier, M. G., White, C. E., Fowler, J. E., Finan, T. M., and Oresnik, I. J. 2019. Galactitol catabolism in *Sinorhizobium meliloti* is dependent on a chromosomally encoded sorbitol dehydrogenase and a pSymB-encoded operon necessary for tagatose catabolism. Mol. Genet. Genomics. 294:739–755

Krol, E., and Becker, A. 2004. Global transcriptional analysis of the phosphate starvation response in *Sinorhizobium meliloti* strains 1021 and 2011. Mol. Genet. Genomics. 272:1–17

Kurumbang, N. P., Dvorak, P., Bendl, J., Brezovsky, J., Prokop, Z., and Damborsky, J. 2014. Computer-assisted engineering of the synthetic pathway for biodegradation of a toxic persistent pollutant. ACS Synth. Biol. 3:172–181

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 227:680–685

Leigh, J. A., and Lee, C. C. 1988. Characterization of polysaccharides of *Rhizobium meliloti exo* mutants that form ineffective nodules. J. Bacteriol. 170:3327–32

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

Livak, K. J., and Schmittgen, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 25:402–8

Long, C. P., Gonzalez, J. E., Feist, A. M., Palsson, B. O., and Antoniewicz, M. R. 2018. Dissecting the genetic and metabolic mechanisms of adaptation to the knockout of a major metabolic enzyme in *Escherichia coli*. Proc. Natl. Acad. Sci. U. S. A. 115:222–227

Luo, L., Yao, S.-Y., Becker, A., Ruberg, S., Yu, G.-Q., Zhu, J.-B., and Cheng, H.-P. 2005. Two

new *Sinorhizobium meliloti* LysR-type transcriptional regulators required for nodulation. J. Bacteriol. 187:4562–4572

Maillet, F., Fournier, J., Mendis, H. C., Tadege, M., Wen, J., Ratet, P., Mysore, K. S., Gough, C., and Jones, K. M. 2020. *Sinorhizobium meliloti* succinylated high-molecular-weight succinoglycan and the *Medicago truncatula* LysM receptor-like kinase MtLYK10 participate independently in symbiotic infection. Plant J. 102:311–326

Martínez-De Drets, G., and Arias, A. 1972. Enzymatic basis for differentiation of *Rhizobium* into fast- and slow-growing groups. J. Bacteriol. 109:467–70

Meade, H. M., Long, S. R., Ruvkun, G. B., Brown, S. E., and Ausubel, F. M. 1982. Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn*5* mutagenesis. J. Bacteriol. 149:114–122

Mendis, H. C., Madzima, T. F., Queiroux, C., and Jones, K. M. 2016. Function of succinoglycan polysaccharide in *Sinorhizobium meliloti* host plant invasion depends on succinylation, not molecular weight. MBio. 7:e00606–e00616

Mendrygal, K. E., and González, J. E. 2000. Environmental regulation of exopolysaccharide production in *Sinorhizobium meliloti*. J. Bacteriol. 182:599–606

Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., Mausset, A.-E., Barloy-Hubler, F., Galibert, F., Kondorosi, A., and Kondorosi, E. 2006. Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*-legume symbiosis. Proc. Natl. Acad. Sci. U. S. A. 103:5230–5

Milunovic, B., DiCenzo, G. C., Morton, R. A., and Finan, T. M. 2014. Cell growth inhibition upon deletion of four toxin-antitoxin loci from the megaplasmids of *Sinorhizobium meliloti*. J. Bacteriol. 196:811–24

Mitra, R. M., Gleason, C. a, Edwards, A., Hadfield, J., Downie, J. A., Oldroyd, G. E. D., and Long, S. R. 2004. A Ca2+/calmodulin-dependent protein kinase required for symbiotic nodule development: Gene identification by transcript-based cloning. Proc. Natl. Acad. Sci. U. S. A. 101:4701–4705

Moffatt, B. A., and Ashihara, H. 2002. Purine and pyrimidine nucleotide synthesis and metabolism. Arab. B. 1:e0018

Moffatt, B., and Somerville, C. 1988. Positive selection for male-sterile mutants of *Arabidopsis* lacking adenine phosphoribosyl transferase activity. Plant Physiol. 86:1150–1154

Morris, D. L. 1948. Quantitative Determination of Carbohydrates With Dreywood's Anthrone Reagent. Science. 107:254–5

Mulukutla, B. C., Yongky, A., Daoutidis, P., and Hu, W.-S. 2014. Bistability in glycolysis pathway as a physiological switch in energy metabolism. PLoS One. 9:e98756

Nagpal, P., Khanuja, S. P. S., and Stanfield, S. W. 1992. Suppression of the *ndv* mutant phenotype of *Rhizobium meliloti* by cloned *exo* genes. Mol. Microbiol. 6:479–488

Newman, J. D., Schultz, B. W., and Noel, K. D. 1992. Dissection of nodule development by supplementation of *Rhizobium leguminosarum* biovar *phaseoli* purine auxotrophs with 4-aminoimidazole-5-carboxamide riboside. Plant Physiol. 99:401–408

O'Brien, E. J., Monk, J. M., and Palsson, B. O. 2015. Using genome-scale models to predict biological capabilities. Cell. 161:971–987

Oldroyd, G. E. D. 2013. Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. Nat. Rev. Microbiol. 11:252–63

Oresnik, I. J., Charles, T. C., and Finan, T. M. 1994. Second site mutations specifically suppress the Fix- phenotype of *Rhizobium meliloti ndvF* mutations on alfalfa: identification of a conditional *ndvF*-dependent mucoid colony phenotype. Genetics. 136:1233–43

Oresnik, I. J., Pacarynuk, L. A., O'Brien, S. A. P., Yost, C. K., and Hynes, M. F. 1998. Plasmid-encoded catabolic genes in *Rhizobium leguminosarum* bv. *trifolii*: evidence for a plant-inducible rhamnose locus involved in competition for nodulation. Mol. Plant-Microbe Interact. 11:1175–1185

Osterås, M., Driscoll, B. T., and Finan, T. M. 1997. Increased pyruvate orthophosphate dikinase activity results in an alternative gluconeogenic pathway in *Rhizobium* (*Sinorhizobium*) *meliloti*. Microbiology. 143:1639–48

Osterås, M., Driscoll, B. T., and Finan, T. M. 1995. Molecular and expression analysis of the *Rhizobium meliloti* phosphoenolpyruvate carboxykinase (*pckA*) gene. J. Bacteriol. 177:1452–60

Pain, A. N. 1979. Symbiotic properties of antibiotic-resistant and auxotrophic mutants of *Rhizobium leguminosarum*. J. Appl. Bacteriol. 47:53–64

Panda, J., and Sarkar, P. 2012. Bioremediation of chromium by novel strains *Enterobacter aerogenes* T2 and *Acinetobacter* sp. PD 12 S2. Environ. Sci. Pollut. Res. Int. 19:1809–17

Pankhurst, C. E., and Schwinghamer, E. A. 1974. Adenine requirement for nodulation of pea by an auxotrophic mutant of *Rhizobium leguminosarum*. Arch. Microbiol. 100:219–238

Park, J. O., Tanner, L. B., Wei, M. H., Khana, D. B., Jacobson, T. B., Zhang, Z., Rubin, S. A., Li, S. H.-J., Higgins, M. B., Stevenson, D. M., Amador-Noguez, D., and Rabinowitz, J. D. 2019. Near-equilibrium glycolysis supports metabolic homeostasis and energy yield. Nat. Chem. Biol. 15:1001–1008

Parke, D., and Ornston, L. N. 1984. Nutritional diversity of Rhizobiaceae revealed by auxanography. Microbiology. 130:1743–1750

Pellicer, M. T., Badía, J., Aguilar, J., and Baldomà, L. 1996. *glc* locus of *Escherichia coli*: characterization of genes encoding the subunits of glycolate oxidase and the *glc* regulator protein. J. Bacteriol. 178:2051–9

Peters, N. K., Frost, J. W., and Long, S. R. 1986. A plant flavone, luteolin, induces expression of *Rhizobium meliloti* nodulation genes. Science. 233:977–80

Peters, N. K., and Long, S. R. 1988. Alfalfa root exudates and compounds which promote or inhibit induction of *Rhizobium meliloti* nodulation genes. Plant Physiol. 88:396–400

Pickering, B. S., and Oresnik, I. J. 2008. Formate-dependent autotrophic growth in *Sinorhizobium meliloti*. J. Bacteriol. 190:6409–6418

Pini, F., East, A. K., Appia-Ayme, C., Tomek, J., Karunakaran, R., Mendoza-Suárez, M., Edwards, A., Terpolilli, J. J., Roworth, J., Downie, J. A., and Poole, P. S. 2017. Bacterial biosensors for *in vivo* spatiotemporal mapping of root Secretion. Plant Physiol. 174:1289–1306

Platt, R., Drescher, C., Park, S. K., and Phillips, G. J. 2000. Genetic system for reversible integration of DNA constructs and lacZ gene fusions into the *Escherichia coli* chromosome. Plasmid. 43:12–23

Poole, P., Ramachandran, V., and Terpolilli, J. 2018. Rhizobia: from saprophytes to endosymbionts. Nat. Rev. Microbiol. :1–13

Poysti, N. J., Loewen, E. D. M., Wang, Z., and Oresnik, I. J. 2007. *Sinorhizobium meliloti* pSymB carries genes necessary for arabinose transport and catabolism. Microbiology. 153:727–736

Poysti, N. J., and Oresnik, I. J. 2007. Characterization of *Sinorhizobium meliloti* triose phosphate isomerase genes. J. Bacteriol. 189:3445–3451

Prasad, C. K., Vineetha, K. E., Hassani, R., Gupta, R., and Randhawa, G. S. 2000. Isolation and symbiotic characterization of aromatic amino acid auxotrophs of *Sinorhizobium meliloti*. Indian J. Exp. Biol. 38:1041–1049

Price, M. N., Wetmore, K. M., Waters, R. J., Callaghan, M., Ray, J., Liu, H., Kuehl, J. V., Melnyk, R. A., Lamson, J. S., Suh, Y., Carlson, H. K., Esquivel, Z., Sadeeshkumar, H., Chakraborty, R., Zane, G. M., Rubin, B. E., Wall, J. D., Visel, A., Bristow, J., Blow, M. J., Arkin, A. P., and Deutschbauer, A. M. 2018. Mutant phenotypes for thousands of bacterial genes of unknown function. Nature. 557:503–509

Quandt, J., and Hynes, M. F. 1993. Versatile suicide vectors which allow direct selection for gene replacement in gram-negative bacteria. Gene. 127:15–21

Radutoiu, S., Madsen, L. H., Madsen, E. B., Felle, H. H., Umehara, Y., Grønlund, M., Sato, S., Nakamura, Y., Tabata, S., Sandal, N., and Stougaard, J. 2003. Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. Nature. 425:585–592

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:R106

Rattray, N. J. W., Deziel, N. C., Wallach, J. D., Khan, S. A., Vasiliou, V., Ioannidis, J. P. A., and Johnson, C. H. 2018. Beyond genomics: understanding exposotypes through metabolomics. Hum. Genomics. 12:4

Reed, J. L., Famili, I., Thiele, I., and Palsson, B. O. 2006. Towards multidimensional genome annotation. Nat. Rev. Genet. 7:130–41

Reeves, R. E., South, D. J., Blytt, H. J., and Warren, L. G. 1974. Pyrophosphate:D-fructose 6-phosphate 1-phosphotransferase. A new enzyme with the glycolytic function of 6-phosphofructokinase. J. Biol. Chem. 249:7737–41

Reinhold, B. B., Chan, S. Y., Reuber, T. L., Marra, A., Walker, G. C., and Reinhold, V. N. 1994. Detailed structural characterization of succinoglycan, the major exopolysaccharide of *Rhizobium meliloti* Rm1021. J. Bacteriol. 176:1997–2002

Resendis-Antonio, O., Reed, J. L., Encarnación, S., Collado-Vides, J., and Palsson, B. Ø. 2007. Metabolic reconstruction and modeling of nitrogen fixation in *Rhizobium etli*. PLoS Comput. Biol. 3:1887–95

Reshetnikov, A. S., Rozova, O. N., Khmelenina, V. N., Mustakhimov, I. I., Beschastny, A. P., Murrell, J. C., and Trotsenko, Y. A. 2008. Characterization of the pyrophosphate-dependent 6-phosphofructokinase from *Methylococcus capsulatus* Bath. FEMS Microbiol. Lett. 288:202–210

Richardson, J. S., Carpena, X., Switala, J., Perez-Luque, R., Donald, L. J., Loewen, P. C., and Oresnik, I. J. 2008. RhaU of *Rhizobium leguminosarum* is a rhamnose mutarotase. J. Bacteriol. 190:2903–10

Richardson, J. S., Hynes, M. F., and Oresnik, I. J. 2004. A genetic locus necessary for rhamnose uptake and catabolism in *Rhizobium leguminosarum* bv. *trifolii*. J. Bacteriol. 186:8433–8442

Robinson, J. L., Kocabaş, P., Wang, H., Cholley, P.-E., Cook, D., Nilsson, A., Anton, M., Ferreira, R., Domenzain, I., Billa, V., Limeta, A., Hedin, A., Gustafsson, J., Kerkhoven, E. J., Svensson, L. T., Palsson, B. O., Mardinoglu, A., Hansson, L., Uhlén, M., and Nielsen, J. 2020. An atlas of human metabolism. Sci. Signal. 13

Rosenblueth, M., Hynes, M. F., and Martínez-Romero, E. 1998. *Rhizobium tropiciteu* genes involved in specific uptake of *Phaseolus vulgaris* bean-exudate compounds. Mol. Gen. Genet. MGG. 258:587–598

Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989. *Molecular cloning: a laboratory manual.* Cold Spring Harbor Laboratory Press.

Scherrer, A., and Dénarié, J. 1971. Symbiotic properties of some auxotrophic mutants of *Rhizobium meliloti* and of their prototrophic revertants. Plant Soil. 35:39–45

Shaw, J. A., Henard, C. A., Liu, L., Dieckman, L. M., Vázquez-Torres, A., and Bourret, T. J. 2018. *Salmonella enterica* serovar Typhimurium has three transketolase enzymes contributing to the pentose phosphate pathway. J. Biol. Chem. 293:11271–11282

Shaw, S. L., and Long, S. R. 2003. Nod factor elicits two separable calcium responses in *Medicago truncatula* root hair cells. Plant Physiol. 131:976–84

Siebers, B., and Schönheit, P. 2005. Unusual pathways and enzymes of central carbohydrate metabolism in Archaea. Curr. Opin. Microbiol. 8:695–705

Soberón, M., Lopez, O., Miranda, J., Tabche, M. L., and Morera, C. 1997. Genetic evidence for 5-aminoimidazole-4-carboxamide ribonucleotide (AICAR) as a negative effector of cytochrome terminal oxidase cbb3 production in *Rhizobium etli*. Mol. Gen. Genet. 254:665–73

Soberón, M., Morera, C., Kondorosi, A., Lopez, O., and Miranda, J. 2001. A purine-related metabolite negatively regulates *fixNOQP* expression in *Sinorhizobium meliloti* by modulation of *fixK* expression. Mol. Plant. Microbe. Interact. 14:572–6

Solovjeva, O. N. 2021. The mechanism of a one-substrate transketolase reaction. Part II. Anal. Biochem. 613:114022

Solovjeva, O. N., Kovina, M. V., Zavialova, M. G., Zgoda, V. G., Shcherbinin, D. S., and Kochetov, G. A. 2020. The mechanism of a one-substrate transketolase reaction. Biosci. Rep. 40

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

Stowers, M. D., and Elkan, G. H. 1983. The transport and metabolism of glucose in cowpea rhizobia. Can. J. Microbiol. 29:398–406

Suvorova, I. A., Korostelev, Y. D., and Gelfand, M. S. 2015. GntR Family of bacterial transcription factors and their DNA binding motifs: structure, positioning and co-evolution. PLoS One. 10:e0132618

Swamynathan, S. K., and Singh, A. 1995. Pleiotropic effects of purine auxotrophy in *Rhizobium meliloti* on cell surface molecules. J. Biosci. 20:17–28

Swamynathan, S. K., and Singh, A. 1992. *Rhizobium meliloti* purine auxotrophs are *nod* + but defective in nitrogen fixation. J. Genet. 71:11–21

Swayambhu, G., Moscatello, N., Atilla-Gokcumen, G. E., and Pfeifer, B. A. 2020. Flux balance analysis for media optimization and genetic targets to improve heterologous siderophore production. iScience. 23:101016

Thiele, I., and Palsson, B. Ø. 2010. A protocol for generating a high-quality genome-scale metabolic reconstruction. Nat. Protoc. 5:93–121

Timmers, A. C., 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. 126:3617–28

Trinick, M. J. 1973. Symbiosis between Rhizobium and the Non-legume, *Trema aspera*. Nature. 244:459–460

Triplett, E. W. 1990. Construction of a symbiotically effective strain of *Rhizobium*

leguminosarum bv. *trifolii* with increased nodulation competitiveness. Appl. Environ. Microbiol. 56:98–103

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

Vanderlinde, E. M., Hynes, M. F., and Yost, C. K. 2014. Homoserine catabolism by *Rhizobium leguminosarum* bv. *viciae* 3841 requires a plasmid-borne gene cluster that also affects competitiveness for nodulation. Environ. Microbiol. 16:205–217

Vegge, C. S., Jansen van Rensburg, M. J., Rasmussen, J. J., Maiden, M. C. J., Johnsen, L. G., Danielsen, M., MacIntyre, S., Ingmer, H., and Kelly, D. J. 2016. Glucose Metabolism via the Entner-Doudoroff Pathway in *Campylobacter*: A rare trait that enhances survival and promotes biofilm formation in some isolates. Front. Microbiol. 7

Van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., Farkas, A., Mikulass, K., Nagy, A., Tiricz, H., Satiat-Jeunemaître, B., Alunni, B., Bourge, M., Kucho, K., Abe, M., Kereszt, A., Maroti, G., Uchiumi, T., Kondorosi, E., and Mergaert, P. 2010. Plant peptides govern terminal differentiation of bacteria in symbiosis. Science. 327:1122–6

Vincent, J. M. 1970. *A manual for the practical study of the root-nodule bacteria.* IBP Handbk 15 Oxford and Edinburgh: Blackwell Scientific Publications.

Vincent, J. M. 1941. Serological properties of the root-nodule bacteria. Strains of *Rhizobium meliloti*. Proc. Linn. Soc. NSW. 66:145–154

Vitousek, P. M., Cassman, K., Cleveland, C., Crews, T., Field, C. B., Grimm, N. B., Howarth, R. W., Marino, R., Martinelli, L., Rastetter, E. B., and Sprent, J. I. 2002. Towards an ecological understanding of biological nitrogen fixation. Biogeochemistry. 57:1–45

Wang, L. X., Wang, Y., Pellock, B., and Walker, G. C. 1999. Structural characterization of the symbiotically important low-molecular-weight succinoglycan of *Sinorhizobium meliloti*. J. Bacteriol. 181:6788–6796

Watson, R. J. 1990. Analysis of the C4-dicarboxylate transport genes of *Rhizobium meliloti*: nucleotide sequence and deduced products of dctA, dctB, and dctD. Mol. Plant. Microbe. Interact. 3:174–81

Wheatley, R. M., Ford, B. L., Li, L., Aroney, S. T. N., Knights, H. E., Ledermann, R., East, A. K., Ramachandran, V. K., and Poole, P. S. 2020. Lifestyle adaptations of *Rhizobium* from rhizosphere to symbiosis. Proc. Natl. Acad. Sci. U. S. A. 117:23823–23834

Yamada, M., Adachi, K., Ogawa, N., Kishino, S., Ogawa, J., Kataoka, M., Shimizu, S., and Isobe, K. 2015. A new aldehyde oxidase catalyzing the conversion of glycolaldehyde to glycolate from *Burkholderia* sp. AIU 129. J. Biosci. Bioeng. 119:410–415

Yao, S.-Y., Luo, L., Har, K. J., Becker, A., Rüberg, S., Yu, G.-Q., Zhu, J.-B., and Cheng, H.-P. 2004. *Sinorhizobium meliloti* ExoR and ExoS proteins regulate both succinoglycan and flagellum production. J. Bacteriol. 186:6042–6049

Yarosh, O. K., Charles, T. C., and Finan, T. M. 1989. Analysis of C4 -dicarboxylate transport genes in *Rhizobium meliloti*. Mol. Microbiol. 3:813–823

York, G. M., and Walker, G. C. 1998. The succinyl and acetyl modifications of succinoglycan influence susceptibility of succinoglycan to cleavage by the *Rhizobium meliloti* glycanases ExoK and ExsH. J. Bacteriol. 180:4184–91

Yost, C. K., Rath, A. M., Noel, T. C., and Hynes, M. F. 2006. Characterization of genes involved in erythritol catabolism in *Rhizobium leguminosarum* bv. *viciae*. Microbiology. 152:2061–2074

Yurgel, S. N., and Kahn, M. L. 2004. Dicarboxylate transport by rhizobia. FEMS Microbiol. Rev. 28:489–501

Zhang, Y., Morar, M., and Ealick, S. E. 2008. Structural biology of the purine biosynthetic pathway. Cell. Mol. Life Sci. 65:3699–724

Zhao, G., and Winkler, M. E. 1994. An *Escherichia coli* K-12 *tktA tktB* mutant deficient in transketolase activity requires pyridoxine (vitamin B6) as well as the aromatic amino acids and vitamins for growth. J. Bacteriol. 176:6134–6138

Zhao, H., Li, M., Fang, K., Chen, W., and Wang, J. 2012. *In silico* insights into the symbiotic nitrogen fixation in *Sinorhizobium meliloti* via metabolic reconstruction. PLoS One. 7:e31287

Zhou, J., Olson, D. G., Argyros, D. A., Deng, Y., van Gulik, W. M., van Dijken, J. P., and Lynd, L. R. 2013. Atypical glycolysis in *Clostridium thermocellum*. Appl. Environ. Microbiol. 79:3000–3008

Zipfel, C., and Oldroyd, G. E. D. 2017. Plant signalling in symbiosis and immunity. Nature. 543:328–336